

RESEARCH ARTICLE

Open Access

# Workplace activity classification from shoe-based movement sensors



Jonatan Fridolfsson<sup>1\*</sup> , Daniel Arvidsson<sup>1</sup>, Frithjof Doerks<sup>2</sup>, Theresa J. Kreidler<sup>3</sup> and Stefan Grau<sup>1</sup>

## Abstract

**Background:** High occupational physical activity is associated with lower health. Shoe-based movement sensors can provide an objective measurement of occupational physical activity in a lab setting but the performance of such methods in a free-living environment have not been investigated. The aim of the current study was to investigate the feasibility and accuracy of shoe sensor-based activity classification in an industrial work setting.

**Results:** An initial calibration part was performed with 35 subjects who performed different workplace activities in a structured lab setting while the movement was measured by a shoe-sensor. Three different machine-learning models (random forest (RF), support vector machine and k-nearest neighbour) were trained to classify activities using the collected lab data. In a second validation part, 29 industry workers were followed at work while an observer noted their activities and the movement was captured with a shoe-based movement sensor. The performance of the trained classification models were validated using the free-living workplace data. The RF classifier consistently outperformed the other models with a substantial difference in in the free-living validation. The accuracy of the initial RF classifier was 83% in the lab setting and 43% in the free-living validation. After combining activities that was difficult to discriminate the accuracy increased to 96 and 71% in the lab and free-living setting respectively. In the free-living part, 99% of the collected samples either consisted of stationary activities or walking.

**Conclusions:** Walking and stationary activities can be classified with high accuracy from a shoe-based movement sensor in a free-living occupational setting. The distribution of activities at the workplace should be considered when validating activity classification models in a free-living setting.

**Keywords:** Accelerometry, Workload, Physical activity, Occupational health

## Background

Industry work is associated with a high physical workload. Although leisure-time physical activity is associated with health, the opposite is true with physically active works [1]. High occupational physical activity (OPA) is associated with more long term sickness absence as well as all-cause mortality [1, 2]. OPA is also a factor contributing to fatigue, which in a work setting could lead to

serious injury or death [3]. Consequently, assessment of OPA would provide indication of detrimental volume, allowing appropriate adjustments of work tasks before fatigue occurs.

Objective monitoring of physical activity is widespread in research; occupational, epidemiological and clinical [4, 5]. However, the most common sensor positions are the hip or thigh, which are unpractical for monitoring OPA over time among industry workers. Since industry work often require safety shoes, because of e.g. toe protection and slip resistance, a sensor built into these shoes would be a practical and relatively easily implementable solution in an occupational setting. Previous

\* Correspondence: [jonatan.fridolfsson@gu.se](mailto:jonatan.fridolfsson@gu.se)

<sup>1</sup>Center for Health and Performance, Department of Food and Nutrition, and Sport Science, University of Gothenburg, Box 300, 405 30 Gothenburg, Sweden

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

work has shown that an accelerometer placed on a shoe has similar accuracy in predicting activities as other placements [6], but the performance of shoe based sensors have not been evaluated in a free-living setting previously [7].

Sensor based physical activity classification is usually based on data recorded by one or more accelerometers [4]. The raw acceleration signal is processed to display specific features. Features comprise for instance mean acceleration, max frequency, correlation between axes that are calculated continuously over a moving window. These features are then used for activity classification, either using simple empirically derived decision trees [8], or using statistical methods, called machine-learning, that is more common today [9]. Machine-learning algorithms can be developed to be extremely accurate in a structured setting reaching 98–100% accuracy [9, 10]. However, when applying the algorithms to other structured datasets (lab-data) the accuracy drops significantly [11], and with free living datasets, the accuracy is reduced even further [12].

The aim of the current study was to investigate the feasibility and accuracy of shoe sensor-based activity classification in an industrial work setting. Therefore, in a first step, shoe acceleration data was captured in a structured setting (standardized lab-activities) to develop an activity classification machine-learning algorithm that will be able to reliably distinguish between different activities. In a second step, the algorithm was validated in a free-living setting (at the workplace).

The following research questions were examined in this study:

1. Is it possible to reliably classify work specific activities with acceleration signals captured from a shoe-based sensor?
2. Does this activity classification work in a free-living (workplace) setting?

## Results

The study consisted of two parts, a calibration part in a controlled setting and a validation part in a free-living workplace setting. Thirty-five subjects participated in the calibration part. Information about the test subjects can be seen in Table 1.

