



Fusion of smartphone sensor data for classification of daily user activities

Gökhan Şengül¹ · Erol Ozcelik² · Sanjay Misra^{1,3} · Robertas Damaševičius⁴  · Rytis Maskeliūnas⁵

Received: 18 July 2020 / Revised: 3 November 2020 / Accepted: 21 May 2021 /
Published online: 20 August 2021

© The Author(s) 2021

Abstract

New mobile applications need to estimate user activities by using sensor data provided by smart wearable devices and deliver context-aware solutions to users living in smart environments. We propose a novel hybrid data fusion method to estimate three types of daily user activities (being in a meeting, walking, and driving with a motorized vehicle) using the accelerometer and gyroscope data acquired from a smart watch using a mobile phone. The approach is based on the matrix time series method for feature fusion, and the modified Better-than-the-Best Fusion (BB-Fus) method with a stochastic gradient descent algorithm for construction of optimal decision trees for classification. For the estimation of user activities, we adopted a statistical pattern recognition approach and used the k-Nearest Neighbor (kNN) and Support Vector Machine (SVM) classifiers. We acquired and used our own dataset of 354 min of data from 20 subjects for this study. We report a classification performance of 98.32 % for SVM and 97.42 % for kNN.

Keywords Human activity recognition · Wearable intelligence · Feature fusion

✉ Robertas Damaševičius
robertas.damasevicius@polsl.pl

Gökhan Şengül
gokhan.sengul@atilim.edu.tr

Erol Ozcelik
ozcelik@cankaya.edu.tr

Sanjay Misra
sanjay.misra@covenantuniversity.edu.ng

Rytis Maskeliūnas
rytis.maskeliunas@vdu.lt

Extended author information available on the last page of the article

1 Introduction

Today, mobile devices such as smartphones and tablet computers have powerful processors, high memory capacities and other sophisticated features, which allow for the development of intelligent context-aware services for smart environments such as smart homes, smart cities, and smart mobility [49]. Activity identification and classification using multi-sensor modalities can be successfully employed for user behaviour analysis, ambient assisted living, elderly care, medical diagnostics, patient rehabilitation after traumas, and surveillance [38, 10, 11, 28]. Fusing user data and information about environment enables to develop context-aware applications for mobile devices. The profile of the user, the information about the usage environment and usage time of the application, and other information about the environment are known as “context”. The applications, which use this contextual information, are defined as “context-aware” applications [2]. Moreover, using data fusion may provide additional benefits such as robustness against noise or external interference, increased reliability and confidence, improved accuracy and reduced ambiguity [19].

Fusion of data from multiple sensors is widely used to aggregate data gathered by heterogeneous devices or sensors [12]. Lower-level contextual spatial and temporal information can be exploited by intelligent mobile applications to improve the quality and usability of provided services. Specifically, the activity of users is one of the more important information in context-aware services. User activities such as walking, standing, and transportation by motorized vehicles provide useful information for the creation and provision of contextual services. Correct determination of user activity enables high-level reasoning over the domain of activities and services in order to create contextual rules such as “unmute my phone when the meeting ends and I start walking”, “forward all the incoming calls if I am driving”, “ring my phone loudly if I am walking”, etc. Such rules can be utilized by contextual reasoning engines to provide support for upper-level applications that provide smart context-aware services to its users. Using personalized context-aware models can significantly improve system performance for users [51], however learning personalized models has high computational cost. Therefore, the development of new methods that use data fusion for the improvement of context-awareness characteristics is of high importance now. Knowing current user activity is of utmost importance to comprehend the users’ context as the user’s task is a key element of the Dey’s notion of context [13]. Current approaches used for human activity recognition (HAR) can be categorized into two groups: the approaches using dedicated devices such as pedometer, and the approaches using smartphone sensors [43].

The contribution of this paper is as follows:

- 1) a novel hybrid data fusion method to estimate daily user activities using the accelerometer and gyroscope data acquired from a mobile phone;
- 2) the application of the matrix time series method for feature fusion, and the modified Better-than-the-Best Fusion (BB-Fus) method with a stochastic gradient descent algorithm for construction of optimal decision trees for classification;
- 3) a new dataset of the accelerometer and gyroscope signals acquired from a smartphone of users performing three types of daily user activities (being in a meeting, walking, and driving with a motorized vehicle);
- 4) classification of the fused accelerometer and gyroscope data using K-NN and SVM classifiers.

2 Related works

There are many studies in the literature dealing with activity recognition on smartphones. For example, Reddy et al. [44] classified user transportation modes (standing still, walking, running, cycling and transportation with a motorized vehicle) using smartphone accelerometer and Global Positioning System (GPS) sensors and obtained an accuracy of 93.6%. Zheng et al. [53] suggested a method to classify walking, cycling, and motorized transportation activities using only GPS data collected from 65 users and 76% performance was achieved. Yang [52] classified six different user activities (sitting, walking, running, cycling, standing, and motorized transportation) using the accelerometer data collected from 12 users and achieved an accuracy of 90%. Martin et al. [33] proposed a study to recognize six activities (slowly walking, fast walking, walking in normal way, running, sitting and standing) of users using the accelerometer data collected from 16 different users, and obtained a correct classification rate of 88%. Liang et al. [32] classified 11 different activities (standing still, lying, driving, sitting, walking, running, going up and down the stairs, cycling, and jumping) of users, and they achieved 85% correct classification performance. Shafique and Hato [54] achieved an overall accuracy of 99.96% when classifying smartphone accelerometer (along *z*-, *y*- and *z*-axes) and orientation (roll and pitch) data among six travel categories (pedestrian walk, bicycle riding, bus, subway, train). Cvetkovic et al. [9] achieved an $87\% \pm 5\%$ average accuracy for activity recognition using a fusion of data from smartphone and wristband sensors. Shdefat et al. [45] performed recognition of 12 human activities using the acceleration and gyroscope sensor data acquired from Android smartphones and reached the average accuracy of 89.79% with SVM and 87.81% with *k*-NN. Ahmed et al. [1] suggested a hybrid filter and wrapper method for feature selection that adopted a sequential floating forward search (SFFS) to extract the most important features for recognition of 12 stationary, motion and combined activities. Classification using multiclass SVM achieved an accuracy of 96.81%. Bragança et al. [6] used symbolic representation algorithms (Symbolic Aggregate Approximation, Symbolic Fourier Approximation, Bag-of-Patterns) to encode sensor data time series as symbolic sequences. The latter are classified by *k*NN classifier. Chen et al. [8] used ensemble Extreme Learning Machine (ELM) approach and achieved an accuracy of 97.35% on a dataset acquired from 12 subjects performing laying, standing, sitting, walking, walking downstairs and upstairs. Jain & Kanhangad [22] adopted the descriptor-based approach for activity classification smartphone sensor data. The histogram of gradient and centroid signature based Fourier descriptor were used to extract feature sets while feature and score level fusion was applied. Classification was performed using multiclass SVM and *k*-NN classifiers, and achieving 97.12 and 96.83% accuracy on the UCI HAR and physical activity sensor datasets. A summary of some other studies can be found in the reviews [46, 47].

