A feasibility study on smartphone accelerometer-based recognition of household activities and influence of smartphone position

Vincenzo Della Mea[1], Omar Quattrin[1], Maria Parpinel[2]

1) *Dept. of Mathematics and Computer Science, University of Udine, Italy*

2) *Dept. of Medical and Biological Sciences, University of Udine, Italy*

CORRESPONDING AUTHOR:

Vincenzo Della Mea
Dept. of Mathematics and Computer Science, University of Udine
Via delle Scienze 206, 33100 UDINE, Italy
Phone: +39-0432-558461
Fax: +39-0432-558499
Email: Vincenzo.dellamea@uniud.it

KEYWORDS

Activity recognition; physical activity; smartphone; accelerometer;household activities; mobile

applications

RUNNING HEAD

Smartphone-based household activities recognition

ABSTRACT

BACKGROUND: Obesity and physical inactivity are the most important risk factors for chronic diseases. The present study aimed at: i) developing and testing a method for classifying household activities based on a smartphone accelerometer; ii) evaluating the influence of smartphone position; and iii) evaluating the acceptability of wearing a smartphone for activity recognition.

METHODS: An Android application was developed to record accelerometer data and calculate descriptive features on 5-second time blocks, then classified with 9 algorithms. Household activities were: sitting, working at the computer, walking, ironing, sweeping the floor, going down stairs with a shopping bag, walking while carrying a large box, and climbing stairs with a shopping bag. Ten volunteers carried out the activities for three times, each one with a smartphone in a different position (pocket, arm, and wrist). Users were then asked to answer a questionnaire.

RESULTS: 1440 time blocks were collected. Three algorithms demonstrated an accuracy greater than 80% for all smartphone positions. While for some subjects the smartphone was uncomfortable, it seems that it did not really condition the activities.

CONCLUSIONS: Smartphones can be used to recognise household activities. A further development is to measure metabolic equivalent tasks starting from accelerometer data only.

1 Introduction

In the last 50 years, the changing of lifestyle has produced an increase in over-nutrition and sedentary habits. Epidemiological studies have demonstrated that obesity and physical inactivity are the most important risk factors for chronic diseases [1,2]. In 2004, 5.5% and 4.8% of deaths

globally were caused by physical inactivity and overweight and obesity, respectively [3]. Recently, the metabolic syndrome (MS) has been defined as a medical condition characterized by excess of adiposity (especially abdominal), glucose intolerance, high levels of haematic lipids, and high blood pressure. In Italy, the prevalence of MS ranges from 3% in subjects between 20 and 29 years to 25% in subjects over 70 years [4]. Two major international associations (the International Diabetes Association and the American Heart Association) agree that the development of MS is primarily due to obesity and limited or no physical activity and that MS is a cluster of metabolic risk factors for chronic diseases [5]. Behaviours associated with MS risk factors are dietary habits and physical activity, which can be evaluated with specific questionnaires to indirectly estimate energy intake and expenditure [6,7]. In the present paper we deal only with physical activity and related energy expenditure.

The gold-standard methods to calculate energy expenditure are doubly labelled water methodology and direct/indirect calorimetry, but they are feasible only in the research settings because they are expensive and require specific equipment [7]. Thus, a shortcut to evaluate an individual physical activity level is to use questionnaires, interviews and activity diaries [8].

The most common and validated instrument to collect information on physical activity is the International Physical Activity Questionnaire (IPAQ) [9,10], which is aimed at recollecting the activities of the last week to measure their intensity through metabolic equivalent task (MET) minutes. One MET is considered as the resting metabolic rate obtained during quiet sitting; energy expenditures of other activities are expressed as multiples of the resting level MET and they range from 0.9 (when sleeping) to 18 (when running at 17.54 km/h). The IPAQ method allows to calculate two scores related to the physical activity carried out in a typical week: a categorical score (low, moderate, high) and a continuous score (MET minutes per week), corresponding to the METs spent summing up all activities times per their MET value.

A first classification of physical activities accompanied by their METs was published in the Compendium of Physical Activities in 1993 [11]. The Compendium was subsequently updated, the last update being published in 2011 [12]. The complete Compendium is available online [13], and is currently used to standardize activities in many epidemiological studies.

Scores calculated using IPAQ represent only a standardized estimate, because the Compendium was not developed to determine the precise energy cost of activities within an individual. However, Byrne et al. [14] and Kozey et al. [15] have developed methods to consider personal variations in sex, body mass, height, and age for correcting MET values.

While questionnaires are typically administered to users by professionals, they are indeed a suitable tool for patient empowerment [16], i.e., active involvement of the person/patient in their own care, including directly recording their own health data, also backed by evidence on significance of patient provided data [17].

More recently, the availability of smartphones and tablets has led to the so-called m-Health (mobile health), based on phones, short text messaging, mobile web access, up to the most recent smartphone applications [18]. In particular, the effectiveness of mobile systems when used for collecting patient diaries has been demonstrated [19]. All these tools may be seen as a new way to strengthen the patient-physician relationship [20], in particular in the case of chronic diseases, which develop on the long term and have consequences on ageing [21].

In fact, many studies have been published in the literature on the use of wearable accelerometers for recognizing activities being done by the smartphone owner.

Most of the research available until now has been carried out using more than one accelerometer in different body parts. In one of the seminal studies, Foerster et al. [22] exploited 5 accelerometers and showed that two were needed to recognize the basic situations (sitting, lying,

standing, and moving) and more were needed for more specific movement subtypes. Five accelerometers, but of a biaxial type, were also used by Bao et al. [23], who demonstrated that those placed on the thigh and the dominant wrist allowed to distinguish among a number of basic activities. A similar study has been recently reproduced with a different classifier [24], confirming previous results.

