Real-Time Activity Monitoring with a Wristband and a Smartphone

Božidara Cvetković[a,b,*], Robert Szeklicki[c], Vito Janko[a,b], Przemyslaw Lutomski[c], Mitja Luštrek[a,b]

[a] *Jožef Stefan Institute, Department of Intelligent Systems, Jamova cesta 39, 1000 Ljubljana, Slovenia*
[b] *Postgraduate School Jožef Stefan, Jamova cesta 39, 1000 Ljubljana, Slovenia*
[c] *Poznan University of Physical Education, Krolowej Jadwigi 27/39, 61-871 Poznan, Poland.*

Abstract

Activity monitoring is a very important task in lifestyle and health domains where physical activity of a person plays an important role in further reasoning or for providing personalized recommendations. To make such services available to a broader population one should use devices which most users already have, such as smartphone. Since trends show an increasing popularity of wrist-worn wearables we also consider sensor-rich wristband as an optional device in this research. We present a real-time activity monitoring algorithm which utilizes data from smartphone sensors, wristband sensors or their fusion for activity recognition and estimation of energy expenditure of the user. The algorithm detects which devices are present and uses an interval of walking for gravity detection and normalization of the orientation of the devices. The normalized data is afterwards used to for detection of the location of the smartphone which serves as a context for selection of location-specific classification model for activity recognition. The recognized activity is finally used for the selection of one or multiple regression models for the estimation of human energy expenditure. To develop the machine-learning models, which can be deployed on the smartphone, we optimized the number and type of extracted features via automatic feature selection. We evaluated each step of the algorithm and each device configuration, and compared the human energy expenditure estimation results against the Bodymedia armband and Microsoft Band 2. We also evaluated the benefit of decision fusion where appropriate. The results show that we achieve a 87% ± 5% average accuracy for activity recognition and that we outperformed both competing devices in the estimation of human energy expenditure by achieving the mean absolute error of 0.6 ± 0.1 MET on average.

Keywords: Wristband sensors, Smartphone sensors, Activity recognition, Estimation of energy expenditure, Machine learning

1. Introduction

Activity monitoring using body sensor network is a mature and important area of research due to vast range of domains that can benefit from it. Accurate activity monitoring is required in domains where further reasoning or person-specific recommendations rely on the user's physical activity. These range from lighter topics such as sports and lifestyle to more sensitive topics such as health [1, 2]. In sports and lifestyle, the activity monitoring gives the user an insight into the amount of their physical activity either to support better sports training, improve self-awareness and help maintaining a healthier lifestyle, whereas in health this information is used to provide better recommendations on how to manage pathologies of a particular disease and thus improve the quality of life of the patient.

It is a known fact that engagement in physical activity has a positive influence on body and mind [3]. The ones that are the most aware of this fact are already using the service of activity monitoring to track their sports activities such as running and cycling either with a smartphone application [4] or dedicated devices [5]. The physical activity self-awareness is higher among young active population due the knowledge about the risk imposed by the sedentary lifestyle (e.g., office work), and are keen to use activity monitoring to improve their lifestyle habits. Elderly population is more difficult to reach with technology even though they needed it more due to chronic diseases such as diabetes and coronary heart diseases [6] and generally less robust health. Moreover, the services which include activity monitoring mostly target the younger population and do not include specific features and knowledge that applies for the elderly, especially not the ones who have been already diagnosed with some disease. Aim of this research is to develop activity monitoring models specialized for the elderly, which can be used either for lifestyle or health applications, so we develop our methods on data of elderly active population [7] consisting of most common every day and sports activities.

To support the lifestyle self-awareness or disease self-management in terms of activity monitoring, two individual tasks have to be solved: activity recognition to recognize the current activity which is being performed, and the estimation of human energy expenditure to quantify the performed activity. Previous research has shown that both of these tasks can be

*Corresponding author
Email addresses: boza.cvetkovic@ijs.si (Božidara Cvetković), szeklicki@awf.poznan.pl (Robert Szeklicki), vito.janko@ijs.si (Vito Janko), lutomski@awf.poznan.pl (Przemyslaw Lutomski), mitja.lustrek@ijs.si (Mitja Luštrek)
URL: http://www.bozacvetkovic.net (Božidara Cvetković), http://www.awf.poznan.pl/ (Robert Szeklicki), http://dis.ijs.si (Vito Janko), http://www.awf.poznan.pl/ (Przemyslaw Lutomski), http://dis.ijs.si/mitjal (Mitja Luštrek)

solved fairly accurate when dedicated devices have static placements (are attached to predefined location on the body, with predefined orientation) [8, 9]. The activity recognition is usually set as a classification machine-learning problem, where the recognized activities include sedentary activities (e.g., sitting, standing, lying), ambulatory activities (e.g., walking, jogging, running) and sports activities (e.g., cycling). The energy expenditure is usually set as a regression machine-learning task where the regression model estimates the expended energy in a unit called MET (Metabolic equivalent of task). METs range from 0.9 MET, which is equal to the energy expended while sleeping, to over 20 MET, which is expended during extreme exertion [10]. The advantage over the more common kilo-calories is that MET does not depend on users weight, and can be easily translated into kilo-calories.

To make the activity monitoring more widely applicable, and to make it available to more people, we need to design a body sensor network from the devices most people already have. The most natural choice is a smartphone, which can serve as a processing unit and as a graphical interface for interaction with the user. There are already many smartphone applications available on the market, whose scope is limited to step counters or to activity-specific monitoring. Application which provide step counters and step-related estimation of energy expenditure [11] are suitable for tracking general ambulatory physical activities which do not require very accurate activity quantification. The activity-specific applications are intended for sports-active people. Such applications are the run trackers with run-related estimation of energy expenditure [4], which monitor the specific activity on users request and are therefore very accurate for the quantification of that exact activity. In general, the step counters are the most popular since we carry our smartphone only when we are headed somewhere (walking, running) and not at home so it captures the expended energy only during ambulatory activities and count the time when smartphone is not carried as rest. To catch every day activities and non-ambulatory exercise in addition to ambulatory activities, we need to improve the recognition and estimation of the models used by the smartphone or to use a commodity device that is worn by the user most of the day. In the past several years, the development of sensor-rich wrist-worn devices (smartwatches, sensor-equipped wristbands) has increased, and so did their popularity. Similarly to smartphones, the wrist-worn devices have activity monitoring methods already integrated into the device and can monitor physical activity even while sedentary. The accuracy of such activity monitoring is usually low but satisfactory for getting an insight into the movement pattern of the day. To get more accurate activity quantification, the user still has to explicitly input the type of workout performed (walking, running, cycling, etc.). Additionally, none of the devices fuse data/information from multiple devices and merely override the activity data with the analysis of preferred device.

The objective of our research was to develop a real-time continuous activity monitoring algorithm that utilizes sensor data from a smartphone or wrist-worn wearable alone and can fuse the data and decisions of both devices if both are present on the body (active in the body sensor network). The algorithm should be optimized to run on the smartphone to perform continuous data analysis in real-time. To develop such a method the following requirements should be fulfilled:

1. The method has to have the ability to detect which devices are present on the body, thus active in the body sensor network
2. The smartphone and wrist-worn device can be worn freely on the body at most common locations (trousers pocket, jacket pocket, bag) and in any orientation
3. The method has to have the ability to normalize the orientation of the present devices and recognize their location
4. To achieve maximum classification and regression accuracy with any combination of devices, they should be fused on multiple levels: the data level, feature level and decision level [12]
5. Procedures for decreasing the computational complexity should be used to optimise the machine-learning models to be deployable on the smartphone

In this paper we present a method that automatically detects whether single or multiple devices are present on the body (smartphone or wristband alone, both devices), normalizes the orientation of the devices and detects the location of the smartphone relative to the user's body (trousers pocket, jacket pocket, bag) if present. The activity recognition model is selected according to the currently present devices and the recognized location (if smartphone is present). The output of the activity recognition is used for selecting an appropriate energy expenditure estimation regression model. Since our algorithm needs to run on the smartphone, we used feature selection procedure to find a trade-off between accuracy and number of features to decrease computational complexity of the feature vector construction.

The rest of the paper is structured as follows: related work on activity recognition and estimation of energy expenditure is presented in Section 2. Section 3 introduces the datasets and the methods we use to develop the activity monitoring algorithm. The evaluation of the methods is presented in Section 4. We discuss the research in Section 5 and Section 6 concludes the paper.

2. Related Work

Activity monitoring as seen in this research is composed of human activity recognition and estimation of human energy expenditure, both very popular in applications of body sensor networks. We focus on accelerometer-based wearables such as smartphones and wristbands and their combinations, where we also account for varying orientations and locations of accelerometer sensors, which is the key distinction of our approach compared to the related work.

2.1. Body Sensor Networks

Body sensor networks are composed of sensor nodes attached to the persons body which are able to sense one or more physiological or motion signals. These are usually dedicated sensor enclosures or wearables with single or multiple sensors and often perform preprocessing and storing before transmitting the data forward to the base station [13] where the data analysis is performed. There are many challenges and opportunities in the research area that can be roughly divided into sensor design, network communication and data fusion [14]. Sensor design tackles the problems of sensor hardware design (power consumption, fault detection, etc.) and enclosure ergonomics [15], network research focuses on the network challenges (topology, security, routing algorithms etc.) [16], and data fusion research is focused on data manipulation (filtering, feature extraction, classification, computational complexity, etc.) [12]. Since in this research we use consumer devices (average smartphone and wristband) which have final design and their own communication protocol we will focus on the research tasks of the data fusion.

One of the main concerns is the development of efficient algorithms in terms of computational complexity and power consumption when preprocessing and classification is performed on the sensor node. Ghasemzadeh et al. [17] evaluated the trade-off between number of sensors (using only 1-D signal from the accelerometer to preserve power) and accuracy for detection of transition events (sit to stand, sit to lie, jump, etc.). The goal was to set the minimum number of sensors (attached to predefined locations in predefined orientation) which can still accurately recognize these events. They conclude that for different events different sensors should be active (from one to 17) and show that decreasing the number of active sensor nodes can decrease the power consumption for up to 98 %. To perform continuous activity monitoring it is crucial to extract complex features rather then raw signals and to sample the data with sufficient frequency. Fortino et al. [18] gathered the requirements for developing an power-efficient body sensor network and proposed a SPINE framework for sensor node configuration is terms of sampling frequency, feature extraction and used window size which can speed-up the prototyping of the applications and evaluate the trade-off between power-consuming tasks (feature extraction) and accuracy.

In our research, we decreased the size of body sensor network to one (smartphone or wristband) or two (smartphone and wristband) active sensor nodes (wearables) which can be worn freely in any orientation and in case of the smartphone on three locations on the body. These brings additional challenges which we review in the next subsections. To further reduce the power consumption of the designed algorithm we performed feature selection to reduce the number of features to be calculated, thus reduced computational complexity and related power consumption.

2.2. Sensor Placement

Sensor placement is composed of location relative to the users body and orientation of the sensor. Most often the research in body sensor network uses single or multiple accelerometers attached to predefined locations and in predefined orientation. Attal et al. [8] thoroughly reviewed the research done until 2015 in the activity recognition domain and showed that the number of recognized activities increases with number of sensors attached to the users body. Since all reviewed work used static placements, the authors doubt about the acceptance of such approach since it is required to carefully follow the instructions about the placement and orientation of the sensors.

Research about sensors orientation and location varying during use was probably pioneered by Kunze et al. [19] in 2005 who used supervised machine-learning on accelerometer data to first recognize walking in a manner independent of the sensors location and afterwards use the walking segment to recognize the sensor placement on the body. In later work [20] they explored the detection of the orientation of an accelerometer-equipped smartphone around the vertical axis by relying on the walking segment. Other researchers explored a possibility to use rotation to normalize the orientation. Mizell [21] proposed a vector calculation to rotate acceleration axes to a canonical orientation and Tundo et al. [22] used quaternion rotations for the task. The results on the impact of the orientation normalization on activity recognition vary, Tundo et al. [22] reports improvement only for the recognition of sitting activity. In the same year Ustev et al. [23] published results where orientation normalization increased overall accuracy of the activity recognition.

Since smartphone (one of the sensor nodes in our research) can be worn in any orientation and in various locations on the body we combined and extended the work of Kunze et al. [20] and Tundo et al. [22] to develop a procedure for real-time orientation normalization and location detection.