Recently, neural networks, including deep learning models, have begun to be used in the HAR domain. Wan et al. [50] explored the use of convolutional neural networks (CNN), long short-term memory (LSTM), and bi-directional LSTM (BiLSTM) models, and Multilayer perceptron (MLP) for recognizing 18 daily physical activities for 9 subjects. Gjoreski et al. [18] used a combination of classical and deep learning methods for recognition of eight locomotion activities (bike, bus, car, run, still, subway, train, walk) using smartphone sensor data and achieved 94.9% accuracy. Pires et al. [40] achieved the accuracy of 85.89% using deep neural networks (DNN) for recognizing five activities (standing, walking, running, walking upstairs and walking downstairs. Qi et al. [41] achieved 95.27% accuracy using a custom DNN for recognizing 12 activities, including dynamical exercises (jogging, going

upstairs and downstairs, jumping, walking, etc.), six static postures (lying to the right and left side, lying supine and prone, standing, sitting), and action transitions. Li et al. [30] used frequency domain and temporal difference domain data from two sensors as inputs of the CNN, which was used as a feature extractor. Then one-class SVM was used for user authentication, achieving a 5.14 % equal error rate (EER).

Most of the analysed works use the accelerometer data, and a few of them use gyroscope, GPS, and Wi-Fi data in addition to accelerometer data. On the other hand, the features used in these studies differ as follows: raw data [39]; statistical features [15], autoregressive coefficients, signal magnitudes, linear discriminant analysis (LDA) and Kernel discriminant analysis [26]; average, variance, correlation coefficients, FFT (Fast Fourier Transformation) energy coefficients and Fourier domain entropies [48]; average, standard deviation, zero crossing rate, frequency domain entropy [52]. In case, when there is not enough data for training the classifier, the data augmentation strategy is employed [29].

Fusion of data can be performed in several ways such as data-level, feature-level, and decision-level [36]. For example, Li et al. [31] used serial feature fusion and parallel feature fusion to aggregate features from three smartphone sensors (accelerometer, gyroscope, and magnetometer). Most related works extract features from sensors and combine them to train a prediction model. However, most of them use aggregation may not produce the desired result as each sensors' data have different statistical characteristics, which do not allow producing a reliable classification model [16].

In this paper we propose a novel hybrid method to classify user activities of being in a meeting, walking and transportation with a motorized vehicle using the data fusion of the gyroscope and accelerometer data.

3 Proposed method

3.1 Outline of the methodology

In this study, a statistical pattern recognition approach was employed for activity classification, which uses a typical classification scheme used by other authors (such as [22]) in the domain of human activity recognition, too. In the training stage, the data whose classification results are known are used. In the testing stage, the performance of the classifier is measured [23]. The block diagram of the approach is given in Fig. 1.

Finally, the proposed method is summarized as a flow chart presented in Fig. 2. First, we acquired data from the accelerometer and gyroscope sensors of the smartphone. Next, we preprocess the data using Kalman filter and perform feature extraction. Next we apply data fusion using matrix eigenvalue based feature fusion method and apply Better-than-the-Best fusion on data originated from accelerometer and gyroscope sensors. Finally, we perform classification using commonly used machine learning methods (K-NN and SVM). These stages are described in more detail in the following subsections.

3.2 Raw data filtering

For data filtering, we applied the Kalman filter. The discrete Kalman filter estimates the state of the system and then measures and corrects its estimation. We use Kalman filter, because it has been successfully applied for denoising and state estimation of human activity signals

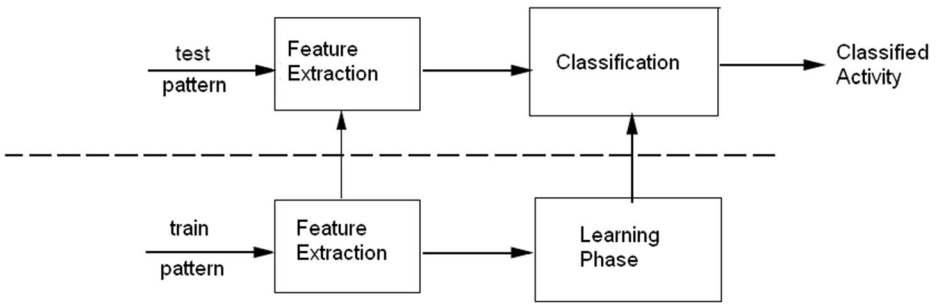


Fig. 1 Classification scheme used in this paper

before [21]. This is a cyclical process where one set of equations predicts the state of the system \hat{x}_t and the other set corrects the predictions. The time t update equations predict the state and covariance estimates from the time $t - 1$ as follows:

$$\hat{x}_t^- = A\hat{x}_{t-1} + Bu_{t-1} \tag{1}$$

$$P_t^- = AP_{t-1}A^T + Q \tag{2}$$

here A is the state transition model, B is the control-input model applied to the control vector u_t , P_t is the estimate error covariance, and Q is the process noise that may change from one time update to another, but is assumed to be constant for most calculations, and R is the observation noise. The measurement update (correction) equations are:

$$K_t = P_t^- H^T (HP_t^- H^T + R)^{-1} \tag{3}$$

$$\hat{x}_t = \hat{x}_t^- + K_t(z_t - H\hat{x}_t^-) \tag{4}$$

$$P_t = (I - K_t H)P_t^- \tag{5}$$

here H is the observation model, and I is the identity matrix.

First, the Kalman gain K_t is calculated. Then Eq. 4. uses the sensor signal measurement z_t to generate a state estimate. Finally, Eq. 5 calculates the error covariance P_t .

3.3 Feature extraction

We have extracted 16 features from smartphone sensors as follows: average power of the time window data, and minimum and maximum of signal power, and variance of signal power (a total of 4 features for gyroscope and accelerometer data, separately), the maximum values of each axis (3 features), the difference between the maximum and minimum values of each axis (3 features, one feature for the x, y and z axis, respectively), the variances of each axis (3 features), and entropies of axis (3 features). These features were extracted for gyroscope and accelerometer data separately. A time window of 1 s was used, as suggested by Reddy et al. (2010), as the time window size that provided the best performance.

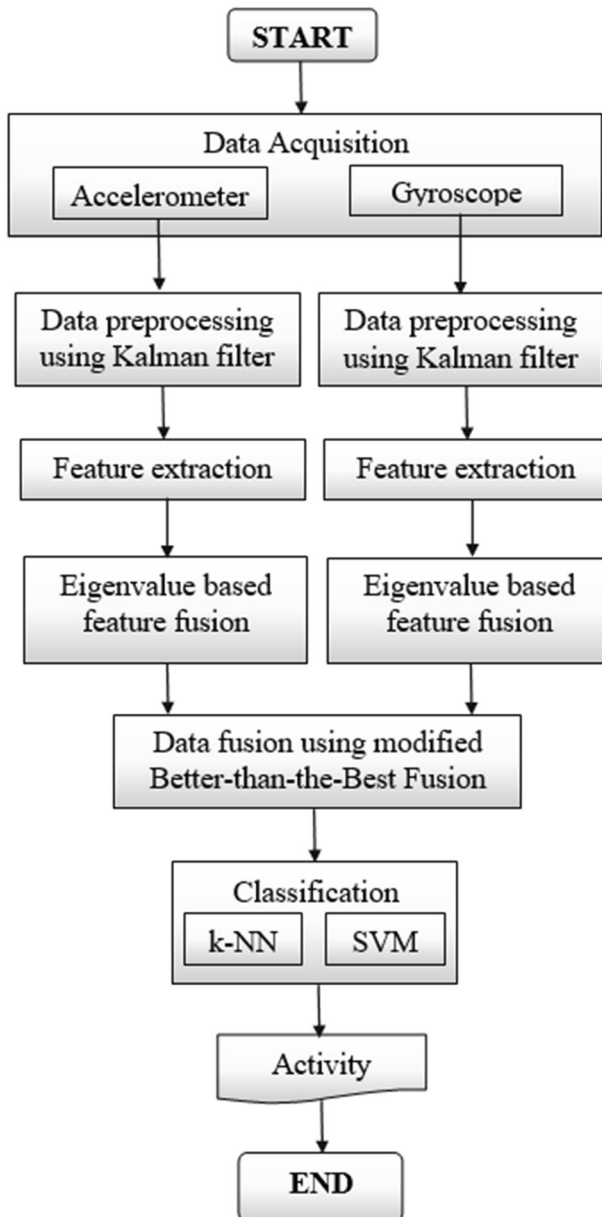


Fig. 2 Flow chart of the proposed method

3.4 Feature fusion

In this stage, fusion was performed on the features derived from a single data source (an accelerometer sensor or a gyroscope sensor). For feature fusion, we have applied the matrix time series method [5]. Given two synchronous numerical time series $x_i, \forall i = 1, 2, \dots, N$ and