In the last few years, a growing number of papers have focused on the use of a single accelerometer. Khan et al. [25] used one triaxial accelerometer on the subject's chest and obtained good recognition accuracy, even for activities such as climbing and going down stairs. Lee et al. [26] obtained similar results placing an accelerometer on the left waist. Long et al., in addition to walking and running, considered also bicycling, driving and generic sports [27]. Driving was easily recognizable, while the others were more complex due to the accelerometer position (on the side) and the variety of movements. Ravi et al. worked also on vacuum cleaning and teeth washing, with a triaxial accelerometer on the subject's pelvis [28]. Of course, activities related to mouth or hands were difficult to recognize.

Some studies have used other sensors in addition to the accelerometer. In the study by Parkka et al. [29], 20 different sensors placed along the body and including two accelerometers were used to successfully recognize basic activities.

Other experiments have embedded various sensors in one single multimodal device. Maurer et al. demonstrated that such a device was able to recognize basic activities independently of the position, although climbing and going down stairs remained more complex to discriminate [30]. Lester et al [31] used 7 sensors in a single device and were able to recognize also stairs and elevator movements, due to pressure and audio sensors.

However, in recent years, a new technology –smartphones- has appeared that has changed the perspectives on the concrete applicability of sensor-based (in particular accelerometer-based) activity recognition. In fact, smartphone accelerometers are being used more and more in experiments aimed at recognizing activities, due to the wide use of smartphones, the growing availability of embedded sensors, and the computational power provided.

In one of the first experiments, Miluzzo et al. [32] used some internal smartphone sensors (accelerometer, microphone, and GPS) to recognize basic activities, and reported difficulties in distinguishing standing and sitting when the phone was placed in the trousers pocket or on the belt.

Kwapisz et al. [33] used a smartphone accelerometer to automatically recognize six daily activities (walking, jogging, sitting, standing, climbing and going down stairs). In their study stair movements were the most difficult to classify, while the other activities were easily classified with good precision. Wu et al. exploited also a gyroscope, but still stair movements were difficult to recognize [34]. In a recent study, Dernbach et al. [35] introduced among the activities also some typical household activities: cleaning, cooking, plant watering, washing hands, and taking a drug. These activities were difficult to recognize because made up of a series of different movements, sequential or parallel, and thus not easily representable as a single action.

Finally, some studies have evaluated the performance of smartphone accelerometers and have confirmed that their accuracy is similar to that of professional triaxial accelerometers [36,37].

To our knowledge, household activities have been studied only by means of a wearable system consisting of SmartShoe sensor and a wrist accelerometer [38]. Household activities have indeed been under scrutiny for their energy expenditure [39] because, according to international recommendations on physical activity levels [1,2], they contribute to the 30 minutes per day of

moderate-intensity activity required to confer health benefits, but their impact is difficult to estimate.

To summarize, previous studies show that using more than one sensor, either in different body parts or in a single device, allows to identify with good accuracy a variety of activities, some of which are more difficult to recognize than others. Fortunately, basic activities can be easily recognized also with a single triaxial accelerometer, even if embedded in a smartphone because equivalent to professional devices. The latter evidence has helped to switch from less practical setups, where the subject had to wear unusual sensors and devices, to some more natural exploitation of nowadays common smartphones. The issue of where to position the sensor or smartphone remains partially open. While some studies have identified a position-independent model for non-smartphone sensors (e.g., [30]), most of the times, some activities are harder to recognize because the sensor position is somewhat biasing the kind of activities that could be recognized.

However, since the most investigated activities include walking, running, bicycling, etc., a complete characterization of the physical activity carried out by a person who works but does no sport is yet to be achieved. Furthermore, to our knowledge, there is not yet an evaluation of the influence of the smartphone position on activity classification, which should also be evaluated from a user acceptability point of view.

Thus, the present study aimed at: i) developing and testing a method for classifying household activities based on smartphone accelerometer output; ii) evaluating the influence of smartphone position on activity recognition; and iii) preliminarily evaluating the acceptability of continuously wearing a smartphone for automated recognition of activities.

2 Methods

The process of recognising activities can be organised in three steps:

- collection of accelerometer data during activities;

- computation of descriptive features from raw data;

- classification of activities based upon their descriptive features.

For each of these steps, actions were taken in order to:

- individuate relevant features to be computed (see Section 2.1);

- support data collection and feature computation through a mobile application (Section 2.1);

- select one or more classification algorithms with good performance from a pool of candidates obtained from the relevant literature (Section 2.2).

For the latter point, we designed an evaluation methodology based on the execution of a number of activities (Section 2.3) by subjects (Section 2.4) carrying a smartphone in three different positions. During the activities, data were collected and summarized using a mobile application. After the activities, the data were classified by means of selected classification algorithms. The performance of the algorithms for the three smartphone positions was finally compared with the methodology described in Section 2.5.

Acceptability of wearing the smartphone was evaluated through a short user survey (Section 2.4).

Finally, we attempted to understand whether the selected algorithms were computationally suitable for low-resource devices by empirically measuring their speed (Section 2.5).

We conducted the study in agreement with the declaration of Helsinki and collected informed consent from participating subjects. However, since it was an observational study that did not

involve drug testing and was carried out on healthy people, according to our Institution regulations, it did not need ethical approval by our Ethics Committee.

2.1 Accelerometer data collection

A prototype Android application was designed and developed to record accelerometer data on the three axes and calculate descriptive features. The application saves two files for further usage: one containing all raw data collected, and one containing descriptive features calculated as depicted in the following.

Recorded data were segmented in 5-second blocks according to a methodology proposed by Kwapisz et al. [33] and adopted in many other studies. Each block was then described by means of 22 features that had been individuated among those proposed in other studies:

- mean and standard deviation on each of the three axis [33], for a total of 6 features;

- minimum and maximum on each of the three axis [35], for a total of 6 features;

- zero-crossing rate and absolute average deviation on each of the three axis [30], for a total of 6 features;

- correlation on each pair of axis (XY, XZ, YZ) [28], for a total of 3 features;

- resulting average acceleration [33].

2.2 Classification

To classify the activities, the Weka software [40] was used, as in other similar studies, because of the large amount of different classification algorithms it provides. The software was run on a desktop computer (2.66GHz Intel® Core™ i7 620m processor, 4 GB RAM), thus classification occurred offline and not directly during the activities. Algorithms were compared using the Experimenter module of Weka on 10 runs of 10 cross-fold validation for each algorithm and smartphone position.