2.3. Activity Recognition

Most of the research done on activity recognition with smartphone in different locations was done using machine-learning techniques with predefined knowledge about the placement of the smartphone[24] [25]. In 2016 Shaoib et al. [26] reviewed the research done on activity recognition with the smartphone in terms of used smartphone-embedded sensors, recognized activities, extracted features, sampling frequency and orientation and location dependency. Lu et al. [27] and Thiemjarus et al. [28] used the Mizell [21] approach to orientation normalization. First did not take into consideration varying location of the smartphone and required the user to leave the smartphone on the table for calibration of the orientation. The latter one used location independent features and required the person to perform a set of predefined activities for the calibration. Anjum et. al [29] and Guo et. al [30] used the Tundo et al. [22] approach to orientation normalization. Former used location independent features, and the latter used predefined locations of the smartphone. Martin et al. [31] is listed as the only research that takes into consideration varying orientation and location of the smartphone. In addition to smartphone-embedded accelerometer, they also use the proximity, gyroscope, magnetometer, gravity and linear acceleration data from the smartphone to estimate the orientation as a relative angle and location of the smartphone with rule based approach. For each estimated location they trained a location specific activity recognition model which increased the accuracy of the recognition. In our previous research [32], we proposed a real-time method that first normalizes the orientation, recognizes the location and afterwards uses a location specific machine-learning model to recognize five activities with 91% average accuracy over all locations. We build upon this approach in this paper.

Research in activity recognition with wrist-worn devices has started with the accelerometer sensor placed on the persons wrist as one of the sensor nodes in the body sensor network [8]. The activity recognition using only the wrist-worn sensor is more popular for recognition of activities which involve sensor-equipped-hand movement [33]. In 2009 Siirtola et al. [34] explored the recognition of sports activities with a single wrist-worn accelerometer. The activity recognition was performed in two steps. In first step, the sample is clustered into one of six clusters (one activity can be included in more clusters). In the second step, a decision tree model is used to recognise the sports activity. This approach achieved 85 % accuracy. In 2011 Chernbumroong et al. [35] achieved accuracy of 94 % for five activities (sitting, standing, lying, walking, running) using a decision tree machine-learning algorithm.

Weiss et al. [36] explored the activity recognition accuracy when recognising the same activity with smartwatch or smartphone alone (placed in the trousers pocket with predefined orientation). The scenario of activities they used was divided into non-hand oriented activities and hand-oriented activities. Overall accuracy was better with the smartwatch alone, well over 70%, since majority of the activities included hand movement (eating, folding clothes, handwriting, etc.) which can't be recognized with the smartphone placed in the trousers pocket. The smartphone activity recognition achieved rather poor accuracy of 30% since it was able to recognise only the non-hand oriented movements (walking, jogging, etc.). Ramos et al. [37] combined the smartphone (trousers pocket) and the wristband to achieve 80% accurate activity recognition of four activities (walking, sitting, standing, driving), but do not report if the orientation of the phone was predefined.

In summary, several researchers investigated the problems of varying orientation and location of the smartphone (or a dedicated sensor) on the users body. For the most part, they found that taking this into account increases the activity recognition accuracy, although the gains varied widely. Only a few of these researchers developed activity recognition systems that automatically take the orientation or the location of the smartphone into account. To the best of our knowledge, the system described in this paper is the first that normalizes the orientation of the smartphone and the wristband, detects the location of the smartphone (trousers, jacket and bag) and fuses the data of both devices if present for activity recognition.

2.4. Estimation of Energy Expenditure

Methods for accurate measurement of the human energy expenditure are expensive and cumbersome (direct and indirect calorimetry or doubly labeled water approach) and not applicable in everyday life [38]. The alternative is to estimate the expended energy using the smartphone and/or other sensor-rich wearables.

The pioneers of pervasive technology in estimation the energy expenditure used a single accelerometer attached to the user's body, usually waist, and tried to correlate motion intensity (activity counts) with the energy expenditure using a single count-based regression equation. This proved insufficient for the energy expenditure estimation of light and vigorous activities [39], so Crouter et al. [40] refined the count-based approach by using different regression equations according to the performed activity (sedentary, ambulatory, lifestyle) recognized from the number of activity counts. For the sedentary activity they assign static value of 1 MET and use regression equations for others. The shortcoming of this approach, apart from the simplicity of the equations, is that the sensor placement (waist) was unsuitable for energy expenditure estimation of activities involving only upper or lower limbs (cycling was omitted from the evaluation). Later works introduce machine-learning based activity recognition as an essential part of the estimation of energy expenditure [41] [42] [43], also known as activity-specific estimation of energy expenditure.

Research utilizing smartphone or wristband accelerometers mostly rely on the research done with dedicated sensors. Pande et al. [44] developed a machine-learning approach which fuses the orientation-independent features from accelerometer sensor attached to the persons waist with demographic data. Altini et al. [45] fused the smartphones' accelerometer data with heart rate data which most of the wristbands have to achieve more accurate estimations.

Research in which multi-sensor fusion is used together with the activity information for the energy expenditure estimation showed a decrease of estimation error. Tapia [43] fused the heart rate information to decrease the estimation error. Altini et al. [46] proposed a combined approach where they use two groups of activities, the sedentary (three activities) and active group (four activities). For sedentary group they used the MET lookup table [10], the values of which were adjusted using the users heart rate. For other activities they used one regression model per activity. Vyas et al. [47] fused the data from accelerometer and temperature-related physiological sensors, where each sensor represented one or more contexts upon which context-specific regression model was used. In 2016 we proposed an approach which fuses the data from accelerometer, heart rate and near-body temperature sensor [48] for estimation of EE. We proposed to use three regression models, each to be used for a set of recognized activities. This approach is revisited in this paper and upgraded to be used with the smartphone and wristband.

Duclos et al. [49] developed an approach that utilities smartphone and smartwatch accelerometers for activity-specific estimation of energy expenditure. They differentiate between four activity categories according to intensity (sedentary, low-, medium-, vigorous-intensity) and use acceleration vector variance based method, essentially a predefined equation, to estimate the energy expenditure. They do not account for varying location of the smartphone nor the orientation of the both devices. They compare the results against static MET values [10] and not the real expenditure which is usually measured with indirect calorimeter.

There are dedicated devices for estimation of human energy expenditure in the market in form of a wristband [5] [50] and armband [51] but they mostly monitor the sports activities and lack on the information about other everyday lifestyle activities. According to Lee et al. [52] who compared several consumer devices in free-living conditions, BodyMedia's armband clearly outperformed other devices, which we use for comparison against our results.

In summary, research in this topic mostly covers the estimation of energy expenditure with accelerometers attached to different parts of the body. The researches evaluated how the number of accelerometers and additional modalities influence the accuracy of energy expenditure. Similar to activity recognition, we did not find any research which would continuously fuse the data of smartphone worn freely on the body and wrist-worn device and use indirect calorimeter measurements for the task.

3. Materials and Methods

To develop and evaluate the activity monitoring algorithm we used the datasets presented in Section 3.1 and methods presented from Section 3.2 on.

3.1. Datasets

We acquired one dataset to develop and evaluate activity monitoring methods and one dataset to evaluate the presence of the smartphone on the body.

3.1.1. Activity monitoring dataset

The difference between elder and younger population (which was the topic of our previous research) is reflected in the range of popular activities, the manner of performing the activities and most importantly in the metabolic rate (the energy cost of physical activities). The metabolic rate of a person depends on the person's body mass index (BMI) and age. To model the activities and metabolic equivalent of task for the target population we had to collect an appropriate dataset which contains their representative activities.

Prior to designing the scenario for the data collection, we surveyed the population characteristics regarding the health and fitness status through a questionnaire. We acquired the response from 277 persons, aged between 50 and 75 (five European countries), which gave as an insight into the most popular activities we should include into the data collection scenario. The final scenario is composed of ten tasks as presented in Table 1. The scenario contains normal everyday activities (lying, eating, cleaning, etc.) as well as ambulatory and exercise activities (walking slowly, walking normally, Nordic walk, running, etc.). The sequence of tasks was ordered by increasing energy expenditure, from light lying at the beginning building up to more intensive activities such as running and cycling at the end. This ensured that the body processes stimulated by the more intense activities did not distort the data collected during the less intense activities. For the same reason mandatory 3 minute rest was imposed between the tasks containing moderate and vigorous activities.

Figure 1: Volunteers equipped with the wearable devices and the indirect calorimeter performing everyday activities.

Data collection was done in the laboratory environment at the Faculty of Physical Education, Sport and Rehabilitation in Poznan University of Physical Education, Poland, under the supervision of physiology and sports experts. We have recruited ten healthy volunteers: six male and four female, aged from 51 to 66 (59 ± 4.6) with different fitness levels, BMI from 22 to 29 (25.8 ± 2.3). All volunteers refrained from eating and drinking (except for water) in the 12 h prior the experiment.

The volunteers were explained the process of data collection and instructed to perform the tasks as they would do them in real life, so we could capture the variability of the tasks. They were also instructed to indicate if performed task is too vigorous (fast walking was omitted from the scenario for safety reasons). The volunteers were equipped with four smartphones (2 Samsung Galaxy S4, 2 Samsung Galaxy S2), two wristbands (Microsoft Band 2 [50] and Empatica E4 [53]), BodyMedia Fit Advantage armband [51] and Oxycon mobile [54] indirect calorimeter. The smartphones were put into the volunteers' trousers pocket, jacket pocket and bag (each with a random orientation), while the fourth smartphone was put in torso pocket and was used for collecting the data from Microsoft Band 2, which was worn on the left hand. The second wristband Empatica E4 was worn on the right hand. The Bodymedia FIT Advantage [51] data was recorded for comparison of our results and the Oxycon mobile for the ground truth measurements of the human energy expended. The Oxycon Mobile indirect calorimeter has good indicators of measurement reliability [55, 56] and is commonly used in validation studies of other devices [57, 58]. Since it is mobile it can be used both in laboratory environment as well as in free-living conditions during every-day activities which made it suitable for our research.

The users were instructed to wear the Microsoft Band 2 on the left hand as they thought correct. Two users wore it in the opposite direction as we considered correct orientation and two people changed the orientation in the middle of the scenario. This data was used to test the normalization of the orientation of the wristband. Figure 1 shows two volunteers equipped with

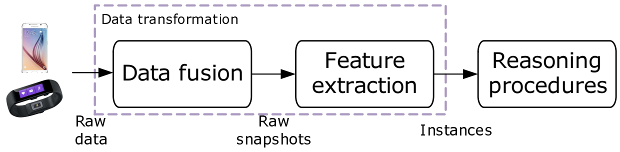

Figure 2: Data transformation procedure

all devices performing activities from the scenario.

We acquired approximately 110 minutes of data from each sensor for each volunteer.

3.1.2. Presence detection dataset

Since the volunteers did not remove the smartphones from their location during the data collection procedure, we had to acquire separate dataset for evaluation of the smartphone's presence. We prepared a fifty-minute scenario designed to test corner cases of the smartphone use. It was carried out by five volunteers with Samsung Galaxy S6 smartphone.

The scenario included wearing the smartphone in all relevant locations for this research: trousers pocket, jacket pocket, breast pocket and bag while the volunteer was moving and while the volunteer was sedentary. We also collected the data when the smartphone was in the hand, with screen either on or off, where it was in a bag but the bag was not worn, and where it was placed on some surface. Both horizontal and non-horizontal surfaces were tried, with smartphones screen facing either up or down, and for both still and shaking table (someone was typing and moving mouse nearby).

3.2. Data fusion

The goal of data fusion is to combine and preprocess the data into a form suitable for further reasoning. The data fusion is the first step of data transformation procedure presented in Figure 2 in which the raw data is fused into time-aligned raw snapshots, which is not trivial due to different sensor frequencies. Table 2 presents the modality and frequency details per device sensor. To fuse the data into structure which contains raw data (raw snapshots) we first divided the sensors in two groups with different modes: the high frequency sensor data (accelerometers) are assigned the "necessary" mode, and the low frequency sensor data (other sensors) are assigned the "duplication" mode.

The construction procedure of the raw snapshots expects at least one sensor reading from the sensors in the necessary mode to release a snapshot as complete for further processing. The sensors in the duplication mode attach the last received sensor reading to the raw snapshot under the condition that it is time-aligned with the data in the necessary mode currently contained by the raw snapshot. The time between the duplication and necessary data should not exceed one second to be included into the same raw snapshot. The exceptions are the Oxycon mobile data which is used in the energy expenditure estimation method, the proximity data, which is used for smartphone presence detection, Bodymedia data and Microsoft Band 2 calories data, which are used for comparison of the results. The Oxycon mobile data is duplicated if the time between the duplication and necessary data does not exceed ten seconds, which matches the needs of the energy expenditure estimation method (Section 3.5.6). Proximity sensor reports only when the change in proximity is detected, this data is duplicated without any specific conditions. The Bodymedia data is duplicated if the time between duplication and necessary data does not exceed one minute, and is not used for final real-time monitoring but merely for the comparison of the energy expenditure estimation method performance. The Microsoft Band 2 calories were duplicated without any specific conditions.