$y_i, \forall i = 1, 2, \dots, N$, which are normalized to $(-1, 1)$. Then the 2nd order matrix time series $A_i, \forall i = 1, 2, \dots, N$ is formed as follows:

$$A_i = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} x_i & x_{i-1} - y_{i-1} \\ x_{i+1} - y_{i+1} & y_i \end{bmatrix}. \tag{6}$$

The features of this matrix time series can be calculated using various methods known from the matrix analysis theory such as determinant. Here we use the eigenvalues of the matrices A_i , in order to derive new fused features as follows:

$$x_i \oplus y_i = |\lambda_i^+| \tag{7}$$

here \oplus is the fusion operation, and $|\lambda_i^+|$ is the absolute value of the first eigenvalue calculated from solving the equation $det(A_i - \lambda_i I) = 0$, here I is the identity matrix, as follows:

$$\lambda_i^\pm = \frac{1}{2} \left[(a_{11} + a_{22}) \pm \sqrt{4a_{12}a_{21} + (a_{11} - a_{22})^2} \right] \tag{8}$$

To evaluate the discriminating power of fused features, we used the absolute value of the two-sample statistical t-test with pooled variance estimate, aka the Z-value. We calculate fused features for each combination of any feature derived from accelerometer and gyroscope x, y and z axis data, and select top three fused features with the largest Z-value for further classification.

3.5 Data fusion

Data fusion was performed to fuse features obtained from both data sources (accelerometer and gyroscope sensors). In this paper, we have adopted a modified Better-than-the-Best Fusion (BB-Fus) algorithm [34]. The method fuses data from different sensors using an optimal decision tree for classification. The optimal tree is created by consecutively discovering the best class and the best sensor data to isolate it at each level of classification decision. The class set is reduced each time a decision is made.

The method uses the confusion matrices $M_i, \forall i = 1, \dots, w$, that are found by examining the sensor-classifier pairs $\gamma_i = (s_i, c_i)$, here $s_i, \forall i = 1, \dots, w$ are sensors, and $c_i, \forall i = 1, \dots, w$ are classifiers. The task of finding the best sensor-classifier combination $*$ can be described formally as

$$\Gamma^* = arg \max_{T_j, \forall i = 1, \dots, n} \gamma_i, \forall i = 1, \dots, w \tag{9}$$

here $(T_j, T_j|\gamma_i)$ is the accuracy metric of the sensor-classifier combination $\gamma_i = (s_i, c_i)$, which is used to separate class T_j from the remaining classes.

In the training stage, the method examines all possible combinations of sensors and classifiers to get the best one-vs-all decision tree D^* as follows:

$$D^* = arg D_v \in Dmax(D_v) \tag{10}$$

here (D_v) is the fitness function, and $D_v, \forall v = 1, 2, \dots, V$ are the decision trees.

To find the optimal decision tree D^* , the original BB-Fus algorithm uses a greedy search algorithm. We however modified the BB-Fus algorithm to use stochastic gradient descent

algorithm that enables effective training with large datasets, which is relevant for the activity recognition domain, which has to deal with large amounts of recorded data. We have adopted an algorithm from [37]. This algorithm is not efficient for deep trees, especially as we need to perform inference once for every stochastic gradient computation. However, this is not a problem in our case, since we do not have many activities, therefore, decision trees are shallow.

3.6 Classification

After data fusion, the next step was selecting the classifier. The classifiers were selected by analysing the data type, data size, and computation time of the classifier. We have selected the *k*-Nearest Neighbor (*k*-NN) and Support Vector Machines (SVM), which are general-purpose classifiers commonly used for activity recognition [43, 7, 35, 20]. These classifiers are up to now widely used in the human activity recognition (HAR) domain with good results (see, e.g., [17]).

3.7 Evaluation of accuracy

The error is computed by measuring the proportion between the incorrectly classified data and the total number of data. For cross-validation, we use the Leave-One-Out Cross-Validation (LOOCV) approach. For a total of N data (fused feature vectors), the $N - 1$ data was employed as the training set and the remaining data was utilized as the testing set. This was repeated iteratively to select every data to be in the testing set once. As a result, every data is used as both for training and testing. The correct classification rate (CCR) is computed as the total number of correctly classified activities divided by the total number of activities. The F-measure is computed as the harmonic mean of the precision and recall, where precision is the ration of correct activities among the classified instances, while recall is the ratio of the total amount of relevant activities with respect to the total count of activities.

4 Experimental validation and results

4.1 Implementation of mobile app

A user-friendly mobile application was developed for the gyroscope and accelerometer data acquisition. A model of the application is given in Fig. 3. In the proposed system, the sensors in a smart watch are used to collect the necessary data about the person in question. Moreover, the smart watch is used as an agent for collecting data and transferring them to the cloud via a wireless connection. Cloud service is implemented on as a web service. The uploaded data is stored temporally on the web server and passed to the classifier. For classification, a supervised machine learning approach is developed and implemented, whose details are presented in Section 3.

Since smart watches have very limited battery power and running machine learning algorithms on them would require high processing power, we opt for installing these methods on a cloud server. We used Apache Web server and a PHP based web application is developed in order to get sensor data from the smart watch, transfer the data to classifier, receive the

predicted activity from the classifier, and to send modifications to the smart watch, if required. The operation of the system is as follows: the accelerometer and gyroscope data of the person in question are acquired by a smart watch application. The acquired data is transferred to the Web Server in JSON format, and the sensor data is fed to the machine learning based classifier. The classifier predicts the activity (note that many different kind of activities can be analysed provided enough data from appropriate activities have been collected), and returns it to the web server. The details of the smart watch used in this study are presented in Table 1. In this study Sony SmartWatch 3 SWR50 is used as the smartwatch and a mobile application is developed and installed on a Samsung Galaxy A7 smartphone to collect and transfer sensor data Via Bluetooth connection and get modifications from the cloud server. Although, the smart watch has several sensors, in this paper only accelerometer and gyroscope data are used for activity recognition.