Subjects wore an accelerometer attached to their right shoe while performing seven standardized activities. The

accelerometry data was processed to signal features used to train three different machine learning classification models to classify the standardized activities. The classification models used were random forest (RF), support vector machine (SVM) and k-nearest neighbour (KNN). The lab accuracy of the initial classification models are presented in Table 2 (Lab calibration 7 activities). The RF model accuracy was slightly higher than the other two models, but all within one percentage point. The activity specific performance of the RF model is presented as a confusion chart in Fig. 1. The distribution of accuracy between the different activities were similar between the three models. The RF accuracy of the different activities (blue marked) range from 45.8% (sitting) to 100% (kneeling). The model had difficulties differentiating standing and sitting (64.9 and 45.8% respectively). The classification accuracy of weight carrying was also less accurate (80.8%). 54.2% of sitting was misclassified as standing, 33.2% of standing was misclassified as sitting and 14.9% of weight carrying was misclassified as walking.

Twenty-nine subjects took part in the validation part of the study. The age and body mass index (BMI) of the subjects in the validation part was significantly different from the subjects in the in-lab calibration part (Table 1). Fifteen subjects were working at a logistics centre warehouse and 14 subjects were working in industrial production. An average of 48 min of free-living data was captured for each participant using an accelerometer attached to their right shoe. The performance of the three classification models in the free-living setting is presented in Table 2. Similar to the lab results, the RF model had the best accuracy, but in the free-living validation the differences between methods are much larger with 7.3 and 13.4 percentage points for SVM and KNN respectively. The activity specific performance of the RF model is presented as a confusion chart in Fig. 2. Numbers inside the chart are normalized to the total number of samples and shows the distribution of activities in addition to the accuracy. The activity specific sensitivity (the proportion of observed samples classified correctly) ranged from 4.0% (kneeling) to 60.0% (stair descending) and the activity specific specificity (the proportion of classified samples in agreement with the observation) ranged from 0.7% (weight carrying) to 87.7% (walking). Similar to the lab-performance, sitting and standing was difficult to differentiate. In contrast to the lab-results, most of the observed stair ascending were classified as stair descending. The specificity of walking was high (87.7%) but the sensitivity was low (22.4%). This was because many of the observed walking samples were classified as either stair descending or weight carrying but samples that were classified as walking mainly consisted of observed walking.

**Table 1** Subjects details

	N (% female)	Age (SD)***	BMI (SD)***
Lab calibration	35 (49%)	25.4 (6.0)	23.1 (2.3)
Free-living validation	29 (10%)	38.7 (11.5)	26.8 (3.5)

BMI: Body Mass Index, SD: Standard Deviation, \*\*\* indicates significant group differences  $p < 0.001$ .

**Table 2** Accuracy of models

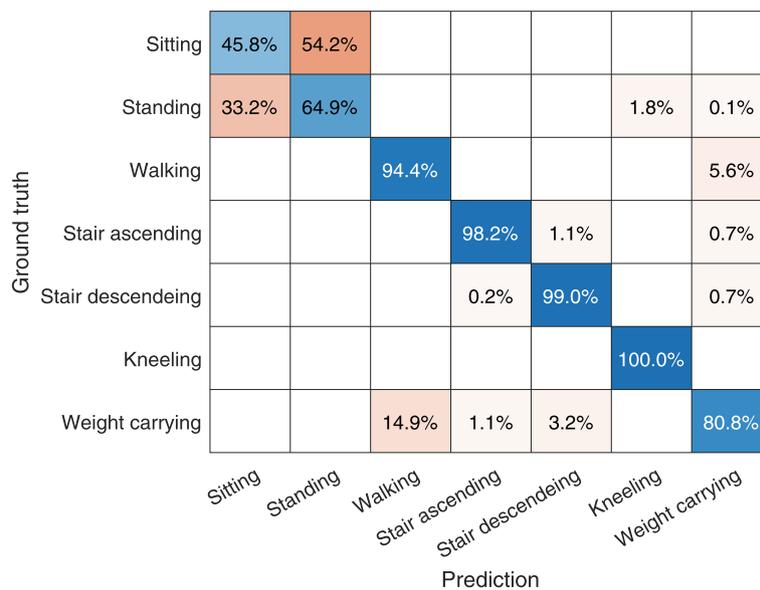
	Random forest	Support vector machine	K-nearest neighbour
Lab calibration 7 activities	83.3%	82.3%	82.5%
Lab calibration 5 activities	96.3%	95.0%	95.7%
Free-living validation 7 activities	43.0%	35.7%	29.6%
Free-living validation 5 activities	71.2%	67.1%	63.4%

Three additional calibration models based on five activities were developed by using the same calibration data but combining sitting and standing into stationary and combining stair ascending and descending into stair walking. Similar to the initial seven activities models, the RF model outperformed the SVM and KNN models in lab and free-living validation (Table 2). The activity specific performance of the second RF model is presented in Fig. 3 (lab) and Fig. 4 (free-living). The classification accuracy of the in lab weight carrying was still less accurate (80.4%). In free-living, the predictions of the second model was highly accurate (88–91%) with regard to stationary and walking activities. These activities made up 99% of the collected samples. However, the performance on the rest of the activities were very low. To investigate the effect of activity type in the validation data, a sub analysis was performed on the industrial production and logistics warehouse data separately. With the industrial production data the accuracy of the seven and five activities models was 40.4 and 66.9% respectively. The accuracy was higher with the logistics warehouse data at 47.2 and 77.8% with the seven and five activities models respectively.