Weka provides for a very large number of classification algorithms. In our experiment, we decided to rely at first on the algorithms proposed in other studies.

In particular, as in the case of descriptive features, we relied again on the paper by Kwapisz et al. [33]. In that paper, the authors used *J48*, *Logistic Regression,* and *Multilayer Perceptron*. We also added the algorithms used by Dernbach et al. [35]: *NaiveBayes*, *BayesNet*, *Decision Table,* and *Kstar*. Finally, we completed the set of algorithms with another two algorithms, *LogitBoost* and *RandomForest*, which we chose to cover Weka categories not considered by other Authors.

2.3 Activities

Activities were selected in order to represent different household activity levels according to their MET: sitting (1 MET), sitting and working at the computer (1.8 MET), walking (2 MET), ironing (2.3 MET), sweeping the floor (3.3 MET), going down stairs with a shopping bag (5 MET), walking while carrying a large box (6 MET), and climbing stairs with a shopping bag (7.5 MET).

Activities were organized in a continuous path that subjects had to follow, with around 30 seconds time for each activity and 10-second breaks between activities in order to facilitate *a posteriori* segmentation. A complete path had a total duration of 320 seconds.

2.4 Subjects

Ten subjects were chosen among the available personnel at the Department of Biological and Medical Sciences of the University of Udine, Italy, in order to have five female and five male subjects. Their age ranged between 26 to 40 years.

Subjects were asked to carry out the above-mentioned activity path for three times, each one with the smartphone in a different position (pocket, arm, and wrist). When on the arm or wrist, the smartphone was fixed using an armband (such as those used for sport activities).

At the end of the path, users were finally asked to answer a short survey about their experience during the experimentation. The questions were aimed at evaluating the acceptability of wearing the smartphone during the activities.

Thus, for each of the three positions, users were requested to tell whether the smartphone was uncomfortable, an impediment (meaning that it blocked the execution of the activities), or a conditioning factor (meaning that it caused the activities to be carried out in an unusual way). Answers were on a 1-5 scale (from best to worst).

2.5 Analysis

Due to the low number of subjects, we only calculated descriptive statistical features.

Classification performance was described by means of accuracy, precision, recall and F-measure; insights on classification errors were examined by looking at confusion matrices.

We also attempted to consider computational complexity for the sake of real-time use on the smartphone by measuring the time needed for model building and activity classification (the latter on 350 instances).

3 Main outcomes

3.1 Data collection

480 time blocks for each smartphone position were collected for a total of 1440, that is, 6 (30 seconds) for each of the 8 activities, all multiplied by 10 subjects. Figure 1 shows an overall view of the recorded accelerometer data for the whole path as done by one subject.

(FIGURE 1 HERE)

3.2 Classification accuracy

Table 1 shows the performance of the algorithms for each smartphone position, expressed in terms of accuracy, precision, recall and F-measures.

(TABLE 1 HERE)

Three algorithms demonstrated a classification accuracy greater than 80% for all three smartphone positions: *RandomForest* (84.6-87.4%), *Kstar* (82.6-89.1%), and *Multilayer Perceptron* (82.5-85.4%). Figure 2 shows a comparison among the three best algorithms.

(FIGURE 2 HERE)

In general, the pocket position showed a slightly higher accuracy (5 algorithms obtained more than 80% accuracy vs 4 for wrist position and 3 for arm position). In absolute terms, the best recognition rate (89.4%) was obtained with *Kstar* and the smartphone in the pocket.

Most of the classification errors were due to activities that appear similar from a specific smartphone position (e.g., sweeping and ironing when the smartphone is on the arm). Some errors were not crucial because the difference in MET was not great (sweeping: 2.3 MET, ironing: 3.3 MET), some others though were more severe (e.g., walking with box: 6 MET, and without box: 2 MET). In the latter case, with the smartphone in the pocket, 103 blocks were correctly recognized (53 as walking with box, 50 without), whereas 16 were incorrectly classified (9 as walking with box, 7 without). Table 2 shows the two confusion matrices for the *Kstar* algorithm, pointing out the above-mentioned cases.

(TABLE 2 HERE)

In Table 3, a summary of the results of related studies is reported for comparison with our results obtained with the *Kstar* algorithm and the smartphone in the pocket. The studies listed are only those with comparable results, i.e., with explicit accuracy results or with accuracy obtainable from confusion matrices. When more than one algorithm or position was used, the best algorithm and the pocket position are shown. Activities listed are those considered in our study.

(TABLE 3 HERE)

Two main considerations can be made. First, the accuracy of our method is generally comparable with that of other smartphone-based systems and slightly lower than that of systems adopting multiple sensors or accelerometers. Differences on specific activities may depend on the set of activities considered in each individual study, which might account for more or less difficulty in classification. Secondly, for specific household activities such as ironing, walking with a box, and sweeping, no previous studies were available, although other authors examined doing the dishes, vacuuming, folding laundry, scrubbing, and eating [23, 38].

We also investigated whether the gender of the subject might influence activity classification. While the results were not significantly different, some differences between males and females were found in algorithm accuracy. Recognition of women activities gave better results (on average: 6%), probably due to the more regular execution of activities by women, as shown in Figure 3 for the three best algorithms. In absolute terms, the best recognition rate (93.6%) was obtained for women activities measured when the smartphone was on the arm. The worst recognition rate (81.2%) was instead obtained for male activities carried out with the smartphone on the wrist.

(FIGURE 3 HERE)

3.3   Computational issues
In our preliminary evaluation of computational complexity, the time for model building ranged from 0.001 (*Kstar*) to 4.34 seconds (*Logistic Regression*), while the classification of 350 samples was always immediate, except for *Kstar (3.30s)*. Table 4 shows the details regarding the time needed for building the model.

Since the computer used for the test is more powerful than a generic smartphone (in particular for floating point operations), maybe *Kstar* is less suitable for real-time recognition of activities

directly on a smartphone than other algorithms. However, this should be proven with an ad-hoc implementation on a smartphone, optimized for its computational resources.