4.2 User activities

In this study, we included three daily user activities for classification using the accelerometer and gyroscope data acquired from smartphones. These activities are being in a meeting, walking, and transporting with a motorized vehicle. For each activity, a subset of sub-activities, which can be different for each user, are defined. The reason to define these sub-activities is to ensure the variability of the activities during data acquisition. The sub-activities of being in a meeting are; keeping the phone in a fixed position (e.g., on a table), holding the phone by hand while sitting, rotating to left or right on a swivel chair, standing still, moving legs while sitting, crossing legs, standing up and sitting down. The sub-activities of walking are walking with normal speed, walking quickly, and climbing up and down stairs. Participants

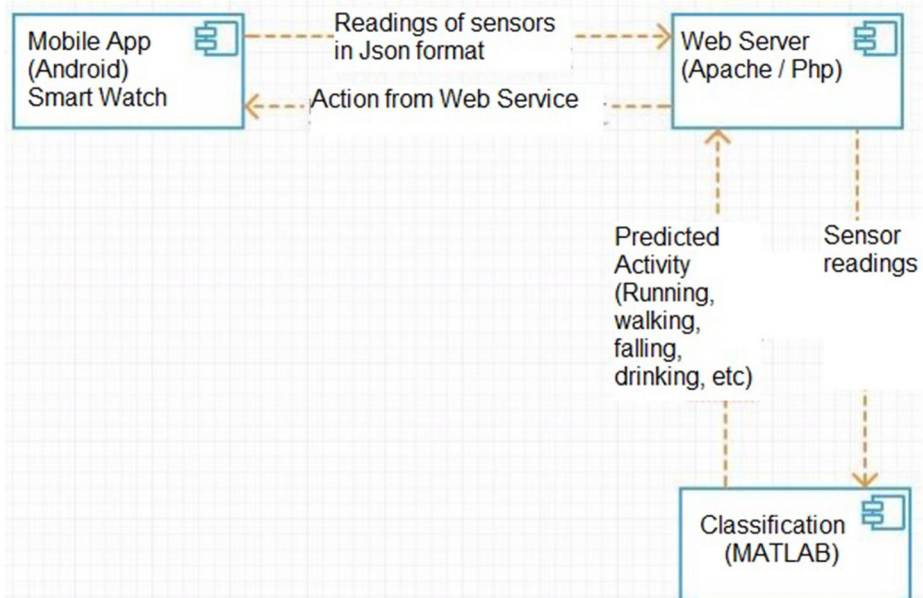


Fig. 3 The model of the developed application

Table 1 The specifications of Sony SmartWatch 3 SWR50

| Property | Capacity |
|----------------------|--|
| CPU | Quad ARM A7, 1.2 Ghz |
| Main Memory | 512 MB RAM |
| Battery power / time | 420 mA /up to 2 days for normal use |
| Water protected | IP68 |
| Android release | Android 4.3 and onwards |
| Weight | 45 g |
| Sensors | Ambient light sensors, Accelerometer, Compass, Gyro, GPS |
| Connections | Bluetooth® 4.0, NFC, Micro USB, Wi-Fi |

were given no instructions on walking with normal and fast speed. These speeds were decided by the participants. The sub-activities of motorized transportation are transportation in a heavy traffic with low speed, in a city centre with varying speeds, and on an express way with high speed. Although data are acquired for each activity, in this study the classification is performed for the main activities and classification of sub-activities are left for future work.

4.3 Data acquisition

The users used the developed mobile app for the collection of data. Firstly, the users selected which activity they would perform. Three seconds after the confirmation of the selected activity, the application automatically started recording data from the sensors. The users were allowed to define the data acquisition time before the program starts recording data. The sampling frequency of the data acquisition was set to 50 Hz. The amount of the data collected for each activity (and sub-activity also) was given in Table 2.

The data was collected from 20 volunteers. Each participant was asked to record their smartphone sensor data for 2 min while being in corresponding real-world situations (i.e., while driving or participating in a meeting). However, in some cases they ended the activities before the end of the time interval. For instance, when participants reached the end of the stairs before 2 min, they ended that activity by pressing a corresponding button. As a result, the total amount of collected data is different for each sub-activity, which can be seen in Table 1. In total, 354 min of data were acquired in the data acquisition phase. An example of sensor data acquired from the accelerometer and gyroscope data are presented in Figs. 4 and 5, respectively. For processing and visualization of data, and classification, we used MATLAB 9.6.0.1072779 (R2019a) on an Intel (R) Core (TM) i5-8635U CPU (x64), running at 1.80 GHz with 8 GB of RAM in Windows 10 operating system.

4.4 Measurement quality

For the acquisition of data, we have employed the participatory sensing approach [3]. The measurements were performed and sensor data was acquired by different subjects at different locations and environments. As a result of low control for experimental condition, the quality of acquired data may be a problem. The smartphone sensors data is very much influenced by sensor imprecision and inaccuracy. We measure the bias and variance parameters of acquired dataset following the methodology described in [27]. Sensor bias is an average of the sensor output that is assessed by averaging N samples of

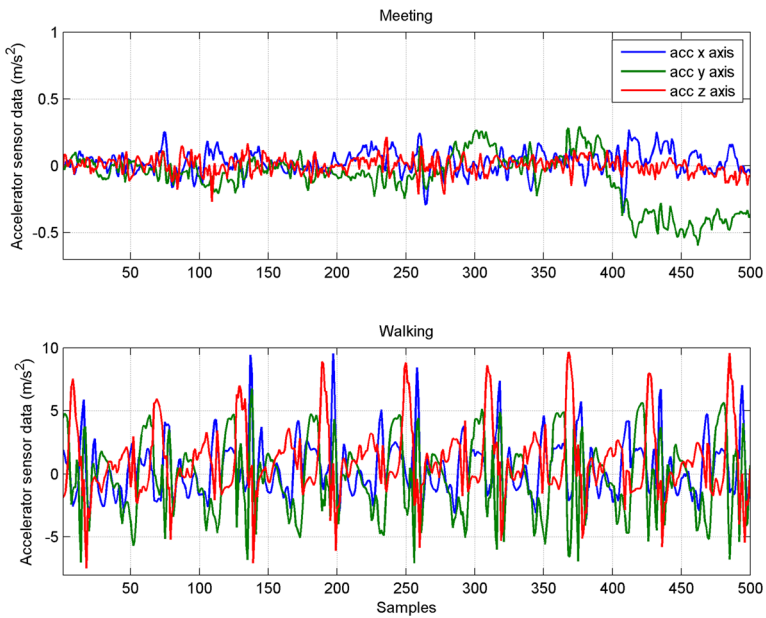


Fig. 4 Example of accelerometer sensor data for the meeting and walking activities

sensor signal as follows:

$$x_{bias} = \frac{1}{N} \sum_{i=0}^{N-1} x[i] \tag{11}$$

Characteristics of sensor noisiness can be analysed by calculating the Allan variance. First, the successive estimates of sensor bias are calculated:

$$e[m] = \frac{1}{N} \sum_{i=0}^{N-1} x[i + mN] \tag{12}$$

Table 2 Duration of data collected for each activity

| Main Activity | Sub-Activity | Acquisition duration (s) | Total duration (s) |
|--------------------------|-----------------------------------|--------------------------|--------------------|
| Being in a meeting | Moving the legs | 682 | 6871 |
| | Moving the phone | 1042 | |
| | Phone is fixed | 2242 | |
| | Moving towards to right or left | 828 | |
| | Standing still | 1105 | |
| | Cross one’s legs | 515 | |
| Walking | Standing up and sitting down back | 457 | 7162 |
| | Walking in a normal way | 3585 | |
| | Quick stepping | 2262 | |
| | Going down the stairs | 564 | |
| | Climbing up the stairs | 751 | |
| Motorized Transportation | Heavy traffic | 921 | 7208 |
| | Being in a city centre | 5589 | |
| | Express way | 698 | |

