

## Research Article

# Activity Recognition on Smartphones via Sensor-Fusion and KDA-Based SVMs

Adil Mehmood Khan,<sup>1</sup> Ali Tufail,<sup>2</sup> Asad Masood Khattak,<sup>3</sup> and Teemu H. Laine<sup>1</sup>

<sup>1</sup> School of Information and Computer Engineering, Ajou University, San 5, Woncheon-dong, Suwon-si, Gyeonggi-do 443-749, Republic of Korea

<sup>2</sup> Department of Computer Engineering, Yildirim Beyazit University, Altindag Ulus, 06030 Ankara, Turkey

<sup>3</sup> Department of Computer Engineering, Kyung Hee University, 1 Seocheon-dong, Giheung-gu, Yongin-si, Gyeonggi-do 446-701, Republic of Korea

Correspondence should be addressed to Teemu H. Laine; [teemu@ubilife.net](mailto:teemu@ubilife.net)

Received 27 December 2013; Revised 21 February 2014; Accepted 10 April 2014; Published 5 May 2014

Academic Editor: Jesus Martinez del Rincon

Copyright © 2014 Adil Mehmood Khan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Although human activity recognition (HAR) has been studied extensively in the past decade, HAR on smartphones is a relatively new area. Smartphones are equipped with a variety of sensors. Fusing the data of these sensors could enable applications to recognize a large number of activities. Realizing this goal is challenging, however. Firstly, these devices are low on resources, which limits the number of sensors that can be utilized. Secondly, to achieve optimum performance efficient feature extraction, feature selection and classification methods are required. This work implements a smartphone-based HAR scheme in accordance with these requirements. Time domain features are extracted from only three smartphone sensors, and a nonlinear discriminatory approach is employed to recognize 15 activities with a high accuracy. This approach not only selects the most relevant features from each sensor for each activity but it also takes into account the differences resulting from carrying a phone at different positions. Evaluations are performed in both offline and online settings. Our comparison results show that the proposed system outperforms some previous mobile phone-based HAR systems.

## 1. Introduction

It was about two decades ago that Mark Weiser introduced to the world the concept of ubiquitous computing, a computing paradigm with a goal of making computing an integral and invisible part of people's lives [1]. The ability of the computing devices to provide services and information properly and automatically, while vanishing into the background at the same time, requires the use of context. Context is any information that can be employed to describe the situation of entities that are considered relevant to the interaction between users and application themselves [2]. Though context comes in different kinds, one such kind is the activity being performed by a user at any given time.

Activity recognition means recognizing the actions of one or more entities using a series of observations on entities' actions and environmental conditions [2]. In the case of

human activity recognition (HAR), activities can be divided into two categories [3]: high-level activities, such as having a meeting and taking a shower, or simple low-level physical activities, such as walking and running.

HAR emerged as an important research area over the past decade because a variety of applications rely on sensing and recognizing users activities, including health and environment monitoring applications, home and industry automation applications, and security and surveillance applications.

The task of HAR starts with sensing the physical world, and two main techniques have been employed for this purpose: external and wearable sensors. The former case is where the devices, such as simple sensors embedded in everyday objects [4–6] and video cameras [3], are used for the sake of HAR. On the other hand, the latter case deals with HAR by means of small sensors attached to a user's body or clothing [7]. Though capable of recognizing activities with a high

accuracy, the external approach to HAR lacks pervasiveness because it forces users to stay within a parameter defined by the position and the capabilities of the sensor. Therefore, the focus of this research is on HAR by means of sensors that can be carried anywhere by the user.

Among the wearable sensors, the most widely used sensor for HAR is the accelerometer sensor. The advancement in technology in the past few years has resulted in miniature, low-cost yet highly reliable accelerometers that can be used to collect information about the physical activities of a user in a highly pervasive and invisible manner. Consequently, a large number of research studies have investigated the use of triaxial accelerometers for the sake of HAR with varying success rates [8–14].

These studies have demonstrated the fact that accelerometers could be used as an inexpensive and reliable means for capturing and analyzing physical activities. However, this technique is still obtrusive, as it relies on attaching a sensor device to a human body, and a very few people would like to have sensors attached to their bodies or wear special t-shirts or belts for the sake of HAR in real life. Furthermore, sometimes noise sources (such as motion artifacts and communication error) affect the recognition accuracy in wearable accelerometer-based HAR systems. Therefore, some studies have investigated the use of multisensor or sensor fusion approach to maximize the information content and reduce both systematic and random errors [15, 16]. But like previous works, these systems also rely on attaching multiple sensors to a human body, which limits their use in real life scenarios.

These days, smartphones come equipped with a rich set of sensors, including accelerometer, pressure, compass, gyroscope, proximity, light, GPS, microphone, and camera. These devices have become a part of our daily lives, as we carry smartphones nearly everywhere we go. Thus, the ubiquity and unobtrusiveness of the phones and the availability of different wireless interfaces, such as Wi-Fi, 3G, and Bluetooth, make them an attractive alternative platform for multisensor based HAR [17].

Recently many studies have used such phones for activity classification [18–28]; however, most of the previous works have used smartphones only for data collection, which were then transferred to an accompanying device (such as PC) for further processing. Also, mobile phones are generally energy constrained and extending their battery life is an essential requirement. Therefore, using these devices for activity recognition requires a lightweight recognition system in order to preserve battery life.

Accordingly, this paper introduces a smartphone-based HAR system based on a sensor fusion approach. The contributions of this work are fivefold. Firstly, instead of focusing on a small number of similar activities, a large variety of activities are recognized by using three most commonly available sensors on smartphones, that is, the accelerometer sensor, the pressure sensor, and the microphone. Secondly, in order to preserve the battery life, only time domain features are used from all the three sensors. Thirdly, to ensure a fast response, small time/data windows are employed. Fourthly, the use of only time domain features, coupled with small data windows, to recognize a large variety of activities while carrying phones

freely results in a complex classification problem. This is solved by employing a hybrid classification strategy where a nonlinear discriminant analysis approach is coupled with a classifier to find the most optimum decision boundaries. Lastly, system evaluations are carried out in both offline and online settings, for both subject-dependent and subject-independent scenarios, using a large number of subjects.

The rest of paper is organized into the following sections. Section 2 discusses some related work. Section 3 explains in detail the proposed sensor fusion approach for activity recognition. Section 4 talks about experiments and presents the experimental results with some discussion. Finally in Section 5, we conclude our work and briefly talk about the future directions.

## 2. Related Work

As mentioned earlier, the presence of a large variety of sensors in a single device makes smartphone a highly suitable device for HAR in particular and context recognition in general. As for HAR, smartphones have been used in many ways. For example, [18, 19, 21, 24] have used the smartphone accelerometer to recognize movements, such as walking and running. In [29, 30], the researchers employed the microphone to classify acoustic environments using sound analysis. Some studies used the smartphone GPS sensor to recognize transportation related activities [31–33]. One common thing between these approaches is that all of them employed one particular kind of sensor for the sake of HAR, or context recognition. It should be noted that the use of one sensor limits the scope of the recognition problem, that is, the number of activity classes that can be recognized, or the accuracy of the recognition [34]. This problem can be solved by employing or fusing multiple sensors together.

Though it did not use a smartphone, the first work that employed the multisensor approach to HAR was reported in [15]. In this work, the researchers employed body-worn microphones and accelerometers to recognize assembly and maintenance tasks in a wood-workshop. The sensors were mounted at different positions on users' arms. Activity classification was performed using linear discriminant analysis (LDA) on the sound data and hidden Markov models (HMMs) on the acceleration data, and four different methods at classifier fusion were compared. For sound classification, they used fast Fourier transform (FFT) as the feature extraction approach, whereas, for acceleration classification, the number of peaks within a frame and the mean amplitude of these peaks were used as features. This work showed that the combination of audio with acceleration helps in improving the accuracy of HAR.

In [35], the authors presented the design, implementation, evaluation, and user experiences of CenceMe application. Their system employs a large variety of sensors, that is, Bluetooth, accelerometer, microphone, camera, and the GPS, to recognize user's context (such as dancing at a party), and then shares this information on social network portals (such as Facebook). To recognize user's context, they used a multiclassifier approach where each classifier recognizes only

one specific type of context, that is, the accelerometer classifier to recognize ambulation (such as walking and running) and the audio classifier to determine whether a person is in a conversation or not. For accelerometer classification, the mean, standard deviation, and the number of peaks of the accelerometer readings were used as features, whereas, for audio classification, discrete Fourier transform (DFT) was used as features.

