

Article

An Ensemble of Condition Based Classifiers for Device Independent Detailed Human Activity Recognition Using Smartphones [†]

Jayita Saha ¹ , Chandreyee Chowdhury ^{1,*} , Ishan Roy Chowdhury ¹ , Suparna Biswas ² 
and Nauman Aslam ³

¹ Department of Computer Science and Engineering, Jadavpur University, Kolkata 700032, India; gjai.2000@gmail.com (J.S.); ishan.r.chowdhury@gmail.com (I.R.C.)

² Department of Computer Science and Engineering, Maulana Abul Kalam Azad University of Technology, Kolkata 700064, India; mailtosuparna@gmail.com

³ Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, UK; nauman.aslam@northumbria.ac.uk

* Correspondence: chandreyee.chowdhury@gmail.com

[†] This paper is an extended version of our paper published in Proceedings of e-Health Pervasive Wireless Applications and Services (e-HPWAS'17), Rome, Italy, 9 October 2017.

Received: 25 January 2018; Accepted: 12 April 2018; Published: 16 April 2018



Abstract: Human activity recognition is increasingly used for medical, surveillance and entertainment applications. For better monitoring, these applications require identification of detailed activity like *sitting on chair/floor, brisk/slow walking, running*, etc. This paper proposes a ubiquitous solution to detailed activity recognition through the use of smartphone sensors. Use of smartphones for activity recognition poses challenges such as device independence and various usage behavior in terms of where the smartphone is kept. Only a few works address one or more of these challenges. Consequently, in this paper, we present a detailed activity recognition framework for identifying both static and dynamic activities addressing the above-mentioned challenges. The framework supports cases where (i) dataset contains data from accelerometer; and the (ii) dataset contains data from both accelerometer and gyroscope sensor of smartphones. The framework forms an ensemble of the condition based classifiers to address the variance due to different hardware configuration and usage behavior in terms of where the smartphone is kept (right pants pocket, shirt pockets or right hand). The framework is implemented and tested on real data set collected from 10 users with five different device configurations. It is observed that, with our proposed approach, 94% recognition accuracy can be achieved.

Keywords: human activity recognition; detailed activity; ensemble; device independence; smartphones

1. Introduction

Human physical activity refers to any body movement produced by skeletal muscles or different position of the limbs with respect to time upstanding against gravity that results in an energy expenditure [1,2]. Activity recognition and monitoring system concurrently identifies, evaluates the actions carried out by a person on a daily basis in real conditions of the surrounding environment and provides context aware feedback for healthcare and elder care. Daily activity is a complex concept; it depends on many factors, including physiological, anatomical, psychological, and environmental effects. Human daily activity tracking was traditionally solved by an image processing approach and vision-based techniques [3,4]. However, these techniques may violate user privacy, mostly require infrastructure support like installing video cameras in the monitoring areas, and depend heavily on

lighting conditions. Several works consider wearable sensors individually and combined with ambient sensor for activity recognition [5–7]. Many of the early efforts focused on detecting fall and daily-life activities, mainly using one/or more wearable accelerometers. However, it may not be convenient for patients to carry out daily activities with sensors worn in hands and/or limbs. However, inertial sensors of smartphones can be a convenient option for activity recognition as most users almost always carry smartphones. Most smartphones are equipped with accelerometer, gyroscope, compass and proximity sensors. In addition, these devices also have communication facilities like Wi-Fi and Bluetooth by which sensor readings can be transferred to a server. Continuous raw data are collected from several sensors of the smartphone during monitoring. The data are processed for extracting useful features, fed to some classification algorithm for training an appropriate activity model, in order to recognize a variety of activities.

Daily activities can be categorized in two ways: coarse-grained or simple activity and fine-grained or detailed activity. Coarse-grained (*Sit, Stand, Walk*, etc.) is simplified larger sub-component of basic activity, whereas fine-grained, that is, detailed activity, refers to smaller distinguishable subcomponents that can be composed together to get a coarse-grained activity. Fine-grained or detailed activity contains activities like *Sit on floor/chair*. Identifying detailed activity can be beneficial for many medical applications such as elderly assistance at home, post trauma rehabilitation after a surgery, detection of gestures, motions and fitness of diabetic patients, etc. The elderly population has increased significantly who are living alone and mostly suffering from chronic diseases. Stroke patients need assistance and require regular monitoring during rehabilitation. Increased walking ability is the focus of rehabilitation. Accurate information on daily activity has the potential to improve the regular monitoring and treatment in several diseases and sometimes it reduces high burden of hospitalization costs.

Existing works [8–10] mostly focus on coarse grained activity recognition. Few works could be found on detailed activity [11,12] using several inertial sensors that may not be present in many smartphone configurations, thus the system is not ubiquitous. In literature, most of the works [9,13,14] consider a single classifier approach to study activity recognition system with smartphones. In [15,16], the authors use an ensemble learning technique for activity recognition. In real life, the training and testing environment for activity recognition are not same all the time. Generally, raw data from several devices are collected to monitor activities. Due to several hardware configuration and calibration problem, sensor readings vary from one device to another. Even orientation of the smartphone, which depends on the usage behavior with respect to human subjects, is also a factor affecting classification accuracy of the system. Several sensors are used for activity recognition in order to make the recognition system device independent in [17]. A recent work in [18] addresses different usage behavior like smartphones kept at a coat pocket or bag. However, no work could be found that enables detailed activity recognition even when training data is collected using one device at one position (say, right trouser pocket) and activity is recognized for test data collected from a different device kept at the same or different position (say, shirt pocket). Consequently, our main contribution in this work are as follows:

- (a) We propose an activity recognition framework using an ensemble of condition based classifiers to identify detailed static activities (*Sit on chair, Sit on floor, Lying right, Lying left*) as well as detailed dynamic activities (*Slow walk, Brisk walk*). The proposed technique works irrespective of device configuration and usage behavior.
- (b) The process utilizes accelerometer and gyroscope sensor of smartphones that are available in almost all smartphones by most of the manufacturers, thus making the framework ubiquitous.
- (c) The proposed technique can identify the effect of accelerometer and gyroscope for identifying individual detailed activity.

The rest of this paper is organized as follows: Section 2 describes state-of-the-art techniques for activity recognition through smartphones. Definition of the problem is discussed in Section 3. Design of the proposed system is detailed in Section 4. Section 5 describes the experimental setup and summarizes the results. Finally, we conclude in Section 6.

2. Related Work

A typical activity identification framework mostly follows four phases including (1) Data collection and preprocessing; (2) Feature extraction; (3) Feature selection and (4) Classification as shown in Figure 1. Data are collected through several sensors with respect to human usage and behavior. Preprocessed data are sent to a server for further processing. Several time and frequency domain features are extracted and selected from preprocessed data. The classification techniques are applied on the server side to recognize an activity.

Data are collected from wearable sensors and smart handhelds. One of the most important issues in data collection is the selection of sensors and the attributes to be measured, which play an important role in the activity recognition system's performance. Incorrect selection of sensors may adversely affect the recognition performance. Inertial sensors such as accelerometers and gyroscopes are used for activity recognition. The accelerometer measures non-gravitational acceleration of a smart device while the gyroscope senses the rate of change of orientation or angular velocity.



Figure 1. Overall data flow for detailed activity recognition system using Smartphone.

Human Activity Recognition can be broadly classified on the basis of medium of data collection into two categories of “Using Wearable Sensor” and “Using Smartphone”. Some relevant state-of-the-art works are summarized in Table 1. Most of these works use wearable accelerometers, or accelerometer sensors of smartphones. It is evident that works are done in different directions, not only detecting detailed daily life activities and fall [7,19], but also on online activity recognition [20], publishing benchmark datasets [21] as well as analyzing different usage behaviour as in [18,22].

Table 1. Comparison between state-of-the-art works.

Existing Work, Year	Activities Considered	Device and Position	Sensors	Remarks
[7] in 2004	Walking, Walking carrying items, Sitting & relaxing, Working on computer, Standing still, Eating /drinking, Watching TV, Reading, Running, Bicycling, Stretching, Strength-training, Scrubbing, Vacuuming, Folding laundry, Lying down & relaxing, Brushing teeth, Climbing stairs, Riding elevator, Riding escalator	Wearable sensors worn at right hip, dominant wrist, non-dominant upper arm, dominant ankle, and non-dominant thigh	Five Biaxial Accelerometers	Using Decision tree classifier they achieved overall accuracy of 80%. They concluded that accelerometer at user's thigh and dominant wrist are relatively best position for distinguishing between activities.
[19] in 2009	Climb Stairs, walk, sit, jump, lay down, run, run on stairs, fall-like motions (quickly sit-down upright, quickly sit-down reclined), flat surface falls (fall forward, fall backward, fall right, fall left), inclined falls (fall on stairs)	Tempo Sensor Nodes 3.0 (Wearable Thigh, Chest)	Accelerometer, Gyroscope	In addition to Activities of Daily Living, fall is detected. It differentiates intentional and unintentional transitions.
[23] in 2010	Lying, Cycling, Climbing, Walking, Running, Sitting, Standing	Wearable Sensor	Accelerometer	Uses wearable accelerometer, introduces classification by Hidden Markov Model
[24] in 2010	Walking, Jogging, Climbing Stairs, Sitting, Standing	Smartphone placed in pocket	Accelerometer	Achieved 90% accuracy, uses J48, logistic regression, multi layer perceptron.
[20] in 2012	Idle, Walking, Cycling, Driving, Running	Smartphone placed in pocket	Accelerometer	The entire system is implemented offline and online. However, in the online mode, the recognition on the device was performed using only a limited number of randomly chosen instances from training data due to limited computational power of the smartphones. Although the online results are almost comparable with the offline results, the system is not entirely user independent.
[21] in 2013	Walking, Climbing Upstairs, Climbing Downstairs, Sitting, Standing, Laying Down	Smartphone, Waist Position	Accelerometer	The finest public dataset for HAR using smartphone till now. An accuracy of 96% is achieved on the dataset using multiclass SVM. Features such as energy of different frequency bands, frequency skewness, and angle between vectors are employed
[22] in 2014	Running, Slow Walk, Fast Walk, Aerobic Dancing, Stairs Up, Stairs Down	Smartphone, Pocket and Held in hand	Accelerometer	Introduces "fine grained" activities for walk class, rest are coarse grained activities. Position independence is introduced with two positions. Accuracy of 91.5% achieved
[18] in 2016	Climb Upstairs, Climb Downstairs, Walking, Running, Standing	Smartphone, Coat Pocket, Trouser Pocket, Hand, Bag	Accelerometer, Gyroscope	Introduces four positions and proposes a position independent system through parameter tuning

A lot of research has been done in Human Activity Recognition (HAR) using Wearable Devices [25] mostly using acceleration data independent of orientation [26]. Combination of accelerometer sensor and gyroscope sensor placed on the neck of users are used in [27] to identify activities. They have evaluated the effect of appending gyroscope with accelerometer data and maintained individual threshold value for identifying several activities. The authors in [19] used gyroscopes and accelerometers placed on the thigh position and chest position of the user to identify several activities as well as fall using inclination angle and accelerometer value. An unintentional transition to a lying posture is regarded as a fall, where large changes in accelerometer and gyroscope readings can be observed. Authors in [19] differentiate intentional and unintentional transitions by applying thresholds to peak values of acceleration and rate of angular velocity from gyroscope. A certain change of angular velocity determines the fall of the subject.