(TABLE 4 HERE)

3.4 User survey

All users filled in the questionnaire as requested. Figure 4 shows the details of their evaluation of the experience with the smartphone.

Smartphone comfort was perceived with a wide range of levels, with a slight predominance of positive opinions. Summing up 1 and 2 as positive opinions and 4-5 as negative opinions, the wrist position turned out to be the most comfortable (8 positive vs 1 negative) and the arm position the least comfortable (6 positive vs 3 negative).

Impediment and conditioning of activities were evaluated in very similar ways, and, in general, it seems that carrying the smartphone did not hinder or condition activities, because no negative opinions (4 and 5) were expressed. As an example, with the smartphone in the wrist position, all opinions were positive for both impediment and conditioning. The arm position, which was considered the least comfortable, did not impede or condition activities more than the other positions.

(FIGURE 4 HERE)

4 Conclusion

The smartphone, placed in a usual position such as in a pocket, can be used to recognize several activities through its accelerometer, including some household activities [23,38].

Although some activities were studied for the first time in our work, our contribution is preliminary, due to the limited number of subjects involved and to the fact that the subjects were not asked about their confidence with the studied household activities.

Since most of the current mobile and wearable devices and applications are devoted to fitness and sports activities, our results may open towards a better understanding of the physical activity behaviour of populations that are not directly involved in activities explicitly aimed at physical wellness, but that do some physical activity in their daily life.

To exploit physical activity recognition, we envisaged two implementation modalities. One modality provides for a pre-calculated generic classification model (subject-independent training), which the user can exploit without the need for specific training and personalization. The other modality is based on an initial subject-dependent training phase, in which the user has to label his/her own activities to help the construction of a personalized classification model, which thus can exploit knowledge on the subject (e.g., preferred smartphone position, gender, etc.). While the pre-calculated modality is of immediate use, the second one is initially more complicated for the user, but it certainly provides greater accuracy in recognizing activities. However, in the first modality, the model building occurs offline, whereas in the second modality the model has to be built on the smartphone, and this can take some time. Classification has to be done in real time or close to real time directly on the smartphone, and algorithms have thus to be optimized for that (e.g., [41]). While this has to be confirmed by a specific experimentation, the best algorithm in our experimentation (*Kstar)* might not be the most adequate algorithm from this point of view. However, since differences in accuracy among the algorithms were minimal, either *RandomForest* or *Multilayer Perceptron* could be instead implemented directly on the smartphone.

The demonstrated capability of recognizing common activities can be exploited in all those situations were physical activity is crucial [1-3], for example in metabolic syndrome [4-5]. Since measurement of physical activities is one of the components to evaluate the possible risk of metabolic syndrome, a smartphone capable of classifying and recording activities during the day could complement the IPAQ questionnaire, thus providing a way for estimating physical activity in

a non-invasive way. However, since there are reports of variable reliability of IPAQ in specific populations [9,42-44] and, more generally, of limitations in measuring physical activity through questionnaires [45], a further development could be to effectively measure METs starting from smartphone sensors only, although some evidence exists that accelerometer alone cannot provide a reliable estimate of energy expenditure [46]. In addition to accelerometer data, which might eventually come from electronic activity monitors [47,48], other sensible sources of information could be the internal gyroscope of some high-end smartphones, or also the vital sign sensors being introduced in forthcoming smartwatches. All of this, through data fusion [49], may concur to a better recognition of daily life activities, although the most sophisticated and sensor-rich systems are also more expensive and thus less adequate for reaching a wide audience.

Declaration of interest: The author reports no conflicts of interest.

5   References

[1] World Health Organization. Diet, nutrition and the prevention of chronic diseases. Geneva: World Health Organization, 2002.

[2] World Cancer Research Fund / American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washington DC: AICR, 2007.

[3] World Health Organization. Global health risks: mortality and burden of disease attributable to selected major risks. Geneva: World Health Organization, 2009.

[4] Miccoli R, Bianchi C, Odoguardi L, Penno G, Caricato F, Giovannitti MG, Pucci L, Del Prato S. Prevalence of the metabolic syndrome among Italian adults according to ATP III definition. Nutr Metab Cardiovasc Dis 2005;15:250-4.

[5] Cornier MA, Dabelea D, Hernandez TL, Lindstrom RC, Steig AJ, Stob NR, Van Pelt RE, Wang H, Eckel RH. The metabolic syndrome. Endocr Rev 2008;29:777-822.

[6] Willett W. Nutritional epidemiology. New York: Oxford University Press, 2013.

[7] WD McArdle, FI Katch, VL Katch. Exercise Physiology: Energy, Nutrition, and Human Performance. 6th ed. Philadelhia: Lippincott Williams & Wilkins, 2006.

[8] Corder K, Ekelund U, Steel RM, Wareham NJ, Brage S. Assessment of physical activity in youth. J Appl Physiol 2008;105:977-987.

[9] Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, et al. International physical activity questionnaire: 12 country reliability and validity. Med Sci Sports Exerc 2003;35:1381–1395.

[10] International Physical Activity Questionnaire Web site. [cited 2016 Feb 10]. Available from: https://sites.google.com/site/theipaq/

[11] Ainsworth B, Haskell W, Leon A, Jacobs DR Jr., Montoye HJ, Sallis JF, Paffenbarger RS, Jr. Compendium of physical activities: Classification of energy costs of human physical activities Med Sci Sports Exerc 1993;25:71-80.

[12] Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR Jr, Tudor-Locke C, Greer JL, Vezina J, Whitt-Glover MC, Leon AS. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 2011;43:1575-81.

[13] Compendium of Physical Activities. [cited 2016 Feb 10]. Available from: https://sites.google.com/site/compendiumofphysicalactivities/

[14] Byrne NM, Hills AP, Hunter GR, Weinsier RL, Schutz Y. Metabolic equivalent: one size does not fit all. J Appl Physiol 2005;99:1112-1119.