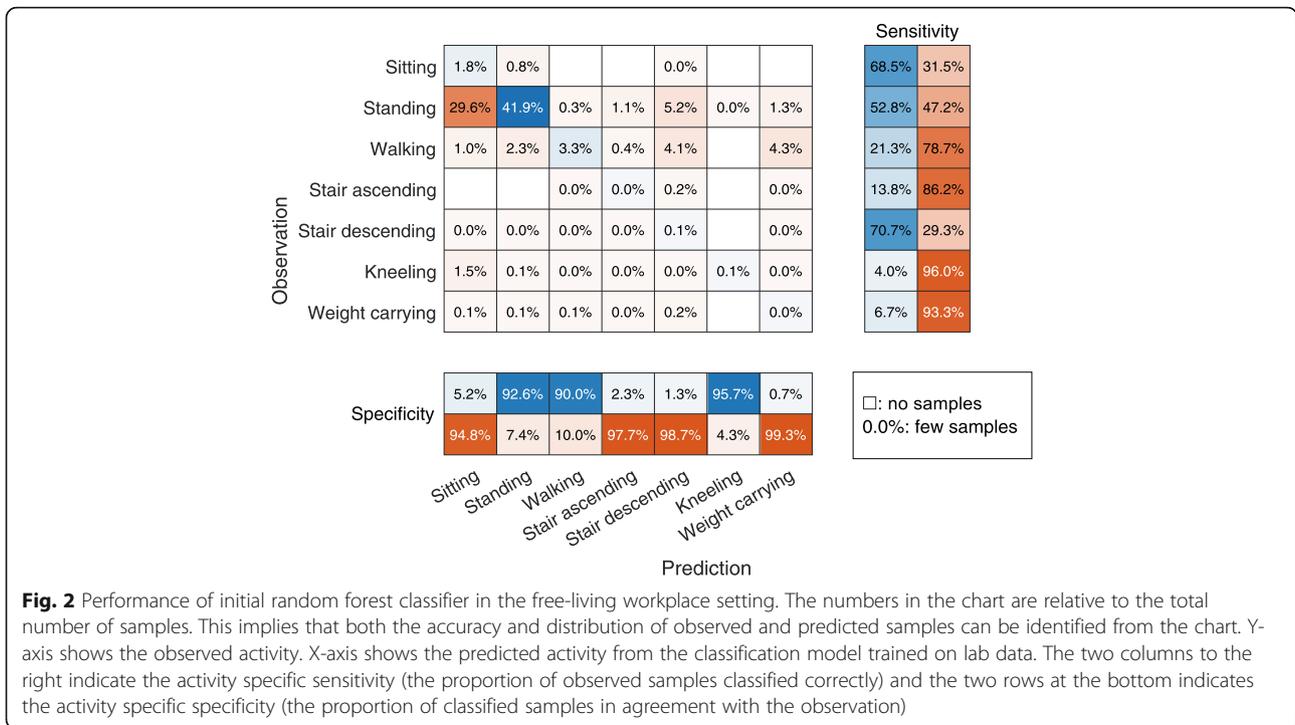
**Discussion**

The RF classification model consistently outperformed the KNN and SVM models. The difference between models were negligible in the calibration setting but increased drastically in free living validation. The lab-setting RF model classification accuracy of activities at the workplace was consistently high except for standing and sitting (Fig. 1). In the free-living setting on the other hand, the classification accuracy was initially low across all activities (Fig. 2). After combining standing and sitting to stationary activity as well as combining stair ascending and descending to stair walking, the level of accuracy for both activities increased in the lab and free-living environment (Fig. 3-4). The good overall performance of the second RF model in the free living (71%) can be explained as 99% of the samples captured consisted of either walking or stationary activity.

Combining sitting and standing and not being able to distinguish between the two might be considered a major shortcoming of the classification model. On the other hand, although there is a small increase in energy consumption from standing compared to sitting [13, 14], standing is still considered a sedentary activity [15] and