Table 2: Frequencies per device [50, 51, 54] sensor.

Device	Sensor	Frequency
Smartphone	Accelerometer	45 – 50 Hz
	Proximity sensor	Reports a change
Microsoft Band 2	Accelerometer	40 – 45 Hz
	Heart rate	1 Hz
	RR interval	2 Hz
	Skin temperature	12 Hz
	GSR	5 Hz
	Calories	0.1 Hz
BodyMedia	Estimated MET	1 per minute
Oxycon mobile	Measured MET	1 per 10 seconds

This data fusion procedure decreases the amount of missing data compared to an approach without duplication and gives us approximately 45 raw snapshots per second if only smartphone data is used, and 40 raw snapshots per second if only wristband is used or both devices are used. The raw snapshots are sent to the feature extraction module for further processing, where additional fusion in done as explained in the next subsection.

3.3. Feature Extraction

The feature extraction procedure is the second step of data transformation procedure presented in Figure 2. It transforms the raw snapshots into feature vectors which are used by machine-learning and rule-based algorithms. The raw snapshots are collected into windows, the length of which depends on the task to be solved. In this research we address four machine-learning tasks for which we use two window lengths. We use 2-second windows for activity recognition and walking detection and 10-second window for smartphone location detection and estimation of energy expenditure (explained in detail in Section 3.5). For each window a number of features are extracted which form a feature vector ready to be used with the machine-learning algorithm.

Raw acceleration data is low-pass (removing noise) and band-pass (removing noise and gravity) filtered, giving us three values for each axis. Low-pass filtered values are used in features describing orientation, while band-pass filtered values are used in features describing movement. The magnitude of the three-axis vector is also computed. Intuitively, magnitude gives us the intensity of motion, useful for distinguishing running

Table 1: Data collection scenario per task, its duration and average measured MET with the indirect calorimeter. In task Walking carrying a burden, female were carrying 2 kg burden and male 4 kg burden and in Walking uphill the inclination was 5% inclination.

Task	Activities	Time	Avg MET
Lying	Lying left side	1′	1.2
	Lying front	1′	1.5
	Lying right side	1′	1.4
	Lying back	7′	1.2
Basic activities	Walking slowly	10′′	3.5
	Sitting down at the desk		
	Sitting still	4′	1.2
	Sitting doing light activities (reading, writing, leafing through a book, using computer, knitting, Rubik's cube, playing cards)	4′	1.2
	Standing up		
	Walking slowly	10′′	3.5
	Standing still	4′	1.3
	Standing talking & gesticulating	2′	1.7
	Walking slowly	10′′	3.5
	Standing washing hands	2′	2.3
	Walking slowly	10′′	3.5
	Home chores (cooking, serving food, washing dishes, sweeping floor, washing windows)	6′	2.5
Eating	Eating with cutlery	2′	1.9
	Eating with hands	2′	1.5
Gardening	Planting seedlings, digging, raking, weeding	6′	2.2
	Rest	3′	
Walking	Walking slowly (4 kmh)	4′	3.5
	Walking normally (6 kmh)	4′	4.2
	Rest	3′	
Nordic walking	Walking normally (6 kmh)	6′	4.5
	Rest	3′	
Walking carrying a burden	Walking slowly (4 kmh)	6′	4.2
	Rest	3′	
Walking uphill	Walking slowly (3 kmh)	6′	4.4
	Rest	3′	
Running	Running normally (8 km/h)	6′	7.1
	Rest	3′	
Stationary cycling	Cycling lightly (60W)	6′	4.2
	Cycling normally (100W)	6′	5.0
	Rest	3′	

from walking and sedentary activities. On the other hand, features that use one of the axis components give us more information about orientation and direction, which is especially useful at distinguishing between sedentary activities, such as sitting, standing and lying. We extract 90 features from acceleration data when single device is present, 13 of which are orientation-independent. Some features come from statistics and describe the intensity and "shape" of the signal: the mean, variance, Pearson's correlation between axes, their covariance, skewness, kurtosis, quartile values and range between them. Other have a more physics-based interpretation, such as velocity and kinetic energy. The rest came from expert knowledge of the domain: the number of peaks in the signal, their height, their mean and their sum of the acceleration size, the number of times the signal crosses its mean value, sum of all data from an axis and squared sum of this data. Note that since most of the listed methods can work with low-pass or band-pass filtered data, with any axis or with magnitude, they each generally generate more than one feature (mean of the x-axis, mean of the y-axis, etc.).

Some features are calculated from the filtered raw signals and some are constructed with feature fusion, for example the mean kinetic energy is calculated from the velocity feature and the person's BMI; skewness, kurtosis, crossing rates and correlations are calculated using the feature representing sum for each axes, etc.

When both devices are present we calculate 192 acceleration features: 90 from each device and 12 features that use data fused from both accelerometer sensors. They are the mean of all data, its sum, area below the magnitudes signal (marked with keyword "total" on Figure 3 and Figure 4) and correlations between axes of the devices. Additionally, we add decision features where possible. Decision features are the outputs of machine-leaning models and are used for decision fusion. When both devices are present, we use the outputs of activity recognition model per individual device as two decision features (phone activity and band activity) in the activity recognition feature vector for the device combination and in the energy estimation task we add the recognized activity as a feature into the energy expenditure estimation feature vector. These are also used for decision fusion in the activity monitoring procedure.

Physiological sensor data retrieved from the wristband are transformed into the average, maximum, minimum and standard deviation value for the current window, which gives us altogether 16 features from the four sensors in the wristband.

One energy expenditure feature is extracted from Bodymedia and one from the Oxycon mobile device. Former is used for comparison and the latter is used as the ground truth.

All features for a given device combination represent a feature vector which is used to build models with machine learning. Since the goal of this research is to design models to be used on a smartphone, the number of features should be decreased and the ones with high computational complexity omitted. The set of features for a given task is selected with the feature selection procedure.

Note that we use raw values of proximity sensor and calories as estimated by the Microsoft Band 2, but these are not included into the feature vector.

3.4. Feature Selection

The feature selection procedure is composed of two steps: a single-feature evaluation, and a wrapper-based feature selection, which evaluates combinations of features and chooses the final feature set. Feature selection is performed only once for each device combination to design the final feature vector used for machine learning.

The single-feature evaluation procedure starts with entire feature set. It uses Gain Ratio measure to evaluate the information about the class carried by each feature in the classification tasks. In regression tasks, it evaluates the importance of features as Pearson's correlation with the class.

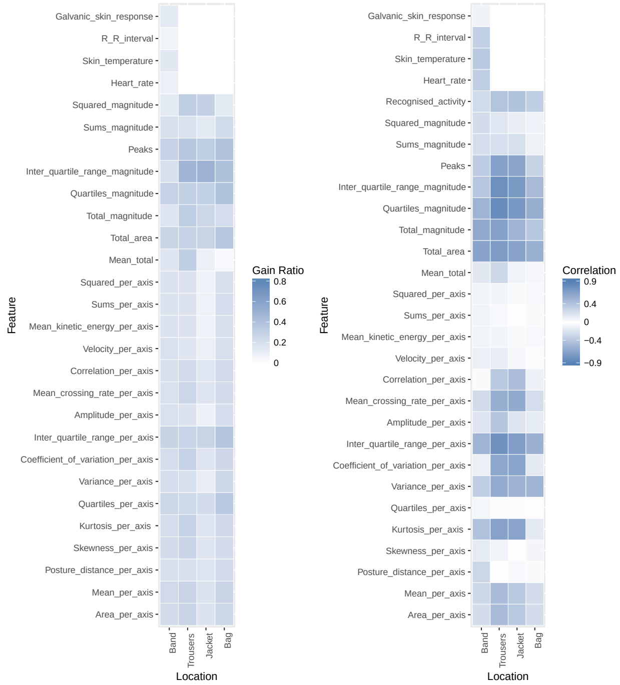

Figure 3: Gain ratio and correlation heatmaps of all extracted features when single device is used.

Figure 3 and Figure 4 show heatmaps of the results of the evaluation procedure for the activity recognition (classification) and energy expenditure estimation tasks (regression) for individual devices and for the combination of devices. For a more compact representation we merged the features which are calculated for individual axis or are similar into single features with averaged mutual information value. Left side of Figures 3 and 4 presents the gain ratio value for the features used in activity recognition. The higher the value, the more information is carried by a feature. The right side of Figures 3 and 4 present the Pearson's correlation coefficient for features used in estimation of energy expenditure task. The correlation values around 0 indicate low correlation and values towards 1 and -1 high positive and negative correlation respectively. Both, gain ratio and correlation were calculated with procedures implemented in the Weka machine-learning suite [59].

The features in the figures are listed as follows. In Figure 3 the physiological sensor features are at the top, next are the decision features (in the case of the energy expednditure estimation estimation), followed by the magnitude features combining all axes, and per-axis features are at the bottom. It is similar in Figure 4: the physiological sensor features are at the top, next are the decision features followed by the features extracted from the wristband's accelerometer (first magnitude features, then per-axis features) and features extracted from the smartphone's accelerometer (first magnitude features, then per-axis features).

The single-feature evaluation procedure in activity recogni-

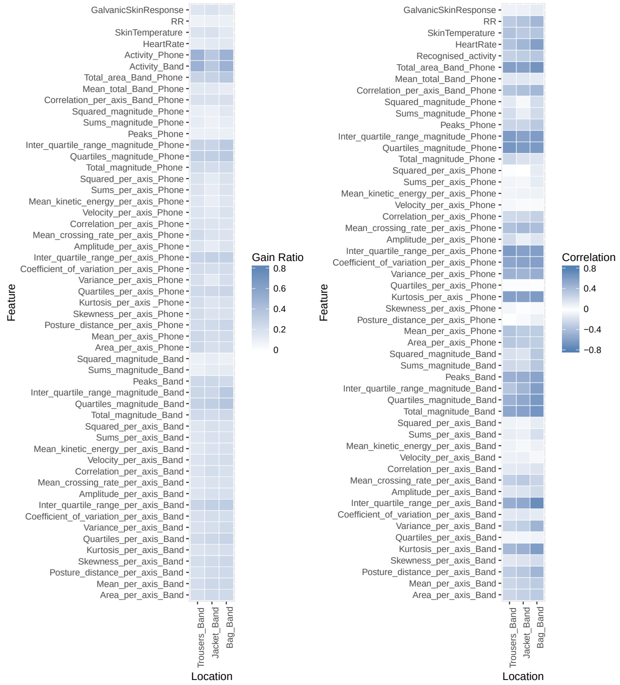

Figure 4: Gain ratio and correlation heatmaps of all extracted features when both devices are used and the smartphone is worn in the trousers pocket.

tion (left sides of Figure 3 and Figure 4) prefers the magnitude features and decision features, some significance is shown for the interquartile range and quartiles features which are orientation-dependent. Common sense tells us that per-axis features can provide more information about the sedentary activities, where not much motion is performed and the orientation of the smartphone is important (sitting, lying, standing, etc.). This justifies our decision to use the wrapper approach, which can remove redundant magnitude features and add important per-axis features to be used in the feature vector.

The single-feature evaluation procedure in energy expenditure estimation (right sides of Figure 3 and Figure 4) shows low correlation for some per axis features and high (positive or negative) correlation for magnitude and decision features as well as for the features extracted from the physiological sensors. Using only the single-feature evaluation procedure would result in to many features, which may be redundant and would increase computational complexity so we perform the wrapper-based feature selection to reduce the number of features.

