DOCTORAL THESIS

Advancing Novel Physical Fatigue Assessment and Human Activity Monitoring Methods towards Perzonalized Feedback with Wearable Sensors

Ardo Allik

TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY
TALLINN 2022

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
62/2022

Advancing Novel Physical Fatigue Assessment and Human Activity Monitoring Methods towards Perzonalized Feedback with Wearable Sensors

ARDO ALLIK

TALTECH PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies
Department of Health Technologies
This dissertation was accepted for the defence of the degree 17/10/2022

Supervisor: Prof Ivo Fridolin
Department of Health Technologies
Tallinn University of Technology
Tallinn, Estonia

Co-supervisor: PhD Kristjan Pilt
Department of Health Technologies
Tallinn University of Technology
Tallinn, Estonia

Prof Gert Jervan
Department of Computer Systems
Tallinn University of Technology
Tallinn, Estonia

Opponents: Ingrid Eitzen, Research Manager
Smart Sensors and Microsystems
SINTEF Digital
Oslo, Norway

Prof Pasi Liljeberg
Digital Health Technology
University of Turku
Turku, Finland

Defence of the thesis: 16/11/2022, Tallinn

Declaration:

Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted for doctoral or equivalent academic degree.

Ardo Allik

----
signature

European Union
European Regional
Development Fund

Investing
in your future

Copyright: Ardo Allik, 2022
ISSN 2585-6898 (publication)
ISBN 978-9949-83-913-1 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9949-83-914-8 (PDF)
Printed by Koopia Niini & Rauam

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
62/2022

Kantavatel seadmetel põhinevate füüsilise väsimuse hindamise ning inimese aktiivsuse monitoorimise meetodite arendamine personaalseks tagasisideks

ARDO ALLIK

TALTECH KIRJASTUS

Contents

List of Publications

The list of author's publications, on the basis of which the thesis has been prepared:

I Allik, A., Pilt, K., Karai, D., Fridolin, I., Leier, M., & Jervan, G. (2019). Optimization of Physical Activity Recognition for Real-Time Wearable Systems: Effect of Window Length, Sampling Frequency and Number of Features. *Applied Sciences, 9*(22), 4833. DOI: 10.3390/app9224833 (IF = 2.838)

II Allik, A., Mägi, S., Pilt, K., Karai, D., Fridolin I., Leier, M., Jervan, G. (2018). Comparison of Predictive Equations for Basal Metabolic Rate. *Proceedings of the 7th International Conference on Wireless Mobile Communication and Healthcare (MobiHealth 2017)*, Vienna, Austria, 261−264. DOI: 10.1007/978-3-319-98551-0

III Allik, A., Pilt, K., Viigimäe, M., Fridolin, I., & Jervan, G. (2022). A Novel Physical Fatigue Assessment Method Utilizing Heart Rate Variability and Pulse Arrival Time towards Personalized Feedback with Wearable Sensors. *Sensors, 22*(4), 1680. DOI: 10.3390/s22041680 (IF = 3.847)

Author’s Contribution to the Publications

Contribution to the papers in this thesis are:

I The author was the main manuscript writer and the corresponding author of the paper. The author performed and wrote the scripts for data preprocessing, classification feature extraction and feature selection, classifier training and classifier performance evaluation. The author optimized the human activity recognition by analyzing the effects of different classification window lengths, sampling frequencies and feature sets on the classifier performance. The author partly conceived the study. The author partly planned and performed the experiments for the human activity recognition assessment.

II The author was the main manuscript writer and the corresponding author of the paper. The author analyzed the data from the indirect calorimetry device. The author performed the required calculations and analyzed the results of the energy expenditure predictive equations. The author partly conceived the study. The author partly planned and performed the experiments for the energy expenditure assessment.

III The author was the main manuscript writer and the corresponding author of the paper. The author analyzed the results and proposed the method for real-time physical fatigue assessment. The author planned and performed the experiments for the physical fatigue assessment. The author set up the devices and systems for measuring the reference parameters (fatigue questionnaire, reaction time, hand grip strength and countermovement jump) and for measuring the cardiovascular parameters during exercise and recovery phases (heart rate, heart rate variability and pulse arrival time). The author managed the data and performed data processing. The author partly conceived the study.

Other Related Publications

IV Allik, A., Pilt, K., Karai, D., Fridolin, I., Leier, M., & Jervan, G. (2016). Activity classification for real-time wearable systems: Effect of window length, sampling frequency and number of features on classifier performance. *Proceedings of the IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES 2016)*, Kuala-Lumpur, Malaysia, 460−464. DOI: 10.1109/iecbes.2016.7843493

V Leier, M., Jervan, G., Allik, A., Pilt, K., Karai, D., & Fridolin, I. (2018). Fall detection and activity recognition system for usage in smart work-wear. *Proceedings of the 16th Biennial Baltic Electronics Conference (BEC 2018)*, Tallinn, Estonia, 1–4. DOI: 10.1109/bec.2018.8600959

VI Allik, A., Pilt, K., Karai, D., Fridolin, I., Leier, M., & Jervan, G. (2018). Classification Algorithm Improvement for Physical Activity Recognition in Maritime Environments. *Proceedings of the World Congress on Medical Physics and Biomedical Engineering 2018 (IUPESM 2018)*, Prague, Czech Republic, 13–17. DOI: 10.1007/978-981-10-9023-3_3

VII Allik, A., Pilt, K., Viigimäe, M., & Fridolin, I. (2019). Pilot Study for Estimating Physical Fatigue Based on Heart Rate Variability and Reaction Time. *Proceedings of the XV Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON 2019)*, Coimbra, Portugal, 193–200. DOI: 10.1007/978-3-030-31635-8_23

Abbreviations

HAR	Human Activity Recognition
BMR	Basal Metabolic Rate
EE	Energy Expenditure
IC	Indirect Calorimetry
CV	Cardiovascular
RS	Rested-state
PFS	Physically-fatigued-state
CMJ	Countermovement Jump
SOFI	Swedish Occupational Fatigue Inventory
RT	Reaction Time
HR	Heart Rate
HRV	Heart Rate Variability
ECG	Electrocardiography
SDNN	Standard Deviation of All NN Intervals
RMSSD	Square Root of the Root Mean Square of the Sum of All Differences Between Successive NN intervals
SFS	Sequential Forward Selection
RMR	Resting Metabolic Rate
PAT	(Blood-pressure Normalized) Pulse Arrival Time
PA	Physical Activity
PFA	Physical Fatigue Assessment

1 Introduction

1.1 Motivation

The circumstances of current human existence are far different from remote past. Physical exertion is no longer a requirement for daily living and today's conditions allow an unprecedentedly sedentary lifestyle. This discordance between our contemporary lives and our genetic makeup has important health implications on skeletal density, cardiovascular diseases, obesity, body composition and insulin resistance. It is important to propagate active lifestyle, since research studies confirm that routine physical activity (PA) has multiple benefits by lowering the risk of diabetes, cardiovascular disease and obesity, while increasing psychological well-being.

Advancement of technology has brought a surge of popularity for devices that help their users keep track of their PA, training schedule, exercises and lost calories. Since this makes training more interactive and allows users to have better overview of their progress, it often motivates the users to have a more active lifestyle. This is achieved by using wearable systems to conveniently measure, collect and analyze the user's physiological data. For convenient use wearables need to be small and unobtrusive, which in turn puts significant demand on optimizing different aspects of these system such as reducing power consumption. The general aim of the thesis is to advance novel physical fatigue assessment (PFA) and human activity monitoring methods that could be applied in real-time by using wearable sensors and systems.