A few works can also be found on wearable sensors using classifiers to recognize activities. In [28], authors considered accelerometers kept in three positions (wrist, ankle, chest) of the human body to monitor daily activities and applied Decision Table on the preprocessed data for activity classification. In [29], several activities are monitored using a customized device that is configured with an accelerometer attached to wrist position. However, the gyroscope sensor alone has not achieved a significant place in human activity recognition, and it can work in combination with other sensors. Using many wearable sensors for activity detection may hamper the movement of a person itself. In [30], “Vital Signs” are detected in addition to acceleration data, where vital signs vary in each activity, for example, when an individual begins running, it is expected that their heart rate and breath amplitude increase. This then becomes the vital sign for that activity and provides better accuracy. However, the advent of smartphones and the exponential increase in their usage in the past decade has resulted in a growing interest in HAR using smartphones for data collection as smartphones provide somewhat more convenient wearable computing environment.

Several works extensively study the use of smartphone based inertial sensors like accelerometers and gyroscopes in activity recognition. The Activity Recognition API by Google [31] provides insights into what users are currently doing and is used by several Android applications to enhance their user experience. The API automatically detects activities by periodically reading short bursts of sensor data and processing them using machine learning models. However, the set of activities that it can recognize is limited to coarse grained activities such as “sit”, “stand”, “walk”, “run”, “biking”, “device in vehicle”, etc. Few works could be found that focus on using minimal sensors like only accelerometers [10,22,24] for making the framework energy efficient and ubiquitous. Both of the works use state-of-the-art classifiers including MultiLayer Perceptron (MLP); however, the work in [22] aims at detecting detailed activities like slow walk and fast walk, while, in [24], mainly coarse grained activities are covered. In [22], the average accuracies of combination of classifiers is used for recognizing activity resulting in around 91% accuracy. Few works consider gyroscope along with accelerometer for gait analysis [32,33] and fall detection [34]. In [33], K-Nearest Neighbor (KNN) is used and achieved around 80% accuracy while different supervised learning algorithms are explored in [34], and Support Vector Machine (SVM) giving the highest accuracy.

In [14], authors consider all sensors available in smartphones like accelerometer, gyroscope and magnetometer for identifying human activity and explain the role of each sensor. The combination of accelerometer and gyroscope [14,35] is found to yield better results in some aspects. In [14], several machine learning algorithms are applied for activity recognition while the SVM is used to classify activities in [35]. The authors in [36] show the potential of using only magnetometer and how it affects activity recognition. In [37], the authors use a combination of an accelerometer and a gyroscope, and the recognition accuracy for some of the activities increases from 3.1–13.4%. The authors in [15] consider Decision tree, Logistic regression and Multilayer neural networks algorithms as base classifiers and designed a majority voting based ensemble [38] to identify human activity. It is found to increase accuracy up to 3.6% from a single classifier based approach. In [16], the authors combine multiple classifiers to improve the accuracy of activity recognition up to 7% and overall accuracy of 93.5% using

a 5-fold cross validation technique. In this way, the output of different classifiers can be combined using several fusion techniques to improve classification accuracy and efficiency.

In [39], authors identify activity applying KNN on the combination of various data like accelerometers, magnetometers, gyroscopes, linear acceleration and gravity, and it performs better than the accelerometer alone. However, most of these works above focus on coarse grained activities [8,9]. Few works could be found on [7,11,12] detailed activity recognition. In [7], the authors classified a number of detailed daily life activities with the help of several wearable sensors. They even predicted possible classification of a slow walk and brisk walk on the basis of speed. The authors in [12] uses wearable sensors to monitor detailed activity along with fall detection using Hidden Markov Model (HMM). In [40], HMM is applied to get activities subject to smartphone and ambient sensors. User ambience is also used in [11], where authors use accelerometers and gyroscopes for body locomotion, temperature and humidity sensors for sensing ambient environment, and location (via communication with Bluetooth beacon location tags). Two-level supervised classification is performed to detect the final activity state. A modified conditional random field based supervised activity classifier is designed by the authors for this purpose. However, the use of several sensors makes the systems more expensive and inconvenient for users.

In reality, for smartphone based ubiquitous activity recognition, we cannot impose constraints like the same device being used for training and testing or smartphones needing to be kept at a fixed position (the same as the one used for training the system). Thus, detailed activity recognition works should also consider the usage of different devices for training and testing (device independence) and usage behavior in terms of where the smartphone is kept for training and testing (position independence). The work in [17] considers device independence issues with multiple sensors and, in [41], position independent activity recognition framework is presented. In [17], the authors focus on several challenges like different users, different smartphone models and orientation. They have used several smartphone based sensors like accelerometers, gyroscopes and magnetic field sensors to remove gravity from accelerometer signals and converted accelerometer signal data from the body coordinate system to the earth coordinate system. Frequency domain features are extracted and the KNN (K Nearest Neighbor) classification algorithm is used in the work. In [1], device independent activity monitoring is achieved using Logistic Regression (LR) based two phase classifier where the best training device gets selected in the first phase while the second phase tunes the classifier for better recognition of activities. However, only coarse grained activities are recognized by this technique. Consequently, in this work, a detailed activity recognition framework is proposed that attempts to recognize detailed activities irrespective of the hardware configuration of smartphones and how the smartphones are kept during training and testing phases. We have not made use of Google's API as the class of activities that we are trying to classify are comprised of finer distinctions of a coarse grained class of activities, such as, for "walk", we have finer distinctions as: "brisk walk" and "slow walk". Similarly, for every coarse grained activity, finer distinctions exist and we propose a system here that learns such finer class of activities.

3. Problem Definition

Activity recognition problem can be defined as follows. Let the set of activities that can be recognized by the Human Activity Recognition System be represented by $A = \{a_1, a_2, a_3, \dots, a_n\}$, where A comprises of both static and dynamic activities. The Dataset $DS = \{ds_{d_1 p_1}, ds_{d_1 p_2}, ds_{d_2 p_1}, \dots, ds_{d_m p_k}\}$ is a set of datasets ($ds_{d_i p_j}$ represents set of data points), each being a function of device position (p_j) used for data collection and the device used (d_i), that is, $DS = f(Device\ Used, Device\ Position)$ where m denotes the number of devices used and k denotes the number of positions used for collecting the data. The dataset DS , when preprocessed results in DS' . $F = \{f_1, f_2, f_3, \dots, f_j\}$ denotes the feature space consisting of all the features extracted from the preprocessed dataset DS' , and each feature vector X_i of dataset DS' has an activity label y_i of the form $(x_1, x_2, x_3, \dots, x_j, y_i) : x_1, x_2, \dots, x_j \in ds_k, y_i \in A$. Given a learning algorithm C , the Human

Activity Recognition problem is to learn to recognize the activity set A from the dataset DS' using the feature space F , by using a function $g : DS' \rightarrow A$, where g is a member of the hypothesis space and it best fits the Dataset DS' to A using a loss function $L : A \times A \rightarrow R$ such that if, for an instance i of training the model, the activity label is y_i and predicted label is y and then the loss is computed as $L(y_i, y)$. The trained model is then tested on an unseen test dataset DS'' , the trained model C using the function $g : DS'' \rightarrow A$, predicts the activity being performed as y , and the accuracy of the model is then computed.

4. Detailed Activity Recognition Framework

The objective of this work is to identify six individual detailed human activities from raw data produced by accelerometer and gyroscope of a smartphone. Four static activities (Sit on floor/chair, Lying left/right) and two dynamic activities (Slow/Brisk walk) are considered for this work. New activities can also be recognized by the system by appropriately updating the training dataset.

Accelerometer and gyroscope sensor readings are collected from individual smartphones (chosen as training and test devices) being kept in either the Shirt PockeT (*SPT*), the Right (front) pants PockeT (*RPT*) or the Right Hand (*Hand*) position. The data collected by holding in one's hand was done in a way that replicates day to day usage; therefore, the subjects were asked to hold the device in their hands as if they are using them during *static fine grained activities*, and, during *dynamic fine grained activities*, subjects were asked to perform the activity while holding the device in their hands as per their preference. Thus, the device held in one's hand does not replicate other positions, that is, *SPT* or *RPT*. The raw data plots of three acceleration axes (A_x, A_y, A_z) for the above-mentioned set of activities are shown in Figure 2a for a device. It reveals that static and dynamic activities grossly show different patterns, which can be easily distinguished using threshold based techniques that measure changes of sensor readings. However, it is difficult to distinguish between two static activities like sitting on the floor and sitting on a chair. The problem becomes complicated when different devices are involved as can be observed from Figure 2a,b. Interestingly, sensor readings of one smartphone also vary depending on how it is kept. Figure 3 shows such patterns for different activities using the same device when it is kept at three different positions—*SPT*/*RPT*/*Hand* respectively. Thus, threshold based techniques are not sufficient to distinguish between static and dynamic detailed activities, especially when device and position heterogeneity are considered. Hence, data transformation, feature extraction, selection and the classification techniques should be designed in a way that can mitigate these challenges.