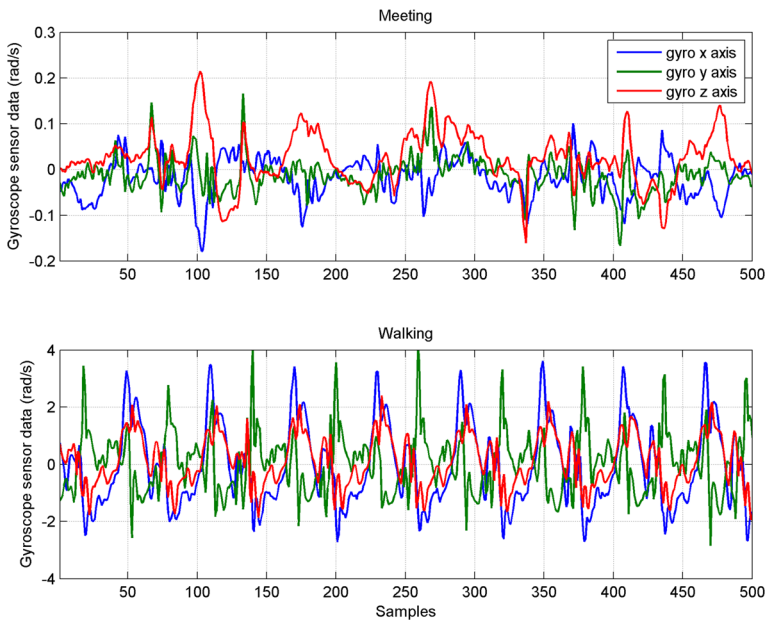


Fig. 5 Example of gyroscope sensor data for the meeting and walking activities

Next, the Allan variance is computed from a block of N signal samples as the mean squared difference between the successive estimates of bias as follows [14]:

$$\sigma_A^2[N] = \frac{1}{2(M-1)} \sum_{m=1}^M (e[m] - e[m-1])^2 \quad (13)$$

The bias estimate and Allan variance values for the sensor measurement data are presented in Figs. 6 and 7, respectively. Both results show that the quality of smartphone sensor data is acceptable, although the gyroscope sensors may require additional calibration due to positive bias estimate and larger than expected x and y axis values.

4.5 Feature fusion

Feature fusion was performed using the method described in subsection 2.3 and it yielded fused features with a much higher discriminatory power as demonstrated by their Z-values as compared with the corresponding features of the X, Y, and Z axis data of the accelerometer and gyroscope sensors. See a comparison of the fused features and original features presented in Fig. 8.

4.6 Results obtained using accelerometer data only

Accelerometer sensor was used to measure the acceleration in X, Y, and Z axis of the device. If speed in any axis increases, then accelerometer gives positive values for that axis. If speed decreases, then it yields negative values. Classification was performed for a combination of two activities (out of three) and lastly for all three activities using the k-NN and SVM classifiers. Leave-one-out cross validation (LOOCV) was administered for error calculation. The results (CCR and F-measure) obtained in these cases are presented in Table 3.

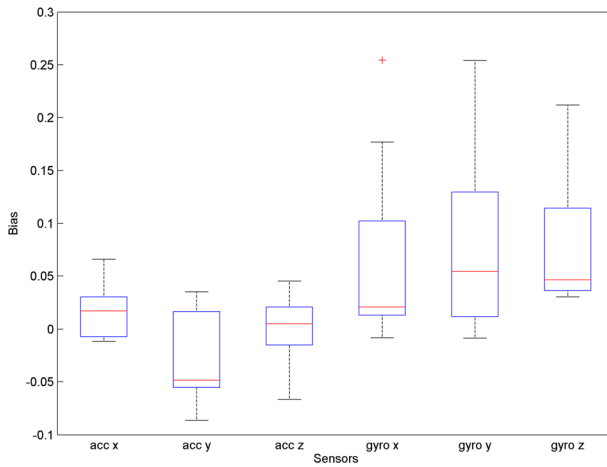


Fig. 6 Bias estimate values for accelerometer and gyroscope sensor data

As we can see from Table 3, the approach using SVM classifier and accelerometer data gave the highest performance for meeting-walking and meeting-motorized transportation activity combinations (98.23 % and 99.23, respectively). On the other hand, the performance was 83.51 % for walking-motorized transportation activity combination. The accuracy of the k-NN classifier was lower than that of SVM. It is concluded that meeting-walking and meeting-motorized transportation activity combinations could be classified with a high-performance rate, but the approach failed whenever walking-motorized transportation activity combination was performed with accelerometer data.

4.7 Results obtained using gyroscope data only

The gyroscope captures the rotation of the sensor towards its own axis using the gravity of the Earth. The rotation of the sensor in clockwise direction with respect to an axis yields a positive value, whereas a counter-clockwise rotation provides a negative value. The performance

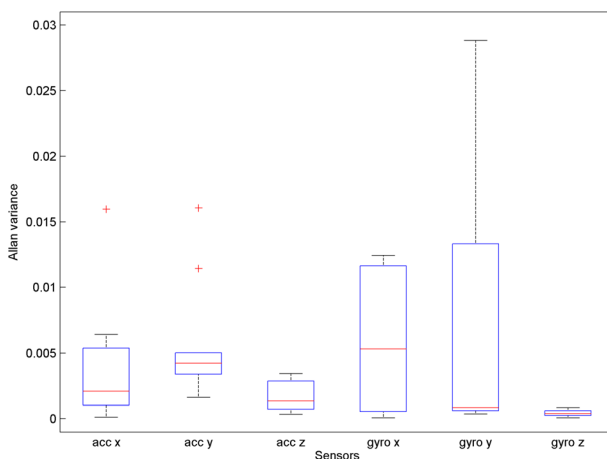


Fig. 7 Allan variance values for accelerometer and gyroscope sensor data

results (CCR and F-measure) of activity combination classifications and of all three activities using the SVM and k-NN classifiers and gyroscope data is presented in Table 4. Leave-one-out cross validation was employed for error calculation.

When only gyroscope data and SVM were used, a correct classification rate of 93.70 % was obtained for the meeting-walking activity combination, and 99.28 % performance was obtained for the walking-motorized transportation activity combination. On the other hand, 77.71 % performance was obtained for the meeting-motorized transportation activity combination, which was lower than the performance observed in the other combinations.

The best performance for the k-NN classifier was 97.42 %, which was obtained for the walking-motorized transportation combination. The worst performance in this scenario was 69.53 % for the meeting-motorized transportation classification. The results presented in Table 4 suggest that walking-motorized transportation activities could be classified with a high correct rate, whereas the performance of other activity combinations were low when only gyroscope data was used.

The activities of motorized transportation involved movements in a city centre with frequent turns. These kinds of activities have been better recognized using the gyroscope sensor than the accelerometer, because the gyroscope sensor captures orientation and angular velocity.

4.8 Results obtained using fused accelerometer and gyroscope data

Classification for all the activities was performed using the proposed method and the results (CCR and F-measure) are presented in Table 5. Leave one out cross validation (LOOCV) was followed, i.e., the classifiers were tested on the data of users, which was not used for training.

The overall classification performance was 98.32 % for the SVM classifier, and 97.42 % for the k-NN classifier. This shows that the SVM classifier gave approximately 1 % better performance than the k-NN classifier. When the activities are analysed separately, it is seen that the motorized transportation activity can be classified with a success rate of 99.35 %. The results obtained using both the SVM and k-NN classifiers are better than the ones given in the literature.