1.2 Problem Formulation

Human activity recognition (HAR) allows automatic recognition of physical activities. Real-time activity recognition provides valuable information for improving online feedback of the activity trackers or for providing extra safety by monitoring the status of the users working in high-risk environments (Leier et al., 2018; Svertoka et al., 2021). Power consumption required for HAR is determined by multiple different components. Some of these components are based on the processing of the acceleration values, such as sampling rate of the signal and filtering (Yan et al., 2012; Straczkiewicz et al., 2021). Other elements are based on classification mechanics, such as classification window length, feature calculation, and the used machine learning algorithm. While studies have explored classification aspects such as training times of different HAR algorithms (Altun et al., 2010; Feng et al., 2015), they do not provide valuable information for real-time classification, since classifier training can be done previously on a desktop computer and later implemented into the wearable system. For classification systems working in real time, it is important to focus on the processing time of the calculations the system has to do online (Altun et al., 2010; Tapia, 2008).

Several studies have evaluated how different window lengths affect HAR performance (Tapia, 2008; Bulling et al., 2014; Straczkiewicz et al., 2021), but the lack of gold standard in physical activity classification makes it difficult to compare these results (Awais et al., 2015; Straczkiewicz et al., 2021). Researchers have used a wide range of various sampling frequencies, typically between 10 Hz to 100 Hz (Yan et al., 2012; Khusainov et al., 2013; Lee et al., 2016; Wang et al., 2019). Various filter methods, wrapper methods and embedded methods have been used for feature selection (Wang et al., 2019), such as the ReliefF algorithm (Moncada-Torres et al., 2014), principal component analysis (Altun et al., 2010), or information gain (Tapia, 2008), but not in connection with window length and sampling frequency.

Energy expenditure (EE) is an important parameter for the studies of PA and is often used as a correlate of its level (Wang et al., 2012). EE determination is an important tool for adjusting the individuals' nutritional supply or to assess the health of a larger population. Modern technologies that are gradually integrated into everyday life are able to non-invasively monitor the PA level and health behavior of their users. Monitoring the PA has moved towards activity specific EE models that first recognize the activity and then apply a suitable EE algorithm for the specific activity (Altini et al., 2012; Farrahi et al., 2019), which relies on accurate assessment of basal metabolic rate (BMR).

BMR is usually clinically measured using indirect calorimetry (IC), but it requires expensive equipment and trained personnel. Therefore for dietetics purposes BMR is commonly estimated using predictive equations, that use simple anthropometric variables such as the weight, height, age and gender of the person (Frankenfield et al., 2005; Amaro-Gahete et al., 2018).

Fatigue is a term used to describe an altered physiological state, which may result in decreased mental or physical performance. Fatigue may be caused by various effects such as sleep loss, circadian changes, or high workload (Mohanavelu et al., 2017; Shortz et al., 2017). The ability to effectively monitor fatigue is highly sought due to multiple reasons: the complaint of fatigue is high in general population (Dawson et al., 2011); it may adversely affect employees' performance, safety, and health (Völker et al., 2015); and the high prevalence of fatigue has been reported in many operational settings as potential hazard (Shortz et al., 2017; Thompson, 2019). Manifestation of high prevalence of fatigue in the working population has spawned growing concern due to reduced performance, high sick leave and work disability (Thompson, 2019). The main topics in the study of fatigue are significance of fatigue tests in different (work) settings, evaluation of muscular fatigue, subjective symptoms of fatigue, indicators of nervous strain, and the practical application of fatigue tests (Yu et al., 2019).

In sport, fatigue manifests as a reduction in the ability to perform the desired movement, exercise, or skill (Hughes et al., 2019). The capacity to effectively monitor fatigue provides coaches and scientists with the ability to optimize training and improve competition performance (Hughes et al., 2019). For a coach it is both useful to have an index of the level of fatigue induced as a prolonged increase in training load over a longer time period as well as to determine how well an athlete is tolerating an acute increase in exercise load in one day (Thomson et al., 2016).

In operation settings, many respected professional organizations, such as healthcare organizations American Nurses Association (ANA), The National Association of Neonatal Nurses (NANN) and Washington State Nurses Association (WSNA), have put forwards position statements to draw awareness to worker fatigue and its consequences, with the aim to legitimize fatigue as a serious and pressing issue (Thompson, 2019). Worker fatigue can adversely impact personal health and safety as well efficiency and safety of the operation (Lerman et al., 2012). In healthcare, in addition to the greater injury risk to the fatigue-impaired worker, fatigue of the healthcare provider is a primary contributor to negative patient outcomes (Thompson, 2019). Adverse effects of physical fatigue on the health and safety of the construction workers has been well researched (Anwer et al., 2021).

There is no single instrument which is used as a gold standard for fatigue measurement, because of the definitional difficulties and multiple causes of fatigue. The multifactorial nature of fatigue may suggest that a single universal test to measure fatigue may not exist (Hughes et al., 2019; Saito, 1999). Fatigue assessment studies have usually compiled different test batteries of various measures for this purpose (Hughes et al., 2019; Thompson, 2019). In high performance sports several performance markers are often used to assess fatigue such as questionnaires, jump tests, sprints, heart rate parameters, hormone levels and postural sway measurements (Hughes et al., 2019). Occupational specific fatigue research has almost exclusively implemented subjective assessments in the form of questionnaires, which is not representative of actual human performance-based functionality and can easily be manipulated by the employee (Thompson, 2019).

It is important to note that most of these require specific administered tests. This type of comprehensive testing of fatigue may not be feasible as time, space, financial resources, testing personnel and the willingness to be tested are all scarce (Völker et al., 2016). For this reason, a new approach utilizing physiological signals (markers, measures) is needed, which would allow fatigue assessment to be conducted autonomously in real-time and unobtrusively for the user.

1.3 Aims of the Thesis

The thesis aims to provide novel methods for physical fatigue assessment and human activity monitoring, which could be applied in real-time and unobtrusively using wearable sensors and systems. To accomplish this, aim different aspects of fatigue and suitable parameters were analyzed in multiple studies. More specifically, the aims of the thesis are:

1. Improve Human Activity Recognition by evaluating the window length, sampling frequency and feature selection in order to optimize the classifier for real-time wearable systems. (**Publication I**)
2. Validate multiple Basal Metabolic Rate predictive equations in order to provide more precise input and validation for fatigue estimation system. (**Publication II**)
3. Improve and develop towards autonomous real-time fatigue estimation system by validating various promising test battery measures which could be measured using wearable sensors and provide fatigue estimation models based on the results. (**Publication III**)

1.4 Contribution

This thesis contributes to advancing and providing novel knowledge about multiple promising methods for real-time fatigue assessment and human activity monitoring. All the analyzed measures may be implemented in real-time models, which may be used in a wearable system.

In **Publication I**, it was evaluated how the effects of sampling frequency of the acceleration signal, window length and number of features (listed in Appendix 1) affect the performance of the HAR algorithm. The main findings were: (i) classification F1-scores with window lengths of 5 s and 3 s were similar, while results with 1 s were lower; (ii) all sampling frequencies performed similarly for most activity types, with an exception of outdoor cycling; (iii) Similar or better results were achieved with the feature sets with 9 to 14 features, achieved with either feature reduction scheme, compared to the initial full feature set of 110 features. The results of the study have been used for developing more efficient real-time physical activity classifiers.

In **Publication II**, different basal metabolic rate predictive equations were compared and validated with the data measured using indirect calorimetry. From the eight different BMR predictive equations explored in this study, Mifflin-St Jeor formula had the best performance. Based on regression analysis most equations had similar results, with Owen and Kleiber formulas being the outliers, which respectively had the lowest and highest average BMR results. The average BMR values with Mifflin-St Jeor formula (1447 ±204 kcal/day) were the closest with IC results (1485 ±255 kcal/day) and also had the lowest RMSE of 175 kcal/day compared to IC.