4.1. Data Preprocessing and Feature Extraction

The raw sensory data may contain noise or abnormal spikes, due to a certain change of position or fall of device, unintentional change of sensor orientation, etc. Filtering techniques remove accelerometer signal noise, outliers like low frequency acceleration (gravity), which capture orientation of the smartphone sensors with respect to ground level data, and noise generated by the dynamic motion of humans, and preserves medium frequency signal components. Data transformation is a significant process of validating and normalizing filtered data. Data transformation is applied to make a linear fit of one dataset against another. The nonlinear transformation generally increases the linear relationship after applying Tr (function for transformation) to each data point. The square root of the value, the inverse of the value, converting into logarithmic scale, etc. are different nonlinear transformation procedures that are used for statistical analysis. The logarithm function is applied when the data cover different orders of magnitude. Logarithm transformation [42] with base 10 is applied on (A_x, A_y, A_z) and (G_x, G_y, G_z) in test and training datasets to improve linear relationship for this work. An orientation insensitive dimension Signal Vector Magnitude (*SVMag*) is added in

order to achieve usage behavior independent recognition [14] along with existing three dimensions (A_x, A_y, A_z) of accelerometer readings and gyroscope readings (G_x, G_y, G_z):

$$SVMagA = \sqrt{(A_x^2 + A_y^2 + A_z^2)}, \quad (1)$$

$$SVMagG = \sqrt{(G_x^2 + G_y^2 + G_z^2)}. \quad (2)$$

Figure 4 shows the accelerometer readings collected from smartphones when it is in the right pants pocket and is faced upright or turned upside down. The plots show variations in particularly A_x and A_y , though data values for A_z do not show much variation in direction, but a slight variation of magnitude can be observed. However, as is evident in Figure 5, SVMag is found to mitigate the change of orientation of smartphones due to minor changes in usage behaviour, such as turning of the device. As is evident from the figure, limb movements can occur even for static activities while maintaining the posture resulting in momentary spikes in the trace.

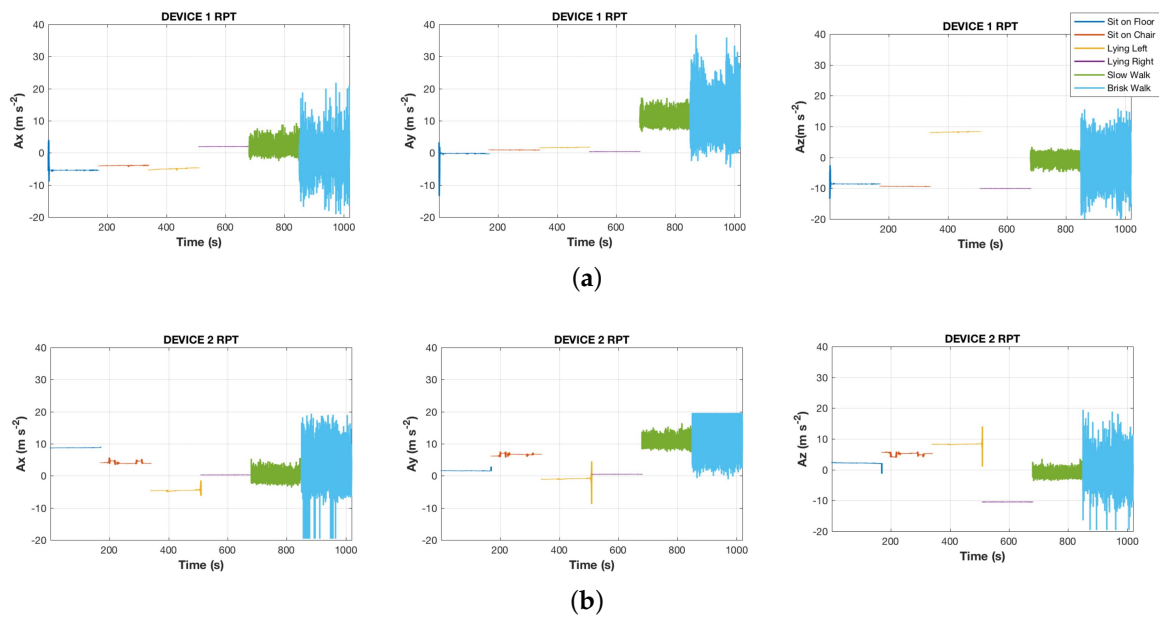


Figure 2. Raw accelerometer readings collected keeping devices in the Right Pants pocket (RPT) for detailed activities (a) A_x , A_y and A_z traces of Device 1 for different activities performed for 1000 s; (b) A_x , A_y and A_z traces of Device 2 for different activities performed for 1000 s.

The transformed data are partitioned into small segments and it is known as segmentation [35]. Proper selection of segment size is necessary to reduce classification complexity of the system and compute features from a small set of values. The short length of the window does not provide sufficient information of individual activities, and more than one activity may be present in the same window if the window size is too big. The sliding window approach is considered to effectively capture cycles in activities. Here, we have considered a 2 s window with 50% (1 s) overlap following [14] to reduce loss of information at the edge of the window. In [43], a 3 s window is found to achieve a minimal gain in classification accuracy in comparison to a 2 s window for short daily activities. The features are extracted from preprocessed data in the next phase of sensor data processing. Discovering meaningful representation of data and formulating the relation of raw sensor data with the expected knowledge for decision-making are the objectives of feature extraction. Feature vectors F_i 's are extracted on the set of segments S of the preprocessed dataset, Dt , by applying $ft()$. Extracted features constitute feature space:

$$F_i = ft(s_i) \forall s_i \in Dt. \quad (3)$$

Several statistical techniques are applied for feature extraction. Time and Frequency domain features are extracted as summarized in Table 2 for three dimensions of acceleration (A_x , A_y , A_z) and gyroscopes (G_x , G_y , G_z) along with orientation of independent dimensions $SVMagA$ and $SVMagG$.

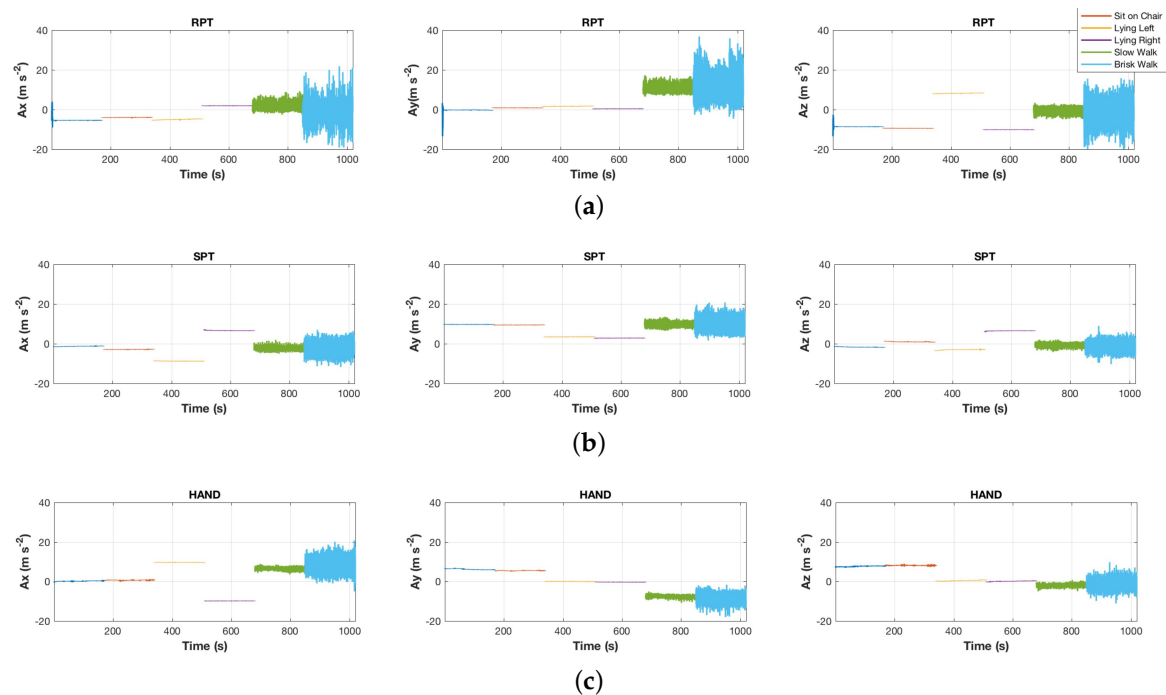


Figure 3. Raw accelerometer readings collected from a smartphone kept in three different positions for detailed activities. (a) device kept in Right Pants pocket (RPT); (b) device kept in the Shirt Pocket (SPT); (c) device kept in the right hand (Hand).