The feature selection wrapper is rather simple but efficient. The input to the procedure are the list of features ranked by the value retrieved from the single-feature evaluation procedure, the machine-learning algorithm for which the feature selection is performed and manually created folds for testing. In our case, we evaluate the models with the leave–one–person–out approach (LOPO) so we have n folds, where the data of $n-1$ people are used as the training set and one person as the testing set (fold). The loop iterates over the list of features and adds them one by one, by decreasing rank, and performs the LOPO evaluation in terms of accuracy for classification and in terms of mean absolute error (MAE) for regression (any other measure

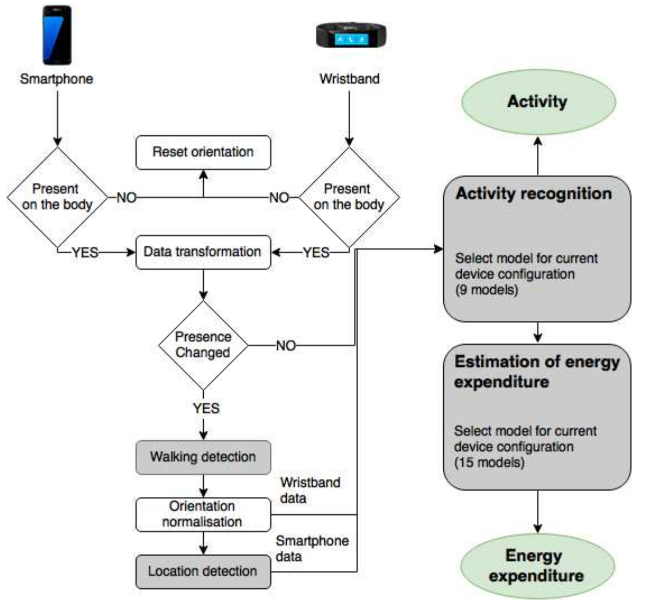

Figure 5: Activity monitoring algorithm work flow. The modules on the left side manage the device configuration and the right side performs the activity analysis (activity recognition and energy expenditure estimation).

can be evaluated). For each new feature set the accuracy/error is compared to the prior one and in case it is higher/lower the last added feature is kept.

3.5. *Activity Monitoring*

Activity monitoring is designed to track the users activity in real-time with the devices currently present on the body. To achieve that, we developed the activity monitoring algorithm presented in Figure 5 with six consecutive modules: four modules to manage the device configuration (left side in Figure 5) and two modules to analyst the activity (right side in Figure 5).

In brief, the sensor data received is first evaluated for the presence of the devices with simple heuristics, then the data is fused and features extracted in the data transformation module (Sections 3.2 and 3.3). The algorithm waits for ten seconds of walking for gravity detection, the data of which is used for normalizing the orientation of each device. While the algorithm waits for the normalization data, the modules for activity analysis use machine-learning models trained to be used without orientation and location information for AR. Once the sensor data is normalized, it is used in all further machine-learning tasks. The first is to detect the location of the smartphone, which is then used for the selection of the machine-learning models for activity recognition and estimation of energy expenditure.

Note that the tasks with gray boxes in Figure 5 use machine learning. All machine-learning tasks use the feature set which was defined with feature selection. The evaluation is in Section 4.

3.5.1. *Presence of the devices*

Presence of the devices is needed for identifying which data is available to be used for activity monitoring. The presence of the Microsoft Band 2 is straightforward since it has an ability

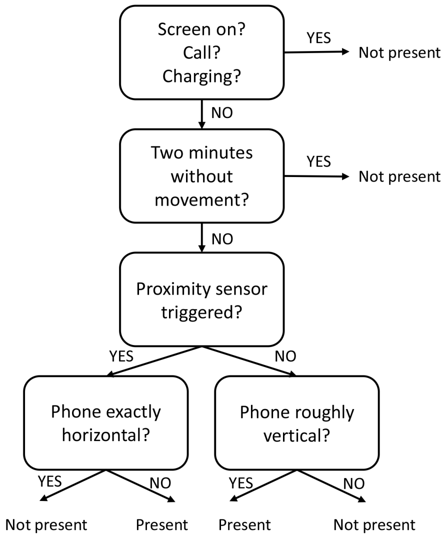

Figure 6: Heuristics for detection of the smartphone's presence.

to self-report its presence, however the detection of the smartphone presence is a bit more difficult.

We assume that the smartphone is present if it is placed in a pocket or a bag (while the bag is worn). To detect that, we use the smartphone-embedded proximity sensor, the accelerometer data and the information about the phone state as shown in Figure 6. First, we check if the smartphone's screen is on, if there is an on-going call or if the phone is charging, all of which indicate that the smartphone is not worn. Second, we detect the amount of movement in last two minutes (looking at the accelerometer data), no movement indicates that the smartphone is not present on the body. Even if the user is sedentary, we expect some movement in two-minute time interval. Third, we check the value of the smartphone proximity sensor, which tells us if there is something close to the smartphone which could mean that it is in a pocket. In case three is something in proximity, we check if the smartphone is completely horizontal (easily determinable from accelerometer data), to eliminate the common false positive of the smartphone being on the table face down. Nothing in proximity implies that the smartphone is not in a bag or a pocket, but we again check for one exception - the smartphone being in a breast pocket with its top end looking out of the pocket. We do so by checking if the smartphone's orientation is roughly vertical.

Data from the present devices is sent into the data transformation module (Section 3.2). If the presence was not changed, the data from the data transformation module is transformed into feature vectors for activity analysis and are feed directly into activity analysis modules, otherwise the orientation is reset and data is transformed into feature vector needed for the walking detection.

3.5.2. Walking detection

Walking detection is essential for normalization of the orientation of the devices. It gives us information about the direction of gravity and under assumption that the average acceleration during walking corresponds to the Earths gravity we can use it to normalize the orientation.

The walking detection is a machine-learning tasks done essentially in the same way for both devices. The data transformation procedure segments the raw data of each present device into separate 2-second window and extracts the orientation-independent features creating a feature vector. The walking detection model is a binary classifier trained to distinguish between walking activity and all other activities. While the model classifies 2-second intervals, the final decision – based on which orientation normalization is performed – is made for ten second interval. We assume that walking is detected if the four out of five consequent classifications are walking. This satisfies the smartphone' prerequisites for orientation normalization and prerequisite to detect the wristband gravity along the x-axis (buttons up/down).

3.5.3. Orientation normalization

The normalization of the orientation is based on the gravity detection. Prerequisite for the normalization is thus the detection of walking, which is described in the previous subsection. Let $\vec{a}_{\text{walk}} = (x_{\text{walk}}, y_{\text{walk}}, z_{\text{walk}})$ be the acceleration vector consisting of the average accelerations along the three accelerometer axes during ten seconds of walking.

Orientation normalization is needed for both devices. We have seen from practical experience that even if the wristband is always worn on the left wrist, it sometimes occurs that the users wear it upside down (buttons down), which switches acceleration x-axis in opposite direction and therefore inflicts errors in extracted orientation-depended features and consequently classification and regression errors.

Wristband orientation normalization uses simple heuristics to detect vertical direction. We assume that the vector $\vec{a}_{\text{band}} = (g, 0, 0)$ is the preferred acceleration vector when the wristband is correctly worn in the vertical direction, while walking. To normalize the orientation it is sufficient to check if the x-axis is negative and if absolute value exceeds $0.8g$ and if it is, all further x-axis data from the wristband are multiplied by -1, otherwise the orientation in vertical direction is considered as correct.

In Figure 7 we present the wrist postures with perfect gravity detected. On left is the posture while standing/walking (A), middle is the posture while the hand is flat on the table (B) and on the right is the posture while hand is on the side on the table (C). Table 3 presents the values per-axis in different orientation (buttons up/buttons down and inside/outside) and different wrist postures presented in Figure 7. We can observe, that it is also possible to normalize z- and y-axis. Since our dataset did not include such data we only evaluated the x-axis orientation normalization.

The normalization of the smartphone is more difficult since the smartphone can be worn in any orientation. Let $\vec{a}_{\text{smartphone}} = (0, 9.81, 0)$ be the preferred acceleration vector, which would be obtained if the smartphones longest side was perfectly aligned with the gravity. To normalize the smartphone's orientation it has to be rotated so that the y-axis of $\vec{a}_{\text{walk}}$ corresponds the

Table 3: Wristband axis at different orientations.

	Buttons up		Buttons down	
Position	Inside	Outside	Inside	Outside
A	(g, 0, 0)	(g, 0, 0)	(−g, 0, 0)	(−g, 0, 0)
B	(0, 0, −g)	(0, 0, g)	(0, 0, −g)	(0, 0, g)
C (Left)	(0, −g, 0)	(0, −g, 0)	(0, g, 0)	(0, g, 0)
C (Right)	(0, g, 0)	(0, g, 0)	(0, −g, 0)	(0, −g, 0)

Figure 7: Wrist postures. Left is standing posture (A), in the middle is the hand down flat on the table posture (B) and on the right is hand down on the side posture (C).

$\vec{a}_{\text{smartphone}}$. The rotation between these two vectors is represented with the quaternion matrix R, such that $\vec{a}_{\text{smartphone}} = R * \vec{a}_{\text{walk}}$. To compute R, we adopted the approach by Tundo et al. [22]. The procedure is as follows. First, we compute the axis-angle pair $\vec{a}_{\text{pair}}$ as the cross product between the preferred vector and the walking vector, i.e. $\vec{a}_{\text{pair}} = \vec{a}_{\text{smartphone}} \times \vec{a}_{\text{walk}}$. Second, we normalize $\vec{a}_{\text{pair}}$ by dividing it by its magnitude giving us $\vec{a}_{\text{norm}}$, which is needed for the quaternion construction. Third, we use the dot product to compute the angle α between the vectors $\vec{a}_{\text{smartphone}}$ and $\vec{a}_{\text{walk}}$ as presented in Equation 1.

$$\alpha = \arccos \frac{\vec{a}_{\text{smartphone}} \cdot \vec{a}_{\text{walk}}}{\left\|\vec{a}_{\text{smartphone}}\right\| \left\|\vec{a}_{\text{walk}}\right\|} \tag{1}$$

Finally, the quaternions are calculated according to Equation 2 and the matrix R is calculated according to Equation 3.

$$\begin{aligned} q_0 &= \cos(\frac{\alpha}{2}) \\ q_1 &= \sin(\frac{\alpha}{2})\vec{a}_{\text{norm}} \cdot x \\ q_2 &= \sin(\frac{\alpha}{2})\vec{a}_{\text{norm}} \cdot y \\ q_3 &= \sin(\frac{\alpha}{2})\vec{a}_{\text{norm}} \cdot z \end{aligned} \tag{2}$$

$$R = \begin{bmatrix} 1 - 2(q_2^2 + q_3^2) & 2(q_1 q_2 - q_0 q_3) & 2(q_0 q_2 + q_1 q_3) \\ 2(q_1 q_2 + q_0 q_3) & 1 - 2(q_1^2 + q_3^2) & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_0 q_1 + q_2 q_3) & 1 - 2(q_1^2 + q_2^2) \end{bmatrix} \tag{3}$$

Once the matrix R is computed, each sensed acceleration vector $\vec{a}_{original}$ can be normalized to the preferred orientation in real time, thus creating the normalised acceleration vector $\vec{a}_{normalised}$ using the Equation 4.

$$\vec{a}_{\text{normalised}} = R * \vec{a}_{\text{original}} \tag{4}$$

3.5.4. Location Detection

Once the data is normalized we use machine learning to detect the location of the smartphone.

The data transformation procedure segments the normalized data from walking detection into 10-second window and extracts the orientation-dependent and orientation-independent features creating a feature vector ready to be used in the machine-learning task. The location model is trained to distinguish between three classes representing the location of the smartphone relative to the user's body. These are the trousers pocket, jacket pocket and bag. Location is assumed immediately.

The wristband location detection is not a subject of this paper, but we can conclude that it is feasible to do so according to the y-axis when worn on left or right wrist (Table 3).

3.5.5. Activity Recognition

Activity recognition is a classification machine-learning task. The normalized data is processed with data transformation procedure and segmented into 2-second windows from which the orientation-independent and orientation-dependent features are extracted forming a feature vector which is fed into the machine-learning classification model. We trained nine location- and combination-specific classification machine-learning models, one for each device configuration (7 models) and one for each device (2 models) to be used before orientation is normalized. The model selection therefore depends on the currently present devices and on the recognized location of the smartphone if present (device configuration). The main reason for differentiating between the locations is that the type of motion detected by the accelerometer embedded into the smartphone depends on the location where it is attached, e.g., the smartphone in trousers pocket does not detect the motion of the upper body as much as if the smartphone is in the jacket pocket. It is similar when combinations of devices are used. It is important to adapt the set of activities to be recognized to the location and combination accordingly and to exploit as much information as possible.