In **Publication III**, a novel method for real-time PFA was proposed, which uses a set of real-time and easily measurable cardiovascular (CV) parameters, that could be continuously and unobtrusively monitored. Evaluated CV parameters were heart rate (HR), measures of heart rate variability (HRV), and blood pressure normalized pulse arrival time (PAT). The main findings were: (i) from the assessed CV parameters, the statistically significant change between the rested-state and physically-fatigued-state was noted in the average HR and HRV measures SDNN and RMSSD; (ii) the strongest linear correlation was found between the reference parameter hand grip strength and PAT. (iii) the best performing CV parameters for separating the mildly fatigued and significantly fatigued groups were based on HRV parameter SDNN between the rested-state and the physically-fatigued state and PAT changes during the physically-fatigued state. The results of the study provide a significant improvement among existing PFA methods.

2 Literature Review

2.1 Human Physical Activity and Fatigue

Physical inactivity is a growing problem in the world, which has been found to cause 6–10% of the major non-communicable diseases of coronary heart disease, type 2 diabetes, and breast and colon cancers. Furthermore, this unhealthy behaviour causes 9% of premature mortality (Lee et al, 2012). The recent development in sensor technologies and decrease in the cost of sensor based devices have driven the implementation of health monitoring and human activity detection using mobile and wearable sensors (Nweke et al., 2019).

Physical activity monitoring has been found to have a positive effect in increasing PA (Larsen et al., 2019). Wearable systems are used to conveniently measure, collect and analyze the user's physiological data and provide their users extra information based on it (Kumari et al., 2017). In work environments wearables can be used to monitor employees' psychological and physiological factors, enhance operational efficiency, promote work environment safety and security, and improve workers' health through monitoring, supporting, training and tracking the personnel (Khakurel et al., 2018, Svertoka et al., 2021). Day by day, new trends can be seen in the field of wearable systems that require wearables to be small and unobtrusive, which in turn puts significant demand on reducing power consumption of the system (Senevirante et al., 2017). With the proliferation of wearable technologies clinicians, researchers, patients and technology developers need to know the current state of what works and what limitations exist (Hilty et al., 2021).

Human activity recognition (HAR) allows automatic recognition of physical activities and provides valuable information for improving online feedback of activity trackers (**Publication I**). HAR may also be used for fatigue estimation using various methods. First, HAR has been proposed for improving energy expenditure (EE) estimation using activity-specific models (Altini et al., 2012). This also requires accurate estimation of basal metabolic rate (BMR), which could be done with EE predictive equations (**Publication II**). HAR can also be used for classifying the work (or exercise) periods and the resting states, which could allow automatic analyzing of the informative post-exercise cardiac recovery phase (Peçanha et al., 2017). This could be a basis for continuous and unobtrusive physical fatigue assessment (PFA), where feedback could be given in real-time by measuring and analyzing multiple cardiovascular parameters (**Publication III**). These are the main topics of this thesis, with the aim to provide results for creating novel methods for physical fatigue assessment and human activity monitoring, which could be monitored in real-time and unobtrusively using wearable sensors and systems (Table 1).

Table 1. Literature review summary with current research gaps and the contributions of this work in human activity recognition (HAR), energy expenditure (EE) and physical fatigue assessment (PFA).

Study type	Related studies and reviews	Research gap	Contributions of this work
HAR optimization	(Wang et al., 2019) (Khusainov et al., 2013) (Straczkiewicz et al., 2021)	Optimization not focused on real-time HAR	Analyzing the performance of a HAR model which would be suitable for use in real-time systems
		The combined effect of the parameters is not analyzed	Three different HAR model aspects (sampling frequency, window length and feature selection) were analyzed concurrently
EE with predictive equations	(Frankenfield et al., 2005), (Amaro-Gahete et al., 2018)	Not all predictive equations were included or validated on similar population	Eight different predictive equations were compered and validated with indirect calorimetry results
Methods and measures for PFA	(Thompson, 2019), (Mohanvelu et al., 2017), (Peçanha et al., 2017)	Assessed parameters are not suitable for use in real-time PFA	Study focused on various cardiovascular parameters that could be used in real-time PFA
		Pulse arrival time not considered and analyzed as a potential measure	Pulse arrival time based parameter was one of the best for separating the mildly fatigued and significantly fatigued groups

2.1.1 Human Activity and Energy Expenditure

The capacity of the body to exercise or do physical work depends on its ability to produce, use, and regulate energy. Energy expenditure (EE) determination is an important tool for adjusting the individuals' nutritional supply or to assess the health of a larger population. The body's 24-hour EE can be divided into three components (Hills et al., 2014): (i) the thermal effect of food, (ii) the resting metabolic rate (iii) the energy cost of PA.

Thermal effect of food is the amount of energy required to digest, absorb, and to process the nutrients in food, such as fat, protein, carbohydrate and constitutes from 5 to 10% of the total EE (Poehlman, 1989). Resting metabolic rate (RMR) is the amount of energy expended to sustain the basic body functions (Pinheiro Volp et al., 2011) and constitutes from 60 to 75% of the total EE (Poehlman, 1989).

Modern technologies that are gradually integrated into everyday life are able to non-invasively monitor the PA level and health behavior of their users. Monitoring the PA is moving towards activity specific EE models that first recognize the activity and then apply a suitable EE algorithm for the specific activity (Farrahi et al., 2019), which relies on accurate assessment of basal metabolic rate (BMR).

Doubly labelled water is considered the gold standard for the measurement EE; however, the considerable costs and analytical requirements limit its feasibility in large cohort studies (Racette et al., 2012). More common alternatives are indirect calorimetry (IC) methods which represent the criterion measure for assessment of the energy cost of an activity but are again limited to structured activities usually within a laboratory (Hills 2014). Heat is released as a by-product in cellular metabolism. The rate of heat release is directly proportional to the rate of metabolism. Therefore, the metabolic rate can be determined by measuring heat release. Direct calorimetry is termed as the process of measuring metabolic heat release (Pinheiro Volp et al, 2011). Direct calorimeters are relatively expensive and used mostly in hospitals, clinics and university research labs

(Webster et al., 1986, Schutz, 1995). The respiratory calorimetry is used in clinics instead of expensive direct calorimetry method for EE estimation. This IC method involves direct measurement of oxygen consumption (VO2) in metabolism through the measurement of respiratory gases. Firstly, the VO2 is measured and then converted into an equivalent EE in kilocalories (kcal) (Pinheiro Volp et al., 2011).

Direct and indirect calorimetry methods are relatively expensive, complex and time consuming. Therefore, a lot of effort has been put on developing a predictive equation for EE estimation (Cunningham, 1991), which are used to calculate an estimation of RMR using anthropometric parameters such as height, weight, gender and age. Harris and Benedict (Harris et al., 1918) and Kleiber (Kleiber, 1932) are most common equations for RMR prediction. However, it has been noted that both of these equations are less predictive for obese subjects (Daly et al., 1985). It is due to the fact that obese people were not included in the data sets of the equation development. The different body sizes and body compositions were taken into account in the development of the RMR prediction equations from Mifflin (Mifflin et al., 1990) and Livingston and Kohlstadt (Livingston et al., 2005). Both of them are best suited for obese subjects, but also valid for normal weight people.

Humans use more energy when performing more rigorous and exhausting activities (Ainsworth et al., 2011) and thus EE is directly linked to the amount of fatigue induced in humans. Precise estimation of EE allows us to use it to validate PFA methods or use it as an input variable (Amor et al., 2015).