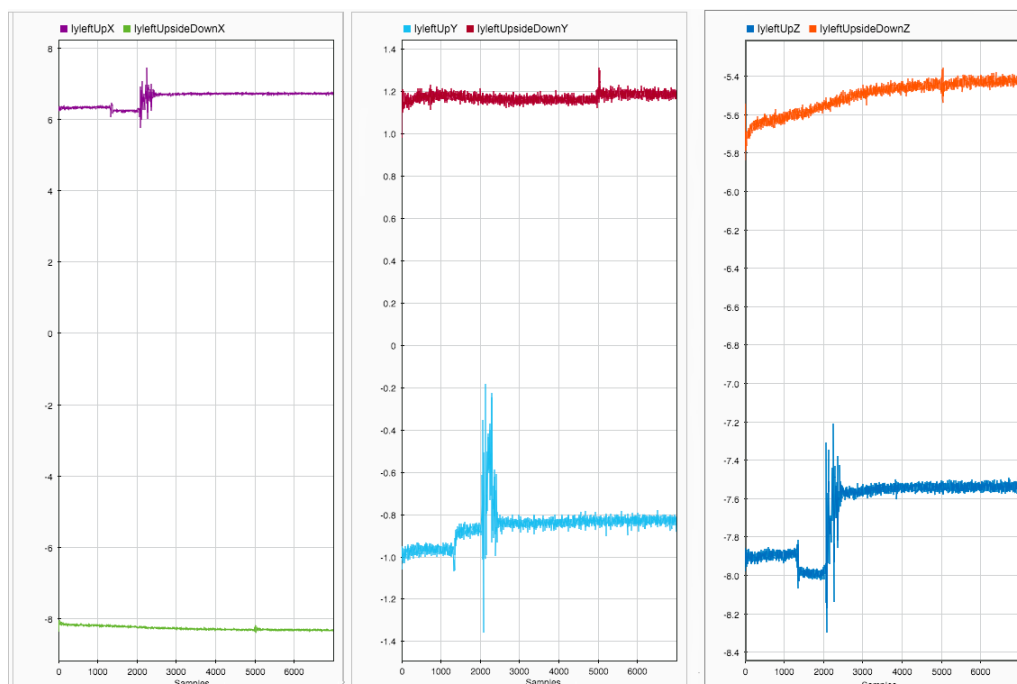


Figure 4. Accelerometer readings for activities lying left collected by keeping the smartphone in RPT, holding the device upright and turning it upside down.

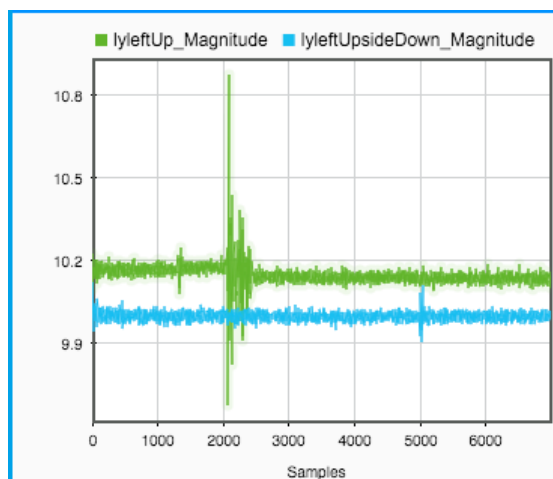


Figure 5. SVMag of Accelerometer readings for activity lyingleft corresponding to the raw data plotted in Figure 4.

Table 2. List of time and frequency domain features.

All Features Considered for Analysis. Each of the Feature is Calculated for $A_x, A_y, A_z, G_x, G_y, G_z, SVMagA$ and $SVMagB$		Important Features Obtained through Feature Selection	
Time Domain Features	Frequency Domain Features	Time Domain Features	Frequency Domain Features
Mean	Mean	$\text{Min}(A_y)$	$\text{Mean}(A_x)$
Standard Deviation	Median	$\text{Max}(A_y)$	$\text{Mean}(A_z)$
Variance		$\text{Max}(G_x)$	$\text{Mean}(G_y)$
Mean Absolute Deviation		$\text{Max}(G_y)$	$\text{Median}(A_y)$
Entropy		$\text{Max}(G_z)$	$\text{Median}(G_y)$
Min			
Max			

Initially, a total of 28 (seven features for four dimensions as mentioned in Table 2) time domain features and eight (two features for four dimensions as mentioned in Table 2) frequency domain features are applied to the preprocessed data. However, all features may not be relevant and informative. Thus, we have used information gain [44] to identify important features for the problem. Information gain value is measured for each attribute (feature) and the Ranker Search method is used to rank attributes by their individual evaluations. Features with low information gain value are removed as they do not add much information. Consequently, for the collected dataset, the following features are found to be informative for the problem considered.

Min and Max of A_y and max of G_x, G_y, G_z from time domain features, respectively, for accelerometers and gyroscopes; median of A_y and G_y , mean of A_z, A_x and G_y from frequency domain features, respectively, as listed in Table 2. However, only a mean is not sufficient to get accurate reflection of several activities on the skewed data. Min and Max are applied to define minimum and maximum values on each segments, respectively, as the acceleration is expected to be restricted to a certain range for each class of activity. The median arranges the observations in order from smallest to largest value and represents an average of the two middle values.

With these selected feature sets, in this paper, we design a condition based ensemble as part of the proposed detailed activity recognition framework. This is detailed in the next subsection.

4.2. Classification of Detailed Activity

Training position selection is a crucial factor for recognition of detailed activity. Here, the training position selection process is done using several base classifiers. Data are collected for each position using a training device. Data collected for one position is supplied as the training set while data pertaining to all other positions are treated as a test set. Representative positions can be found in this way based on accuracy of activity recognition. This is detailed in the experimental setup discussed in the next section.

It is difficult to identify individual activity with reasonable accuracy using a single classifier (C_i) with default parameters. As different devices are configured with different sensors having varying sensitivities, even for the same activity, the sensor readings differ from one device configuration to another as shown in Figure 2. Even the position where the smartphone is kept also influences the sensor readings (due to change in device orientation with respect to body) as reflected in Figure 3. Hence, keeping a device (d_i) at a specific position (p_i) for data collection is designated as a condition denoted by $d_i - p_i$. Data is collected for different conditions. Specifically, for each training device, data are collected, keeping the phone at each of the representative positions. A base classifier is applied to each feature set obtained corresponding to each of the conditions. However, parameter tuning is needed for individual conditions even with a selected base classifier. Moreover, including every possible condition into one classifier and retaining its power of generalization are not feasible tasks. However, classifiers could be individually tuned to effectively classify data for each training condition so that, when one specific classifier fails to achieve the desired result, an ensemble of such classifiers may prove to be a reasonable choice. Ensemble model is a combination of several condition based classifiers ($C_1, C_2, \dots, C_k$) to increase the performance of prediction. Here, we used Logistic Regression (LR) [10] as the base classifier and take individual training datasets, $ds_{d_i p_i}$ (collected from a device kept at a representative position) to train each of the classifiers with different parameter values. Hence, each individual classifier (C_i) is tuned to effectively classify data collected in a specific condition. The k trained classifiers are represented by $C' = \{C_1, C_2, C_3, \dots, C_k\}$. Given a test set (that may be collected using a different device keeping at RPT/SPT or held in hand), each C_i classifies the instances of the test set. Then, C' performs a weighted majority voting of the decisions made by the individual classifiers to come up with a decision for each of the test instances. The relative performance accuracy for C' is computed using a cost function $R': C' \times DS' \rightarrow W$, where $W = \{w_1, w_2, \dots, w_k\}$. R' computes the relative weights for the classifiers based on their performance accuracy for each activity and these weights are represented by w_i . This is defined as follows:

$$w_i = R'(C', DS') = \frac{\text{accuracy}(C_i)}{\sum_i \text{accuracy}(C_i)}, \text{ where } \sum_i w_i = 1 \quad (4)$$

Here, $\text{accuracy}(C_i)$ denotes the classification accuracy of the i th classifier that is obtained experimentally. This is calculated on a training dataset. Thus, the problem of detailed human activity recognition is to form a k -condition based ensemble classifier $EC: DS'' \times C' \rightarrow A'$, where EC performs weighted majority voting from C' using W on the test dataset DS'' and returns the activity being performed. Thus, the weighted majority voting scheme considers the classification of test instances that is predicted by the weighted majority of the classifiers as shown in Figure 6. The accuracy of ensemble EC is then computed by comparing the predictions with the true labels.

The basic block diagram of this detailed activity recognition framework is illustrated in Figure 6 where data are collected for k different conditions and, correspondingly, the selected base classifier is tuned for each of the k feature sets to form k condition based classifiers. We have considered two types of arrangements for collecting datasets—(i) using smartphone accelerometer and (ii) using both accelerometers and gyroscopes. Experimentations are conducted in these two modes to identify the role of each sensor for detailed activity recognition. The experimental results for validating the proposed framework are detailed in the next section.

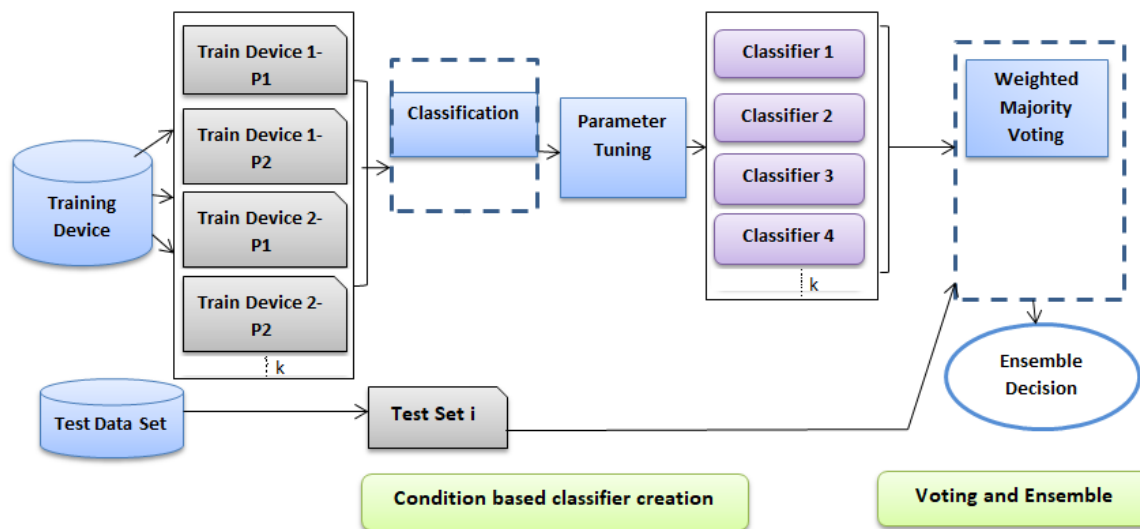


Figure 6. Block diagram of proposed ensemble based detailed activity recognition framework.