[15] Kozey S, Lyden K, Staudenmayer J, Freedson P. Errors in MET estimates of physical activities using 3.5 ml x kg(-1) x min(-1) as the baseline oxygen consumption. J Phys Act Health 2010;7:508-516.

[16] Bruegel RB. Patient empowerment--a trend that matters. J AHIMA 1998;69:30-3.

[17] Sloan JA, Cella D, Hays RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. J Clin Epidemiol 2005;58:1217-9.

[18] World Health Organization. mHealth: New horizons for health through mobile technologies. 2011. [cited 2016 Feb 10]. Available from: http://www.who.int/goe/publications/ehealth_series_vol3/en

[19] Dale O, Hagen KB. Despite technical problems personal digital assistants outperform pen and paper when collecting patient diary data. J Clin Epidemiol. 2007;60:8-17.

[20] Fortney JC, Burgess JF Jr, Bosworth HB, Booth BM, Kaboli PJ. A re-conceptualization of access for 21st century healthcare. J Gen Intern Med 2011;26:639-47.

[21]   Ascherio A, Rimm EB, Giovannucci EL, Colditz GA, Rosner B, Willett WC, Sacks F, Stampfer MJ. A prospective study of nutritional factors and hypertension among US men. Circulation. 1992;86:1475-84.

[22]   Foerster F, Fahrenberg J. Motion pattern and posture: correctly assessed by calibrated accelerometers. Behav Res Methods Instrum Comput 2000;32:450-7.

[23]   Bao L, Intille S. Activity Recognition from User-Annotated Acceleration Data. Lect Notes Comput Sc 2004; 3001:1-17.

[24]   Mannini A, Sabatini AM. Machine learning methods for classifying human physical activity from onbody accelerometers. Sensors 2010;10:1154-1175.

[25]   Khan AM, Lee YK, Lee SY, Kim TS. A triaxial accelerometer-based physical activity recognition via augmented-signal features and a hierarchical recognizer. IEEE Trans Inf Technol Biomed 2010;14:1166–1172.

[26]   Lee M, Kim J, Kim K, Lee I, Jee SH, Yoo SK.. Physical activity recognition using a single tri-axis accelerometer. World Congress on Engineering and Computer Science 2009, San Francisco, USA: 14-17

[27]   Long X, Yin B, Aarts RM. Single-accelerometer-based daily physical activity classification. Conf Proc IEEE Eng Med Biol Soc. 2009;2009:6107-10..

[28]   Ravi N, Dandekar N, Mysore P, Littman ML. Activity recognition from accelerometer data. 17th conference on Innovative applications of artificial intelligence, Pittsburgh 2005; Volume 3:1541-1546.

[29]   Parkka J, Ermes M, Korpipaa P, Mantyjarvi J, Peltola J, Korhonen I. Activity classification using realistic data from wearable sensors. IEEE Trans Inf Technol Biomed 2006;10:119-128.

[30]   Maurer U, Smailagic A, Siewiorek D, Deisher M. Activity recognition and monitoring using multiple sensors on different body positions. International Workshop on Wearable and Implantable Sensor Networks, Cambridge; 2006. Washington DC: IEEE Computer Society; 2006: 113-116.

[31]   Lester J, Choudhury T, Borriello G. A practical approach to recognizing physical activities. Lect Notes Comput Sc 2006; 3968:1–16.

[32]   Miluzzo E, Lane N , Fodor K, Peterson R, Lu H, Musolesi M, Eisenman S, Zheng X, Campbell A. Sensing meets mobile social networks: The design, implementation and evaluation of the CenceMe application. 6th ACM Conference on Embedded Networked Sensor Systems, Raleigh, USA 2008. New York: ACM; 2008:337-350.

[33]   Kwapisz J, Weiss G, Moore S. Activity Recognition using Cell Phone Accelerometers. Human Factors, 2010.

[34]   Wu W, Dasgupta S, Ramirez EE, Peterson C, Norman GJ. Classification accuracies of physical activities using smartphone motion sensors. J Med Internet Res 2012;14:e130.

[35]   Dernbach S, Das B, Krishnan NC, Thomas BL, Cook DJ. Simple and Complex Activity Recognition through Smart Phones. 8th International Conference on Intelligent Environments, Guanajuato, México, 2012. Washington DC: IEEE Computer Society; 2012; 214-221.

[36]   Nishiguchi S, Yamada M, Nagai K, Mori S, Kajiwara Y, Sonoda T, et al. Reliability and validity of gait analysis by android-based smartphone. Telemed J E Health 2012;18:292-296.

[37]   Saha IK, Dirik AE, Topkara U, Memon N, Gutierrez G, Rao S. Reliability and validity of accelerometer-based smartphones to assess physical activity. Proceedings of the Annual Meeting of the American Society of Biomechanics, Providence, USA, 2010:571-2.

[38]   Edgar SR, Fulk GD, Sazonov ES. Recognition of household and athletic activities using SmartShoe. Conf Proc IEEE Eng Med Biol Soc. 2012;2012:6382-5.

[39]   Brooks AG, Withers RT, Gore CJ, et al. Measurement and prediction of METs during household activities in 35- to 45-year-old females. Eur J Appl Physiol 2004;91:638-648.

[40]   Hall M, Frank E, Holmes G, Fahringer BP, Reutemann P, Witten IH. The WEKA Data Mining Software: An Update. SIGKDD Explorations 2009:11(1), 10-18.

[41]   Tapia EM, Intille SS, Haskell W, Larson K, Wright J, King A, Friedman R. Real-Time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. 11th IEEE International Symposium on Wearable Computers, Boston, USA, 2007. Washington DC: IEEE Computer Society; 2007;1-4.

[42]   Wang C, Chen P, Zhuang J. Validity and reliability of International Physical Activity Questionnaire-Short Form in Chinese youth. Res Q Exerc Sport. 2013;84:S80-6.

[43]   Hansen AW, Dahl-Petersen I, Helge JW, Brage S, Grønbæk M, Flensborg-Madsen T. Validation of an Internet-based long version of the International Physical Activity Questionnaire in Danish adults using combined accelerometry and heart rate monitoring. J Phys Act Health 2014;11:654-64.