All recognized activities are listed in Table 4. Each row corresponds to one activity we can recognize, and each column to one device configuration. First are the no orientation models followed by the seven models according to the device, location and combination. We can observe that walking and running are recognized with all device configurations, whereas Nordic walking is recognized only with the wristband or in combination with the wristband when the orientation is known. Upright activity is one of the activities recognized with a single device, the smartphone located in the jacket pocket. It is composed of sitting and standing which cannot be distinguished only with an accelerometer attached to the persons torso due to the same orientation and amount of movement. When we add the wristband to the jacket-located smartphone, we can distinguish sitting from standing, and the number of recognized activities increases from five to eight. Since bag is usually not worn indoors (while lying, sitting), we can observe that the combination of the wristband and the smartphone located in a bag can recognize fewer activities as than the wristband alone. However with this combination we can recognize cycling which cannot be recognized with the wristband only. The data in the dataset presented in Section 3.1 contains 33 activity labels which were first merged into 10 activities (row labels in Figure 8) and afterwards we merged these activities (colors in Figure 8) to optimize the recognition according to the device configuration (columns in Figure 8). For example, for the device configuration T (smartphone placed in trousers pocket) sitting and lying were merged into rest because the position of the smartphone is the same in both activities, thus impossible to distinguish accurately.

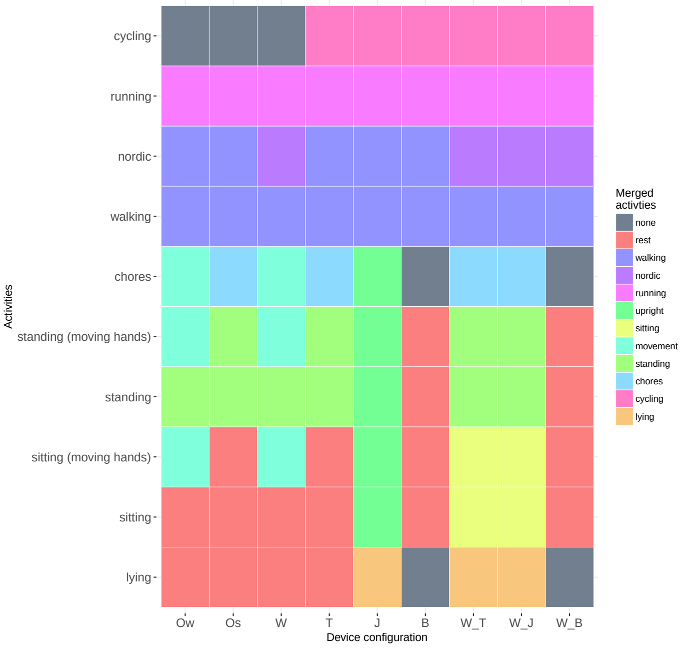

Figure 8: Merged activities per device configuration.

Table 4: Recognized activities (most left) per wristband and smartphone when orientation information and location is missing (second group of columns) and wristband and smartphone location (third group of columns) and combination of devices (most right) when orientation normalized and location detected. O_w = wristband without orientation, O_s = smartphone without orientation and no location, W = wristband, T = trousers location, J = jacket location, B = bag location. In combinations it is assumed that the wristband is also present.

	Single device						Combination		
Activity	O_w	O_s	W	T	J	B	T	J	B
Rest	✓	✓	✓	✓	×	✓	×	×	✓
Standing	✓	✓	✓	✓	×	×	✓	✓	×
Lying	×	×	×	×	✓	×	✓	✓	×
Sitting	×	×	×	×	×	×	✓	✓	×
Upright	×	×	×	×	✓	×	×	×	×
Movement	✓	×	✓	×	×	×	×	×	×
Chores	×	✓	×	✓	×	×	✓	✓	×
Walking	✓	✓	✓	✓	✓	✓	✓	✓	✓
Nordic	×	×	✓	×	×	×	✓	✓	✓
Running	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cycling	×	×	×	✓	✓	✓	✓	✓	✓
#	5	5	6	6	5	4	8	8	5

3.5.6. Estimation of Energy Expenditure

Estimation of energy expenditure is a regression machine-learning task. The data is segmented into 10-second windows, which was adopted as the minimal sensible interval based on our previous research. The data transformation module extracts the orientation-independent and orientation-dependent features from the data stream and includes the recognized activity to form a feature vector which is fed into a regression machine-learning algorithm.The output is the estimated energy expenditure expressed in MET.

We have developed and evaluated two approaches:

1. Single model approach
2. Multiple-model approach which utilities recognized activity for decision fusion about model selection

In single model approach we developed seven regression models, one for each device configuration with features extracted from all available data.

In the multiple-model approach we developed multiple regression models for the device configurations where different sensor modalities are available. Our previous research [9] showed that energy expenditure estimation performance is improved if similar activities with a large range of possible MET values (ambulatory activities) and activities with similar range of possible MET values have their own regression models, so we trained three models per device configuration where different modalities are present. The device configurations where only smartphone is present use single model approach

All together we trained 15 models which are as follows:

- Ambulatory regression models: four models are trained to be used when the recognized activity is either walking or running
- Low regression models: four models were trained to be used when low-intensity activities such as lying, sitting, rest are recognized
- Other regression models: seven models were trained to be used with all other recognized activities and in case only smartphone is present

Feature vectors of the ambulatory regression models and the other regression models are composed of features extracted from all available modalities. The feature vector of the low regression models are composed of features extracted from the accelerometers only and not physiological sensors, even if the wristband is present.

The evaluation of both approaches is presented in the Section 4.7.

4. Evaluation Results

For the comprehensive understanding of the performance of our algorithm, we evaluated each task separately in the same sequence as implemented. The errors in orientation normalisation and AR are carried over to subsequent tasks, while we evaluated each location separately. The tasks were evaluated on the dataset presented in Section 3.1 in the LOPO manner. We experimented with six machine-learning algorithms for classification tasks (J48 - decision trees, Support Vector Regression, JRip - decision rules, Random Forest, Naive Bayes and IBK - clustering) and five for regression task (Support Vector Regression, Multi-Layer Perceptron - neural network, Linear Regression, Random Forest and RepTree - decision tree) all as implemented in the Weka machine-learning suite [59]. We report the results of the best performing algorithm only, which was the Random Forest algorithm in all classification and regression tasks.

The results for classification are presented in terms of the classification accuracy (the percentage of correctly classified feature vectors), the kappa statistic (the homogeneity of the classifications) and the F-score measure. The results of regression are presented in terms of mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean squared error (RMSE).

4.1. Presence detection

We evaluated the presence detection on the separate dataset presented in Section 3.1. In 90% of cases, the smartphone was correctly identified as being present/not present. The error occurred only in two cases. The first case was when a person left the smartphone on the table and the heuristics needed two minutes to decide that the smartphone is not present. These two minutes contributed to the error which could be eliminated by identifying the two-minute period as not present retroactively, which we decided not to do since we wanted a real-time algorithm. The second case was when the smartphone was screen-down on a non-horizontal surface on a shaking table (while someone is typing), as it is very similar to the position when smartphone is in a trousers pocket while the person is sitting.

4.2. No orientation and no location

These models are used every time the presence of the devices changes until orientation is normalized and location detected. The results of the classifier trained with Random Forest are presented in Table 5. The accuracy for activity recognition is around 79 % for the recognition of the five activities as given in Table 4 in columns O_w and O_s. The results for kappa shows that there is high intra-class variability and that the classification is not very stable. This argues for the normalization of the orientation and detection pf location of the devices. We can also observe that after feature selection the number of features decreased from 13 orientation-independent features to 8 features in the case of the wristband and 6 features in the case of the smartphone.

At this stage we are most interested in recognition of walking and the results of the per-class f–score (around 0.8 for individual device) indicate that this is feasible with specialized classifier.

4.3. Walking Detection

The walking detection is performed for both devices separately with objective to detect the gravity needed for orientation normalization of both devices. The classifier was trained on the entire dataset on orientation-independent features. The non-walking samples were re-labeled with “other”, to create a binary classification problem. The results of walking detection for individual device are presented in Table 6. We can observe that we achieve accuracy over 90% with only 6 orientation independent features in case of a wristband and 91% accuracy with 9 orientation independent features in case of a smartphone. One average we achieve 0.8 kappa and 0.9 f-score.

Table 5: Results of the AR without orientation and location (W=wristband, S=smartphone, # = number of features after feature selection).

Device	Accuracy	Kappa	F–score	#
Wristband	79.2	0.67	0.63	8
Smartphone	79.4	0.67	0.50	6

Table 6: Results of the walking and location detection. Walking detection is performed for both devices separately and location detection only when the smartphone is present (W=wristband, S=smartphone, # = number of features after feature selection).

Detection	Device	Accuracy	Kappa	F-score	#
Walking	W	90%	0.76	0.89	6
	S	91%	0.83	0.91	9
Location	S	91%	0.88	0.91	11

4.4. Orientation normalization

In this experiment we could evaluate only the accuracy of the orientation detection for the wristband. The true smartphone orientation could not be clearly defined so we cannot evaluate it.

We evaluated the heuristics for wristband orientation normalization on the entire dataset. First we waited for the ten seconds of recognized walking, which was detected using the model explained in the previous subsection. When walking was detection we used the walking data to evaluate the up-down orientation of the wristband and normalize x-axis. We achieve 100% accuracy on our dataset. The accuracy might decrease in such cases where the hand is raised above a head and the movements are miss-misclassified as walking. However, we did not have such events in our dataset.

4.5. Location Detection

The location detection is performed for smartphone only. The classification model is trained on the normalised data classified as walking from the walking detection (Section 4.3). The feature vector is constructed from both orientation-independent and orientation-dependent features. The data was labelled with the three labels: trousers, jacket and bag which comply with the location of the smartphone from which the data was acquired. The results for location detection are presented in the bottom row of Table 6. We can observe that feature selection decreased the number of features from 90 to 11 which resulted in 91% accuracy with high kappa and f-score. Interestingly, only one orientation-independent feature remained in the final feature set, other 10 are orientation-dependent.

The confusion matrix is presented in Figure 9 showing that the smartphone at trousers location is best recognised and that some miss-classification occurs between jacket and bag location, probably because both locations are on the torso.

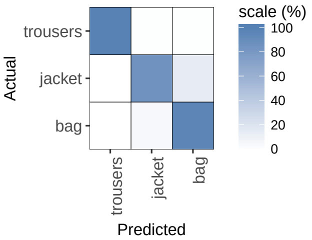

Figure 9: Confusion matrix for smartphone location detection.

4.6. Activity recognition

Activity recognition was evaluated for each device configuration separately. The activity recognition classification models were trained on the normalized data from all scenarios in the dataset with different set of merged activities as labels (presented in Table 4 and Figure 8). The feature selection procedure returned a set of the best performing features to be used in the feature vector. The details on the feature vector structure is presented in Table 8.The results for each device configuration is presented in Table 7 and in form of the confusion matrices in Figure 10 and Figure 11.

Each model was trained to recognize merged set of activities to find a trade-off between the acceptable accuracy and number of recognized activities. The number of activities per device configuration is also presented in Table 7. The results show that the accuracies are mostly over 85% and that kappa and f-score also show stability for the evaluated configurations. The results of the device configurations are not directly comparable since they recognize different number of activities.

We can observe that for single device the highest accuracy (6 activities) is achieved with the classification model for smartphone in a trousers pocket. The confusion matrix for the same model shows that there is a slight miss-classification of chores into standing and walking which is reflected in lower overall f-score. This miss-classification is understandable since most of the home chores are composed of these two activities and the smartphone placement makes the acceleration measurements of upper limbs difficult. The wristband classification model achieves 80% accuracy and does miss-classifications of standing (mostly standing still data) into rest due to no movement in the wrist. The running and rest are sometimes miss-classified into movement due to wrist movement while performing these activities (gesticulating while running, gesticulating while at rest). The jacket classification model achieves 80% accuracy and does miss-classification of cycling into walking and upright activity. These errors occur because of the placement on the torso it is impossible to measure the movement of the lower limbs accurately. The classification model for bag achieves 95% accuracy since it recognizes only those activities that people do when they carry a bag (we excluded lying, eating and gardening from the training and testing data).

The combinations achieve accuracies towards 90% and can recognize from five to eight activities. We can see in the confusion matrices, that all models perform well and that there is no large miss-classification going on.

Table 7: Results of AR per device configuration.

	Single device				Combination		
	W	T	J	B	T	J	B
Accuracy (%)	80	92	80	95	89	85	89
Kappa	.75	.89	.70	.90	.87	.80	.83
F-score	.75	.78	.65	.92	.89	.83	.71
# activities	6	6	5	4	8	8	5

Table 8: Number of orientation independent (O-independent), orientation dependent (O-dependent), physiological (Physiological) and machine-learning (Recognized activity) for each activity recognition model after feature selection.