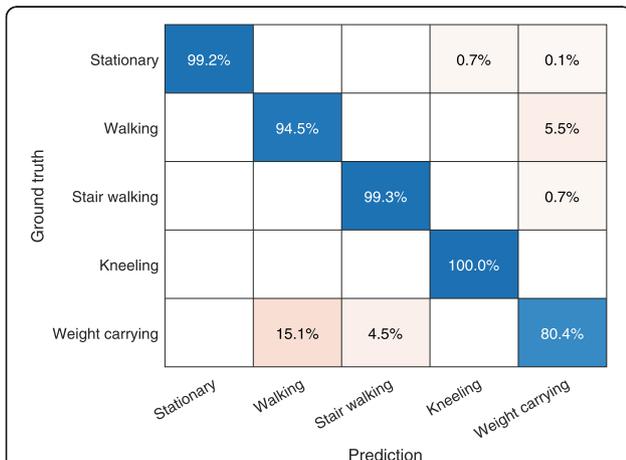
**Fig. 1** Performance of initial random forest classifier in the lab calibration setting. The numbers in the chart are row normalized, showing the distribution of the classification of the samples from each activity according to the test protocol. Y-axis shows the activity performed according to the test protocol. X-axis shows the predicted activity from the trained classification model



**Fig. 2** Performance of initial random forest classifier in the free-living workplace setting. The numbers in the chart are relative to the total number of samples. This implies that both the accuracy and distribution of observed and predicted samples can be identified from the chart. Y-axis shows the observed activity. X-axis shows the predicted activity from the classification model trained on lab data. The two columns to the right indicate the activity specific sensitivity (the proportion of observed samples classified correctly) and the two rows at the bottom indicates the activity specific specificity (the proportion of classified samples in agreement with the observation)

there are no cardiovascular health benefits with standing compared to sitting [16]. During both sitting and standing the feet are usually parallel to the ground and no movement occurs. Since the inclination and movement of the sensor is used for classification, this explains the difficulty discriminating the two stationary activities.

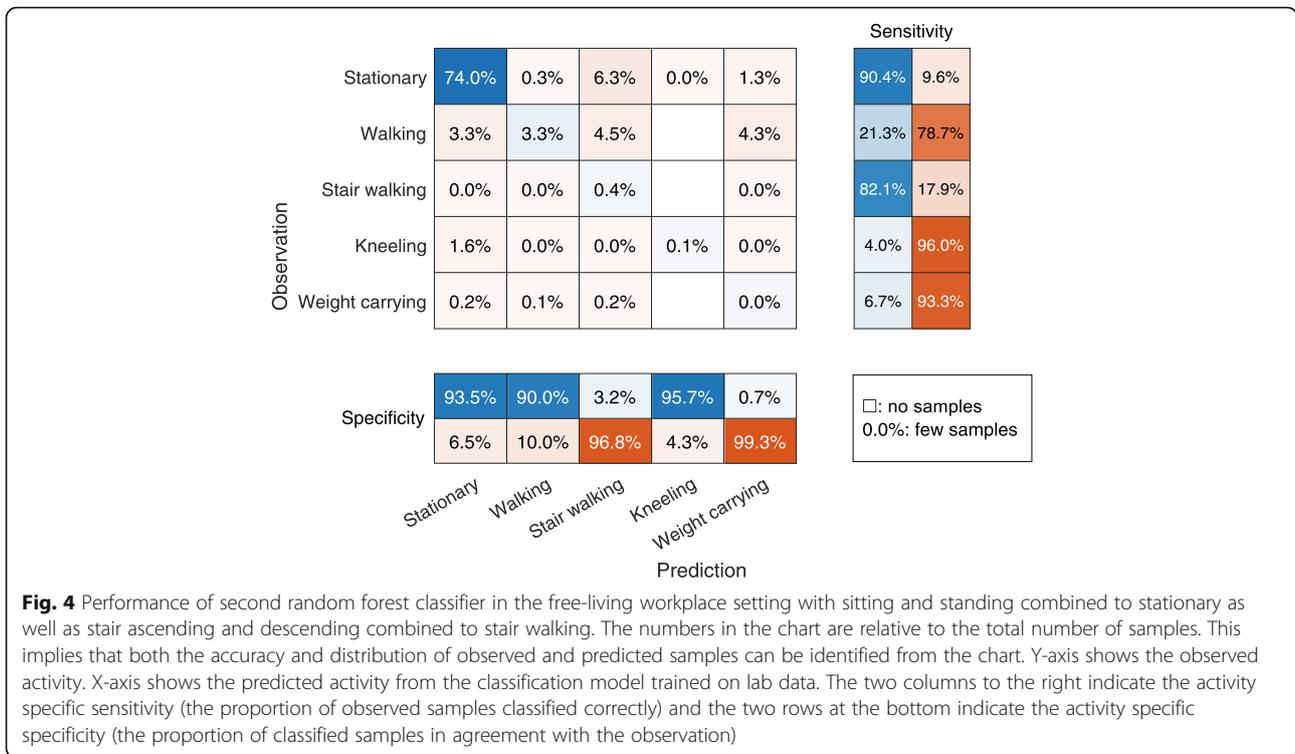
Stair ascending and descending were also combined to a single activity in the second model. However, these activities are associated with significantly different energy expenditure as opposed to sitting and standing [14]. The estimated workload from stair walking might therefore be underestimated. Although, in most cases, stair descending and ascending could be assumed to be equally distributed.