4.9 Evaluation of results

In a statistical pattern recognition approach, it is difficult to estimate the best classifier to use [23]. Due to this reason two different classifiers were used in the study, while fusion was performed using the modified Better-than-the-Best Fusion (BB-Fus) algorithm with a stochastic gradient descent algorithm. The performance of both classifiers (98.32 % for SVM and

Table 3 Classification performance obtained using accelerometer data only

| Activity combination | Correct classification rate (%) / F-measure of k-NN classifier | Correct classification rate (%) / F-measure of SVM classifier |
|------------------------------------|---|--|
| Meeting – Walking | 95.53 / 0.9364 | 98.23 / 0.9721 |
| Meeting – Motorized transportation | 89.63 / 0.8593 | 99.23 / 0.9774 |
| Walking – Motorized transportation | 72.21 / 0.7042 | 83.51 / 0.8081 |

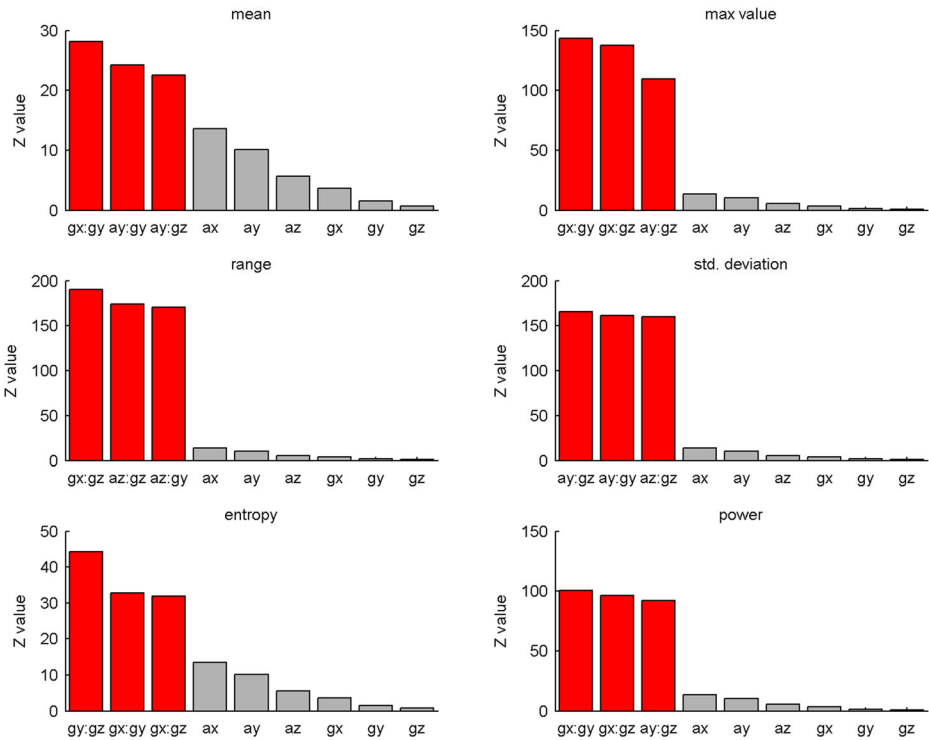


Fig. 8 Comparison of the discriminatory power of fused features with respect to original features of accelerometer (a) and gyroscope (g) data

97.42 % for k-NN) was higher than the ones given in the literature. So, any of these classifiers can be used. Moreover, the high performance of classification results obtained in this study was as a result of the data fusion method, rather than due to the used classifier. The performance of the current study can be compared with the performances of the existing studies in the literature. A summary of these studies are given in Table 6. The performance of the current work is higher than the existing studies. When the classified activities are considered, there are only three studies similar to current work. In these studies, standing still, walking and motorized transportation activities have been classified. As can be seen in the results section, the performance of the current work is higher than performance of the other studies. Moreover, the number of people that the data was collected in current work was also higher than those studies, which increases the generalizability of the results of the current study. However, note that those results are not directly comparable considering the different in input dataset and the experiment setting or environment.

Table 4 Classification performance obtained using gyroscope data

| Activity combination | Correct classification rate (%) / F-measure of k-NN classifier | Correct classification rate (%) / F-measure of SVM classifier |
|------------------------------------|--|---|
| Meeting – Walking | 89.57 / 0.8674 | 93.70 / 0.9082 |
| Meeting – Motorized transportation | 69.53 / 0.6134 | 77.71 / 0.7349 |
| Walking – Motorized transportation | 97.42 / 0.9423 | 99.28 / 0.9863 |

Table 5 Results of the proposed data fusion method

| Activity | Correct classification rate (%) / F-measure of k-NN classifier | Correct classification rate (%) / F-measure of SVM classifier (%) |
|--------------------------|---|--|
| Meeting | 96.35 / 0.9376 | 97.60 / 0.9587 |
| Walking | 96.95 / 0.9401 | 97.55 / 0.9532 |
| Motorized transportation | 98.95 / 0.9684 | 99.35 / 0.9822 |
| Overall performance | 97.42 / 0.9529 | 98.32 / 0.9783 |

A higher number of activities were considered by Zheng et al. [53]. However, increasing the number of activities may decrease the classification performance, so the results of the current study can not be directly compared with other studies. Similar (including accelerometer and gyroscope) data has been employed to separate static (such as the meeting) activities from dynamic activities (such as walking or driving) by other studies as well (see, e.g., [33, 25]). However, high performance obtained in this study

Table 6 Summary of activity classification studies found in the literature

| Study | Sensor | Classified activities | No. of subjects | Accuracy (%) |
|--------------------------|---|---|-----------------|--------------|
| Reddy et al. [44] | Accelerometer, GPS | Standing still, walking, running, bicycling, motorized transportation | 16 | 93 |
| Zheng et al. [53] | GPS | walking, bicycling, motorized transportation | 65 | 76 |
| Yang [52] | Accelerometer | Sitting, standing, walking, running, motorized transportation, bicycling | 12 | 90 |
| Martin et al. [33] | Accelerometer | Slowly walking, normally walking, fast walking, running, sitting, standing still | 16 | 88 |
| Kau and Chen [25] | Accelerometer | Falling down | 9 | 92 |
| Liang et al. [32] | Accelerometer | Standing still, sitting, lying, motorized transportation, walking, running, going down and up to the stairs, bicycling, jumping | 24 | 85 |
| Bayat et al. [4] | Accelerometer | Dancing, going down and up to the stairs, slowly walking, walking, fast walking, running | 4 | 91 |
| Ahmed et al. [1] | accelerometer and gyroscope | standing, sitting, lying, walking, walking downstairs, walking upstairs; stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie, and lie-to-stand | 30 | 96.81 |
| Jansi et al. [24] | accelerometer, gyroscope, magnetometer and orientation sensor | lie-down, jump, jog, sit, stand, walk, upstairs and downstairs. | 15 | 97.13 |
| Quaid and Jalal [42] | wrist worn accelerometer | brush teeth, climb stairs, comb hairs, descend stairs, drink glass, eat meat, eat soup, getup bed, liedown bed, pour water, sitdown chair, standup chair, use telephone, and walk | 30 | 93.10 |
| Jain and Kanhan-gad [22] | accelerometer and gyroscope sensor | slow walking, normal walking, brisk walking, jogging, sitting, normal upstairs, normal downstairs, brisk upstairs and brisk downstairs | 16 | 96.83 |

can be attributed by the application of the data fusion algorithm, which allowed for achieving better accuracy of classification.