Model	W	T	J	B	T	J	B
O-independent	9	4	4	7	10	6	1
O-dependent	9	12	11	7	15	18	13
Physiological	3	0	0	0	1	0	1
Recognized activity	0	0	0	0	2	2	2
# features	21	16	15	14	28	26	17

Additionally, the number of remaining features after feature selection do not surpass 28 (decreased from over 90 in single device and over 200 in combination). Table 8 presents the number of orientation-independent, orientation-dependent features, physiological features and in case of combination of devices also the number of features which already need machine-learning (recognized activity by each device). We can observe, that orientation-dependent features form the majority of the feature vector which is the information gained by normalizing the orientation of the device which would otherwise be lost.

4.7. Estimation of energy expenditure

Energy expenditure estimation was evaluated for each device configuration separately. The energy expenditure estimation regression model was trained on the normalized data from all scenarios in the dataset labeled with the expended energy as measured by the indirect calorimeter Oxycon mobile [54]. The feature selection procedure returned a set of the best performing features to be used in the feature vector. The details on the feature vector structure is presented in Table 10.

We evaluated and compared both approaches, the single model approach and the multiple-model approach. The results for each device configuration is presented in Table 9. We first compared the results against Bodymedia Fit armband, which is one of the most accurate EE estimation devices on the market [52]. Since smartphone only device configurations lack the presence of different sensor modalities, they do not meet the requirements of multiple-model design. In Figure 12 we present the box-charts where multiple-model approach was applicable.

The results show that the errors decrease when device configuration combines devices and fuses different modalities from

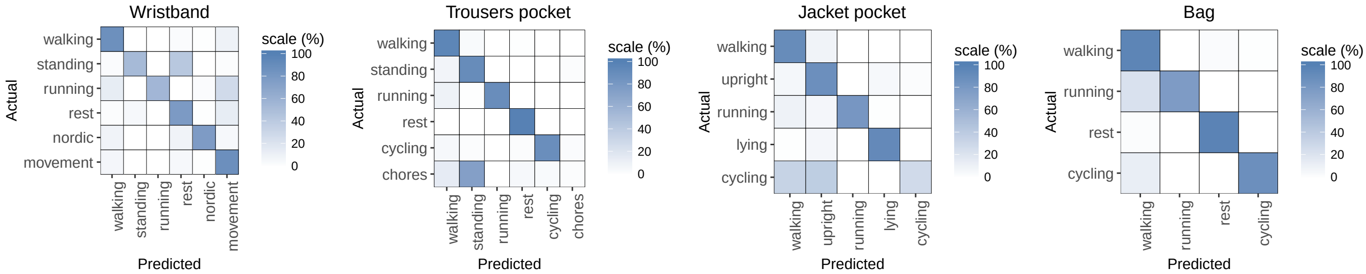

Figure 10: Confusion matrices when activity recognition is done using single device.

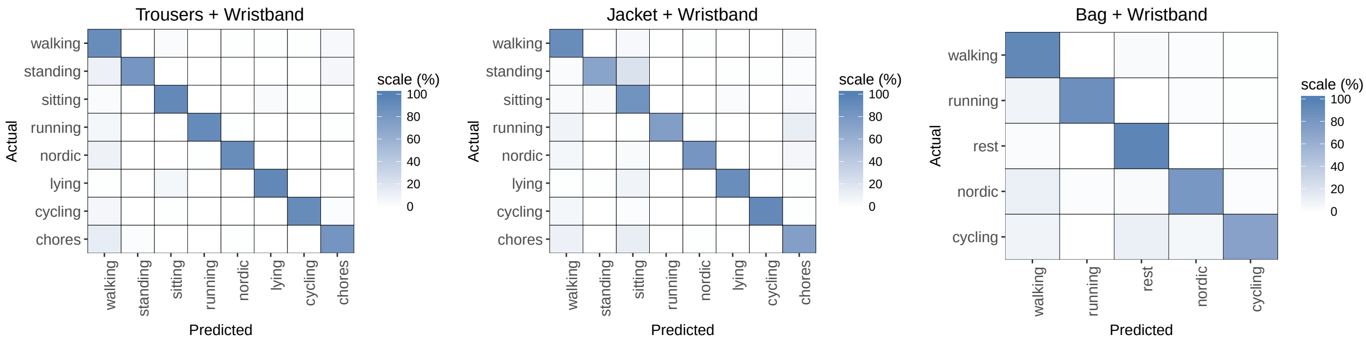

Figure 11: Confusion matrices when activity recognition is done using two devices.

the data stream. We can observe that all device configurations outperform Bodymedia armband, probably because Bodymedia armband is specifically designed for sport exercises and not every day activities. The best estimation of energy expenditure is achieved with wristband and smartphone in trousers pocket, for which we also present a chart in Figure 13 for one person and all estimations over the duration of the entire scenario (estimations each 10 seconds). The Figure 13 presents the energy expended in MET as measured by the indirect calorimeter (blue line) and estimations as estimated by our multiple-model approach (pink line). We can observe that the the errors occur when there is significant amount of hand movement while light-intensive activity is performed (e.g., washing hands) and at the beginning of vigorous-intensive activity, since the physiological signal need time to reach the exercise value.

The box-carts in Figure 12 present the distribution of error for both single-model and multiple-model approach. We can observe that when multiple-model approach is used error decreases for every device configuration and that the range of error distribution narrows, which brings more stable estimations for the activities.

We compared our results against the the Bodymedia armband and the Microsoft Band 2 in kilo-calories (kcal). For fair comparison we only compare against the results where only wristband is present. We chose to compare them in kilo-calories because Microsoft Band 2 reports the energy expenditure estimations only in this unit and the Oxycon mobile indirect calorimeter and Bodymedia armband also report kilo-calories. Since our models estimate the energy expenditure in METs, we used the Equation 5 for the conversion of estimated METs of our approach to kilo-calories. The results are reported in Table 11

Table 9: Results of energy expenditure estimation errors per device configuration. MAE is expressed in MET and MAPE in %. In combination of devices the wristband is always present.

			MAE	MAPE	RMSE
Single device	No orientation no location	O_w	.94	29	.98
		O_s	1.12	34	1.21
	Single model	W	.64	27	.86
		T	.67	26	.92
		J	.72	34	1.02
		B	.75	30	1.05
	Multiple model	W	.58	25	.79
Combination	Single model	T	.59	25	.79
		J	.71	31	.93
		B	.57	22	.74
	Multiple model	T	.55	23	.76
		J	.59	27	.82
		B	.50	18	.70
Bodymedia			1.03	37	1.60

for each scenario (Section 3.1) and device. We took the Oxycon mobile measurements as the ground truth and calculated the sum of errors and an average error for the entire scenario for each device (right columns in Table 11). From the results we can conclude that our approach outperforms the compared devices and that Bodymedia approach is comparable to our ac-

Table 10: Number of orientation independent (O-independent), orientation dependent (O-dependent), physiological (Physiological) and machine-learning (Recognized activity) for each energy expenditure estimation model after feature selection. The configuration which utilizes wristband data use multiple-model approach (a=ambulatory model, l=low level activity model, o-other model) and other use single model approach (s).

Model	W			T	J	B	T			J			B		
	a	l	o	s	s	s	a	l	o	a	l	o	a	l	o
O-independent	4	4	4	3	5	6	9	6	7	10	6	8	3	4	1
O-dependent	14	12	16	10	15	8	15	18	23	16	13	23	10	10	9
Physiological	3	0	4	0	0	0	3	0	2	3	0	3	2	0	4
Recognized activity	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
# features	22	17	25	14	21	15	28	25	33	30	20	35	16	15	15

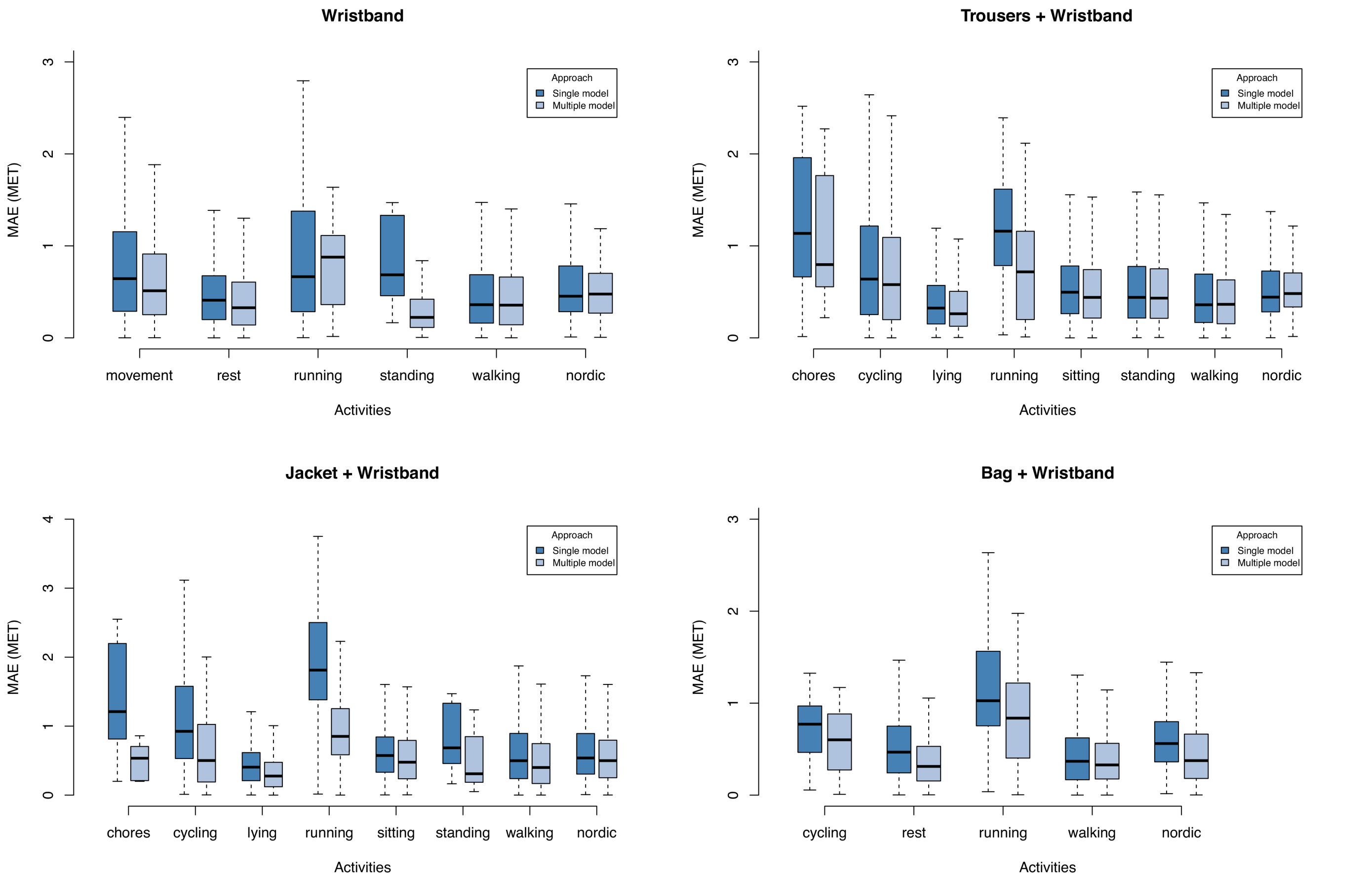

Figure 12: MAE comparison between single model approach and multiple model approaches for device configurations where multiple modalities are present per recognized activity.

cording to the average error in calories. We can observe that Bodymedia perfectly estimates the burned kilo-calories while running and Nordic walking and has significantly lower error compared to Microsoft Band 2, which heavily underestimated all activities. This also proves, that the commodity devices available on the market are not very accurate and that there is a place for improvement.

$$kcal = METs \times wight(kg) \times time(hours) \quad (5)$$

Additionally, the number of remaining features after feature selection do not surpass 35 (decreased from over 90 in single device and over 200 in combination). Table 10 presents the number of orientation-independent, orientation-dependent features, physiological features and in case of combination of devices also the number of features which already need machine-learning (recognized activity by each device). The configurations in which we could not use the multiple-model approach list the number of features per single model (s= single model) and the configurations in which we use multiple-model approach list the features per model (a=ambulatory model, l= low level activity model, o=other model).