**Fig. 3** Performance of second random forest classifier in the lab calibration setting with sitting and standing combined to stationary as well as stair ascending and descending combined to stair walking. The numbers in the chart are row normalized, showing the distribution of the classification of the samples from each activity according to the test protocol. Y-axis shows the activity performed according to the test protocol. X-axis shows the predicted activity from the trained classification model

The initial classification models' performance were poor for all free-living activities (Fig. 2). The reason only standing/sitting and stair ascending/descending was combined was that these activities was clearly mixed up with each other but not with any other activity. With the other activities, the misclassification was more spread out. The differentiation between walking and stationary activities could probably just as well have been performed using an acceleration intensity metric alone [4]. However, activity type might be a more applicable output for the workplace than the abstract intensity measures.

Other weaknesses of the study are the significant sex, age and BMI differences between subjects in the two study parts. The participants in the validation group consisted of more men, were older and had higher BMI than the participants in the calibration group, which might have affected the classification performance. It should also be considered that the validation was performed indoors only whereas parts of the calibration was performed outdoors. Although the outdoor walking in the calibration part was also done at slow pace, most of



**Fig. 4** Performance of second random forest classifier in the free-living workplace setting with sitting and standing combined to stationary as well as stair ascending and descending combined to stair walking. The numbers in the chart are relative to the total number of samples. This implies that both the accuracy and distribution of observed and predicted samples can be identified from the chart. Y-axis shows the observed activity. X-axis shows the predicted activity from the classification model trained on lab data. The two columns to the right indicate the activity specific sensitivity (the proportion of observed samples classified correctly) and the two rows at the bottom indicate the activity specific specificity (the proportion of classified samples in agreement with the observation)

the indoor walking in the validation part might have been done at even slower pace. Calibration of stair walking and weight carrying was performed indoors and at slower speeds than the normal walking speed, which could explain the misclassification of free-living walking into these activities (Fig. 4). The workers in the logistics warehouse were covering larger areas while walking, whereas the workers in productions were mainly walking a few steps between machines. Covering larger areas could make the difference between walking and stationary more prominent and explain the higher accuracy in with the logistics warehouse data.

Although direct observation is considered the criterion method for activity classification in a free-living setting [4], this method is not perfect. It has been shown that direct observation has an accuracy of 87% where the activity classification of senior researchers was considered the reference [17]. In a free-living setting, the activities might not be equivalent to the standardized lab-activities. Most of the validation data consisted of standing work that was stationary most of the time with walking a few steps in between (Fig. 2), which could be difficult to define with the current classification scheme.

Many studies on accelerometer based machine-learning classification models have been published previously, most of them using similar techniques as the current study [9, 18]. We have only found one other study that investigated the performance of a lab calibrated machine-learning method in a free living setting

and there the accuracy was 49–55% [12]. The accuracy of the current study is substantially higher at 71% (Fig. 4) although it is very low with some activities. However, the activities classified are different in the two studies. Lab calibration of activity classification may be prone to overfitting, even when validating the model using leave one subject out cross validation [18]. Nevertheless, RF classification models are in general relatively robust to over fitting, but on the other hand may perform poorly on data that deviate much from the training data [19]. The main limitation in generalization of lab developed activity classification models is thought to be the diverse activity types, different characteristics within each activity type and individual variation [18]. The limited number of samples with other activities than stationary and walking at the two workplaces in the validation part limits further analysis of the accuracy of the classification of these activities. The classification model might perform better in a setting where other activities are more common. The difference in accuracy between workers in the logistics warehouse and industrial production also supports this assumption. A more diverse free-living dataset with a more even activity distribution could also be used to improve the classification algorithm further by analysing the temporal structure of activities [20].

Certain work-related physical activity patterns are suggested to have a negative health effect. For example, prolonged physical activity elevates 24-h heart rate and

static postures and lifting is suggested to elevate 24-h blood pressure [21]. Such patterns might be detected by the activity classification system suggested in this paper. Continuous monitoring of workload among industrial workers could be used in many ways for preventive measures and improving health. The monitoring gives basic data on the distribution of physical workload across different tasks at the workplace. This can be utilized when partitioning tasks between workers, both with regard to sharing heavy work between more employees and to lower the physical demand on specific individuals. Continuous monitoring also provides the possibility to follow workload over time, which might enable identification of employees getting fatigued at an early stage. This could potentially result in an overall less long-term sickness absence and better health status among workers [1]. However, constant monitoring of activities during work do raise concerns regarding privacy of the workers [5].