5. Performance Evaluation

In this section, the performance of the proposed framework is evaluated for real data collected from five smart handheld devices (D1–D5) for 10 users. The devices are kept in three positions on/around the body, namely, SPT, RPT and Hand. An android application Sensor Kinetics pro [45] collects the embedded tri-axial accelerometer and gyroscope sensor data for six detailed daily activities *Sit on chair*, *Sit on Floor*, *Slow Walk*, *Brisk Walk*, *Lying left* and *Lying right*. The static fine grained activities are found to be dependent on the user posture. For instance, the activities “Sit on Floor” and “Sit on chair” are not affected by the hardness of the surface of the chair or floor. The data collection was done on both chairs with cushion and chairs with no cushion. What separates these two similar activities is that the posture in which the user sits, and the relative position of the rest of the body parts in a particular posture. On average, the subjects took at least 55–65 steps per minute when “walking slow” and around 105–110 steps minimum when “walking briskly”. A user carried out each of these activities for 3–4 min while keeping each device at each of the three positions considered. The phone is held upright in each of these three positions while collecting the dataset. Each dataset contains around 54,000 accelerometer and gyroscope records. The experimental setup is detailed in Table 3. After removing the low-frequency acceleration (gravity) and noise, preprocessed data are grouped into overlapping windows and features are calculated from the acceleration and gyroscope values using MatlabR2013 (MathWorks, Natick, MA, United States) [46]. A Weka3.7 [47] tool is used for applying classification algorithms where default parameters of classifiers are changed as necessary.

Initially, experiments are conducted to find representative training positions. If the classifiers can accurately identify several activities, even when the training and test position of the smartphone are different, then that position of the training device is considered to be a representative position. Two training devices *D1* and *D2* are considered (Table 3) in order to keep it minimal to show the effectiveness of the ensemble on three different test devices. However, the ensemble would work for any number of devices. Details of the experiments for selecting the base classifier for a condition based ensemble are provided, followed by experiments to show the effectiveness of the ensemble subject to device and position independence. Initially, it is applied only on the accelerometer data set and then a combination of accelerometers and gyroscopes is used. Finally, the overall performance of the framework is also verified.

Table 3. Summary of the experimental setup with default values of parameters.

Devices	MotoG4(D1), Pixel(D2), MotoZ2(D3), RedmiNote4(D4), Redmi2Prime(D5)
Data sampling rate	50 samples/s
Filtering method	Butterworth and median filter [48]
Window type and size	2 s sliding window with 1 s overlapping window
Data transformation technique	logarithm with base 10
Dataset size	540,000 samples approximately
Number of users	10
Feature selection technique	InfoGainAttributeEval, RankerSearch
Training dataset	SPT-D1, RPT-D1, SPT-D2, RPT-D2
Test dataset	SPT-D1-Tst, SPT-D3-Tst, SPT-D4-Tst, SPT-D5-Tst, Hand-D4-Tst

5.1. Training Position Selection

The main objective of this section is to verify whether the classifiers can identify different activities, when the smartphone is kept at one position to collect training data while test data are collected by keeping it at another position. State-of-the-art classifiers, such as Bayesian Network (BN), Decision Tree (J48), lazy learner such as k-Nearest Neighbor (IBK), ensemble learner such as bagging and Logistic Regression (LR) are applied. Logistic function or sigmoid function are used in LR. Bagging, which stands for Bootstrap Aggregation, helps to reduce variance and avoid overfitting.

The results are shown in Figure 7. Here, we have considered three positions (SPT, RPT, Hand), and, individually, each position is considered as a training position and other positions are test positions. For instance, if SPT is the training position, then RPT and Hand are considered as test positions. Several classification algorithms are considered for this experiment that are applied on the selected feature set extracted from the training and test datasets. From Figure 7a, it can be observed that, when a device is kept at SPT for training, and test data are collected by keeping it in RPT or Hand, all classifiers show comparable results. If training data are collected by keeping the device in SPT, accuracy of activity recognition is comparable with state-of-the-art classifiers. However, it becomes difficult to recognize activity when the device is kept in Hand while collecting training data as shown in Figure 7c. Trying to train the activities while holding the device in hand is the most challenging position as the way a device is held varies from person to person, in addition to other factors, such as the amount of gestures someone does, also affect the sensor values. Hence, it is the position for which we received the least relative accuracy in terms of learning various fine grained activities. Thus, SPT and RPT are found to be the two representative training positions to consider for collecting training datasets as shown in Figure 7a,b. The default parameters of the classifiers are tuned for accuracy as detailed in Table 4.

Table 4. Classifiers and their tuning parameters.

Classifier Name	Parameter Name	Parameter Detail	Default Value
Decision Tree (J48)	minObj	Minimum number of instances per leaf	2
K Nearest Neighbor (IBK)	K	The number of neighbors to use	1
Logistic Regression (LR)	MaxIts	Maximum number of iterations to perform	−1
Multilayer Perceptron (MLP)	Seed	The value used to seed the random number generator	0
	Learning rate	Learning Rate for the back-propagation algorithm	0.3
	Momentum	Momentum Rate for the back-propagation algorithm	0.2

5.2. Base Classifier Selection

The base classifier for the ensemble is selected from state-of-the-art classifiers Bayesian network, K nearest neighbor, LR, Multilayer perceptron and Decision tree. The “no free lunch theorem” for optimization states that no optimization technique (algorithm/heuristic/meta-heuristic) is the best for the generic case and all special cases (specific problems/categories) [49]. Thus, we find the classifier that consistently performs best for our collected dataset and select that as the base classifier for the ensemble. Test set SPT-D5-Tst is used for classification with respect to two training data sets (SPT-D1 and RPT-D1) as shown in Figure 8. This experiment is also repeated with other datasets. From these experiments,

it can be observed that LR consistently performs better than the other classifiers. Even from Figure 7, we find that LR performs well in most of the cases. Hence, LR is considered as the base classifier for our proposed ensemble of condition based classifiers. The main benefit of LR is that it is simple and the logistic cost function is convex, and thus finds the global minimum. The maxIts parameter of LR is tuned for several conditional classifiers, maintaining the range 20 to 40 in order to get stable output.

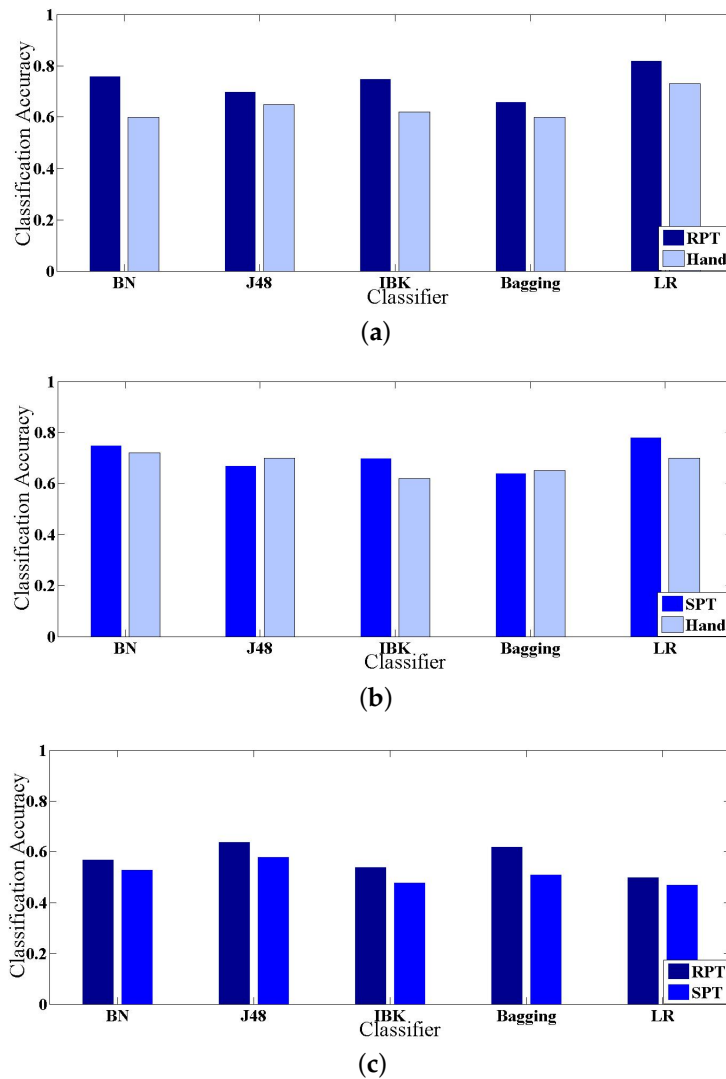


Figure 7. Classification accuracy for a device kept in different position (SPT, RPT and Hand) (a) for training position SPT; (b) for training position RPT; (c) for training position Hand.

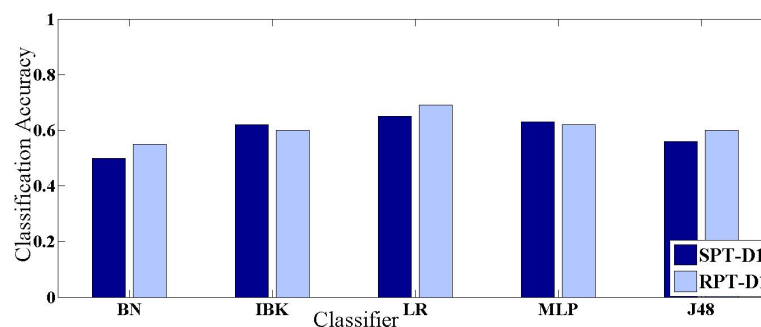


Figure 8. Base classifier selection in detailed activity recognition for SPT-D1 and RPT-D1 as training datasets and SPT-D5-Tst as test dataset.