We can observe, that the same as in activity recognition, the orientation-dependent features form the majority of the feature vector. This indicates that the decrease in error was affected by

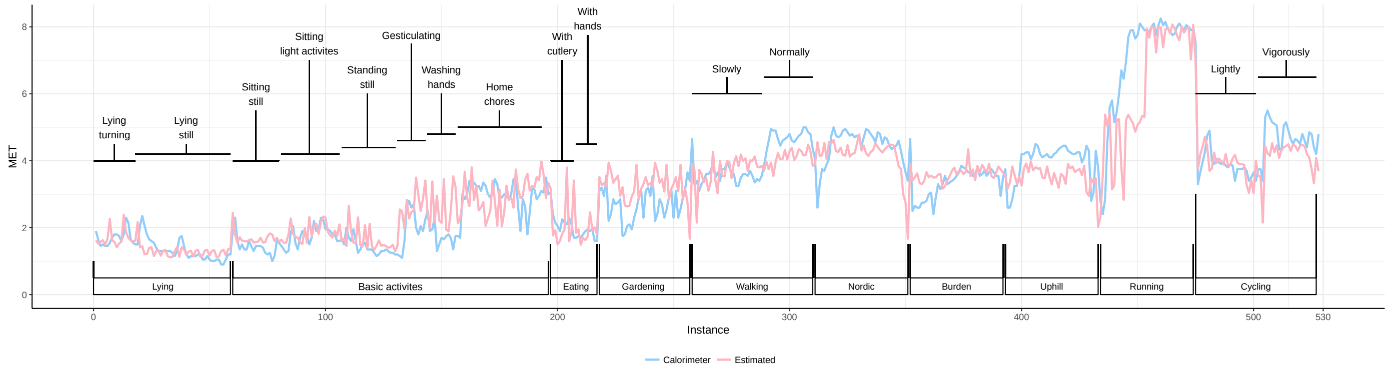

Figure 13: The comparison between the ground truth as measured by the Oxycon mobile indirect calorimeter (blue line) and estimates from our approach (pink line) for the duration of the entire scenario.

Table 11: Comparison of the energy expenditure estimation results in kilo-calories. We evaluate the wristband only device configuration to be comparable to other devices. Oxycon measures expended energy and is taken as a ground truth, Bodymedia and Microsoft Band 2 are the estimates from the devices and our approach presents the estimates of our approach. The error columns present comparison against the Oxycon mobile measurements.

	Scenario										Error	
	A	B	C	D	E	F	G	H	I	J	Sum	AVG
Oxycon	13	41	7	28	34	23	27	26	35	45	/	/
Bodymedia	10	44	7	35	36	23	29	27	35	26	40	4
Microsoft band 2	5	13	2	6	10	7	6	7	11	7	202	20
Our approach	**14**	**44**	**7**	**26**	**34**	**24**	**28**	**24**	**40**	**41**	**19**	**2**

the knowledge about the orientation and consequently knowledge about the location.

5. Discussion

Activity monitoring in body sensor networks research ranges from the ergonomics of the sensor nodes, number of nodes, power consumption on hardware level and software level to data fusion. The senor nodes are most often dedicated sensors with embedded accelerometers and sometimes physiological sensors with predefined static placements. The requirement to wear single or multiple senor nodes in predefined location and orientation makes the application cumbersome and in some cases unattractive, since miss-placement contributes to errors in activity monitoring and therefore useless activity analysis. In recent years, we observed increase in popularity of activity monitoring research in which the senor nodes are replaced with commercial devices such as a smartphone (which most people already have) or more recent wearable technology such as wristbands which makes the service accessible to broader population. However, the the developed activity monitoring with these devices are either very simple and essentially count users steps and estimates the energy expenditure accordingly or are limited to monitoring specific sport activity (run trackers) on users request. The research with broader range of monitored activities using a smartphone analyze the data with intelligent methods, however they either require the smartphone to be worn at exact location and mostly use orientation independent features which limit the recognition and estimation accuracy.

In this research we explore how to use commercial devices such as accelerometer-embedded smartphone, sensor-rich wristband or both to achieve comparable results to using senor nodes in predefined static manner. Additionally, we perform the feature selection to decrease the computational complexity of the algorithm and explore how to optimally define activities to exploit the location and orientation of the smartphone, wristband or both.

We start with no knowledge about the orientation and location of the devices where we use orientation- and location-independent features. With this approach we can recognize five activities with accuracy of 79 %, kappa of 0.67 and f-score from 0.63 down to 0.5 in case of the wristband. Once the algorithm is introduced with the normalized orientation and the detected location we explored the gain of this knowledge to increase the number of recognized activities, increase in activity recognition accuracy and decrease in energy estimation error as presented in Table 12.

Accelerometer placed at different locations on the body can sense the movement of the same activity differently. For example, activities such as sitting and standing have the same orientation if the smartphone is places on the torso and lying and sitting activities have the same orientation when the smartphone is placed in the trousers pocket. We explored the trade-off between merged or split activities and accuracy to find the best set of activities to be recognized according to the recognized location of the smartphone. The procedure of merging and splitting activities is presented in Section 3.5.5 and the achieved results in Section 4.6. The average gain in number of activities over all locations in presented in row marked with # and gained activity recognition in row AR both in Table 12. We can observe that even with increased number of recognized activities we achieve increase in accuracy for up to 10 percentage points in case of smartphone alone. In case of the wristband we get slight increase in recognition accuracy but significant increase in f-score. If we fuse the data of both devices, we can recognize three more activities (up to eight) with higher accuracy, kappa and f-score.

The comparison of energy expenditure estimation model against the no orientation and no location model, shows high overall decrease in all errors (row EEE in Table 12). We start with average MAE of 1.03 MET and decrease it by 0.32 MET in case of the smartphone, by 0.45 MET in case of the wristband and by 0.49 MET when both devices are present. The decrease in MAPE is from 1 to 9 percentage points and RMSE from 0.1 up to 0.34. Note that the decrese in error is in addition to knowledge about the location and orientation also affected by more fine grained activity recognition.

We experimented only with the accelerometer sensor data from the smartphone, since the objective of the application was to develop a power efficient application. We believe that the use of additional sensor data such as GPS data, Wi-fi signal etc. would enable us to recognize even wider range of activities with comparable accuracy. However, using additional senors have an impact on battery life and the algorithm would require additional mechanism which would manage the senors frequency and activity which is not a subject of this research.

Table 12: Gain in number of recognized activities (#), gain in activity recognition (AR) accuracy and decrease in energy expenditure estimation error (EEE) when orientation is normalized and the location of the smartphone detected and used as a context for selection of the appropriate machine-learning model.

	#	AR		EEE	
Smartphone	+1	Accuracy	+10%	MAE	-.32
		Kappa	+.26	MAPE	-1
		F-score	+.18	RMSE	.1
Wristband	+1	Accuracy	+1%	MAE	-.45
		Kappa	+.08	MAPE	-6
		F-score	+.25	RMSE	-.31
Combination	+3	Accuracy	+8%	MAE	-.49
		Kappa	+.16	MAPE	-9
		F-score	+.18	RMSE	-.34

6. Conclusion

We present a real-time activity monitoring algorithm for activity recognition and estimation of energy expenditure with smartphone and wristband. The design of the algorithm enables the activity monitoring with individual device as well as with the combination of both. It first detects which devices are present on the body, then it expects ten seconds of walking to detect the gravity and normalise the orientation of the devices which enables the devices to be worn in any orientation. In case the smartphone is present it detects the location of the smartphone, which can be worn freely on the body (trousers pocket, jacket pocket and bag) and uses this information for selection of activity recognition classification model. The last step utilises the information about the location of the smartphone and the recognised activity for selection of the energy expenditure estimation regression model. The output of the activity monitoring algorithm is the performed activity and energy expenditure expresses in MET. We use indirect calorimeter Oxycon mobile to label the ground truth for estimation of energy expenditure.

We have evaluated each step in the algorithm and compared the results of the energy expenditure estimation model against two commercial devices. We first compared the resulted estimates in MET against Bodymedia armband, which is one of the most accurate EE estimation devices on the market and then we compared the estimates in kilo-calories against Bodymedia armband and Microsoft Band 2. The results have shown that our approach outperforms both devices and that even though there are many commercially available devices on the market for estimation of energy expenditure, they might not perform as accurately as expected and have room for improvement.

Accurate activity monitoring is important in domains where further decisions about the lifestyle or person-specific recommendations rely on the user's physical activity and its intensity. In the past, we have used simpler algorithms in health domains, for example, we have used activity monitoring with dedicated wearable sensors for monitoring physical activities of schoolchildren (eGibalec project), patients with chronic heart failure and for diabetes patients, for which the quantity of the activity is highly relevant to self-manage the disease [60][61]. This paper presents the upgrade of these approaches and the presented algorithm is implemented into the prototype of the AAL project Fit4Work, pilots of which will start soon. The results of the activity monitoring will be used to provide person-specific recommendations to older workers and help maintain their physical condition in good state. Additionally, this algorithm is going to be implemented as an initial algorithm for activity monitoring in the H2020 Heartman project.

The future work first of all includes the development of the orientation normalization method for the wristband to normalize it along the remaining axis (y-axis and z-axis) and to develop a method for detecting whether the person wears the wristband on right or left wrist. We will also stay informed about wearable devices that are coming to the market and adapt the algorithm to use them. We currently own activity monitoring datasets (labeled with calorimeter) of school-children, younger adults and older workers. and we plan to enrich them with disease-specific data and with more versatile population regarding age, gender and ethnicity.

Acknowledgement

The research presented herewith has been conducted within the Fit4Work project co-financed through the AAL Program (project no. AAL-2013-6-060).

References

[1] O. Cristian, S. Poser, R. Bernardini, R. Rinaldo, P.E. Di Prampero, Energy Cost and Metabolic Power in Elite Soccer: A New Match Analysis Approach, Med. Sci. Sport. Exerc. 42 (2010) 170178. doi:10.1249/MSS.0b013e3181ae5cfd.

[2] S. Chatterjee, A. Price, Healthy Living with Persuasive Technologies: Framework, Issues, and Challenges, J. Am. Med. Informatics Assoc. 16 (2009) 171178. doi:10.1197/jamia.M2859.

[3] F.J. Penedo, J.R. Dahn, Exercise and well-being: a review of mental and physical health benefits associated with physical activity., Curr. Opin. Psychiatry. 18 (2005) 189193. doi:10.1097/00001504-200503000-00013.

[4] Runkeeper. https://runkeeper.com/, 2017 (accessed 07.03.17).

[5] FitBit. https://www.fitbit.com/, 2017 (accessed 07.03.17).

[6] WHO, Global Health Observatory (GHO) data. http://www.who.int/gho/, 2017 (accessed 07.03.17).

[7] Fit4Work, AAL project. http://www.fit4work-aal.eu/, 2017 (accessed 07.03.17).

[8] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, Y. Amirat, Physical Human Activity Recognition Using Wearable Sensors., Sensors (Basel). 15 (2015) 3131438. doi:10.3390/s151229858.

[9] B. Cvetković, R. Milić, M. Luštrek, Estimating Energy Expenditure with Multiple Models Using Different Wearable Sensors, IEEE J. Biomed. Heal. Informatics. 20 (2016) 10811087. doi:10.1109/JBHI.2015.2432911.

[10] B.E. Ainsworth, W.L. Haskell, S.D. Herrmann, N. Meckes, D.R. Bassett, C. Tudor-Locke, J.L. Greer, J. Vezina, M.C. Whitt-Glover, A.S. Leon, 2011 Compendium of Physical Activities, Compend. Phys. Act. Track. Guid. Heal. Lifestyles Res. Center, Coll. Nurs. Heal. Innov. Arizona State Univ. (2011) 17. doi:10.1249/MSS.0b013e31821ece12.

[11] Apple Health. http://www.apple.com/ios/health/, 2017 (accessed 07.03.17).

[12] R. Gravina, P. Alinia, H. Ghasemzadeh, G. Fortino, Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges, Inf. Fusion. 35 (2017) 6880. doi:10.1016/j.inffus.2016.09.005.

[13] X. Lai, Q. Liu, X. Wei, W. Wang, G. Zhou, G. Han, A survey of body sensor networks., Sensors (Basel). 13 (2013) 54065447. doi:10.3390/s130505406.

[14] M.A. Hanson, H.C. Powell, A.T. Barth, K. Ringgenberg, B.H. Calhoun, J.H. Aylor, J. Lach, Body Area Sensor Networks: Challenges and Opportunities, Computer (Long. Beach. Calif). 42 (2009) 5865. doi:10.1109/MC.2009.5.