In [34], a comprehensive context recognizer is presented that is capable of recognizing ambulatory activities (such as walking, jogging, or being still) and transportation related activities (such as riding a bus, or a subway) by means of the accelerometer, the microphone, and the GPS sensor. Their approach is also a multilevel classification approach, where at the first level accelerometer data is used to recognize whether a person is in ambulatory mode or transportation mode. For ambulatory mode, the GPS sensor is then employed to verify if the user is walking, jogging, or being still. As for the transportation mode, microphone is activated to capture audio data, which is then utilized to detect if a person is in a bus or a subway. They used HMMs with simple time domain features, frequency domain features, and linear predictive coding (LPC) features for accelerometer data classification, whereas GMMS with Mel Frequency Cepstral Coefficients (MFCCs) were employed for audio data classification. Other similar works include [17, 33, 36].

In summary, some excellent multisensor approaches for smartphone-based activity recognition have been developed by researchers in the past; however, a few problems can be identified. Firstly, most of these works have used a multilevel or a multiclassifier approach for activity classification. This approach, though capable of providing high recognition results, uses long time windows to recognize activities. For example, in [33–35], window sizes of 12.8 seconds, 6 seconds, and 30 seconds were used, respectively. In contrast, non-multi-level approaches use shorter time windows. For example, [17, 18, 24] used a time window of 2.5 seconds, 3 seconds, and 1 second, respectively. Secondly, existing systems are based on both time domain and frequency domain features. In [37], it is shown that the time domain features are less energy/battery hungry as compared to frequency domain features. Lastly, majority of the systems have explored the recognition of a small number of activities. Therefore, it is desirable to investigate if it is possible to recognize a large number of activities using only the time domain features computed from small data windows employing only a single level classifier. And this is what we have investigated in this study.

### 3. The Proposed System

The architecture of the proposed system is shown in Figure 1, which follows a common activity recognition architecture, that is, Data Collection → Preprocessing → Feature Extraction → Classification [7].

The recognition process starts with collecting data from the three sensors, that is, the accelerometer sensor, the pressure, and the microphone, using a 3.5 second window. The

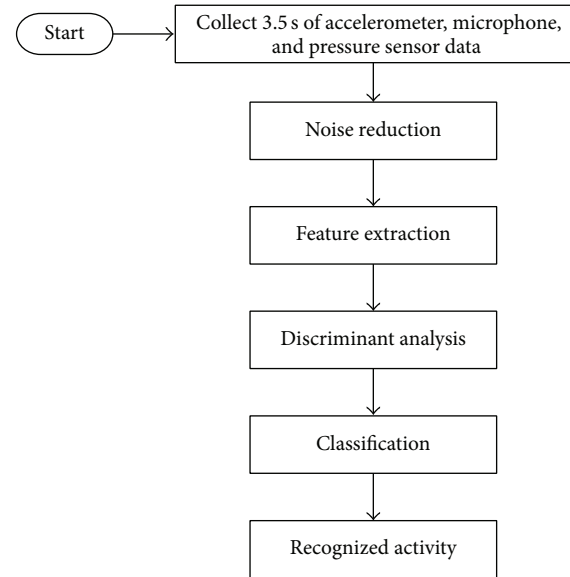


FIGURE 1: Flowchart of the proposed system.

window size was chosen based on the results of [37]. In [37], we implemented a real-time smartphone-based HAR system. Using exploratory data analysis techniques on acceleration signals of different physical activities, it was shown that these signals are generated by an autoregressive (AR) process, and an accurate time series models of these signals can be built using low sampling rates (20 Hz) and small data windows (3.5 seconds).

A total of 15 activities are targeted in this work, which are a combination of physical activities mostly explored in the area of physical activity recognition over the last decade [3, 7]. These include walking, walking on treadmill, running, running on treadmill, going upstairs, going downstairs, riding elevator up, riding elevator down, hopping, riding a bike, idle (sitting/standing), watching TV, vacuuming, driving a car, and riding a bus.

After collecting the activity data, the data are subjected to appropriate noise reduction techniques in the preprocessing step. This step is important because the data can contain noise, such as the gravity component in the case of acceleration data. Once filtered, the data are processed to extract various features.

After feature extraction, discriminant analysis is carried out. This step is important because an accelerometers output is very sensitive to the position of the human body at which it is placed, as shown in Figure 2, which shows the result of an experiment in which a person walked with 5 phones (LG Nexus 4) at 5 different positions (trousers front pockets, trousers back pockets, and jackets inner pocket) for 15 minutes along an L-shaped corridor. Acceleration data were recorded from each phone. The figure clearly shows how different the output of a smartphones accelerometer could be for the same activity when carried at different positions. Such differences result in high within-class variance, which could result in low classification accuracy. It was for this reason that

we employed discriminant analysis after feature extraction, that is, to minimize this variance.

Finally, the newly generated features after the discriminant analysis are fed to the classifier for activity recognition. A brief description of each of these components is given below.

**3.1. Noise Reduction.** A triaxial accelerometer embedded in a phone carried by a user registers two kinds of acceleration along 3 dimensions ( $x$ ,  $y$ , and  $z$ ): a constant acceleration due to gravity, and any acceleration the mobile device is subjected to by the user [38]. The acceleration due to gravity is considered as noise. Thus to calculate the real acceleration of the device the gravity must be eliminated. The technique to get rid of the gravity factor is borrowed from Androids own documentation [38], which employs a low-pass filter as follows.

$$g = (1 - \alpha) * g + \alpha * v_t, \quad (1)$$

where  $\alpha$  is the filter factor. Several values were compared and finally a value of 0.2 was for  $\alpha$ , as it not only got rid of most of the noise but also kept most of the details intact, as shown in Figure 3, which compares the after-filtering results for walking for three different values of alpha and the raw data.  $g$  is the global variable initially set to 0 (you will need three of these corresponding to three axes), and  $v$  is the accelerometer's value for a particular axis at time  $t$ . This will isolate the force of gravity which can then be subtracted from the sensor value as

$$\bar{v}_t = v_t - g. \quad (2)$$

**3.2. Feature Extraction.** Just like any other classification system, feature extraction plays a vital role in any wearable HAR system. Mobile phones are generally energy constrained, so extending phone's battery life is an essential requirement. In other words, using smartphones for HAR requires features that are both light-weight (energy efficient) and accurate (possess high discriminating power) to preserve battery life and ensure high accuracy. The details of feature extraction are as follows.

**3.2.1. Features from the Accelerometer Sensor Data.** A large number of frequency and time domain features have been investigated in the past for accelerometer-based HAR with varying success rates. The most widely used time domain features include mean [39–41], variance or standard deviation [39, 41], energy [39–41], entropy [40], correlation between axes [39–41], signal magnitude area [42], tilt angle [42], and autoregressive (AR) coefficients [9]. The most popular frequency domain features used so far are the FFT [43–45] and discrete cosine transform (DCT) coefficients [46].

In one of our previous research studies [37], we performed exploratory data analysis on the acceleration signals of various activities and learned that simple statistical features (SSFs) and the coefficient of time series analysis are highly suitable for smartphone-based activity recognition, as these features are capable of providing high recognition rates at

lower sampling rates. Based on this finding, this work extracts two kinds of time domain features from the accelerometer sensor data: (1) SSFs: mean, standard deviation, correlation, and signal magnitude area; (2) coefficients of time series analysis, including autoregressive (AR) analysis, and moving average (MA) analysis.

AR models are useful for describing situations in which the present value of a time series depends on its preceding value, whereas MA models are useful in capturing seasonality in time series data. For more details on AR and MA models, please refer to [47].

In order to use the AR and MA models, one must first choose a model order, that is, the number of coefficients of the model. Based on the findings of [37], this work employs a model order of 10, that is, 10 coefficients per axis, for both the AR and MA models. Including SSFs, AR, and MA coefficients, a total of 70 features are extracted from the accelerometer sensor data.