5.3. Activity Classification Using Only an Accelerometer

The classification accuracy of the condition based classifiers along with the classification accuracy calculated by the majority voting ensemble are shown in Table 5, when the data set contains only data collected from an accelerometer sensor. It can be observed that, for most of the cases, a condition based ensemble classifier provides improved results from individual classifiers. Classification accuracy is improved in ensemble classification from an individual classifier by 3–20% as shown in Table 5 and Figure 9. When SPT-D5-Tst is considered as the test dataset, accuracy increases from 75% to 90% with ensemble. In this way, the framework becomes device independent. Reasonable accuracy can be achieved even when test data are collected by holding the device in hand (D4-Hand-Tst as test dataset). This makes the framework not only device independent but position (usage behavior) independent as well. If the test dataset is considered from one of the training devices, then ensemble is also found to provide better results, the overall classification accuracy is 91% as shown in Figure 9. This is the case for a test data set D1-SPT-Tst and that collected from training device D1.

Table 5. Summary of classification accuracy (in %) obtained for different training and test devices with four condition-based classifiers and Ensemble of these classifier. D1 and D2 are considered as training devices and D3, D4 and D5 as test devices. Devices are kept at different positions (SPT, RPT, or Hand). The column headings are: Training Dataset (TrDt), Test Dataset (TsDt), maxIts-Parameter tuning Value for LR (PtVI), Accuracy (in %) (Ac) and Error (in %) (Ec).

TrDt	TsDt	Classifier	PtVI	Ac	Er
D1-SPT	D1-SPT-Tst	Classifier1	33	85%	8%
D1-RPT		Classifier2	40	62%	32%
D2-SPT		Classifier3	−1	71%	21%
D2-RPT		Classifier4	21	65%	32%
		Ensemble	—	91%	5%
D1-SPT	D5-SPT-Tst	Classifier1	20	67%	16%
D1-RPT		Classifier2	22	60%	28%
D2-SPT		Classifier3	23	75%	18%
D2-RPT		Classifier4	21	68%	17%
		Ensemble	—	90%	5%
D1-SPT	D3-SPT-Tst	Classifier1	40	65%	27%
D1-RPT		Classifier2	22	51%	34%
D2-SPT		Classifier3	40	68%	25%
D2-RPT		Classifier4	22	51%	31%
		Ensemble	—	85%	8%
D1-SPT	D4-SPT-Tst	Classifier1	−1	76%	11%
D1-RPT		Classifier2	20	60%	30%
D2-SPT		Classifier3	23	60%	26%
D2-RPT		Classifier4	21	62%	36%
		Ensemble	—	79%	6%
D1-SPT	D4-Hand-Tst	Classifier1	21	64%	20%
D1-RPT		Classifier2	35	55%	29%
D2-SPT		Classifier3	32	56%	25%
D2-RPT		Classifier4	28	54%	27%
		Ensemble	—	78%	12%

Figure 10 shows the performance of the ensemble in identifying individual activities. Most of the activities are found to be effectively classified using ensemble with above 90% accuracy as shown in Figure 10. The figure interestingly indicates the effectiveness of a majority voting scheme employed here. However, *Sit on floor* is an activity that is not identified by the ensemble effectively as all the condition based classifiers are showing almost average classification accuracy. Actually, with accelerometer data alone, it is difficult to differentiate between two closely related static activities. Hence, experiments are also conducted using both accelerometer and gyroscope sensors in the next subsection. The classification errors are also reported.

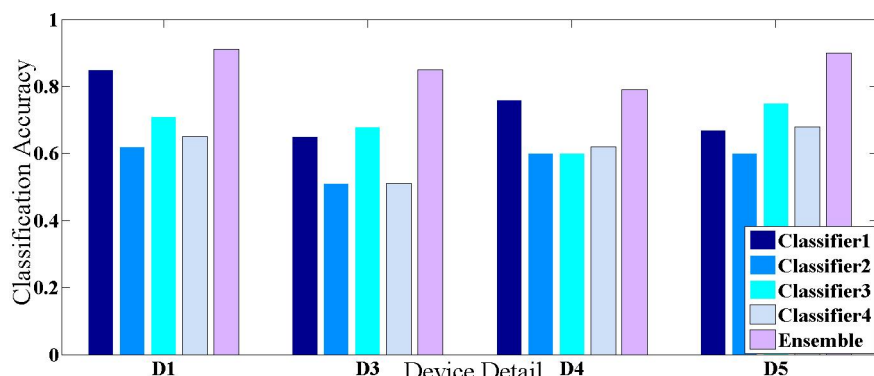


Figure 9. Classification accuracy for device and position independent activity recognition: Test devices (D1, D3, D4 and D5) kept in SPT.

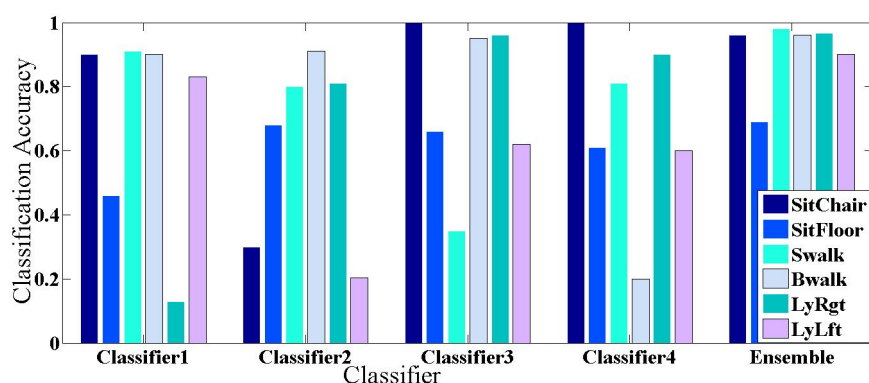


Figure 10. Classification accuracy (activity wise) for device independent activity recognition with D5 as test device using condition-based and ensemble classifiers for accelerometer sensors.

5.4. Activity Classification Using Both an Accelerometer and Gyroscope

Experiments are conducted for device independent activity recognition with an ensemble classifier using data collected from both accelerometer and gyroscope, and the results are reported in this section. The overall experimental procedure is the same as reported in the previous section. The classification accuracy of the condition based classifiers along with the classification accuracy calculated by the majority voting ensemble are shown in Table 6. When SPT-D5-Tst is considered as the test dataset, accuracy increases from 90% to 93% for ensemble, making the framework device independent. If the test data set is collected from one of the training devices, then overall classification accuracy is found to increase from 88% to 94% when test data set is D1-SPT-Tst.

Classification accuracies of the framework for individual activities are shown in Figure 11. Most of the activities are found to be classified with better accuracy compared to the one with an accelerometer as shown in Figure 11. *Sit on floor* was not effectively detected when using only an accelerometer, as is reflected in Figure 10, but can be better recognised when gyroscope readings are added to the dataset (Figure 11).

The work in [22] is compared with our work on our collected training dataset D1-SPT and test dataset D5-SPT-Tst. They have considered the average of probabilities fusion method, which returns the mean of the probability distributions for each of the single classifiers. Multiple combinations of classifiers are used to find the highest accuracy using an average of probabilities. Difference of Min and Max, Correlation, Root mean square, Average count of peak (AP), Variance of AP, Mean, and Standard deviation are considered as features in each window (128 samples) as detailed in the paper [22]. A combination of three classifiers—MLP, Random Forest (RF), and Simple Logistic (SL)—as

mentioned in their paper is considered, with the average probabilities fusion method. Both training and test datasets contain only accelerometer readings. Our proposed ensemble is applied on the selected feature set as detailed in Table 2, but the features are extracted from the same dataset, that is, training dataset D1-SPT and test dataset D5-SPT-Tst. The results are reported in Table 7. Our proposed framework is found to perform better showing 90% accuracy with minimal features. The necessity of having a condition based classifier is also reflected in the table.

Table 6. Summary of classification accuracy (in %) obtained for different training and test devices with four condition-based classifiers and an ensemble of these classifiers. D1 and D2 are considered as training devices; D3, D4 and D5 are considered as test devices. Devices are kept at different positions (SPT, RPT, or Hand). The column headings are: Training Dataset (TrDt), Test Dataset (TsDt), maxIts-Parameter tuning Value for LR (PtVl), Accuracy (in %) (Ac) and Error (in %) (Ec).

TrDt	TsDt	Classifier	PtVl	Ac	Er
D1-SPT	D1-SPT-Tst	Classifier1	21	88%	7%
D1-RPT		Classifier2	−1	64%	32%
D2-SPT		Classifier3	21	70%	24%
D2-RPT		Classifier4	20	61%	31%
		Ensemble	—	94%	4%
D1-SPT	D5-SPT-Tst	Classifier1	23	55%	26%
D1-RPT		Classifier2	35	60%	24%
D2-SPT		Classifier3	28	79%	18%
D2-RPT		Classifier4	30	75%	12%
		Ensemble	—	93%	3%
D1-SPT	D3-SPT-Tst	Classifier1	25	61%	26%
D1-RPT		Classifier2	28	50%	33%
D2-SPT		Classifier3	32	66%	28%
D2-RPT		Classifier4	40	60%	29%
		Ensemble	—	86%	8%
D1-SPT	D4-SPT-Tst	Classifier1	26	64%	25%
D1-RPT		Classifier2	31	60%	30%
D2-SPT		Classifier3	29	70%	12%
D2-RPT		Classifier4	22	68%	25%
		Ensemble	—	81%	7%
D1-SPT	D4-Hand-Tst	Classifier1	−1	56%	31%
D1-RPT		Classifier2	21	50%	34%
D2-SPT		Classifier3	28	65%	24%
D2-RPT		Classifier4	26	67%	27%
		Ensemble	—	80%	11%

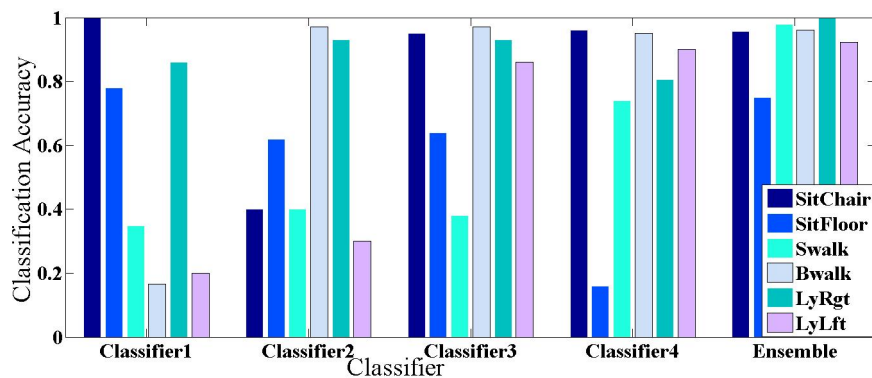


Figure 11. Classification accuracy (activity wise) for device independent activity recognition when D5 is used as a test device using condition-based and ensemble classifiers for both an accelerometer and a gyroscope.