[15] Y.L. Zheng, X.R. Ding, C.C.Y. Poon, B.P.L. Lo, H. Zhang, X.L. Zhou, G.Z. Yang, N. Zhao, Y.T. Zhang, Unobtrusive sensing and wearable devices for health informatics, IEEE Trans. Biomed. Eng. 61 (2014) 15381554. doi:10.1109/TBME.2014.2309951.

[16] U. Mitra, B.A. Emken, Sangwon Lee, Ming Li, V. Rozgic, G. Thatte, H. Vathsangam, D. Zois, M. Annavaram, S. Narayanan, M. Levorato, D. Spruijt-Metz, G. Sukhatme, KNOWME: a case study in wireless body area sensor network design, IEEE Commun. Mag. 50 (2012) 116125. doi:10.1109/MCOM.2012.6194391.

[17] H. Ghasemzadeh, P. Panuccio, S. Trovato, G. Fortino, R. Jafari, Power-aware activity monitoring using distributed wearable sensors, IEEE Trans. Human-Machine Syst. 44 (2014) 537544. doi:10.1109/THMS.2014.2320277.

[18] G. Fortino, R. Giannantonio, R. Gravina, P. Kuryloski, R. Jafari, Enabling Effective Programming and Flexible Management of Efficient Body Sensor Network Applications, IEEE Trans. Human-Machine Syst. 43 (2013) 115133. doi:10.1109/TSMCC.2012.2215852.

[19] K. Kunze, P. Lukowicz, H. Junker, G. Tröster, Where am i: Recognizing on-body positions of wearable sensors, Locat. Context. Lect. Notes Comput. Sci. Springer. (2005) 264275. doi:10.1007/b136418.

[20] K. Kunze, P. Lukowicz, K. Partridge, B. Begole, Which way am I facing: Inferring horizontal device orientation from an accelerometer signal, in: Proc. - Int. Symp. Wearable Comput. ISWC, 2009: pp. 149150. doi:10.1109/ISWC.2009.33.

[21] D. Mizell, Using gravity to estimate accelerometer orientation, in: Seventh IEEE Int. Symp. Wearable Comput. 2003. Proceedings., 2003: pp. 252253. doi:10.1109/ISWC.2003.1241424.

[22] M.D. Tundo, E. Lemaire, N. Baddour, Correcting Smartphone orientation for accelerometer-based analysis, in: MeMeA 2013 - IEEE Int. Symp. Med. Meas. Appl. Proc., 2013: pp. 5862. doi:10.1109/MeMeA.2013.6549706.

[23] Y.E. Ustev, O. Durmaz Incel, C. Ersoy, User, device and orientation independent human activity recognition on mobile phones, in: ACM Conf. Pervasive Ubiquitous Comput. Adjun. Publ., 2013: pp. 14271436. doi:10.1145/2494091.2496039.

[24] M. Shoaib, H. Scholten, P.J.M. Havinga, Towards Physical Activity Recognition Using Smartphone Sensors, 2013 IEEE 10th Int. Conf. Ubiquitous Intell. Comput. 2013 IEEE 10th Int. Conf. Auton. Trust. Comput. (2013) 8087. doi:10.1109/UIC-ATC.2013.43.

[25] C.A. Ronao, S. Cho, Human activity recognition with smartphone sensors using deep learning neural networks, Expert Syst. Appl. 59 (2016) 235244. doi:10.1016/j.eswa.2016.04.032.

[26] M. Shoaib, S. Bosch, O. Incel, H. Scholten, P. Havinga, A Survey of Online Activity Recognition Using Mobile Phones, Sensors. 15 (2015) 20592085. doi:10.3390/s150102059.

[27] H. Lu, J. Yang, Z. Liu, N.D. Lane, T. Choudhury, A.T. Campbell, The Jigsaw Continuous Sensing Engine for Mobile Phone Applications, in: Proc. 8th ACM Conf. Embed. Networked Sens. Syst., ACM, New York, NY, USA, 2010: pp. 7184. doi:10.1145/1869983.1869992.

[28] S. Thiemjarus, A. Henpraserttae, S. Marukatat, A study on instance-based learning with reduced training prototypes for device-context-independent activity recognition on a mobile phone, in: 2013 IEEE Int. Conf. Body Sens. Networks, 2013: pp. 16. doi:10.1109/BSN.2013.6575462.

[29] A. Anjum, M.U. Ilyas, Activity recognition using smartphone sensors, in: 2013 IEEE 10th Consum. Commun. Netw. Conf. CCNC 2013, 2013: pp. 914919. doi:10.1109/CCNC.2013.6488584.

[30] H. Guo, L. Chen, G. Chen, M. Lv, Smartphone-based activity recognition independent of device orientation and placement, Int. J. Commun. Syst. 29 (2016) 24032415. doi:10.1002/dac.3010.

[31] H. Martín, A.M. Bernardos, J. Iglesias, J.R. Casar, Activity logging using lightweight classification techniques in mobile devices, Pers. Ubiquitous Comput. 17 (2013) 675695. doi:10.1007/s00779-012-0515-4.

[32] B. Cvetković, V. Janko, M. Luštrek, Demo abstract: Activity recognition and human energy expenditure estimation with a smartphone, Pervasive Comput. Commun. Work. (PerCom Work. 2015 IEEE Int. Conf. (2015) 193195. doi:10.1109/PERCOMW.2015.7134019.

[33] D. Biswas, A. Cranny, N. Gupta, K. Maharatna, J. Achner, J. Klemke, M. Jöbges, S. Ortmann, Recognizing upper limb movements with wrist worn inertial sensors using k-means clustering classification, Hum. Mov. Sci. 40 (2015) 5976. doi:10.1016/j.humov.2014.11.013.

[34] P. Siirtola, P. Laurinen, E. Haapalainen, J. Rning, H. Kinnunen, Clustering-based activity classification with a wrist-worn accelerometer using basic features, in: 2009 IEEE Symp. Comput. Intell. Data Mining, CIDM 2009 - Proc., 2009: pp. 95100. doi:10.1109/CIDM.2009.4938635.

[35] S. Chernbumroong, A.S. Atkins, Activity classification using a single wrist-worn accelerometer, 2011 5th Int. Conf. Software, Knowl. Information, Ind. Manag. Appl. Proc. (2011) 16. doi:10.1109/SKIMA.2011.6089975.

[36] G.M. Weiss, J.L. Timko, C.M. Gallagher, K. Yoneda, A.J. Schreiber, Smartwatch-based activity recognition: A machine learning approach, in: Proc. IEEE-EMBS Int. Conf. Biomed. Heal. Informatics, 2016: pp. 426429. doi:10.1109/BHI.2016.7455925.

[37] F. Ramos, U. Federal, C. Grande, R. Rolim, U. Federal, C. Grande, H. Oliveira, A.U. Federal, C. Grande, A. Perkusich, U. Federal, C. Grande, Combining Smartphone and Smartwatch Sensor Data in Activity Recognition Approaches: an Experimental Evaluation, Int. J. Eng. Technol. 8 (2016). doi:10.18293/SEKE2016-040.

[38] J.A. Levine, Measurement of energy expenditure., Public Nealth Nutr. 8 (2005) 11231132. doi:10.1079/PHN2005800.

[39] D.R. Bassett Jr., B.E. Ainsworth, A.M. Swartz, S.J. Strath, W.L. OBrien, G.A. King, Validity of Four Motion Sensors in Measuring Moderate Intensity Physical Activity, Med. Sci. Sport. Exerc. 32 (2000) S471-80.

[40] S.E. Crouter, E. Kuffel, J.D. Haas, E.A. Frongillo, D.R. Bassett, Refined two-regression model for the actigraph accelerometer, Med. Sci. Sports Exerc. 42 (2010) 10291037. doi:10.1249/MSS.0b013e3181c37458.

[41] A.G. Bonomi, G. Plasqui, A.H.C. Goris, K.R. Westerterp, Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer., J. Appl. Physiol. 107 (2009) 655661. doi:10.1152/japplphysiol.00150.2009.

[42] F. Albinali, S.S. Intille, W. Haskell, M. Rosenberger, Using wearable activity type detection to improve physical activity energy expenditure estimation, Proc. 12th ACM Int. Conf. Ubiquitous Comput. Ubicomp 10. (2010) 311. doi:10.1145/1864349.1864396.

[43] E.M. Tapia, Using machine learning for real-time activity recognition and estimation of energy expenditure, Mit. PhD Thesis (2008).

[44] A. Pande, J. Zhu, A.K. Das, Y. Zeng, P. Mohapatra, J.J. Han, Using Smart-

phone Sensors for Improving Energy Expenditure Estimation, IEEE J. Transl. Eng. Heal. Med. 3 (2015). doi:10.1109/JTEHM.2015.2480082.

[45] M. Altini, J. Penders, R. Vullers, O. Amft, Personalizing energy expenditure estimation using physiological signals normalization during activities of daily living, Physiol. Meas. 35 (2014) 17971811. doi:10.1088/0967-3334/35/9/1797.

[46] M. Altini, J. Penders, R. Vullers, O. Amft, Estimating energy expenditure using body-worn accelerometers: A comparison of methods, sensors number and positioning, IEEE J. Biomed. Heal. Informatics. 19 (2015) 219226. doi:10.1109/JBHI.2014.2313039.

[47] N. Vyas, J. Farringdon, D. Andre, J.I. Stivoric, Machine Learning and Sensor Fusion for Estimating Continuous Energy Expenditure, AI Mag. 33 (2012) 55. doi:10.1609/aimag.v33i2.2408.

[48] B. Cvetković, R. Milić, M. Luštrek, Estimating Energy Expenditure with Multiple Models Using Different Wearable Sensors, IEEE J. Biomed. Heal. Informatics. 20 (2016) 10811087. doi:10.1109/JBHI.2015.2432911.

[49] M. Duclos, G. Fleury, P. Lacomme, R. Phan, L. Ren, S. Rousset, An acceleration vector variance based method for energy expenditure estimation in real-life environment with a smartphone/smartwatch integration, Expert Syst. Appl. 63 (2016) 435449. doi:10.1016/j.eswa.2016.07.021.

[50] Microsoft Band 2, https://www.microsoft.com/microsoft-band/en-us, 2017 (accessed 07.03.17).

[51] Bodymedia FIT Advantage, http://www.bestfitnesstrackerreviews.com/-bodymedia-fit-review.html,2017 (accessed 07.03.17).

[52] J.M. Lee, Y. Kim, G.J. Welk, Validity of consumer-based physical activity monitors, Med. Sci. Sports Exerc. 46 (2014) 18401848. doi:10.1249/MSS.0000000000000287.

[53] Empatica E4, https://www.empatica.com/, 2017 (accessed 07.03.17).

[54] Oxycon, http://www.carefusion.com/, 2017 (accessed 07.03.17).

[55] M.A. Akkermans, M.J.H. Sillen, E.F.M. Wouters, M.A. Spruit, Validation of the Oxycon mobile metabolic system in healthy subjects, J. Sport. Sci. Med. 11 (2012) 182183.

[56] H. Rosdahl, L. Gullstrand, J. Salier-Eriksson, P. Johansson, P. Schantz, Evaluation of the Oxycon Mobile metabolic system against the Douglas bag method, Eur. J. Appl. Physiol. 109 (2010) 159171. doi:10.1007/s00421-009-1326-9.

[57] Y. Kim, G.J. Welk, Criterion validity of competing accelerometry-based activity monitoring devices, Med. Sci. Sports Exerc. 47 (2015) 24562463. doi:10.1249/MSS.0000000000000691.

[58] G. Zorrilla-Revilla, A. Mateos, O. Prado-Nvoa, M. Vidal-Cordasco, J. Rodrguez, Carrying loads: Validating a portable tri-axial accelerometer during frequent and brief physical activity, J. Sci. Med. Sport. (2017). doi:10.1016/j.jsams.2017.01.236.

[59] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I. Witten, The WEKA data mining software: An update, SIGKDD Explor. 11 (2009) 1018. doi:10.1145/1656274.1656278.

[60] N. Oldridge, D.R. Thompson, A.-D. Zwisler, K. Rees, N. Martin, R.S. Taylor, Exercise-Based Cardiac Rehabilitation for Coronary Heart Disease, J. Am. Coll. Cardiol. 67 (2016) 1211. doi:10.1016/j.jacc.2015.10.044.

[61] S.R. Colberg, R.J. Sigal, J.E. Yardley, M.C. Riddell, D.W. Dunstan, P.C. Dempsey, E.S. Horton, K. Castorino, D.F. Tate, Physical activity/exercise and diabetes: A position statement of the American Diabetes Association, Diabetes Care. 39 (2016) 20652079. doi:10.2337/dc16-1728.