## Conclusion

The results of the study shows that walking and stationary activities can be classified with high accuracy from a shoe based accelerometer. In order to accurately classify other activities (kneeling, stair ascending and descending and weight carrying), workplaces with a higher number of those activities should be considered. However, the present study also addresses issues with activity distribution when classifying activities in a free-living setting. It highlights difficulties with free-living validation of activity classification with regard to generalisation of lab-calibrated classification, observation of activities and activity dispersion in free-living. Despite this, the results suggest that activity classification using shoe-based sensors could give accurate and comprehensive feedback on walking and stationary activities in an occupational setting.

## Methods

Subjects for the calibration part were recruited through e-mail announcements to students and staff at the Department of Food and Nutrition and Sport Science and personal communication. Written informed consent was retrieved from the subjects and the study was approved by the regional ethics committee in Gothenburg (no. 765–18).

For the data collection, subjects wore safety shoes (Ergo-Active Grant, Elten GmbH, Uedem, Germany) in their respective size (EU 36–48) and width (narrow, medium, wide). Accelerometers (AX3, Axivity Ltd., Newcastle upon Tyne, UK) were then firmly attached to the heel-cap of each shoe orthogonal to the outsole using non-elastic adhesive tape. The accelerometers were set to record triaxial acceleration at a sampling frequency of 100 Hz and a range of  $\pm 16g$  where  $1g$  is

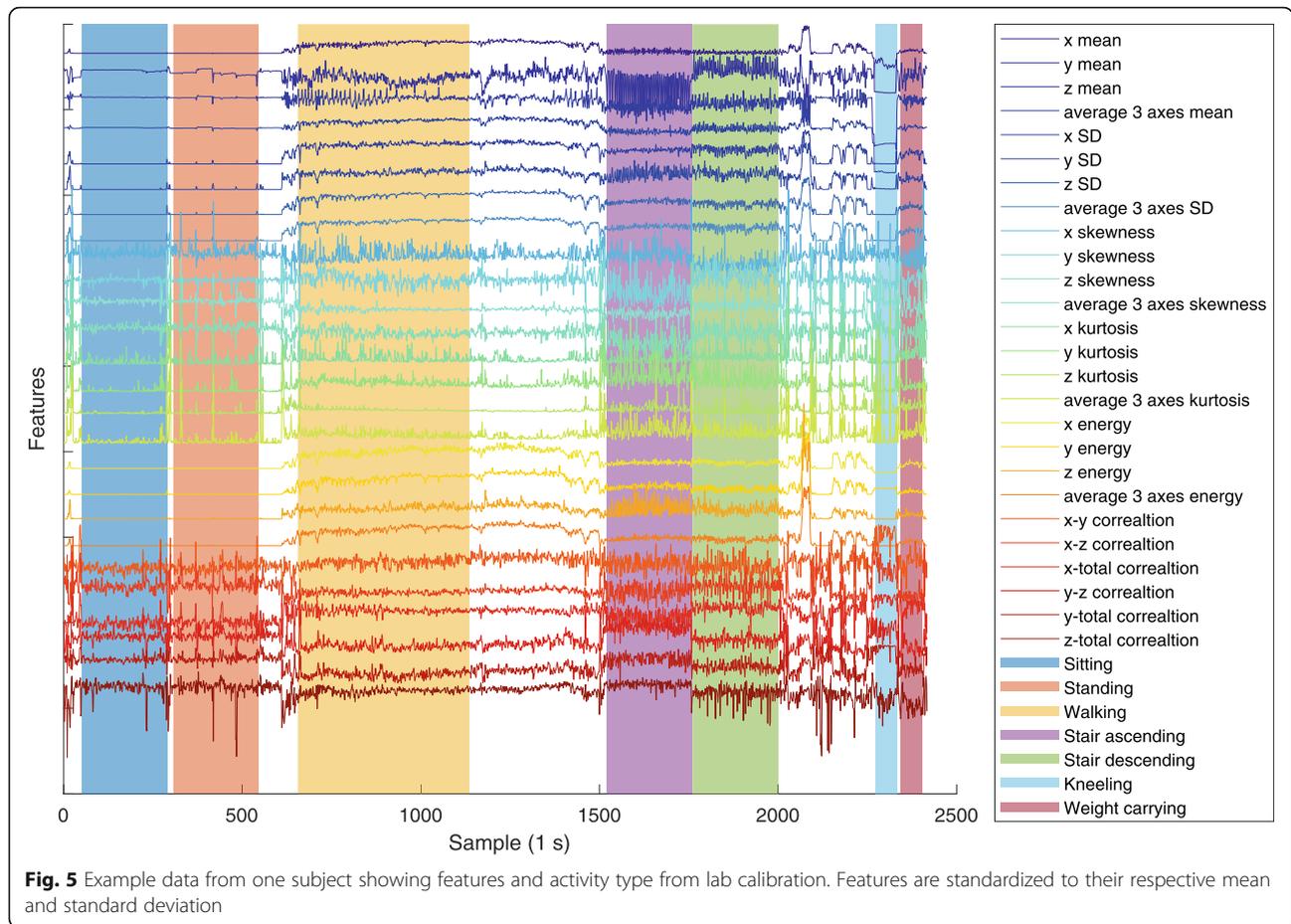
equivalent to the gravitational acceleration. These specifications were sufficient to capture all acceleration related to human movement [22].