Table 7. Comparison of the proposed technique with a multi-classifier combination approach for detailed activity recognition as in [22] on the basis of classification accuracy, where the training dataset is *D1-SPT* and the test dataset is *D5-SPT*.

Classifier Model	Feature Extraction and Selection	Classification Accuracy
Combination of Classifier (Multilayer Perceptron, Random Forest, Simple Logistic) as in [22]	Total 18 Features as in [22]	75%
Ensemble of Condition based Classifiers	Selected Time and Frequency domain features as in Table 2	90%

5.5. Evaluating the Performance of the Proposed System

Confusion matrices for the ensemble classifier considering only accelerometer values and both accelerometer and gyroscope values of D5 are shown in Figure 12. For case (a), the model mostly misclassifies *Sit on floor* as *Sit on chair*, and *Lying on left side* with *Lying on right side*. The misclassification for these classes decreases in (b), as the addition of gyroscope sensor values, in addition to accelerometer values enabling the model to better learn the finer differences in these similarly detailed activity classes.

```

=== Detailed Accuracy By Class ===
          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
          0.950    0.072    0.726     0.950    0.823     0.793    0.956    0.717    SitChair
          0.663    0.010    0.930     0.663    0.774     0.752    0.959    0.852    SitFloor
          0.978    0.010    0.952     0.978    0.965     0.958    0.998    0.978    Swalk
          0.950    0.000    1.000     0.950    0.975     0.970    1.000    1.000    Bwalk
          0.967    0.022    0.897     0.967    0.931     0.917    0.993    0.952    LyRgt
          0.890    0.007    0.964     0.890    0.925     0.912    0.993    0.911    LyLft
Weighted Avg.    0.900    0.020    0.912     0.900    0.899     0.884    0.983    0.902

=== Confusion Matrix ===
  a  b  c  d  e  f  <-- classified as
172  9  0  0  0  0 | a = SitChair
 61 120 0  0  0  0 | b = SitFloor
 4  0 177 0  0  0 | c = Swalk
 0  0  9 172 0  0 | d = Bwalk
 0  0  0  0 175  6 | e = LyRgt
 0  0  0  0 20 161 | f = LyLft

```

(a)

```

=== Detailed Accuracy By Class ===
          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
          0.956    0.050    0.794     0.956    0.867     0.843    0.971    0.776    SitChair
          0.751    0.002    0.986     0.751    0.853     0.838    0.971    0.942    SitFloor
          0.978    0.009    0.957     0.978    0.967     0.961    0.994    0.935    Swalk
          0.961    0.003    0.983     0.961    0.972     0.967    0.996    0.988    Bwalk
          1.000    0.022    0.900     1.000    0.948     0.938    0.996    0.969    LyRgt
          0.923    0.000    1.000     0.923    0.960     0.953    0.999    0.996    LyLft
Weighted Avg.    0.928    0.014    0.937     0.928    0.928     0.917    0.988    0.934

=== Confusion Matrix ===
  a  b  c  d  e  f  <-- classified as
173  2  1  0  5  0 | a = SitChair
 44 136 0  0  1  0 | b = SitFloor
 1  0 177 3  0  0 | c = Swalk
 0  0  7 174 0  0 | d = Bwalk
 0  0  0  0 181  0 | e = LyRgt
 0  0  0  0 14 167 | f = LyLft

```

(b)

Figure 12. Confusion matrices for device independent activity recognition with D5 as the test device using an ensemble classifier, for (a) accelerometer sensor readings and (b) both accelerometer and gyroscope readings.

Performance of the proposed system is evaluated using the following error metric. Here, E_i denotes the error of the i th instance and E_{tot} denotes the total error for the individual dataset. These are calculated as follows:

$$\begin{aligned} E_i &= 0, \text{ if } \text{label}(C_{test}) = \text{label}(C_{classification}) \\ &= 0.5, \text{ if } \text{label}(C_{test}) \neq \text{label}(C_{classification}) \text{ and } \text{CategoryLabel}(C_{test}) = \text{CategoryLabel}(C_{classification}) \\ &= 1, \text{ otherwise,} \end{aligned} \quad (5)$$

$$E_{tot} = \frac{1}{N} \times \sum_{i=1}^N E_i \times 100 \quad \text{where } N \text{ is the total number of instances.} \quad (6)$$

If the actual label is *brisk walk*, and the label is accurately predicted by the classifier, then E_i is 0, according to Equation (5). If the predicted label is *slow walk* instead of *brisk walk*, as both are dynamic of activities. E_i is 0.5. Otherwise, it is 1.

The error of activity recognition (E_{total}) for test datasets using individual classifiers and the majority voting ensemble is shown in Figure 13. A majority voting ensemble is found to produce an average error of 5% only, which is much lower than the errors using individual classifiers as shown in Figure 13. The average error can still be decreased to 2% when we are taking gyroscope readings along with accelerometer readings in the proposed framework as shown in Figure 14.

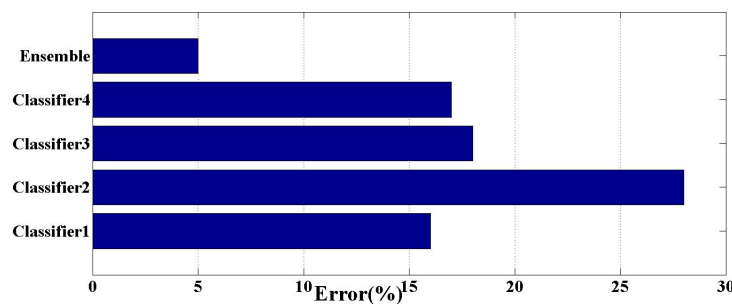


Figure 13. Error (%) for different classifiers when the test dataset is D5-SPT-Tst calculated using Equation (6), considering accelerometer readings only.

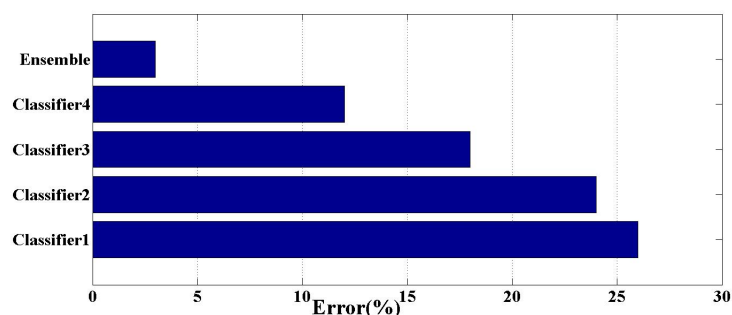


Figure 14. Error (%) for different classifiers when the test dataset is D5-SPT-Tst calculated using Equation (6), considering both accelerometer and gyroscope readings.

In this way, we can identify the activities when training device and test device positions (usage behavior) are different. The system gives better accuracy when training devices are D1 and D2 and the test device is D5. The condition based ensemble approach with selected minimum features is found to improve the overall system accuracy by 20% on an average as depicted in Table 5. Our proposed framework is found to perform well with ensemble, providing 90% accuracy using only an accelerometer for any device. The overall performance is improved when data from two sensors (both accelerometers and gyroscopes) are utilized. Activity wise, accuracy is also found to be increased.

6. Conclusions

In this work, we have proposed a framework for detailed activity recognition that classifies both static activities like *Sit on chair/Floor*, *Lying left/right* and dynamic activities like *Slow Walk*, *Brisk Walk* irrespective of usage behavior and differing hardware configuration of smartphones. Through feature extraction and selection, the framework can perform better for individual classifiers. The proposed weighted majority voting based ensemble of condition based classifiers is found to perform detailed activity recognition with considerable accuracy (90%) better than an individual classifier using only an accelerometer. SPT and RPT are found to be representative positions to keep a training device while collecting data for activity recognition. The framework is found to perform better when both accelerometer and gyroscope sensors work together, and achieves 94% accuracy. Results are taken from 10 users using five devices. The solution is ubiquitous, as it uses accelerometer and gyroscope sensors only, which are widely available in any smart handheld devices and does not need any specific device nor does it need the device to be held at some specific orientation.

We plan to look into more detail on how to train and test detailed activities when a device is held in hand, as this position is influenced by user gestures, with dynamic activities, etc. We also plan to combine datasets containing both coarse grained and fine grained activities, and make it public as part of our future work.