2.1.2 Fatigue Physiology and Classification

Muscle activation begins in the cortex, continues with excitation of lower motor neurons in the spinal cord, to the axon of the lower motor neuron and eventually to the neuromuscular junction of the muscle (Noakes, 2012). In this process, fatigue can potentially arise at any point of the pathway.

When focusing on the processes inside the spinal cord and the brain, fatigue is defined as “central”, and when focusing on the peripheral nerve, neuromuscular junction, and the muscle, fatigue is defined as “peripheral” (Allen et al., 2008).

Central fatigue is described as fatigue coming not from the muscle itself, but rather from the central nervous system (CNS) and the transmission of signals from the brain to the muscle. Therefore, central fatigue is related to the brain and the spinal cord.

Peripheral fatigue is the failure to maintain an expected power output and can be caused by two different actions: (i) depletion of glycogen, phosphate compounds, or acetylcholine within the muscular unit; (ii) accumulation of lactate or other metabolites that are released during activity. Therefore, peripheral fatigue occurs within the muscle.

Skeletal muscle fatigue has been generally defined as “the decrease in force or power production in response to contractile activity” (Kent-Braun, 2012). In vitro studies have shown that the impairment of muscle contraction, and thus the development of muscle fatigue at the cellular level, derives from either (i) alterations in excitability of the muscle fiber, (ii) accumulation of metabolic by-products, (iii) production of reactive oxygen species and (iv) Ca^{2+} movements in the fiber compartments (Allen et al., 2008). All of the above can be grouped in two major mechanisms that are responsible for the inhibition of muscle function witnessed during fatigue: (i) impairment at the level of activation, and (ii) impairment of the actin–myosin interaction.

There is a need to monitor both short-term fatigue, which is typically metabolic in origin, and more prolonged, neuromuscular fatigue. Metabolic fatigue is described as a decrement in muscle force generating capacity as a response to physical exercise that has outstripped the rate of ATP1 replacement. Its effects begin to diminish after a period of five minutes and is generally thought to have dissipated after 3 h (Layzer, 1990). Neuromuscular fatigue is defined as a prolonged decrease in the muscle's ability to generate a force or power output after a period of recovery. Neuromuscular fatigue can be present for upwards of 48 h, and can be identified as a compound system with both central and peripheral origins (Overton, 2013).

Based on the origin and the effect on the body, fatigue is often classified between cognitive, visual and physical fatigue. Cognitive or mental or central fatigue, henceforth referred to as cognitive fatigue, involves decrements in human information processing due to mental workload. It may be conceptualized as an executive failure to sustain attention in order to maintain or optimize performance (Ackerman et al., 2009).

Visual fatigue is a consequence of prolonged visual activity rather than mental workload, which causes changes in arousal level. Visual fatigue might be confused with cognitive fatigue, as there are cases where a decrement in arousal may lead to changes in oculomotor behavior despite no visual discomfort (Megaw, 1995).

Physical fatigue involves the inability to maintain physical performance, and can be attributed to metabolic disturbances, failure of neuromuscular transmission, changes affecting the myosin-actin complex, etc. Physical fatigue might also be attributed to changes in function of the central nervous system and impairments might occur in supraspinal areas, spinal areas, and in the muscle afferent system (Behm, 2004).

Therefore, it might be misleading to term cognitive fatigue as central fatigue, as central mechanisms might contribute to physical fatigue without changes in cognitive workload. Thus, it is commonly accepted that cognitive fatigue impairs physical performance, probably by increasing the effort perception.

2.1.3 Work Fatigue

Work fatigue represents extreme tiredness and reduced functional capacity that is experienced during and at the end of the workday. Work fatigue involves both extreme tiredness (i.e., lack of energy) and reduced functional capacity. This can occur with respect to each of the three energetic resources: (i) physical (involving muscular movement), (ii) cognitive (involving cognitive processing) and (iii) emotional (involving expression and regulation of emotions) (Frone et al., 2015).

The distinction between physical fatigue resulting from depletion of muscular energy and cognitive fatigue resulting from depletion of cognitive energy dates at least 90 years (Pillsbury, 2009). Growing attention has focused on emotional fatigue, resulting from depletion of emotional energy, in addition to physical and cognitive fatigue (Shirom et al., 2006). Considering the three separate energy resources, the following resource-specific definitions of work fatigue are proposed: (i) physical work fatigue, which represents extreme physical tiredness and reduced capacity to engage in PA, (ii) cognitive work fatigue represents extreme mental tiredness and reduced capacity to engage in cognitive activity and (iii) emotional work fatigue represents extreme emotional tiredness and reduced capacity to engage in emotional activity.

Work fatigue is also temporally tied to the workday (Demerouti et al., 2003). It has an onset when energy depletion becomes too great and an offset when energetic demands and energy is restored through rest. Work fatigue can be assessed as an acute/state condition (e.g., the experience of fatigue at the present moment) or a chronic/trait condition (e.g., the experience of work fatigue over the past 12 months). The acute/state experience of work fatigue resolves shortly after the end of every workday, if it occurs frequently over an extended period of time, it may be viewed as a chronic/trait condition.

Based on the above definitions, a measure of work fatigue should be multidimensional, separately assessing physical, mental, and emotional dimensions of work fatigue.

2.2 Human Activity and Fatigue Monitoring

The objective of activity monitoring is to analyse or interpret the ongoing events from data automatically (Kumari et al, 2017). With the development of new technology and wearable devices, such as wrist-wearable smartwatches, monitoring human activity has become more and more popular and accessible. Wearables are smart electronic devices available in various forms that are used to conveniently measure, collect and analyze the user's physiological and behavioral data using a variety of methods, interventions and outcomes (Kumari et al., 2017, Khakurel et al., 2018; Hilty et al., 2021). In addition to specialized wearable systems, there has also been a lot of research effort in monitoring human activity with smartphones, using their numerous built-in sensors (Straczkiewicz et al., 2021). Other researchers have based activity monitoring on various visual data, such as videos from Closed-Circuit Television or even images from social media (Arshad et al., 2022).

In healthcare, activity monitoring can provide objective and reproducible information regarding traditional and emerging risk factors of human populations. Additionally, behavioral risk factors, including sedentary behavior, sleep and physical activity can all be monitored using wearables or smartphones (Straczkiewicz et al., 2021). Activity monitoring can be also used for monitoring the daily activities of hospitalized patients, whose inactivity can lead to functional decline or increased activity could mean readiness for discharge (Fridriksdottir et al., 2020). Sensor-based activity monitoring and recognition is also one of the most promising assistive technologies to support older people's daily life (Wang et al., 2019).

Human activity monitoring includes two processes – first data acquisition, which is followed by classification of the acquired data. The acquisition of data includes acquiring the bio-signals and signal preprocessing. Signal preprocessing includes amplifying, filtering, averaging, extracting relevant features to be used as training data for classifier etc (Kumari et al., 2017). Various methods from the field of signal processing have been used to distill collected sensor data, including k-NN, random forest, decision tree, gaussian models and hidden Markov models or simpler threshold methods (Castro-Garcia et al., 2022).

Data acquisition process has two different approaches – one is the traditional approach which uses external sensors such as cameras or other monitoring devices (Lin, 2009) and the second one is the newly introduced approach which uses wearable wireless sensors. Both approaches use different types of sensors to acquire the physiological signals. However, in the second approach, sensors are attached to the human body (Kumari et al., 2017).