**3.2.2. Features from the Pressure Sensor Data.** The pressure sensor embedded in a smartphone measures air pressure. The main reason behind including this sensor into our system is to be able to track altitude, more importantly the relative altitude between different points, as a user performs a physical activity. This information could help us in recognizing the activities which result in an increase or decrease of altitude, such as going upstairs or downstairs. To the best of our knowledge, it is the first time this sensor has been employed in this way in a smartphone-based HAR system. Using the air pressure captured by the pressure sensor, altitude can be calculated as follows [48]:

$$\begin{aligned} h(p_0, p_t) &= \frac{T_0}{L} \left( 1 - \left( \frac{p_t}{p_0} \right)^{RL/gM} \right) \\ &= 44330 \times \left( 1 - \left( \frac{p_t}{p_0} \right)^{1/5.255} \right), \end{aligned} \quad (3)$$

where  $h$  is altitude,  $p_0$  is the standard air pressure at the sea-level,  $p_t$  is the air pressure estimate provided by the pressure sensor at time  $t$ ,  $T_0$  is sea-level standard temperature,  $L$  is the temperature lapse rate,  $R$  is the universal gas constant,  $g$  is the gravitational acceleration, and  $M$  is the molar mass of dry air. Relative altitude can then be calculated as the difference between the altitudes in two points, such as altitude difference between floors in a building, as follows [48]:

$$h_r(a, b) = h(p_0, p_b) - h(p_0, p_a), \quad (4)$$

where  $h_r$  is the relative altitude between points  $a$  and  $b$ . A positive value reflects an increase in the overall altitude, and vice versa. In this study, for every 3.5 second window of pressure/altitude data, altitude differences are calculated between three points, that is, first, middle, and the last samples. Different points were tested before choosing these three points. In conclusion, the altitude feature  $f_{\text{alt}}$  calculated from the pressure sensor data can be expressed as follows:

$$f_{\text{alt}} = [h_r(f, m), h_r(m, l), h_r(f, l)], \quad (5)$$

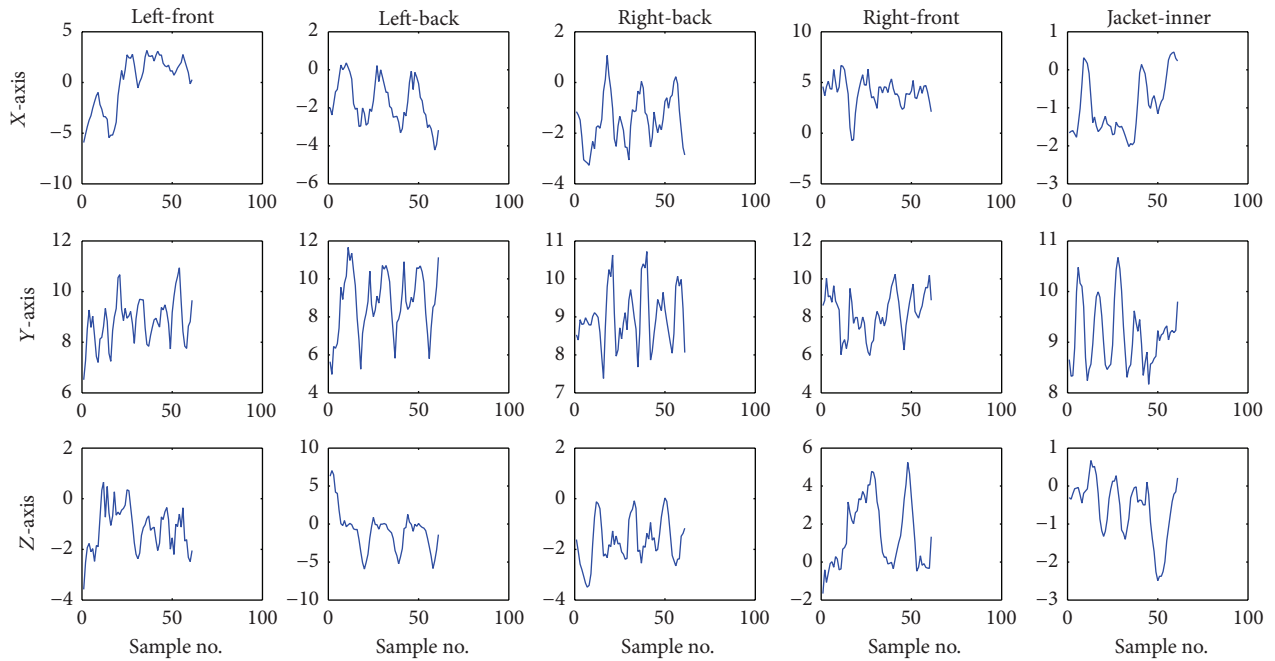


FIGURE 2: Activity acceleration signals for walking from 5 different positions, showing how different the output of a smartphone accelerometer could be for the same activity when carried at different positions.

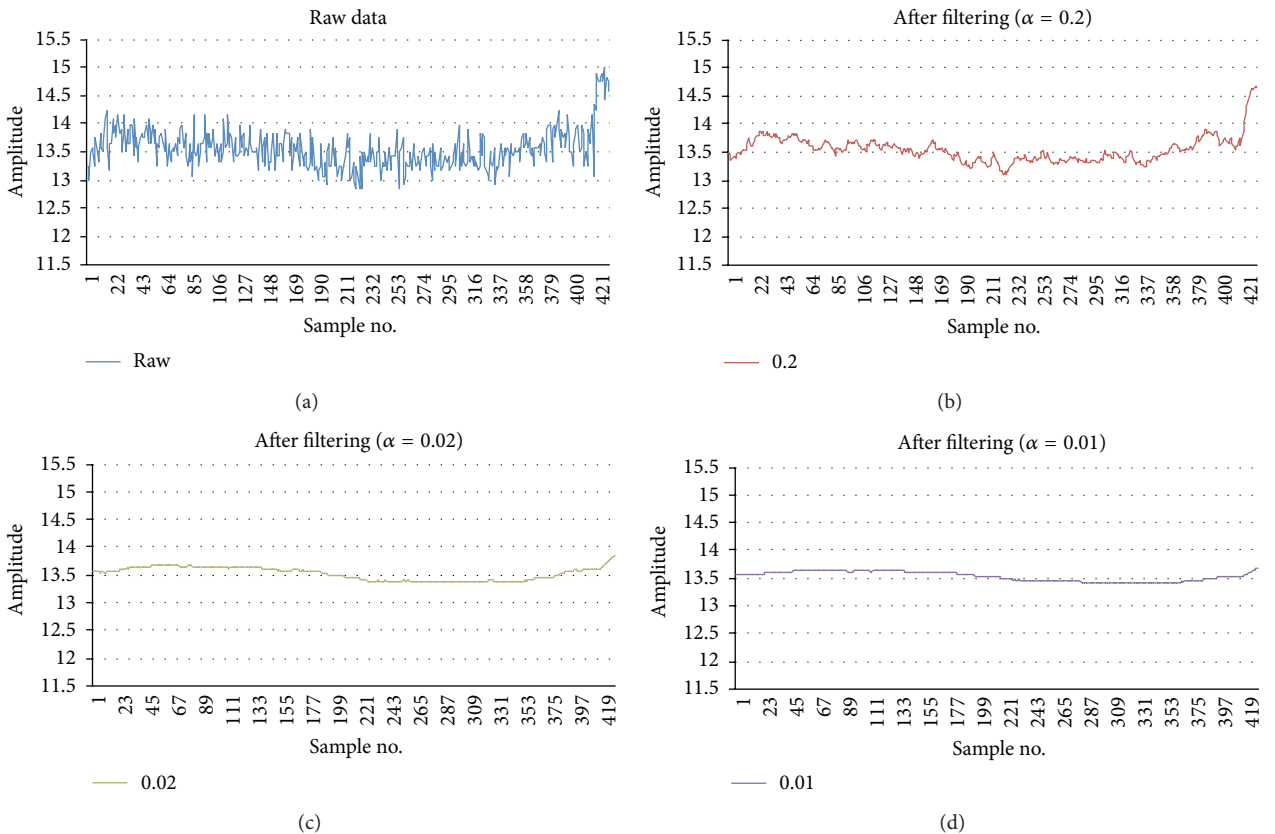


FIGURE 3: After-filtering results for walking (y-axis) for three different values of alpha and their comparison with the original raw data.