Author Contributions: J.S. and C.C. conceived and designed the experiments; J.S. and I.R.C. performed the experiments; J.S., I.R.C. and C.C. wrote the paper, S.B. and N.A. enhanced the quality of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Saha, J.; Chakraborty, S.; Chowdhury, C.; Biswas, S.; Aslam, N. Designing Two-Phase Ensemble Classifier for Smartphone based Activity Recognition. In Proceedings of the IEEE eHPWAS'17 Workshop of WiMob, Rome, Italy, 9 October 2017; pp. 257–264.
2. Shany, T.; Redmond, S.J.; Narayanan, M.R.; Lovell, N.H. Sensors-Based Wearable Systems for Monitoring of Human Movement and Falls. *IEEE Sens.* **2012**, *12*, 658–670.
3. Banerjee, T.; Keller, J.M.; Skubic, M.; Stone, E. Day or Night Activity Recognition From Video Using Fuzzy Clustering Techniques. *IEEE Trans. Fuzzy Syst.* **2014**, *22*, 483–493.
4. MalekiTabar, A.; Keshavarz, A.; Aghajan, H. Smart Home Care Network using Sensor Fusion and Distributed Vision-based Reasoning. In Proceedings of the VSSN'06, Santa Barbara, CA, USA, 27 October 2006; pp. 145–154.
5. Ren, L.; Shi, W. Chameleon: Personalised and adaptive fall detection of elderly people in home-based environments. *Int. J. Sens. Netw.* **2016**, *20*, 142–149.
6. Suryadevara, N.K.; Mukhopadhyay, S.C. Determining Wellness through an Ambient Assisted Living Environment. *IEEE Intell. Syst.* **2014**, *29*, 30–37.
7. Bao, L.; Intille, S.S. Activity recognition from user-annotated acceleration data. *Proc. Pervasive* **2004**, *3001*, 1–17.
8. Liu, T.; Inoue, Y.; Shibata, K. Development of a wearable sensor system for quantitative gait analysis. *Measurement* **2009**, *42*, 978–988.
9. Tran, D.N.; Phan, D.D. Human activities recognition in android smartphone using support vector machine. In Proceedings of the 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), Bangkok, Thailand, 25–27 January 2016; pp. 64–68.
10. Wannenburg, J.; Malekian, R. Physical Activity Recognition From Smartphone Accelerometer Data for User Context Awareness Sensing. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *47*, 3142–3149.
11. De, D.; Bharti, P.; Das, S.K.; Chellappan, S. Multimodal Wearable Sensing for Fine-Grained Activity Recognition in Healthcare. *IEEE Int. Comput.* **2015**, *19*, 26–35.
12. Pham, C.; Diep, N.N.; Phuong, T.M. A wearable sensor based approach to real-time fall detection and fine-grained activity recognition. *Mob. Multimed.* **2013**, *9*, 15–26.

13. Awan, M.A.; Guangbin, Z.; Kim, H.-C.; Kim, S.-D. Subject-independent human activity recognition using Smartphone accelerometer with cloud support. *Int. J. Ad Hoc Ubiquitous Comput.* **2015**, *20*, 172–185.
14. Shoaib, M.; Scholten, H.; Havinga, P.J.M. Towards Physical Activity Recognition Using Smartphone Sensors. In Proceedings of the 2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and IEEE 10th International Conference on Autonomic and Trusted Computing, Vietri sul Mare, Italy, 18–21 December 2013; pp. 80–87.
15. Yuan, Y.; Wang, C.; Zhang, J.; Xu, J.; Li, M. An Ensemble Approach for Activity Recognition with Accelerometer in Mobile-Phone. In Proceedings of the 2014 IEEE 17th International Conference on Computational Science and Engineering, Chengdu, China, 19–21 December 2014; pp. 1469–1474.
16. Mo, L.; Liu, S.; Gao, R.X.; Fredson, P.S. Multi-Sensor Ensemble Classifier for Activity Recognition. *J. Softw. Eng. Appl.* **2012**, *5*, 113–116.
17. Ustev, Y.E.; Incel, O.D.; Ersoy, C. User, device and orientation independent human activity recognition on mobile phone challenges and a proposal. In Proceedings of the 2013 ACM conference on pervasive and ubiquitous computing adjunct publication, Zurich, Switzerland, 8–12 September 2013; pp. 1427–1436.
18. Yang, R.; Wang, B. PACP: A Position-Independent Activity Recognition Method Using Smartphone Sensors. *Information* **2016**, *7*, 72.
19. Li, Q.; Stankovic, J.A.; Hanson, M.A.; Barth, A.T.; Lach, J.; Zhou, G. Accurate, Fast Fall Detection Using Gyroscopes and Accelerometer-Derived Posture Information. In Proceedings of the 2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks, Berkeley, CA, USA, 3–5 June 2009; pp. 138–143.
20. Siirtola, P.; Rönning, J. Recognizing human activities user-independently on smartphones based on accelerometer data. *Int. J. Interact. Multimed. Artif. Intell.* **2012**, *1*, 38–45.
21. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. A Public Domain Dataset for Human Activity Recognition Using Smartphones. In Proceedings of the 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013, Bruges, Belgium, 22–24 April 2013; pp. 24–26.
22. Bayat, A.; Pomplun, M.; Tran, D.A. A Study on Human Activity Recognition Using Accelerometer Data from Smartphones. *Procedia Comput. Sci.* **2014**, *34*, 450–457.
23. Mannini, A.; Sabatini, A.M. Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors* **2010**, *10*, 1154–1175.
24. Kwapisz, J.R.; Weiss, G.M.; Moore, S.A. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explor. Newsl.* **2010**, *12*, 74–82.
25. Lara, O.D.; Labrador, M.A. A survey on human activity recognition using wearable sensors. *IEEE Commun. Surv. Tutor.* **2013**, *15*, 1192–1209.
26. Yurtman, A.; Barshan, B. Activity recognition invariant to sensor orientation with wearable motion sensors. *Sensors* **2017**, *17*, 18–38.
27. Baek, W.-S.; Kim, D.-M.; Bashir, F.; Pyun, J.-Y. Real life applicable fall detection system based on wireless body area network. In Proceedings of the 2013 IEEE 10th Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, USA, 10–13 January 2013; pp. 62–67.
28. Ermes, M.; Parkka, J.; Cluitmans, L. Advancing from offline to online activity recognition with wearable sensors. In Proceedings of the 30th Annual International Conference on 2008 IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, 20–25 August 2008; pp. 4451–4454.
29. Kao, T.-P.; Lin, C.-W.; Wang, J.-S. Development of a portable activity detector for daily activity recognition. In Proceedings of the 2009 IEEE International Symposium on Industrial Electronics, Seoul, South Korea, 5–8 July 2009; pp. 115–120.
30. Lara, O.D.; Perez, A.J.; Labrador, M.A.; Posada, J.D. Centinela: A human activity recognition system based on acceleration and vital sign data. *Pervasive Mob. Comput.* **2012**, *8*, 717–729.
31. Google APIs for Android. Available online: <https://developers.google.com/android/reference/com/google/android/gms/location/DetectedActivity> (accessed on 26 February 2018).
32. Coley, B.; Najafi, B.; Paraschiv-Ionescu, A.; Aminian, K. Stair climbing detection during daily physical activity using a miniature gyroscope. *Gait Posture* **2005**, *22*, 287–294.

33. Brezmes, T.; Gorricho, J.-L.; Cotrina, J. Activity recognition from accelerometer data on a mobile phone. In *Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living*; 5518 of Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2009; pp. 796–799.
34. Lustrek, M.; Kaluza, B. Fall Detection and Activity Recognition with Machine Learning. *Informatica* **2009**, *33*, 197–204.
35. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. Human Activity Recognition on Smartphones Using a Multiclass Hardware-Friendly Support Vector Machine. In Proceedings of the 4th international conference on Ambient Assisted Living and Home Care-IWAAL'12, Vitoria-Gasteiz, Spain, 3–5 December 2012; Springer: Berlin/Heidelberg, Germany; pp. 216–223.
36. Kunze, K.; Bahle, G.; Lukowicz, P.; Partridge, K. Can magnetic field sensors replace gyroscopes in wearable sensing applications? In Proceedings of the 2010 International Symposium on Wearable Computers (ISWC), Seoul, South Korea, 10–13 October 2010; pp. 1–4.
37. Wu, W.; Dasgupta, S.; Ramirez, E.E.; Peterson, C.; Norman, G.J. Classification Accuracies of Physical Activities Using Smartphone Motion Sensors. *J. Med. Int. Res.* **2012**, *14*, e130.
38. Zhou, Z.H. Ensemble Learning. In *Encyclopedia of Biometrics*; Li, S.Z., Jain, A.K., Eds.; Springer: New York, NY, USA, 2009; pp. 270–273.
39. Martin, H.; Bernardos, A.M.; Iglesias, J.; Casar, J.R. Activity logging using lightweight classification techniques in mobile devices. *Pers. Ubiquitous Comput.* **2013**, *17*, 675–695.
40. Roy, N.; Misra, A.; Cook, D. Ambient and smartphone sensor assisted ADL recognition in multi-inhabitant smart environments. *J. Ambient Intell. Humaniz. Comput.* **2017**, *7*, 1–19.
41. Miao, F.; He, Y.; Liu, J.; Li, Y.; Ayoola, I. Identifying typical physical activity on smartphone with varying positions and orientations. *BioMed. Eng. OnLine* **2015**, *14*, 32.
42. Boslaugh, S. *Statistics in a Nutshell*, 2nd ed.; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2012.
43. Preece, S.J.; Goulermas, J.Y.; Kenney, L.P.J.; Howard, D. A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data. *IEEE Trans. Biomed. Eng.* **2009**, *56*, 871–879.
44. Dinakaran, S.; Thangaiah, P.R.J. Role of Attribute Selection in Classification Algorithms. *Int. J. Sci. Eng. Res.* **2013**, *4*, 67–71.
45. Sensor Kinetics Pro. Available online: <https://play.google.com/store/apps/details?id=com.innoventions.sensorkineticspro&hl=en> (accessed on 10 January 2018).
46. Matlab 2013. Available online: <https://www.mathworks.com> (accessed on 10 January 2018).
47. Weka—University of Waikato. Available online: www.cs.waikato.ac.nz/ml/weka/ (accessed on 10 January 2018).
48. Wang, W.Z.; Guo, Y.W.; Huang, B.Y.; Zhao, G.R.; Liu, B.Q.; Wang, L. Analysis of filtering methods for 3D acceleration signals in body sensor network. In Proceedings of the 2011 International Symposium on Bioelectronics and Bioinformatics, Suzhou, China, 3–5 November 2011; pp. 263–266.
49. Wolpert, D.H.; Macready, W.G. No Free Lunch Theorems for Optimization. *IEEE Trans. Evolut. Comput.* **1997**, *1*, 67–82.