[44]   Oyeyemi AL, Oyeyemi AY, Adegoke BO, Oyetoke FO, Aliyu HN, Aliyu SU, Rufai AA. The Short International Physical Activity Questionnaire: cross-cultural adaptation, validation and reliability of the Hausa language version in Nigeria. BMC Med Res Methodol 2011;11:156.

[45]   Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. Br J Sports Med 2003;37:197-206.

[46]   Sasaki JE, Hickey A, Mavilia M, Tedesco J, John D, Kozey Keadle S, Freedson PS. Validation of the Fitbit wireless activity tracker for prediction of energy expenditure. J Phys Act Health 2015;12:149-54.

[47]   Vooijs M, Alpay LL, Snoeck-Stroband JB, Beerthuizen T, Siemonsma PC, Abbink JJ, Sont JK, Rövekamp TA. Validity and usability of low-cost accelerometers for internet-based self-monitoring of physical activity in patients with chronic obstructive pulmonary disease. Interact J Med Res 2014;3:e14.

[48]   Lee JM, Kim Y, Welk GJ. Validity of consumer-based physical activity monitors. Med Sci Sports Exerc.2014;46:1840-8.

[49]   Shoaib M, Bosch S, Incel OD, Scholten H, Havinga PJ. Fusion of smartphone motion sensors for physical activity recognition. Sensors (Basel). 2014;14:10146-76.
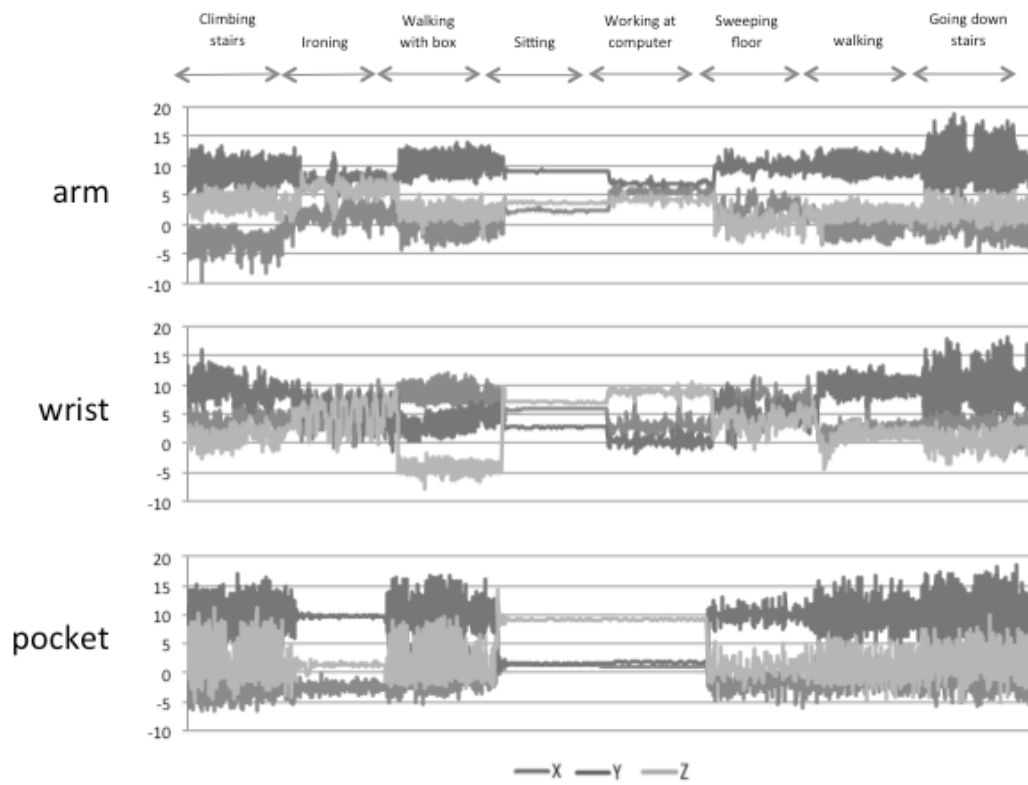
Figure 1 – An example of accelerometer data for all activities and smartphone positions
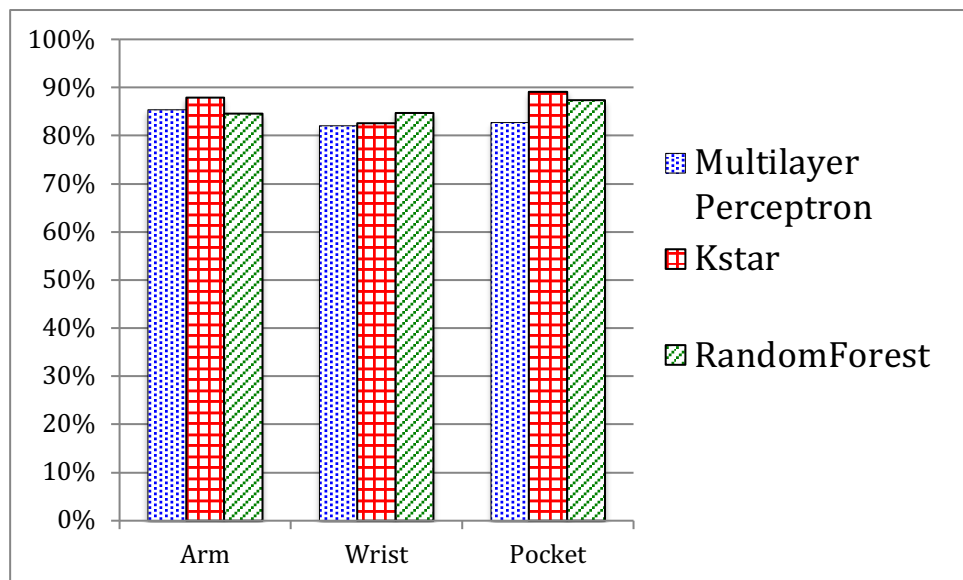