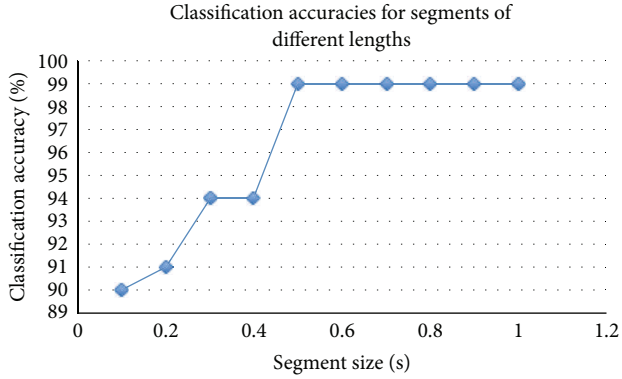


FIGURE 4: Comparison of classification accuracy for segments of different lengths.

where  $f$ ,  $m$ , and  $l$  are the first, middle, and the last samples in a 3.5 second window.

**3.2.3. Feature from the Microphone Data.** The use of microphone for HAR has been studied before [15, 34]. However, these works employed frequency domain features, such as MFC, LPC, and FFT coefficients. The results of [37] showed that in terms of energy consumption, it is less expensive to use time domain features for smartphone-based HAR as compared to frequency domain features. Accordingly, this work uses only the time domain features from the audio data.

For microphone data feature extraction, we followed the same approach as [15]; that is, instead of calculating the features once for the entire window (3.5 seconds), the given window is divided into segments and subsegments of size 0.5 seconds and 50 ms, respectively. The features are then repeatedly calculated for these segments and subsegments. These values were chosen empirically, where classification accuracy was compared for segments and subsegments of different lengths. The results of this study for segments and subsegments of different lengths are shown in Figures 4 and 5, respectively.

In other words, firstly, each 3.5 second window of the audio data is divided into equal segments of 0.5 seconds, with no overlap, which are further divided into subsegments of 50 ms. From each of these subsegments two features are calculated: (1) zero-crossing rate (ZCR) [29], number of zero-crossings within a frame, and (2) Short-time average energy (SAE) [29], computed as the sum of squared amplitudes within a frame. This step generates a total of 140 features from a 3.5 second window. Secondly, mean, variance, min, and max of both ZCR and SAE are calculated for each 0.5 second segment of audio data. This step generates 56 features. In the third and the last step, the same four values are calculated for both ZCR and SAE for the entire 3.5 seconds of audio data. In conclusion, a total of 204 features are calculated from each window of microphone data. The whole process of audio feature extraction is demonstrated in Figure 6.

**3.3. Discriminant Analysis.** It should be noted that accelerometer's and microphone's outputs depend on where the

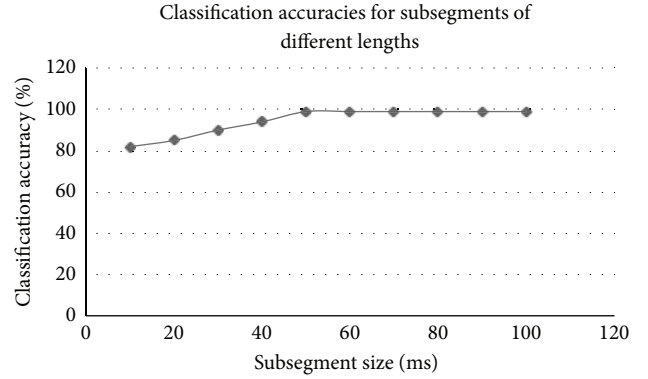


FIGURE 5: Comparison of classification accuracy for subsegments of different lengths.

phone is placed. While performing an activity, a user can carry the phone freely in any pocket or place it on a table or hold it in hands in some cases. This results in high within-class variance. Also, the goal is to recognize a variety of activities using only time domain features calculated from small data windows. Therefore, it is desirable to improve the discriminatory power of these features. Furthermore, as mentioned above, instead of using a particular feature extraction method, the proposed system utilizes various kinds of well-known feature extraction techniques to construct a high number of features. However, using a high number of features does not guarantee better performance. Therefore, it is desirable to reduce the high within-class variance while achieving dimensionality reduction by selecting only the most useful features. Most commonly used techniques for this purpose include principal component analysis (PCA), linear discriminant analysis (LDA) [49], and kernel discriminant analysis (KDA) [49, 50]. The performances of each of these techniques were compared (please see Section 4.2.1), and KDA was chosen for its better performance.

KDA is a nonlinear discriminating approach, which seeks nonlinear discriminating features using kernel techniques. Suppose we have a set of  $m$  feature vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m \in \mathbb{R}^n$  belonging to  $C$  activity classes where  $n$  is the number of dimensions of the feature vector. KDA considers the problem in a feature space  $F$  induced by some nonlinear mapping  $\varphi: \mathbb{R}^n \rightarrow F$ . In this work, after experimenting with various options  $\varphi$  is selected to be the radial basis function (RBF) with Gaussian kernel (please see Section 4.2.1).

For a properly chosen  $\varphi$ , an inner product  $\langle \cdot, \cdot \rangle$  can be defined in  $F$  which makes for the so-called reproducing the kernel Hilbert space. More specifically,  $\langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j) \rangle = K(\mathbf{x}_i, \mathbf{x}_j)$  holds where  $K(\cdot, \cdot)$  is a positive semidefinite kernel function. To find the linear discriminant in  $F$ , the following criterion needs to be maximized:

$$J(\omega) = \frac{\omega^T S_b^\varphi \omega}{\omega^T S_w^\varphi \omega}, \quad (6)$$

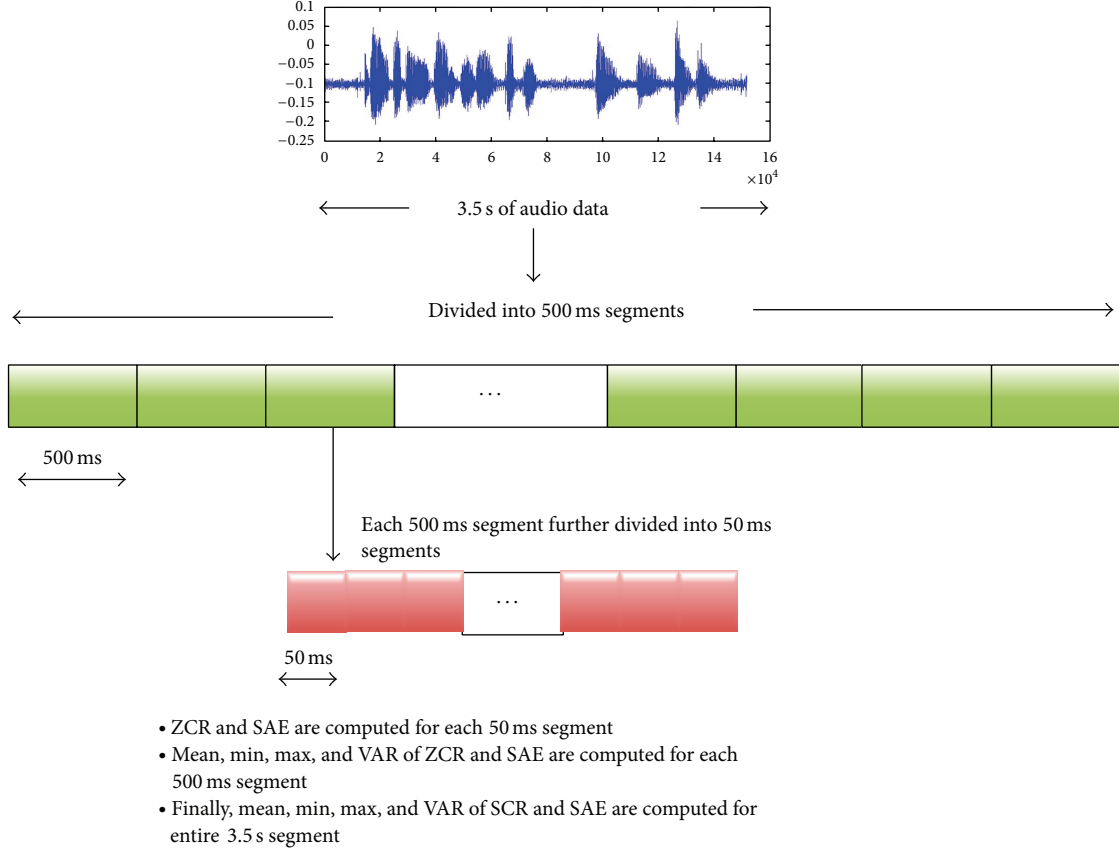


FIGURE 6: Feature extraction from the microphone data. ZCR stands for zero-crossing rate, SAE stands for short-time average energy, whereas VAR stands for variance.