Another point that needs to be considered when the performances are compared among different research is the bias, sensitivity and noise characteristics of sensors used during data acquisition. However, there is no information about the sensitivity of the sensors in the studies given in the literature, therefore, such comparison can not be performed.

5 Conclusions

We presented a data-fusion approach based on the feature fusion using a matrix time series method and the modified Better-than-the-Best Fusion (BB-Fus) algorithm with a stochastic gradient descent algorithm for the construction of optimal decision trees for classification. The approach was validated on three user activities using the accelerometer and gyroscope data acquired from smartphone sensors by 20 subjects. The quality of the measurement data was evaluated using the Allan variance method. For classification, we have used the k-NN and SVM classifiers. The meeting-walking activity combinations and meeting-motorized transportation activity combinations were classified with a high correct classification rate (98.23 and 99.23 %, respectively) when accelerometer data was used. On the other hand, when the gyroscope data was used, walking-motorized transportation activities were classified with a 99.28 % correct classification rate. Future studies will include user activity classification using a larger number of activities, including more fine-grained sub-activities, and the use of data fused from a larger number of sensors available on smartphones such as GPS, Wi-Fi, camera, and microphone by using the proposed data fusion methodology.

Declarations

Conflict of interest The authors declare no conflict of interest.

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

References

1. Ahmed N, Rafiq JI, Islam MR (2020) Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model. *Sensors* 20(1). <https://doi.org/10.3390/s20010317>
2. Alirezaie M, Renoux J, Köckemann U, Kristoffersson A, Karlsson L, Blomqvist E, Tsiftes N, Voigt T, Loutfi A (2017) An ontology-based context-aware system for smart homes: E-care@home. *Sensors* 17(7): 1586. <https://doi.org/10.3390/s17071586>

3. Ballesteros J, Rahman M, Carbanar B, Rishé N (2012) Safe cities. A participatory sensing approach. In: 37th Annual IEEE Conference on Local Computer Networks. IEEE, New York. <https://doi.org/10.1109/lcn.2012.6423684>
4. Bayat A, Pomplun M, Tran DA (2014) A study on human activity recognition using accelerometer data from smartphones. *Procedia Comput Sci* 34:450–457
5. Bikulciene L, Navickas Z, Vainoras A, Poderys J, Ruseckas R (2009) Matrix analysis of human physiologic data. In: ITI 2009 31st International Conference on Information Technology Interfaces. <https://doi.org/10.1109/iti.2009.5196052>
6. Bragança H, Colonna JG, Lima WS, Souto E (2020) A smartphone lightweight method for human activity recognition based on information theory. *Sensors* 20(7). <https://doi.org/10.3390/s20071856>
7. Chen Z, Zhu Q, Soh YC, Zhang L (2017) Robust human activity recognition using smartphone sensors via CT-PCA and online SVM. *IEEE Trans Industr Inf* 13(6):3070–3080. <https://doi.org/10.1109/TII.2017.2712746>
8. Chen Z, Jiang C, Xie L (2019) A novel ensemble ELM for human activity recognition using smartphone sensors. *IEEE Trans Ind Inf* 15(5):2691–2699. <https://doi.org/10.1109/TII.2018.2869843>
9. Cvetković B, Szeklicki R, Janko V, Lutomski P, Luštrek M (2018) Real-time activity monitoring with a wristband and a smartphone. *Inf Fusion* 43:77–93
10. Damaševičius R, Maskeliunas R, Venčauskas A, Woźniak M (2016) Smartphone user identity verification using gait characteristics. *Symmetry* 8(10). <https://doi.org/10.3390/sym8100100>
11. Damaševičius R, Vasiljevas M, Šalkevičius J, Woźniak M (2016) Human activity recognition in AAL environments using random projections. *Comput Math Methods Med.* <https://doi.org/10.1155/2016/4073584>
12. De Paola A, Ferraro P, Gaglio S, Lo Re G (2016) Context-awareness for multi-sensor data fusion in smart environments. In: Adorni G, Cagnoni S, Gori M, Maratea M (eds) AI*IA 2016 Advances in Artificial Intelligence. AI*IA 2016. LNCS vol 10037. Springer, Cham. https://doi.org/10.1007/978-3-319-49130-1_28
13. Dey AK (2001) Understanding and using context. *Pers Ubiquit Comput* 5(1):4–7. <https://doi.org/10.1007/s007790170019>
14. El-Sheimy N, Hou H, Niu X (2008) Analysis and modeling of inertial sensors using Allan variance. *IEEE Trans Instrum Meas* 57:140–149
15. Ferrari A, Micucci D, Mobilio M, Napolitano P (2020) On the personalization of classification models for human activity recognition. *IEEE Access* 8:32066–32079. <https://doi.org/10.1109/ACCESS.2020.2973425>
16. García-Ceja E, Galván-Tejada CE, Brena R (2018) Multi-view stacking for activity recognition with sound and accelerometer data. *Inf Fusion* 40:45–56. <https://doi.org/10.1016/j.inffus.2017.06.004>
17. García-González D, Rivero D, Fernández-Blanco E, Luaces MR (2020) A public domain dataset for real-life human activity recognition using smartphone sensors. *Sensors* 20(8). <https://doi.org/10.3390/s20082200>
18. Gjoreski M, Janko V, Slapničar G, Mlakar M, Reščič N, Bizjak J, Gams M (2020) Classical and deep learning methods for recognizing human activities and modes of transportation with smartphone sensors. *Inf Fusion* 62:47–62. <https://doi.org/10.1016/j.inffus.2020.04.004>
19. Gravina R, Alinia P, Ghasemzadeh H, Fortino G (2017) Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges. *Inf Fusion* 35:1339–1351. <https://doi.org/10.1016/j.inffus.2016.09.005>
20. Hassan MM, Uddin MZ, Mohamed A, Almogren A (2018) A robust human activity recognition system using smartphone sensors and deep learning. *Futur Gener Comput Syst* 81:307–313. <https://doi.org/10.1016/j.future.2017.11.029>
21. He J, Bai S, Wang X (2017) An unobtrusive fall detection and alerting system based on kalman filter and bayes network classifier. *Sensors* 17:1393
22. Jain A, Kanhangad V (2018) Human activity classification in smartphones using accelerometer and gyroscope sensors. *IEEE Sens J* 18(3):1169–1177. <https://doi.org/10.1109/jsen.2017.2782492>
23. Jain AK, Duin RPW, Mao J (2000) Statistical pattern recognition: A review. *IEEE Trans Pattern Anal Mach Intell* 22:4–37
24. Jansi R, Amutha R (2019) Sparse representation based classification scheme for human activity recognition using smartphones. *Multimed Tools Appl* 78(8):11027–11045. <https://doi.org/10.1007/s11042-018-6662-5>
25. Kau LJ, Chen CS (2015) A smart phone-based pocket fall accident detection, positioning, and rescue system. *IEEE J Biomed Health Inform* 19(1):44–56. <https://doi.org/10.1109/jbhi.2014.2328593>
26. Khan AM, Lee YK, Lee SY, Kim TS (2010) Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis. In: 5th international conference on future information