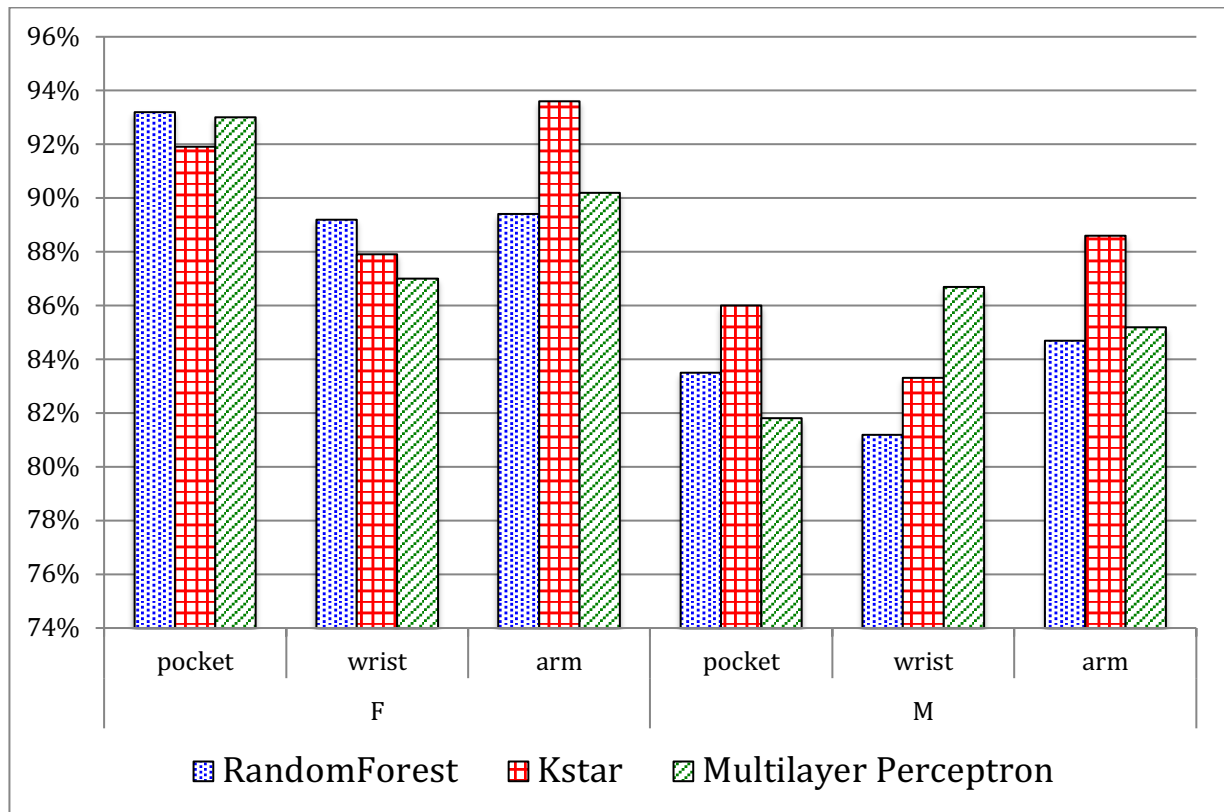*Figure 2 – the accuracy of the best three algorithms by smartphone position*

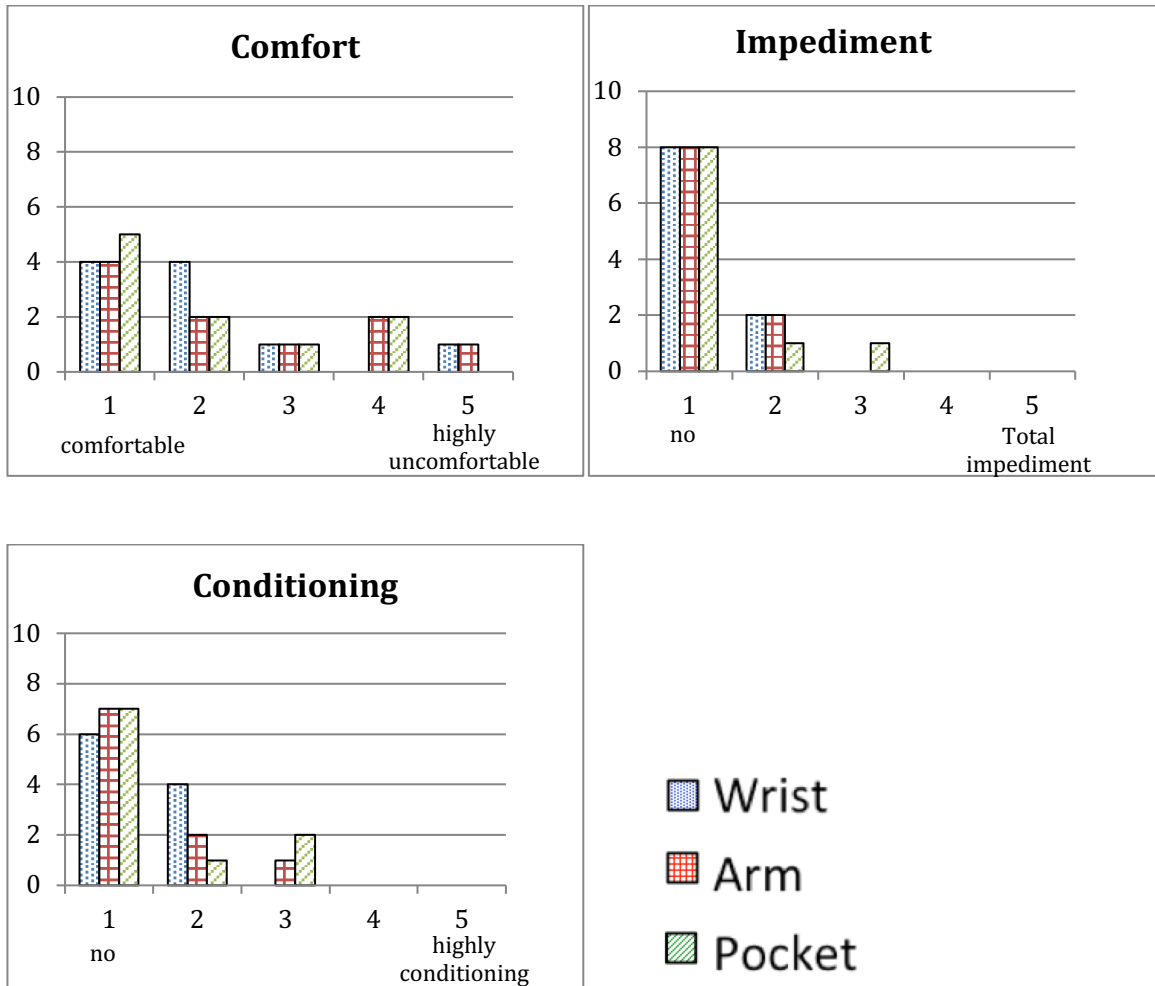*Figure 3 – Accuracy of the best three algorithms by gender and by smartphone position*