where  $S_b^\varphi$  and  $S_w^\varphi$  are the between-class and within-class scatter matrices in  $F$  and are defined as

$$S_b^\varphi = \sum_{k=1}^C m_k (\mu_\varphi^k - \mu_\varphi) (\mu_\varphi^k - \mu_\varphi)^T, \quad (7)$$

$$S_w^\varphi = \sum_{k=1}^C \left( \sum_{i=1}^{m_k} (\varphi(x_i^k) - \mu_\varphi^k) (\varphi(x_i^k) - \mu_\varphi^k)^T \right),$$

where  $\omega$  is the KDA basis vector.  $\mu_\varphi^k$  and  $\mu_\varphi$  are the mean of the  $k$ th class and the global mean, respectively.  $m_k$  is the number of samples in the  $k$ th class. The solution to (6) is a linear combination of  $\varphi(x_i)$  with coefficients  $\alpha_i$  such that

$$\omega = \sum_{i=1}^m \alpha_i \varphi(x_i). \quad (8)$$

Let  $\alpha = [\alpha_1, \dots, \alpha_m]^T$ , and it can be proved that (8) is equivalent to

$$J(\alpha) = \frac{\alpha^T \mathbf{K} \mathbf{W} \mathbf{K} \alpha}{\alpha^T \mathbf{K} \mathbf{K} \alpha} \quad (9)$$

and the optimal  $\alpha(s)$  are given by the eigenvectors with respect to the maximum eigenvalues of

$$\mathbf{K} \mathbf{W} \mathbf{K} \alpha = \lambda \mathbf{K} \mathbf{K} \alpha, \quad (10)$$

where  $\mathbf{K}$  is the kernel matrix ( $\mathbf{K}_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$ ) and  $\mathbf{W}$  is defined as

$$\mathbf{W}_{ij} = \begin{cases} 1/m_k, & \text{if } \mathbf{x}_i \text{ and } \mathbf{x}_j \text{ belong to } k\text{th class} \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

For a new pattern  $\mathbf{x}$ , its projection onto a KDA basis vector  $\omega$  in  $F$  is calculated as

$$(\omega, \varphi(\mathbf{x})) = \alpha^T K(:, \mathbf{x}), \quad (12)$$

where

$$K(:, \mathbf{x}) = [K(\mathbf{x}_1, \mathbf{x}), \dots, K(\mathbf{x}_m, \mathbf{x})]^T. \quad (13)$$

For more details on KDA please refer to [50]. In summary, the discriminant analysis step produces the KDA feature vector that can be considered as a mapping from a linear feature space to a nonlinear feature space with better discriminatory capabilities.

**3.4. Classifier.** The proposed system employs SVMs for activity classification. The selection of SVMs is based on a comparison study, which is explained in Section 4.2. SVMs are a very powerful data classification technique. The working of SVMs is based on finding the best separating hyperplane. It is the plane with maximum margins between the two classes

of the training samples within the feature space. For this, SVMs focus on the training cases placed at the edge of the class descriptors. In this way, not only an optimal hyperplane is fitted but also less training samples are effectively used; thus high classification accuracy is achieved with small training sets [51].

The traditional SVM algorithm can be summarized as the following optimization problem:

$$\begin{aligned} \min_{w,b,\xi} \quad & \left\{ \frac{1}{2} w^T w + C \left( \sum_{i=1}^l \xi_i \right) \right\} \\ \text{subject to:} \quad & y_i (w^T \phi(x_i) + b) \geq 1 - \xi_i, \\ & \xi_i \geq 0, \end{aligned} \quad (14)$$

where  $x_i$  are the training vectors that are mapped into a higher dimensional space using  $\phi(x)$  which is the kernel function.  $w$ ,  $b$ ,  $\xi_i$ , and  $C$  are the weight vector, bias, slack variable, and the penalty error term, respectively. For more details on SVM classification, please refer to [52].

## 4. System Validation

**4.1. Data Acquisition.** As mentioned earlier, a total of 15 activities were selected for this research study. Data on these activities were collected in an unsupervised study using 30 healthy subjects. The subjects consisted of 18 males and 12 females between the ages of 26 and 35 years old, with an average height of 172.4 cm and average weight of 64 kg. The data were collected using an Android operating system based mobile phone called the LG NEXUS 4. Ten different NEXUS 4 mobile phones were used. Each user used a single phone at a time. Table 1 provides a linking for mobile phones to each user.

A custom-built application was used for data collection and annotation. The application and a brief description on how to use the application are available for download at <http://www.ubilife.net/ad/>. Sample data will also be made available at the same web address in near future. The application has a simple user interface that allows subjects to enter their names, choose the activity label, and start/stop the data collection. Subjects were trained on the use of this application before data collection. They were also requested to perform the activities in a natural way, without any fixed duration or sequence. Each subject then collected the activity data at their homes for a period of one month.

For accelerometer and pressure sensor, a sampling rate of 50 Hz was used; whereas, the microphone data were captured at the rate of 8000 Hz. In order to ensure position-free recognition of activities, subjects were allowed to carry phone freely in any pocket, including trousers' (both front and back) pockets or jacket's (both inner and outer) pockets. Furthermore, they were also allowed to place the phone on a table or simply carry it in hands while being idle (sitting/standing) or watching TV.

Initially the data were stored on the SD cards. Later the data were moved to a computer for further analysis. We collected over 33,000 instances on 15 activities. Approximately,

TABLE 1: Linking between mobile phones and users.

Mobile Phone ID	User ID
P_1	U_1, U_11, U_21
P_2	U_2, U_12, U_22
P_3	U_3, U_13, U_23
P_4	U_4, U_14, U_24
P_5	U_5, U_15, U_25
P_6	U_6, U_16, U_26
P_7	U_7, U_17, U_27
P_8	U_8, U_18, U_28
P_9	U_9, U_19, U_29
P_10	U_10, U_20, U_30

3000 instances were collected on each activity, except going upstairs and downstairs (1205 & 1513), elevator up and down (1008 & 1213), hopping (1205), vacuuming (1134), and riding bike (2120) instances. Also, 1500 instances were collected for cases which were not among the 15 activities and labeled as others.

**4.2. Experimental Results and Discussion.** Five experiments were performed: (1) offline recognition via 10-fold cross validations (subject-dependent); (2) subject-independent offline recognition via leave-one-subject-out; (3) offline evaluation of different sensors and combination of sensors; (4) subject-independent online recognition on smartphones using eight new subjects; and (5) comparison with some of the previous mobile phone based HAR systems. The term subject-independent means that the classifier did not see any data from the test subjects during the training phase, and vice versa. The results for these experiments are discussed below.

**4.2.1. Offline Recognition via 10-Fold Cross Validations.** The purpose of this experiment was four-fold: (a) compare the performance of different kernel functions for KDA; (b) compare the performance of different discriminant analysis techniques; (c) compare the performance of employing KDA versus using all the features; and (b) compare the performance of different classifiers.

For this experiment, the training data were divided into 10 subsets. Of these, the data from one subset were retained as the validation data, whereas the data from the remaining nine subsets were used as the training data. The whole process was repeated 10 times, each time picking a different subset as the validation subset.

For (a), four different kernel functions were compared for KDA, that is, linear, RBF (Gaussian), polynomial, and sigmoid. The results are summarized in Figure 7. It can be seen that the RBF provided the best accuracy. For (b), three techniques were compared, that is, PCA, LDA, and KDA (with RBF). The results are summarized in Figure 8; KDA outperformed the other two in this case.

For (c) and (d), each time, two sets of three different classifiers, that is, ANNs, GMMs, and SVMs, were trained. The first set was trained using all the features, whereas the



TABLE 2: Average recognition rates for the three classifiers for subject-dependent offline recognition test, using All-Features and KDA-Features (Unit: %).