- technology (FutureTech 2010), Busan, Korea, pp 1–6. <https://doi.org/10.1109/FUTURETECH.2010.5482729>
27. Kos A, Tomažič S, Umek A (2016) Evaluation of smartphone inertial sensor performance for cross-platform mobile applications. *Sensors* 16(4):477. <https://doi.org/10.3390/s16040477>
 28. Lauraitis A, Maskeliunas R, Damasevicius R, Polap D, Wozniak M (2019) A smartphone application for automated decision support in cognitive task based evaluation of central nervous system motor disorders. *IEEE J Biomed Health Inform* 23(5):1865–1876. <https://doi.org/10.1109/JBHI.2019.2891729>
 29. Li Y, Hu H, Zhou G (2019) Using data augmentation in continuous authentication on smartphones. *IEEE Internet Things J* 6(1):628–640. <https://doi.org/10.1109/jiot.2018.2851185>
 30. Li Y, Hu H, Zhu Z, Zhou G (2020) SCANet: Sensor-based continuous authentication with two-stream convolutional neural networks. *ACM Trans Sens Netw* 16(3):1–27. <https://doi.org/10.1145/3397179>
 31. Li Y, Zou B, Deng S, Zhou G (2020) Using feature fusion strategies in continuous authentication on smartphones. *IEEE Internet Comput* 24(2):49–56. <https://doi.org/10.1109/mic.2020.2971447>
 32. Liang Y, Zhou X, Yu Z, Guo B (2013) Energy-efficient motion related activity recognition on mobile devices for pervasive healthcare. *Mob Netw Appl* 19(3):303–317. <https://doi.org/10.1007/s11036-013-0448-9>
 33. Martín H, Bernardos AM, Iglesias J, Casar JR (2012) Activity logging using lightweight classification techniques in mobile devices. *Pers Ubiquit Comput* 17(4):675–695. <https://doi.org/10.1007/s00779-012-0515-4>
 34. Najjar N, Gupta S (2015) Better-than-the-best fusion algorithm with application in human activity recognition. In: Braun JJ (ed) *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications* 2015. SPIE, Washington. <https://doi.org/10.1117/12.2177123>
 35. Nguyen ND, Truong PH, Jeong G (2017) Daily wrist activity classification using a smart band. *Physiol Meas* 38(9):L10–L16. <https://doi.org/10.1088/1361-6579/aa7c10>
 36. Noori FM, Riegler M, Uddin MZ, Torresen J (2020) Human activity recognition from multiple sensors data using multi-fusion representations and CNNs. *ACM Trans Multimed Comput Commun Appl* 1–19. <https://doi.org/10.1145/3377882>
 37. Norouzi M, Collins MD, Johnson M, Fleet DJ, Kohli P (2015) Efficient non-greedy optimization of decision trees. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'15)*, vol 1, pp 1729–1737
 38. Nweke HF, Teh YW, Mujtaba G, Al-garadi MA (2019) Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions. *Inf Fusion* 46:147–170. <https://doi.org/10.1016/j.inffus.2018.06.002>
 39. Ofstad A, Nicholas E, Szcodronski R, R., and Choudhury RR (2008) AAMPL: accelerometer augmented mobile phone localization. In: *1st ACM international workshop on mobile entity localization and tracking in GPS-less environments (MELT 2008)*, San Francisco, USA
 40. Pires IM, Marques G, Garcia NM, Flórez-Revuelta F, Teixeira MC, Zdravevski E, Coimbra M (2020) Pattern recognition techniques for the identification of activities of daily living using a mobile device accelerometer. *Electronics* 9(3). <https://doi.org/10.3390/electronics9030509>
 41. Qi W, Su H, Yang C, Ferrigno G, De Momi E, Aliverti A (2019) A fast and robust deep convolutional neural networks for complex human activity recognition using smartphone. *Sensors* 19(17). <https://doi.org/10.3390/s19173731>
 42. Quaid MAK, Jalal A (2020) Wearable sensors based human behavioral pattern recognition using statistical features and reweighted genetic algorithm. *Multimed Tools Appl* 79(9–10):6061–6083. <https://doi.org/10.1007/s11042-019-08463-7>
 43. Rahim KNKA, Elamvazuthi I, Izhar LI, Capi G (2018) Classification of human daily activities using ensemble methods based on smartphone inertial sensors. *Sensors* 18(12). <https://doi.org/10.3390/s18124132>
 44. Reddy S, Mun M, Burke J, Estrin D, Hansen M, Srivastava M (2010) Using mobile phones to determine transportation modes. *ACM Trans Sens Netw* 6(2):1–27. <https://doi.org/10.1145/1689239.1689243>
 45. Shdefat AY, Halimeh AA, Kim HC (2018) Human activities recognition via smartphones using supervised machine learning classifiers. *Prim Health Care Open Access* 8(1). <https://doi.org/10.4172/2167-1079.1000289>
 46. Shoaib M, Bosch S, Incel OD, Scholten H, Havinga PJ (2015) A survey of online activity recognition using mobile phones. *Sensors* 15(1):2059–2085. <https://doi.org/10.3390/s150102059>
 47. Sousa Lima W, Souto E, El-Khatib K, Jalali R, Gama J (2019) Human activity recognition using inertial sensors in a smartphone: an overview. *Sensors* 19:3213

48. Sun L, Zhang D, Li B, Guo B, Li S (2010) Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. *Ubiquitous Intell Comput Lect Notes Comput Sci* 6406:548–562
49. Talari S, Shafie-khah M, Siano P, Loia V, Tommasetti A, Catalão JPS (2017) A review of smart cities based on the Internet of Things concept. *Energies* 10:421. <https://doi.org/10.3390/en10040421>
50. Wan S, Qi L, Xu X, Tong C, Gu Z (2020) Deep learning models for real-time human activity recognition with smartphones. *Mob Netw Appl* 25(2):743–755. <https://doi.org/10.1007/s11036-019-01445-x>
51. Weiss GM, Lockhart JW (2012) The impact of personalization on smartphone-based activity recognition. In: *AAAI Workshop on Activity Context Representation: Techniques and Languages*
52. Yang J (2009) Toward physical activity diary: motion recognition using simple acceleration features with mobile phones. In *Proceedings of the 1st international workshop in Interactive multimedia for consumer electronics*. ACM Press, New York, 1–10
53. Zheng Y, Chen Y, Li Q, Xie X, Ma WY (2010) Understanding transportation modes based on GPS data for web applications. *ACM Trans Web* 4(1):1–36
54. Shafique MA, Hato E (2016) Travel Mode Detection with Varying Smartphone Data. [10.3390/s16050716](https://doi.org/10.3390/s16050716)
- Shafique MA, Hato E (2016) Travel mode detection with varying smartphone data collection frequencies. *Sensors* 16(5):716. <https://doi.org/10.3390/s16050716>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Gökhan Şengül¹ · Erol Ozcelik² · Sanjay Misra^{1,3} · Robertas Damaševičius⁴ · Rytis Maskeliūnas⁵

¹ Computer Engineering Department, Atılım University, Ankara Kızılcaşar Mah, İncek, Turkey

² Cankaya University, Yukarıyurtçu Mahallesi, Mimar Sinan Caddesi No:4, Etimesgut, Ankara 06790, Turkey

³ Department of Electrical and Information Engineering, Covenant University, 0123 Ota, Nigeria

⁴ Faculty of Applied Mathematics, Silesian University of Technology, Kaszubska 23, 44-100 Gliwice, Poland

⁵ Department of Applied Informatics, Vytautas Magnus University, Vileikos 8, Kaunas, Lithuania