Human Activity Recognition (HAR) systems based on wearable sensors can be categorized in two stages. First stage is learning stage, which may be supervized, unsupervized or semi-supervized. In the second stage, which may be either offline or online, performed actions are recognized and feedback is given accordingly. While offline schemes require more time to respond to the actions performed. Offline scheme demands high computation and is suitable for applications that do not demand immediate feedback in real-time. (Kumari et al, 2017)

Wearable sensors are typically wireless tiny sensors enclosed in bandages or some patches or something that can be worn. Calorimetric, potentiometric, amperometric, optical, piezo-electric biosensors and immunosensors are different types of wearable sensors. The data acquired from these wearable sensors are processed as per requirement for a particular application. Wearable sensors are completely unobtrusive devices that help physicians to overwhelm the restrictions of traditional technologies. Through wearable systems, biological signals can be continuously acquired wirelessly and thus patients can be monitored remotely. (Kumari et al, 2017)

Before developing a wearable system, it is essential to have a clear idea about the basic requirements and designing challenges for any wearable device. There are always hardware and software constraints beginning from low-energy operations, lightweight and safety requirements. While person is placing the wearable sensor on his/her body, the chances of thermal injury must be considered and should be reduced by controlling the sensing and wireless frequency and radio duty cycle of wearable sensor. Some basic requirements to take into account are: (i) aesthetics, (ii) size, (iii) water tolerance, (iv) power consumption, (v) wireless communication, (vi) operating system. (Kumari et al, 2017).

Day by day, new trends can be seen in the field of wearable systems which has enhanced features. For example, shirt or other clothes with all-fabric keyboard made by conductive thread can be washed in the machine same as ordinary clothes. So, it is water durable which is one of the basic requirements for a wearable device. Computerized clothes can be the next generation for computers and other devices which does not require strap of electronics into our body. This requires wearables to be small and unobtrusive, which in turn puts significant demand on reducing power consumption of the system (Senevirante et al., 2017). Although a huge amount of effort is being made in the wearable sensors, challenges like user-acceptance, low power consumption, interference in wireless systems are still to be resolved for better usability and functionality of these wearable devices. (Kumari et al., 2017).

2.2.1 Human Activity Recognition

Human activity recognition (HAR) allows to recognize the activity or activity type that the user is conducting based on the signals from a wearable sensor. Even though the precise methods for HAR vary, then it is usually done based on accelerometer sensor data and the main algorithm used in all HAR researchers can be divided into following stages (Qi et al., 2018; Straczkiewicz et al., 2021):

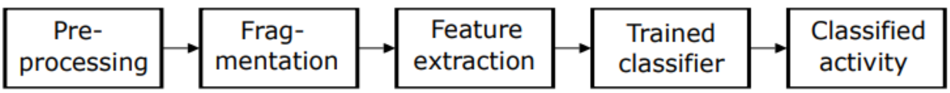

Figure 1. Stages of the HAR algorithm: (i) pre-processing of the raw data, (ii) fragmentation of the filtered data into smaller time segments and labelling them according to the activity class, (iii) choosing the amount and type of features to use in activity classification and extracting them from the data, (iv) training the classifier based using the chosen features based on the training set (v) classification of new signals using the previously trained classifier.

Accelerometer signals are usually measured using triaxial inertial measurement unit (IMU) sensors, which are attached to the human body. Studies have explored and validated the results with one or more accelerometers in multiple different locations, commonly on wrists, ankles, thighs or chest. (Chowdhury et al., 2013; Loh et al., 2015, Castro-Garcia et al., 2022). While combining the data from multiple sensors has been shown to improve the classification performance, then it comes with a trade-off due to the increase of the system complexity and computational power requirements, which are important factors when optimizing the HAR for use in wearables systems working in real-time. This is also significant for real-time HAR, which can be used in wearables for online activity recognition by allowing automatic recognition of the activities that the user is performing (Lee et al., 2018; Wannenburg et al., 2017). Real-time activity recognition provides valuable information for improving online feedback of the activity trackers or for providing extra safety by monitoring the status of the users working in high-risk environments (Leier et al., 2018).

Power consumption required for HAR is determined by multiple different components. Some of these components are based on processing of the acceleration values, such as sampling rate of the signal and filtering (Yan et al., 2012). Other elements are based on classification mechanics, such as classification window length, feature calculation and the used machine learning algorithm. While studies have explored classification mechanics such as training times of different HAR algorithms (Altun et al., 2010; Feng et al., 2015), they do not provide valuable information for real-time classification, since classifier training can be done previously on a desktop computer and later implemented into the wearable system. For classification systems working in real-time, it is important to focus on the processing time of the calculations the system has to do online (Altun et al., 2010; Tapia 2008).

Few previous studies have evaluated how different sampling frequencies affect HAR performance. Lowering the sampling frequency, f_s, decreases the number of samples in the classification fragment, s_f, which is calculated as follows:

$$s_f = f_s \cdot w_f \tag{1}$$

where w_f is the window length of a fragment given in seconds. Based on sampling theorem, for accurate representation of a signal, two conditions must be satisfied: the signal must be band-limited and sampling frequency must be at least twice the maximum frequency in the signal (Khusainov et al., 2013). It has been stated that frequencies above 20 Hz cannot be expected to arise from voluntary movement (Bouten et al., 1997). While researchers have used various sampling frequencies, usually in the range of 10 Hz to 100 Hz (Yan et al., 2012; Lee et al., 2016), for similar HAR measurement scenarios, then around 20 Hz has been found adequate by others (Bouten et al., 1997; Khusainov et al., 2013). This section has been changed accordingly in the manuscript.

Filtering is applied in HAR to separate the recorded acceleration signals into static and dynamic components. The static component in the acceleration signal is mostly affected by gravity and captures the posture information, while the dynamic component is based on motion and captures the human movement information.

For HAR, measured acceleration signals are fragmented into shorter consecutive fragments based on which various features are calculated for classifier training and activity classification. Usually, these fragments are found as consecutive time-windows and some studies opt for an overlap between windows to increase the classification performance. Some previous studies have evaluated how different window lengths, commonly chosen between 1.5 s and 5 s (Altun et al., 2010; Aktaruzzaman et al., 2015), affect HAR performance (Tapia 2008; Bulling et al., 2014), but the lack of gold standard in HAR makes it difficult to compare these results (Awais et al., 2015). In a system with a physical activity classifier working in real-time, the window length determines the delay of the system, since each classification is done after signals have been collected for a whole window. The number of samples in the fragment is determined by both the sampling frequency and the window length according to (1).

When using machine learning methods for HAR, the classifier training is done based on features that are extracted from signal fragments. The feature set has to capture specific and diverse information of posture and human motion to allow precise activity classification. These features are usually found in time-domain, frequency-domain or as wavelets (Liu et al., 2012; Moncada-Torres et al., 2014, Tapia 2008), but for real-time wearable systems the possible performance gain from using frequency-domain and wavelets in addition to time-domain features may not be worth the trade-off in terms of computational power, since the system requires extra resources in order to find the transforms which are needed to calculate these features (Preece et al., 2009).

Another possible optimization is in reducing the number of calculated features, which can be achieved using different feature selection methods. Various methods have been used for feature selection, such as ReliefF algorithm (Moncada-Torres et al., 2014), principal component analysis (Altun et al., 2010) or information gain (Tapia 2008). The study presented in **Publication I** is the first study to assess the HAR performance with different number of features in connection to the window length and sampling frequency.