Activities	ANNs		GMMs		SVMs	
	All Features	KDA Features	All Features	KDA Features	All Features	KDA Features
Walking	70	85	69	84	78	98
Walking on Treadmill	73	86	72	85	81	99
Running	82	99	82	96	87	100
Running on Treadmill	85	99	85	99	86	100
Going Upstairs	70	83	68	80	84	99
Going Downstairs	70	85	65	80	83	99
Elevator Up	80	92	78	92	87	99
Elevator Down	78	93	75	92	86	99
Riding a Bike	72	95	69	93	84	99
Hopping	72	92	72	90	84	99
Idle (Sitting/Standing)	78	95	76	95	89	100
Watching TV	72	94	67	91	82	99
Vacuuming	76	96	69	93	84	100
Driving a Car	72	93	65	90	82	98
Riding a Bus	78	92	72	90	83	99
Others	79	91	72	90	84	99
Total	75	92	72	90	84	99

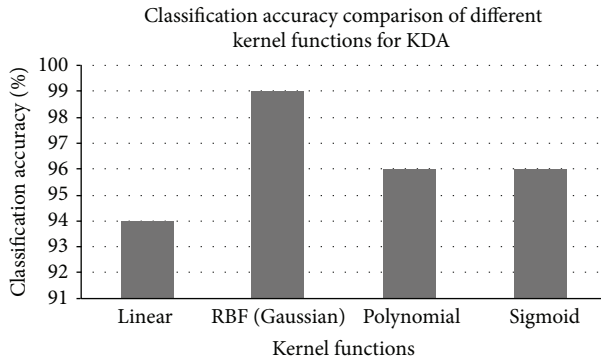


FIGURE 7: Comparison of classification accuracy for different kernel functions for KDA.

second set was trained using the KDA features. The results for this experiment are summarized in Table 2.

For the ANNs, each network had  $n$  input neurons (where  $n$  corresponds to the dimensions of the input feature vector), one hidden layer with three neurons, and 15 output neurons corresponding to the 15 activities. A different number of hidden layers and neurons were tested for each ANN to optimize the accuracy, and at the end, the given settings were chosen. As for the GMMs, we followed the details given in [34].

It can be noticed that using all features did not result in high accuracy with any of the classifiers. However, when coupled with KDA, the accuracy for all the classifiers got better, especially that of SVMs with the average accuracy for 99.1%. In conclusion, KDA was able to learn discriminative feature transformation which was able to help each classifier

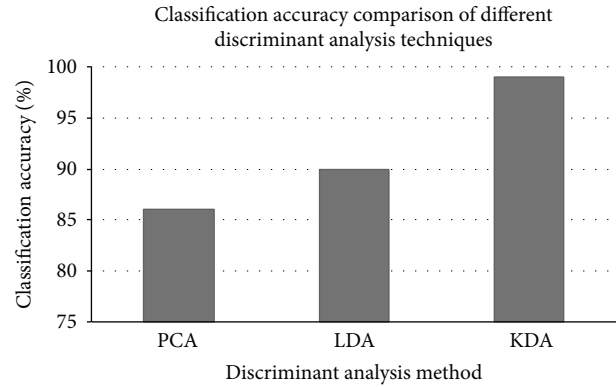


FIGURE 8: Comparison of classification accuracy for different discriminant analysis techniques.

in discriminating 15 activities more accurately, especially the SVMs.

**4.2.2. Offline Recognition via Leave-One-Subject-Out.** One major goal of any context-aware system is to be able to recognize context of those who it has not seen before, in other words, the ability of the system to recognize context of new users. Accordingly, the purpose of this experiment was twofold: (a) compare the performance of employing KDA versus using all the features for subject-independent recognition and (b) compare the performance of different classifiers for subject-independent recognition.

This experiment was repeated 30 times. Each time, the data from one subject were retained as the validation data, whereas the data from the remaining 29 subjects were used as the training data. The whole process was repeated 30 times, each time picking a different subject as the validation subject.

Each time, two sets of three different classifiers, that is, ANNs, GMMs, and SVMs, were trained. The first set was trained using all the features, whereas the second set was trained using the KDA features. The results for this experiment are summarized in Table 3.

It is clear that subject-independent case is more difficult than subject-dependent case, as the performance for all the three classifiers decreased compared to the previous case. However, the KDA-with-SVMs still provided the best performance in this experiment, with an average accuracy of 94%.

**4.2.3. Evaluation of Different Sensors and Their Combinations.** The purpose of this experiment was to evaluate the performance of different sensors and their combinations using the SVM and KDA features, in order to assess which is the best system.

Seven settings were evaluated in total. The overall recognition results are summarized in Figure 9. In summary, when used alone, though none of the sensors produced high recognition accuracy, accelerometer sensor performed better than the other two sensors. It is mainly due to the fact that most of the activities involved body movement. For the two-sensor combination, acc-mic setting produced the best results. However, the optimum performance was achieved when all the sensors were used at the same time.

In conclusion:

- (i) when considering a large number of activities in an activity recognition problem, a single sensor might not produce high recognition accuracy. It is mainly due to the fact that activities appear similar if only one source of information is considered, such as watching TV versus idle, walking/running on treadmill versus normal walking/running, and going upstairs/downstairs versus walking.
- (ii) In other words, the use of one sensor limits the scope of the recognition problem, that is, the number of activity classes that can be recognized, or the accuracy of the recognition.
- (iii) In such scenarios, combining or fusing the information from different sensors can improve the accuracy.

**4.2.4. Online Recognition via Eight New Subjects.** Since KDA-based SVMs provided the best performance in the offline recognition tests; they were selected as the recognition model for online recognition tests. It was a subject-independent recognition test. Under this setting, eight new subjects (four males and four females) were recruited. These subjects belonged to different age groups: two subjects had the same age group as the subjects who collected the training data, two were aged between 18 and 20 years old, whereas the rest of the four subjects were between the ages of 45 and 50 years old.

The subjects carried phones with a custom-built Android application for capturing the acceleration data, computing the features, classifying the activity, and storing the true label, as well as the classified label, in a database. Finally, the recognition accuracy was evaluated by comparing the

recognized labels for the activities with their true labels. These results are summarized in Table 4.

One can argue that the recognition accuracy for our online tests is very similar to [34]; however, it should be noted that this work considers the problem of recognizing 15 human activities, as opposed to [34], where the researchers targeted only three physical activities and two transportation modes. Furthermore, the proposed system uses only time domain features which are less resource hungry in contrast to frequency domain features. Also, our online test was completely subject-independent with eight new subjects, none of whom took part in data collection. Lastly, the proposed method does not use the GPS sensor in any scenario, and it is an obvious fact that the GPS sensor is one of the most battery hungry sensors in a smartphone.

**4.2.5. Comparison with Previous Mobile Phone-Based HAR Systems.** It is difficult to compare our system with other mobile phone-based HAR systems because each system employed a different set of sensors and had its own experimental plan, data collection, and reporting mechanism. Nevertheless, we picked four recent studies that were the most similar to our work for comparison. These are [17, 18, 34, 35].

In [17], authors focused on the recognition of basic locomotion activities using the accelerometer, gyroscope, and magnetic field sensor embedded in smartphones. In [18], the authors proposed an evolutionary fuzzy model to recognize dynamic activities using smartphone accelerometers; whereas accelerometer and microphone were used in [34] and [35] to recognize the basic activities of daily living.

All of these methods were implemented using the instructions given in their respective papers. During the experiment, 10 subjects were used to collect data on the 15 activities using our custom-built Android application. All of the sensors used in these studies as well as in our work were active during the data collection. The data were then transferred to a computer for recognition accuracy comparison and an  $n$ -fold cross-validation rule (based on subjects) was applied. The average recognition results for all of these systems, along with their respective time windows, are shown in Figure 10. It can be seen that the proposed system outperformed the existing mobile phone-based HAR systems in terms of accuracy using a window of just 3.5 seconds.

The recognition accuracies of [17, 18, 34, 35] in this experiment were lower than the values presented by their corresponding authors in the original papers, that is, 70%, 61%, 81%, and 62% compared to 96.82%, 95%, 92.43%, and 78%, respectively. We believe that this drop in accuracy is due to two factors: (a) more complex classification problem, that is, 15 activities in contrast to 5 [17], 7 [18], 8 [34], and 6 activities [35]; and (b) carrying phones on different positions in contrast to attaching or carrying them on a single position.

## 5. Conclusion

This study aimed at implementing an accurate and robust HAR system for smartphones with three characteristics: (1) instead of targeting a small number of activities, the system

TABLE 3: Average recognition rates for the three classifiers for subject-independent offline recognition test, using All-Features and KDA-Features (Unit: %).