Figure 4 – Results of user survey

Position		J48	Logistic Regression	Multilayer Perceptron	Naive Bayes	Bayes Net	Decision Table	KStar	Random Forest	Logit Boost
arm	*Accuracy*	74.8	76.3	84.4	74.2	75.7	62.1	**87.9**	85.6	78.1
	Precision	0.64	0.73	0.82	0.64	0.63	0.42	**0.84**	0.83	0.72
	Recall	0.66	0.68	0.87	0.69	0.67	0.56	**0.93**	0.78	0.73
	F-measure	0.63	0.69	0.83	0.64	0.63	0.46	**0.88**	0.79	0.71
wrist	*Accuracy*	76.7	76.8	84.3	76.8	79.5	64.4	83.1	**85.2**	81.6
	Precision	0.73	0.77	0.80	0.64	0.77	0.46	0.73	**0.81**	0.78
	Recall	0.70	0.73	0.86	0.79	0.73	0.66	**0.91**	0.85	0.79
	F-measure	0.70	0.73	**0.82**	0.70	0.74	0.53	0.80	**0.82**	0.77
pocket	*Accuracy*	80.7	75.6	82.2	66.1	73.6	63.0	**89.4**	**89.4**	83.1
	Precision	0.84	0.86	0.92	0.90	0.86	0.49	**0.96**	0.95	0.89
	Recall	0.82	0.85	**0.91**	0.81	0.76	0.73	**0.91**	**0.91**	0.89
	F-measure	0.82	0.85	0.91	0.84	0.79	0.57	**0.93**	**0.93**	0.88

Table 1 – Performance of classifiers (accuracy, precision, recall and F-measure) by smartphone position; in bold the most accurate classifier for each measure and position

ARM

	climbing stairs	ironing	walking with box	sitting	working at PC	sweeping	walking	going down stairs
climbing stairs	55	0	3	0	0	0	1	0
ironing	0	47	0	2	0	4	0	0
walking with box	0	0	57	0	0	0	2	1
sitting	0	1	0	52	4	2	0	0
working at PC	0	3	0	2	55	0	0	0
sweeping	5	7	0	2	0	44	2	0
walking	3	0	3	0	0	2	50	2
going down stairs	3	0	2	0	0	0	1	54

POCKET

	climbing stairs	ironing	walking with box	sitting	working at PC	sweeping	walking	going down stairs
climbing stairs	54	0	2	0	0	0	2	2
ironing	0	45	0	0	0	5	1	1
walking with box	0	0	53	0	0	0	7	0
sitting	1	0	0	49	10	0	0	0
working at PC	0	0	0	1	59	0	0	0
sweeping	0	3	0	0	0	57	0	0
walking	1	0	9	0	0	0	50	0
going down stairs	2	0	2	0	0	0	2	54

Table 2 – Confusion matrices for Kstar algorithm, on arm and pocket positions

	position	climbing stairs	ironing	walking with box	sitting	working at PC	sweeping	walking	going down stairs
Multiple accelerometers and sensors									
(22) Foerster 2000	5 sensors on sternum, right and left thigh	93.5%			94.2%			96.7%	95.7%
(23) Bao 2004	5 sensors on arm, wrist, knee, ankle, waist	85.6%			94.8%	97.5%		89.7%	
(25) Khan 2010	1 sensor at a position closer to the center of mass	99.0%			90.8%			99.0%	99.0%
(27) Long 2009	1 sensor on wrist							80.3%	
(28) Ravi 2005	1 sensor on a pelvic region	42.9%						97.8%	100%
(29) Parkka 2006	22 sensors on chest, wrist, finger, forehead, upper back, below neck, shoulder, breastbone				96%			79.0%	
(31) Lester 2006	7 sensors on shoulder, wrist and waist	84.5%			50.5%			79.8%	80.3%
(38) Edgar 2012	9 sensors on the inside of the wrist and the foot	77.7%			96.7%			100%	95.7%
Smartphone									
(32) Miluzzo 2008	average among pocket, hip, necklace				68.2%			94.4%	
(33) Kwapisz 2010	pocket	61.5%			95%			91.7%	44.3%
(34) Wu 2012	pocket	69.8%			100%			92.0%	79.4%
(35) Dernbach 2012	not standardized	71.8%			87.3%		60.0%	86.9%	
our study (using Kstar)	pocket	90.0%	86.5%	88.3%	81.7%	98.3%	95%	83.3%	90%

Table 3 – Accuracy comparison of our study using Kstar algorithm with smartphone in the pocket with previous relevant works

Classifier	Model building time (s)
J48	0.14
Logistic Regression	4.34
Multilayer Perceptron	**2.90**
Naivebayes	0.04
Bayesnet	0.09
Decision Table	0.22
Kstar	**<0.001**
Random Forest	**0.16**
Logit Boost	0.24

Table 4 – time needed for model building; in bold the 3 most accurate algorithms