HAR may be used for fatigue estimation using multiple different methods. First, HAR has been proposed for improving energy expenditure (EE) estimation using activity-specific models (Altini et al., 2012). These models first classify the physical activities using IMU sensors and use that information for creating more accurate EE models. HAR and activity tracking has been also used for monitoring and detecting user's behavioral health and stress levels (Hilty et al., 2021, Castro-Garcia et al., 2022) which are important for mental fatigue assessment. Another potential usage for physical fatigue assessment would be classifying the work (or exercise) periods and the resting states. Post-exercise cardiac autonomic recovery has been found to be a practical clinical tool for the assessment of cardiovascular health and has been used for fatigue assessment (Peçanha et al., 2017). Using HAR to automatically detect recovery periods would also allow to use the information from post-exercise cardiac autonomic recovery in real-time physical fatigue estimation with wearable sensors.

2.2.2 Energy Expenditure Estimation and Monitoring

Various methods have been employed with wearable systems for EE estimation. Heart rate has a good linearity with oxygen consumption in a large range of aerobic tasks (Livingstone et al., 1997). However, the poor relationship between HR and EE in resting and low intensity activities is an important limiting factor (Luke et al., 1997). In addition, HR is affected by several factors that are not directly related to metabolism e.g., mental stress, emotions, and medication (Hiiloskorpi et al., 1999).

Accelerometry is also a widely used tool for estimating PA related EE in free-living conditions (Lu et al., 2018). With count-based methods, the activity count is calculated using acceleration, and then directly linked to EE, while the type of activity being performed is not considered. In activity related methods, first the activity recognition is preformed, then the EE is estimated through a look-up table or by using the activity specified EE model (Altini et al., 2015), which requires an accurate assessment of basal metabolic rate.

Predictive equations are commonly used for estimation of the resting energy expenditure (REE) (Table 2). These equations are generally developed for gender, age, body weight, statue and ethnicity, and some of them have been recently formulated for diseases (Marra et al., 2020). Some of the most widely used predictive equations for healthy adults are Harris & Benedict (Harris et al., 1918), Schofield (Schofield 1985), FAO/WHO/UNO (World Health Organ 1985) and Mifflin-St Jeor (Mifflin et al., 1990). These and additional predictive equations were compared and validated with the results of IC in **Publication II** to find the most suitable one for EE model.

*Table 2. Different predictive equations that were assessed in **Publication II**. W – weight (kg), H – height (cm), A – Age (years).*

Authors	Age (y)	Gender	Equation (kcal/day)
Harris-Benedict (Harris et al., 1918)	15 – 74	F	655.0955 + 9.5634·W + 1.8496·H – 4.6756·A
		M	66.4730 + 13.7516·W + 5.0033·H – 6.7550·A
Schofield (Schofield, 1985)	18 – 29	F	14.818·W + 486.6
		M	15.057·W + 692.2
	30 – 59	F	8.126·W + 845.6
		M	11.472·W + 873.1
FAO/WHO/UNU (World Health Organ, 1985)	18 – 29	F	14.7·W + 496
		M	15.3·W + 679
	30 – 59	F	8.7·W + 829
		M	11.6·W + 879
Henry-Rees (Henry et al., 1991)	18 – 29	F	11.472·W + 612.3
		M	13.384·W + 669.2
	30 – 60	F	11.472·W + 585.1
		M	10.994·W + 755.2
Mifflin-St Jeor (Mifflin et al., 1990)	Any	F	9.99·W + 6.25·H – 4.92·A - 161
		M	9.99·W + 6.25·H – 4.92·A + 5
Owen (Owen et al., 1986; 1987)	Any	F	9.99·W + 6.25·H – 4.92·A - 161
		M	9.99·W + 6.25·H – 4.92·A + 5
Livingston-Kohlstadt (Livingston et al., 2005)	Any	F	248·W^0.4356 - (5.09·A)
		M	293·W^0.4330 - (5.92·A)
Kleiber (Kleiber, 1932)	Any	F	65.8·W^0.75
		M	71.2·W^0.75

2.2.3 Physical Fatigue Estimation and Suitable Parameters

The nature of muscle fatigue depends on the characteristics of exercise, i.e., its intensity and duration. Methods for quantifying fatigue include measurements of the drop in peak force, torque or power of muscle contraction, expressed as a "fatigue index", i.e., the percentage or rate of performance decrease over time (Finsterer et al., 2014). That fatigue index may be taken as a measure of resistance to fatigue and may be assessed using various ergometers. On an isokinetic dynamometer, fatigue resistance may be assessed: (a) by the number of maximum effort repetitions until exhaustion; (b) by the number of maximum effort repetitions until a 50% reduction in torque output is reached; (c) by the percent decline in torque from the beginning to the end of a predetermined time period (Kannus, 1994).

Fatigue index may also be assessed using: (a) maximal sprint cycling tests, such as the Wingate test; (b) by calculating the difference between the highest and lowest power output, expressed as a percentage of the highest power (Vanderwalle et al., 1987).

Other fatigue resistance assessment methods include: (a) measurement of the number of repetitions against a submaximal load during resistance exercise (Terzis et al., 2008; Mayhew et al., 2011); (b) measurement of time to exhaustion during steady or varying pace submaximal or maximal intensity running or cycling (Slawinski et al, 2005).

Several questionnaires have been developed for assessing exertion and fatigue. Borg Scale is a category scale which increases linearly with the exercise intensity for work on cycle ergometer (Borg, 1982). Another questionnaire-based tool has been developed for measuring stress tolerance in elite athletes (Rushall, 1990). The Profile of Mood States (McNair et al., 1971) has a subcomponent for assessing fatigue. The Multidimensional Fatigue Symptom Inventory (MFSI) and the short form (MFSI-SD) have demonstrated positive psychometric properties (Donovan et al., 2014). The Swedish Occupational Fatigue Inventory (SOFI) has been evaluated for physical fatigue by a study using cycle ergometer to induce fatigue with different workloads (Åhsberg et al., 1998). The scale items are scored based on a 7-point Likert scale to assess fatigue from 0 (not at all) to 6 (to a very high degree) (Figure 2).

	nothing to maximal						
1. Physical exertion (palpitations, sweaty, out of breath, breathing heavily)	0	1	2	3	4	5	6
2. Physical discomfort (tense muscles, numbness, stiff join, aching)	0	1	2	3	4	5	6
3. Lack of motivation (lack of concern, passive, indifferent, uninterested)	0	1	2	3	4	5	6
4. Sleepiness (falling asleep, drowsy, yawning, sleepy)	0	1	2	3	4	5	6
5. Lack of energy (worn out, spent, drained, overworked)	0	1	2	3	4	5	6
	Total Score: ___						

*Figure 2. The Swedish Occupational Fatigue Inventory (SOFI) short-form questionnaire adopted from (Åhsberg et al., 1998), which was used in **Publication III**.*

In addition to questionnaires, multiple other methods and parameters have been used for fatigue assessment (Table 3). The effect of fatigue on reaction time (RT) has been invested in several studies. RT has been found to be negatively affected by exhaustion when the participants have been prompted to perform physical tasks (Sant'Ana et al., 2017). The study with psychomotor vigilance tasks, using the developed PC-PVT platform (Khitrov et al., 2014; Reifman et al., 2018), also shows that RT has increases with the fatigue level (Thompson, 2019). It has been found that choice reaction time increases with exercise induced fatigue regardless of the type of the exercise (Sabzi, 2012).

Table 3. Various methods and parameters that have been used for fatigue assessment.