Activities	ANNs		GMMs		SVMs	
	All Features	KDA Features	All Features	KDA Features	All Features	KDA Features
Walking	62	74	57	72	68	90
Walking on Treadmill	63	78	62	78	72	94
Running	75	90	71	88	76	94
Running on Treadmill	76	92	74	91	75	95
Going Upstairs	64	75	60	72	73	95
Going Downstairs	60	75	58	70	72	94
Elevator Up	65	82	62	82	75	95
Elevator Down	63	83	60	82	75	95
Riding a Bike	61	86	59	81	74	94
Hopping	60	83	60	81	74	95
Idle (Sitting/Standing)	65	89	62	86	78	100
Watching TV	61	84	70	82	72	94
Vacuuming	60	85	59	82	74	98
Driving a Car	62	80	56	81	72	95
Riding a Bus	65	82	58	82	73	96
Others	64	82	62	81	73	95
Total	64	82	62	81	73	94

TABLE 4: Average recognition rates for KDA-based SVMs for subject-independent online recognition test (Unit: %).

Activities	Subjects								Total
	1	2	3	4	5	6	7	8	
Walking	90	91	88	89	90	88	90	90	89.5
Walking Treadmill	90	91	90	92	90	92	90	90	90.6
Running	92	93	92	94	94	92	92	90	92.3
Running Treadmill	94	95	94	93	95	92	92	92	93.3
Going Upstairs	92	91	92	92	93	91	90	90	91.3
Going Downstairs	92	92	91	91	92	93	91	91	91.6
Elevator Up	94	95	94	92	93	92	94	92	93.2
Elevator Down	94	94	91	92	92	91	93	92	92.3
Riding a Bike	92	92	93	92	93	91	92	91	92
Hopping	94	93	92	93	94	92	91	90	92.3
Idle	98	99	99	100	100	99	98	99	99
Watching TV	91	92	92	93	94	91	91	90	91.7
Vacuuming	94	95	92	93	92	94	93	92	93.1
Driving a Car	92	94	92	93	92	91	92	91	92.1
Riding a Bus	93	94	92	92	91	93	92	92	92.3
Others	93	93	92.7	92.7	93	92	92	91	92
Total	92.8	93.4	92.2	92.7	93	92.1	92	91.4	92.4

should be able to recognize a large variety of activities using a single device; (2) it should be able to provide high recognition accuracy while allowing its users the freedom to keep the device in any pocket or hold it in hands in some cases; (3) the system should be less power hungry and should have a decent response time.

In order to achieve the first goal, the proposed system incorporates a sensor fusion approach to HAR, where three smartphone sensors, that is, the triaxial accelerometer, the pressure sensor, and the microphone, are used to recognize 15

activities. For the second characteristic, the system employs a nonlinear discriminatory approach (KDA) together with a nonlinear classifier (SVMs). This helps in finding nonlinear discriminating features while reducing the high within-class variance that results from carrying phones at different positions. Lastly, the proposed system uses only time domain features from the three sensors, computed from small data windows, in order to preserve phone's battery and achieve a better response time, respectively. Average recognition accuracies of 99.1% in offline subject-dependent test, 94% in

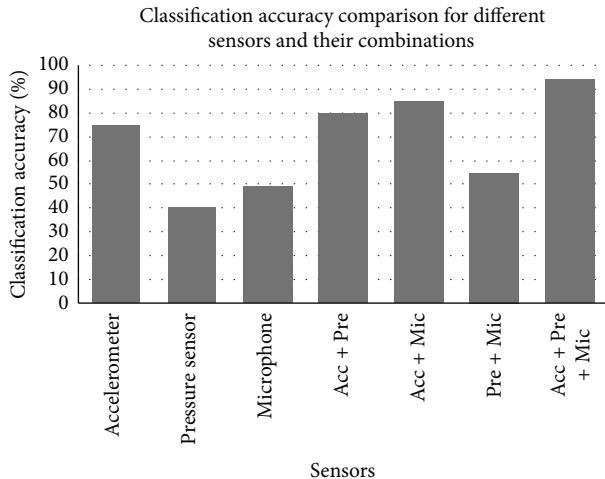


FIGURE 9: Comparison of classification accuracy for different sensors and their combinations.

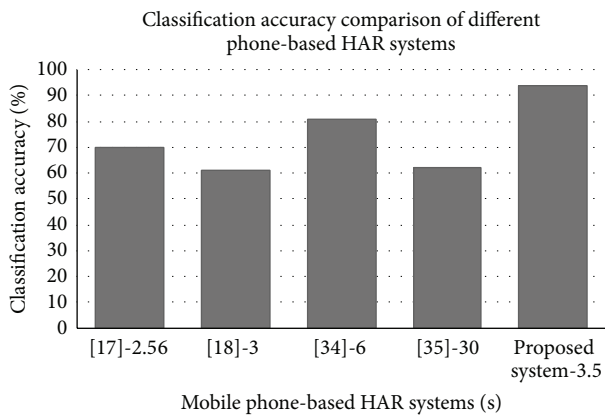


FIGURE 10: Comparison of classification accuracy of different mobile-phone based HAR systems.

offline subject-independent test, and 92.4% in online subject-independent test show the feasibility of using this system for long-term monitoring of human activities using smartphones in free-living conditions.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

This work was supported by the New Faculty Research Fund of Ajou University (2012).

## References

[1] M. Weiser, "The computer for the 21st century," in *Human-Computer Interaction*, R. M. Baecker, J. Grudin, W. A. S. Buxton, and S. Greenberg, Eds., pp. 933–940, Morgan

Kaufmann Publishers Inc., San Francisco, CA, USA, 1995, <http://dl.acm.org/citation.cfm?id=212925.213017>.

- [2] W. Liu, X. Li, and D. Huang, "A survey on context awareness," in *Proceedings of the International Conference on Computer Science and Service System (CSSS '11)*, pp. 144–147, June 2011.
- [3] S. Vishwakarma and A. Agrawal, "A survey on activity recognition and behavior understanding in video surveillance," *The Visual Computer*, vol. 29, no. 10, pp. 983–1009, 2013.
- [4] A. Tolstikov, X. Hong, J. Biswas, C. Nugent, L. Chen, and G. Parente, "Comparison of fusion methods based on DST and DBN in human activity recognition," *Journal of Control Theory and Applications*, vol. 9, no. 1, pp. 18–27, 2011.
- [5] J. Yang, J. Lee, and J. Choi, "Activity recognition based on RFID object usage for smart mobile devices," *Journal of Computer Science and Technology*, vol. 26, no. 2, pp. 239–246, 2011.
- [6] J. Sarkar, L. T. Vinh, Y.-K. Lee, and S. Lee, "GPARS: a general-purpose activity recognition system," *Applied Intelligence*, vol. 35, no. 2, pp. 242–259, 2011.
- [7] O. Lara and M. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.
- [8] M. J. Mathie, A. C. F. Coster, N. H. Lovell, and B. G. Celler, "Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement," *Physiological Measurement*, vol. 25, no. 2, pp. R1–R20, 2004.
- [9] A. M. Khan, Y. K. Lee, and T.-S. Kim, "Accelerometer signal-based human activity recognition using augmented autoregressive model coefficients and artificial neural nets," in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '08)*, pp. 5172–5175, August 2008.
- [10] T. P. Kao, C. W. Lin, and J. S. Wang, "Development of a portable activity detector for daily activity recognition," in *Proceedings of the IEEE International Symposium on Industrial Electronics (ISIE '09)*, pp. 115–120, July 2009.
- [11] L. C. Jatobá, U. Großmann, C. Kunze, J. Ottenbacher, and W. Stork, "Context-aware mobile health monitoring: evaluation of different pattern recognition methods for classification of physical activity," in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '08)*, pp. 5250–5253, August 2008.
- [12] M. Helmi and S. M. T. AlModarresi, "Human activity recognition using a fuzzy inference system," in *Proceedings of the IEEE International Conference on Fuzzy Systems*, pp. 1897–1902, August 2009.
- [13] M. Zhang and A. A. Sawchuk, "A bag-of-features-based framework for human activity representation and recognition," in *Proceedings of the International Workshop on Situation Activity & Goal Awareness*, pp. 51–55, ACM, September 2011.
- [14] P. Casale, O. Pujol, and P. Radeva, "Human activity recognition from accelerometer data using a wearable device," in *Pattern Recognition and Image Analysis*, vol. 6669 of *Lecture Notes in Computer Science*, pp. 289–296, 2011.
- [15] J. A. Ward, P. Lukowicz, G. Tröster, and T. E. Starner, "Activity recognition of assembly tasks using body-worn microphones and accelerometers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1553–1567, 2006.
- [16] L. Gao, A. Bourke, and J. Nelson, "Activity recognition using dynamic multiple sensor fusion in body sensor networks," in *Proceedings of the IEEE Annual International Conference*, pp. 1077–1080, Engineering in Medicine and Biology Society (EMBC), 2012.