The test protocol of the calibration part consisted of eight activities that were performed by the subjects continuously for 1–4 min.

- Sitting on a chair, while solving Sudoku on a table
- Standing, while solving Sudoku on a high table
- Walking slow, self-paced, outdoors
- Walking brisk, self-paced, outdoors
- Stair ascending
- Stair descending
- Weight carrying while walking, 15 kg
- Kneeling

All analyses were performed in MATLAB R2018b (MathWorks, Natick, MA, USA). Acceleration was captured between the 6th and 55th second of the last minute of each activity. Using a 2 s window with 50% overlap [23], 26 signal features [6] were calculated for the extracted acceleration for each shoe. The features from each window were labelled according to the test protocol. Features and activity type from one subject are presented as an example in Fig. 5. Samples with outliers were removed from the data from each activity using a criterion of more than three scaled median absolute deviations (MATLAB `rmoutliers`-function). The labelled data was used as training data to generate three different machine-learning classification algorithms, random forest (RF), support vector machine (SVM) and k-nearest neighbour (KNN). These are among the most commonly used techniques for classifying activity based on accelerometer data [9, 18]. The models were implemented as an ensemble of bagged decision trees (RF), third degree support vector machine (SVM) and the ten nearest neighbours weighted by the inverse distance squared (KNN) using the MATLAB Classification Learner. Validation of the classification algorithms was performed by leave-one-out validation for each subject. This validation technique implies that there will never be data from the same subject in both the training- and validation data set simultaneously which leads to a more realistic accuracy measure [24].

Recruitment of subjects for the free-living validation part took place at an industrial workplace through information to the workers. Written informed consent was retrieved from the workers who agreed to participate. Subjects were fitted with the same kind of shoes as in the calibration part with the same accelerometers attached to the heel-cap. Then the subjects were told to work as usual while followed by an observer. The observer noted the start, end and type of each activity continuously for about 60 min. Since separating standing



and walking could be difficult during standing work, walking was considered continuous movement more than three meters. This way a few sidesteps during standing work would still be considered a standing activity.

The collected accelerometer data was processed as in the calibration part of the study. The different features were input to the developed classification algorithms to get a predicted activity for each window. The observed activities were then compared to the activities predicted by the classification models using confusion charts. The lab-setting performance was analysed using row normalized confusion charts. With the free-living confusion charts, the cells were normalized to the total number of samples since the distribution of samples between the observed activities were not even. The activity specific sensitivity and specificity of the classification models were also added to the confusion charts of the free-living results. To show the strengths and limitations of the classification models, activities that could possibly be difficult to discriminate between were combined in a second analysis (sitting and standing to stationary and stair ascending and descending to stair walking). A sub

analysis of the free-living performance on the logistics warehouse and industrial production workers was performed to investigate the impact of different activities in the validation data. BMI and age differences between the calibration and validation groups were evaluated using a two-sample t-test with a significance level of  $p < 0.05$ .

**Abbreviations**

OPA: Occupational physical activity; RF: Random forest; SVM: Support vector machine; KNN: k-nearest neighbour; BMI: Body mass index

**Acknowledgements**

Not applicable.

**Authors' contributions**

JF, SG and DA designed the study procedure. FD and TJK collected the data. JF performed the processing and analysis of the data and wrote the original draft. SG and DA edited the draft. The authors read and approved the final manuscript.

**Funding**

The study was funded by Elten GmbH (Uedem, Germany). Elten had no role in the study design, data collection, analysis, interpretation or writing of the manuscript. Open access funding provided by University of Gothenburg.

**Availability of data and materials**

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

### Ethics approval and consent to participate

Ethical approval was obtained from the regional ethics committee in Gothenburg (no. 765–18). Written informed consent was retrieved from all participants.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Center for Health and Performance, Department of Food and Nutrition, and Sport Science, University of Gothenburg, Box 300, 405 30 Gothenburg, Sweden. <sup>2</sup>Hochschule Koblenz, University of Applied Sciences RheinAhr Campus, Remagen, Germany. <sup>3</sup>Institute for Applied Movement Science, Chemnitz University of Technology, Chemnitz, Germany.