Method or parameter type	Related studies
Questionnaire	(McNair et al., 1971), (Borg, 1982), (Rushall, 1990), (Åhsberg et al., 1998), (Donovan et al., 2014)
Reaction time	(Sabzi, 2012), (Khitrov et al., 2014), (Sant'Ana et al., 2017), (Reifman et al., 2018), (Thompson, 2019)
Hand grip strength	(Thompson et al., 2017), (Thompson, 2019)
Countermovement jump	(Kennedy et al, 2017), (Wu et al., 2019), (Petrigna et al., 2019)
Heart rate	(Mohanvelu et al., 2017), (Thomson et al., 2016)
Heart rate variability	(Mohanvelu et al., 2017), (Peçanha et al., 2017)
Pulse arrival time	(Liu et al., 2011)

Study on work related fatigue has found hand grip strength to decrease after a rigorous work period (Thompson et al., 2017). While in a later study the same researcher has found hand grip to be the only studied variable not to have a significant decline following a multiple work shift (Thompson, 2019).

Practitioners often use the Countermovement jump (CMJ) test to monitor athlete fatigue, or equivalently recovery status, in terms of neuromuscular and/or metabolic fatigue. In multiple studies, the CMJ is used to characterize fatigue in functional lower body dynamic performance following acute training interventions or as a longitudinal monitoring tool (Wu et al., 2019). CMJ is routinely used in many sporting settings to provide a functional measure of neuromuscular fatigue and suitable testing methods have been described (Petrigna et al., 2019). However, the variables that are most sensitive to fatigue remain somewhat unclear (Kennedy et al, 2017).

Heart rate (HR) has been accepted by many researchers for the assessment of human fatigue (Mohanvelu et al., 2017). Heart rate is usually calculated based on measured ECG signals, where R-peaks are detected using Hamilton-Tompkins algorithm (Hamilton et al., 1986). Exercise induced physical fatigue has been found to increase the average HR, while also decreasing the change in HR when comparing the HR during the exercise with the resting HR (Thomson et al., 2016).

Heart rate variability (HRV) has been found to be inversely proportional to workload and has been used for assessment of human fatigue (Mohanvelu et al., 2017). Both time domain and frequency domain parameters have been used to assess HRV. Often used measures from time domain are SDNN (the standard deviation of all NN intervals) and RMSSD (the square root of the root mean square of the sum of all differences between successive NN intervals). From frequency domain the Low Frequency component (LF), High Frequency component (HF) and their ratio LF/HF have been used. Focusing on time-domain measures when developing for real-time wearable systems is more suitable in order to save on computational power. Work related fatigue study has found HRV parameter RMSSD to decrease with fatigue. In exercise induced fatigue studies HRV has been also found to decrease with exercise and the HRV analysis during the post-exercise period has been proposed to be a surrogate market of the cardiac autonomic recovery (Peçanha et al., 2017).

Pulse wave analysis is a novel method for assessing the cardiovascular health and artery stiffness. Pulse wave analysis has mostly been explored for continuous cuff-less blood pressure monitoring (Mukkamala et al., 2015; Muehlsteff et al., 2008). Pulse arrival time (PAT), which is a measure of pulse wave analysis, is defined as the time-delay between the R-peak of the QRS wave from the ECG and the arrival of the arterial pulse wave at the periphery (Muehlsteff et al., 2008). PAT is the sum of the pre-ejection period (that covers the iso-volumic ventricular contraction phase) and the pulse transit time (purely vascular component) which both represent different underlying cardiovascular mechanisms that could be affected by fatigue. To reduce the blood-pressure induced component in PAT values, it is important to normalize the values with respect to blood pressure (Mukkamala et al., 2015). Research literature shows the relation of every 1 mmHg difference in blood pressure causing 1 ms discrepancy in PAT (Muehlsteff et al., 2008). While prior studies have explored PAT in exercise settings (Liu et al., 2011), then to the best of the authors' knowledge, the study composed in **Publication III** is the first study to evaluate PAT for physical fatigue assessment with promising results.

2.2.4 Current State-of-the-Art Solutions and Possible Developments

The optimal management of fatigue-related risks in different settings requires the capacity to effectively monitor fatigue. Nowadays, the main topics in the study of fatigue are related to fatigue tests in different (work) settings, evaluation of muscular fatigue, subjective symptoms of fatigue, indicators of nervous strain, and the practical application of fatigue tests (Yu et al., 2019). An examination of prior measures suggests that a practical need exists for a new multidimensional measure of fatigue.

There is no single instrument which is used as a gold standard for fatigue measurement, because of the definitional difficulties, multiple causes of fatigue, considerable overlap between different categories of fatigue and their interaction between each other (Saito, 1999; Aaronson et al., 1999). Additionally, fatigue has several confounding factors such as medication, psychological and cognitive conditions, and deconditioning (Finsterer et al., 2014; Stadje et al., 2016).

The multi-factorial nature of fatigue suggests that a single universal test to measure fatigue may not exist (Saito, 1999). Fatigue assessment studies have usually compiled different test batteries of various measures (Hughes et al., 2019; Thompson et al., 2019). These measures can be classified into six different categories: (1) questionnaires or subjective feelings of fatigue, (2) psychological tests, (3) neuropsychological tests, (4) biochemical indexes, (5) physiological tests and (6) autonomic nervous function tests (Saito, 1999) (Table 4). Measures are often collected as part of a test battery which can be administered during work breaks and control for factors that may affect interpretation (e.g., muscle length, movement velocity, magnitude of exerted force).

Table 4. Fatigue measurement test categories and some of the commonly used tests in each category (Saito, 1999).

Fatigue measurement test category	Commonly used tests
Questionnaires on subjective feelings of fatigue	Various tests and questionnaires for subjective symptoms of fatigue
Psychological tests	Blocking test, Kraepelin Test, measurement of perception of time
Physiological tests	Muscular strength, respiratory and circulatory functions, heart rate, near point distance
Neurophysiological tests	Electroencephalography, sensory evoked response, reaction time, galvanic skin response, visual tracing reaction test
Autonomic nervous function tests	Adrenaline test, atropine test, cold pressor test, Czermak test
Biochemical blood and urinary indexes	Urinary excretion of protein, sugar, urobilinogen, creatinine etc. Eosinophilic leucocytes, total gravity of blood, hemoglobin content

It is important to note that most of these require specific administered tests and thus are not suitable to implement in real-time and in real-life physical fatigue assessment. For this reason, a new approach to the utilization of physiological signals (markers, measures) is needed. Continuous measurement during work activity, which might be advantageous in providing information representative of work, may also be disruptive to the work process. Test batteries quantify cumulative fatigue whereas continuous measurement might provide information directly representative of workload.

Previous studies have observed that fatigue development may be dependent on the task, more specifically the intensity, duration, muscle groups involved, and the type of contraction (Allen et al., 2008; Finsterer et al., 2014). Not all measures revealed increasing fatigue over the workday or over the workweek, which may be a result of fatigue measures reflecting different processes of fatigue. It appeared that measures reflecting central mechanisms were responsive within a workday, while measures reflecting both central and peripheral mechanisms were responsive over the workweek. Therefore, fatigue measures, reflecting changes to both central and peripheral processes, may be useful in measuring tasks and exercises of varying parameters.

It has been suggested that appropriate fatigue monitoring should include both objective and subjective measures (Aaronson et al., 1999). Since the existing fatigue tests fail to give the same results, it is essential that for the evaluation of fatigue, data obtained from a single fatigue test or a combination of fatigue tests having no correlation with each other must be considered with extreme care because in some cases the results will be useless.

For the above-mentioned reasons it would be a significant advancement if physical fatigue could be measured continuously and unobtrusively. This novel approach utilizing wearables could measure information continuously and give feedback in real-time. Thus, for this need, **Publication III** proposes a test-battery of cardiovascular parameters, which could be analyzed in real-time for continuous personalized feedback of physical fatigue.