- [17] Y. E. Ustev, O. Durmaz Incel, and C. Ersoy, "User, device and orientation independent human activity recognition on mobile phones: challenges and a proposal," in *Proceedings of the ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, pp. 1427–1436, ACM, New York, NY, USA, 2013.
- [18] M. Fahim, I. Fatima, S. Lee, and Y. T. Park, "Efm: evolutionary fuzzy model for dynamic activities recognition using a smart-phone accelerometer," *Applied Intelligence*, pp. 1–14, 2013.
- [19] O. D. Lara and M. A. Labrador, "A mobile platform for real-time human activity recognition," in *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC '12)*, pp. 667–671, IEEE, 2012.
- [20] M. Berchtold, M. Budde, D. Gordon, H. R. Schmidtke, and M. Beigl, "ActiServ: activity recognition service for mobile phones," in *Proceedings of the IEEE International Symposium on Wearable Computers (ISWC '10)*, pp. 1–9, October 2010.
- [21] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.
- [22] N. Györbiró, Á. Fábrián, and G. Hományi, "An activity recognition system for mobile phones," *Mobile Networks and Applications*, vol. 14, no. 1, pp. 82–91, 2009.
- [23] H. Lu, J. Yang, Z. Liu, N. D. Lane, T. Choudhury, and A. T. Campbell, "The Jigsaw continuous sensing engine for mobile phone applications," in *Proceedings of the 8th ACM International Conference on Embedded Networked Sensor Systems (SenSys '10)*, pp. 71–84, ACM, New York, NY, USA, November 2010.
- [24] L. Sun, D. Zhang, B. Li, B. Guo, and S. Li, "Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations," in *Proceedings of the 7th international conference on Ubiquitous intelligence and computing (UIC '10)*, pp. 548–562, Springer, Berlin, Germany, <http://dl.acm.org/citation.cfm?id=1929661.1929712>.
- [25] S. Wang, C. Chen, and J. Ma, "Accelerometer based transportation mode recognition on mobile phones," in *Proceedings of the Asia-Pacific Conference on Wearable Computing Systems (APWCS '10)*, pp. 44–46, IEEE, Washington, DC, USA, April 2010.
- [26] P. Siirtola and J. Röning, "Recognizing human activities user-independently on smartphones based on accelerometer data," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 1, no. 5, pp. 38–45, 2012.
- [27] J. Frank, S. Mannor, and D. Precup, "Activity recognition with mobile phones," in *Machine Learning and Knowledge Discovery in Databases*, vol. 6913 of *Lecture Notes in Computer Science*, no. 3, pp. 630–633, 2011.
- [28] J. B. Gomes, S. Krishnaswamy, M. M. Gaber, P. A. C. Sousa, and E. Menasalvas, "Mars: a personalised mobile activity recognition system," in *Proceedings of the IEEE 13th International Conference on Mobile Data Management (MDM '12)*, pp. 316–319, IEEE, Washington, DC, USA.
- [29] A. Eronen, V. Peltonen, J. Tuomi et al., "Audio-based context recognition," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, no. 1, pp. 321–329, 2006.
- [30] H. Lu, W. Pan, N. D. Lane, T. Choudhury, and A. T. Campbell, "SoundSense: Scalable sound sensing for people-centric applications on mobile phones," in *Proceedings of the 7th ACM International Conference on Mobile Systems, Applications, and Services (MobiSys '09)*, pp. 165–178, ACM, New York, NY, USA, June 2009.
- [31] L. Liao, D. Fox, and H. Kautz, "Extracting places and activities from GPS traces using hierarchical conditional random fields," *International Journal of Robotics Research*, vol. 26, no. 1, pp. 119–134, 2007.
- [32] "Location-based activity recognition," in *In Advances in Neural Information Processing Systems (NIPS)*, pp. 787–794, MIT Press, Boston, Mass, USA, 2005.
- [33] A. Thiagarajan, J. Biagioni, T. Gerlich, and J. Eriksson, "Cooperative transit tracking using smart-phones," in *Proceedings of the 8th ACM International Conference on Embedded Networked Sensor Systems (SenSys '10)*, pp. 85–98, ACM, New York, NY, USA, November 2010.
- [34] M. Han, L. T. Vinh, Y. K. Lee, and S. Lee, "Comprehensive context recognizer based on multimodal sensors in a smartphone," *Sensors*, vol. 12, no. 9, pp. 12588–12605, 2012.
- [35] E. Miluzzo, N. D. Lane, K. Fodor et al., "Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application," in *Proceedings of the International Conference on Embedded Networked Sensor Systems (SenSys '08)*, pp. 337–350, ACM, 2008.
- [36] O. Banos, M. Damas, H. Pomares, and I. Rojas, "Activity recognition based on a multi-sensor meta-classifier," in *Proceedings of the 12th International Conference on Artificial Neural Networks: Advances in Computational Intelligence*, pp. 208–215, Springer, Berlin, Germany, 2013.
- [37] A. M. Khan, M. H. Siddiqi, and S. W. Lee, "Exploratory data analysis of acceleration signals to select light-weight and accurate features for real-time activity recognition on smartphones," *Sensors*, vol. 13, no. 10, pp. 13099–13122, 2013.
- [38] Android SensorEvent: Isolating the Force of Gravity Using a Low-Pass Filter, 2014, <http://developer.android.com/reference/android/hardware/SensorEvent.html#values>.
- [39] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing*, vol. 3001 of *Lecture Notes in Computer Science*, pp. 1–17, 2004.
- [40] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Proceedings of the national conference on artificial intelligence*, pp. 1541–1546, AAAI Press, MIT Press, July 2005.
- [41] S. Wang, J. Yang, N. Chen, X. Chen, and Q. Zhang, "Human activity recognition with user-free accelerometers in the sensor networks," in *Proceedings of the International Conference on Neural Networks and Brain Proceedings (ICNNB '05)*, vol. 2, pp. 1212–1217, IEEE, October 2005.
- [42] M. J. Mathie, A. C. F. Coster, N. H. Lovell, B. G. Celler, S. R. Lord, and A. Tiedemann, "A pilot study of long-term monitoring of human movements in the home using accelerometry," *Journal of Telemedicine and Telecare*, vol. 10, no. 3, pp. 144–151, 2004.
- [43] T. Huynh and B. Schiele, "Analyzing features for activity recognition," in *Proceedings of the Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-Aware Services: Usages and Technologies*, pp. 159–163, ACM, 2005.
- [44] Y. P. Chen, J. Y. Yang, S. N. Liou, G. Y. Lee, and J. S. Wang, "Online classifier construction algorithm for human activity detection using a tri-axial accelerometer," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 849–860, 2008.
- [45] J. Mäntyjärvi, M. Lindholm, E. Vildjiounaite, S. M. Mäkelä, and H. Ailisto, "Identifying users of portable devices from gait pattern with accelerometers," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 2, pp. ii973–ii976, March 2005.
- [46] Z. He and L. Jin, "Activity recognition from acceleration data based on discrete cosine transform and SVM," in *Proceedings*



- of the *IEEE International Conference on Systems, Man and Cybernetics (SMC '09)*, pp. 5041–5044, October 2009.
- [47] M. Natrella, e-handbook of statistical methods, 2010, <http://www.itl.nist.gov/div898/handbook/eda/eda.htm>.
- [48] A. S. Greg Milette, *Professional Android Sensor Programming*, John Wiley & Sons, New York, NY, USA, 2012.
- [49] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K. R. Muller, “Fisher discriminant analysis with kernels,” in *Proceedings of the 9th IEEE Workshop on Neural Networks for Signal Processing (NNSP '99)*, pp. 41–48, August 1999.
- [50] G. Baudat and F. Anouar, “Generalized discriminant analysis using a kernel approach,” *Neural Computation*, vol. 12, no. 10, pp. 2385–2404, 2000.
- [51] C. Wei Hsu, C. Chung Chang, and C. Jen Lin, “A practical guide to support vector classification,” 2010.
- [52] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “Training algorithm for optimal margin classifiers,” in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory (COLT '92)*, pp. 144–152, ACM, New York, NY, USA, July 1992.