Received: 12 December 2019 Accepted: 17 June 2020

Published online: 24 June 2020

### References

- Holtermann A, Hansen JV, Burr H, Søgaard K, Sjøgaard G. The health paradox of occupational and leisure-time physical activity. *Br J Sports Med.* 2012;46:291–5.
- Holtermann A, Burr H, Hansen JV, Krause N, Søgaard K, Mortensen OS. Occupational physical activity and mortality among Danish workers. *Int Arch Occup Environ Health.* 2012;85:305–10.
- Noy YI, Horrey WJ, Popkin SM, Folkard S, Howarth HD, Courtney TK. Future directions in fatigue and safety research. *Accid Anal Prev.* 2011;43:495–7.
- Arvidsson D, Fridolfsson J, Börjesson M. Measurement of physical activity in clinical practice using accelerometers. *J Intern Med.* 2019;286:137–53.
- Scott KA, Browning RC. Occupational physical activity assessment for chronic disease prevention and management: a review of methods for both occupational health practitioners and researchers. *J Occup Environ Hyg.* 2016;13:451–63.
- Cleland I, Kikhia B, Nugent C, Boytsov A, Hallberg J, Synnes K, et al. Optimal placement of accelerometers for the detection of everyday activities. *Sensors.* 2013;13:9183–200.
- Ngueleu AM, Blanchette AK, Maltais D, Moffet H, McFadyen BJ, Bouyer L, et al. Validity of instrumented insoles for step counting, posture and activity recognition: a systematic review. *Sensors.* 2019;19:2438.
- Skotte J, Korshøj M, Kristiansen J, Hanisch C, Holtermann A. Detection of physical activity types using Triaxial accelerometers. *J Phys Act Health.* 2014; 11:76–84.
- Attal F, Mohammed S, Dedabrishvili M, Chamroukhi F, Oukhellou L, Amirat Y. Physical human activity recognition using wearable sensors. *Sensors.* 2015;15:31314–38.
- Ferrannini E. The theoretical bases of indirect calorimetry: a review. *Metabolism.* 1988;37:287–301.
- Montoye AHK, Westgate BS, Fonley MR, Pfeiffer KA. Cross-validation and out-of-sample testing of physical activity intensity predictions with a wrist-worn accelerometer. *J Appl Physiol.* 2018;124:1284–93.
- Sasaki JE, Hickey AM, Staudenmayer JW, John D, Kent JA, Freedson PS. Performance of activity classification algorithms in free-living older adults. *Med Sci Sports Exerc.* 2016;48:941–50.
- Reiff C, Marlatt K, Dengel DR. Difference in caloric expenditure in sitting versus standing desks. *J Phys Act Health.* 2012;9:1009–11.
- Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR Jr, Tudor-Locke C, et al. 2011 compendium of physical activities: a second update of codes and MET values. *Med Sci Sports Exerc.* 2011;43:1575–81.
- 2018 Physical Activity Guidelines Advisory Committee. 2018 physical activity guidelines advisory committee scientific report. Washington, DC: U.S. Department of Health and Human Services; 2018. [https://health.gov/paguidelines/second-edition/report/pdf/PAG\\_Advisory\\_Committee\\_Report.pdf](https://health.gov/paguidelines/second-edition/report/pdf/PAG_Advisory_Committee_Report.pdf).
- Bailey DP, Locke CD. Breaking up prolonged sitting with light-intensity walking improves postprandial glycemia, but breaking up sitting with standing does not. *J Sci Med Sport.* 2015;18:294–8.
- Sasaki JE, John D, Hickey A, Lyden K, Hagobian T, Freedson P. Feasibility of using a continuous direct observation technique for assessment of free-living physical activity in young adults. *Arq Ciênc Esporte.* 2017;4. <http://seer.uftm.edu.br/revistaelectronica/index.php/aces/article/view/1186>. Accessed 18 Sept 2019.
- Farrahi V, Niemelä M, Kangas M, Korpelainen R, Jämsä T. Calibration and validation of accelerometer-based activity monitors: a systematic review of machine-learning approaches. *Gait Posture.* 2019;68:285–99.
- Segal MR. Machine Learning Benchmarks and Random Forest Regression. UCSF: Center for Bioinformatics and Molecular Biostatistics. 2004. Retrieved from <https://escholarship.org/uc/item/35x3v9t4>.
- Willets M, Hollowell S, Aslett L, Holmes C, Doherty A. Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK biobank participants. *Sci Rep.* 2018;8:1–10.
- Holtermann A, Krause N, van der Beek AJ, Straker L. The physical activity paradox: six reasons why occupational physical activity (OPA) does not confer the cardiovascular health benefits that leisure time physical activity does. *Br J Sports Med.* 2018;52:149–50.
- Bhattacharya A, McCutcheon EP, Shvartz E, Greenleaf JE. Body acceleration distribution and O2 uptake in humans during running and jumping. *J Appl Physiol.* 1980;49:881–7.
- Banos O, Galvez J-M, Damas M, Pomares H, Rojas I. Window size impact in human activity recognition. *Sensors.* 2014;14:6474–99.
- Koskimäki H. Avoiding bias in classification accuracy - a case study for activity recognition. In: 2015 IEEE symposium series on computational intelligence; 2015. p. 301–6.

### Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