3 Methods

In this chapter the study designs and experiments for achieving the aims of the thesis are introduced. Two different experimental setups are presented. The studies of human activity recognition (HAR) and energy expenditure (EE) shared one experimental setup where the data was jointly gathered for both studies. The experiments for the physical fatigue assessment study were conducted separately.

3.1 Study Design

This thesis aims to provide results for creating novel methods for physical fatigue assessment and human activity monitoring, which could be monitored in real-time and unobtrusively using wearable sensors and systems. Firstly, the thesis aimed to improve real-time HAR by optimizing the window length, sampling frequency and feature selection (**Publication I**). HAR is based on acceleration signals which are fragmented into shorter consecutive fragments based on the chosen window length. In a system where HAR is done in real-time higher window length also means the delay of the system is longer, since each classification is done after the signals have been collected for the whole window. The number of samples in the fragment is determined by both the window length and sampling frequency – lowering the sampling frequency also decreases the numbers of the samples in a classification fragment according to (1). In addition, multiple feature selection methods were used to decrease the initial classification feature set into smaller subsets in order to analyze how different number of features affect the HAR and what is the minimal number of features to use without compromising classification performance. These parameters also affect the power consumption and computational power that is required for HAR, which are both particularly limiting aspects when HAR is done using wearable systems and in real-time. While few previous studies have explored their effects separately, then this was the first study that thoroughly explored them in connection to each other. To fulfill the aim, a test study was conducted where subjects performed various physical activities while their body movement was measured and recorded using accelerometer.

Secondly, the thesis aimed to assess different basal metabolic rate (BMR) predictive equations to improve EE estimation (**Publication II**). Monitoring EE, which is an essential tool for assessing individuals' physical activity (PA) and adjusting nutritional supply, is moving towards activity specific models where first the activity is recognized using HAR and then suitable EE algorithm is applied for the specific activity. These algorithms rely on accurate assessment of basal metabolic rate (BMR), which is commonly estimated using predictive equations that use simple anthropometric variables such as the weight, height and gender of the person. For the purpose of this study the previous experiment also included calorimetry measurements in order to compare different predictive equations and validate their results with indirect calorimetry to choose the most suitable one for EE models.

Thirdly, the thesis aimed to propose a method for real-time PFA suitable for wearable systems by utilizing a set of real-time and easily measurable cardiovascular (CV) parameters (**Publication III**). There is no single instrument that can be applied as gold standard for fatigue measurement and many of the popular methods require special conditions and testing environment which makes them unsuitable for real-time assessment of physical fatigue. In this study it was hypothesized that a multi-parametrical model based on an enhanced test-battery of various CV parameters could yield an effective method for estimating physical fatigue possibly in real-time and with wearable systems. During the conducted experiment various CV parameters were measured in both rested-state and physically-fatigued-state of the test subjects to explore how they are affected by physical fatigue.

3.2 Study Groups

The HAR and EE study was conducted on one study group and the PFA study on a separate study group. In both experiments only healthy and active participants were included. The anthropometric parameters of the study groups are shown in Table 5. For HAR study a separate study group of 5 participants was used to measure the signals of outdoor cycling.

Table 5. Anthropometric parameters of the study groups in Experiment 1 (human activity recognition and energy expenditure studies) and Experiment 2 (physical fatigue assessment study).

Experiment #	Count	Age (years) mean±SD; range	Height (cm) mean±SD; range	Weight (kg) mean±SD; range	BMI (kg/m²) mean±SD; range
Experiment 1 HAR + EE	All (25)	32.0±8.8; 21–57	174.4±9.4; 158–193	73.5±10.5; 54–96	24.1±2.3; 20.0–29.4
	Females (13)	31.0±7.7; 21–45	167.4±5.8; 158–176	68.0±8.9; 54–82	24.0±2.5; 20.0–29.4
	Males (12)	32.8±10.0; 21–57	180.8±7.3; 167–193	78.6±9.5; 61–96	24.2±2.2; 21.6–27.6
Experiment 2 PFA	All (16)	28.3±7.9; 18–48	173.9±8.1; 163–190	69.9±12.3; 55–91	23.0±2.9; 18.3–30.1
	Females (8)	28.4±7.0; 18–42	169.1±5.9; 163–180	63.9±10.5; 55–89	22.4±3.5; 18.3–30.1
	Males (8)	28.3±9.2; 18–48	178.6±7.3; 166–190	75.9±11.4; 60–91	23.7±2.2; 20.4–26.4

The studies were conducted according to the guidelines of the Declaration of Helsinki and approved by the Tallinn Medical Research Ethics Committee (protocol no. 1954). Informed consent was obtained from each subject before participating in the study.

3.3 Experimental Setups

3.3.1 Activity Recognition and Energy Expenditure Estimation

The aim of the first experiment was to gather information for the HAR and the EE studies (**Publication I** and **Publication II**). Test subjects performed various physical activities during which acceleration signals were measured and recorded from the left wrist using the Shimmer3 sensor platform (Shimmer Research, Dublin, Ireland) (Figure 3). Each test subject conducted activities based on a precise schedule, where each activity was carried out for a fixed amount of time. Outdoor cycling signals were recorded after indoor measurements with a separate study group.

Figure 3. Shimmer3 IMU, which was used for measuring and recording acceleration signals from the left wrist. Signals were recorded with a dynamic range of ±16 g and numeric resolution of 16-bit with the built-in accelerometer STMicroelectronics LSM303AHTR (Shimmer Research).

The analysis was done based on the wide range accelerometer data with the dynamic range set to ±16 g. The wide range accelerometer uses LSM303AHTR sensor (STMicroelectronics, Geneva, Switzerland), which has a numeric resolution of 16-bit. Activities during which signals were measured are shown in Table 6. Acceleration was measured with a sampling rate of 512 Hz. During preprocessing these signals were resampled using MATLAB function resample, filtered with third order low-pass Butterworth IIR filter (passband and stopband edge frequencies and ripples were respectively 0.1 Hz and 0.5 Hz, and 1 dB and 20 dB) and fragmented into shorter consecutive fragments based on the window length. Following an initial set of 110 features were extracted, which was decreased using various feature selection methods. Decision tree based classifier was trained using MATLAB's function fitctree and the results were evaluated based on sensitivity and F1-score (Powers et al 2011).

Indirect calorimetry (IC) measurements were done using open-circuit indirect spirometry device MasterScreen CPX (CareFusion, Hoechberg, Germany) (Figure 4), which calculates EE based on Weir equation (Weir 1949). System was calibrated before each test subject. Since the IC device was not portable, the energy expenditure was only measured during "indoor test 2" and "indoor test 3", shown in Table 6. The predictive equations assessed in the study for BMR were Harris-Benedict (Harris et al., 1918), Schofield (Schofield 1985), FAO/WHO/UNU (World Health Organ 1985), Henry-Rees (Henry et al., 1991), and Kleiber (Kleiber 1932) and for RMR were Mifflin-St Jeor (Mifflin et al., 1990), Owen (Owen et al., 1986; Owen et al., 1987), Livingston-Kohlstadt (Livingston et al., 2005) (Table 2). The values achieved with RMR equations were divided by 1.1 to compare the results with BMR equations.

Figure 4. Carefusion Masterscreen CPX spirometry device. Alternative set-up with treadmill was used for indirect calorimetry measurements during "indoor test 2" and "indoor test 3". MasterScreen CPX allows for breath-by-breath indirect calorimetry measurements using Weir equation by measuring oxygen consumption and carbon dioxide production (CareFusion).

Table 6. Conducted activities and their duration in minutes during which acceleration (ACC) and energy expenditure (EE) was measured.

Indoor test 1 – ACC	Indoor test 2 – ACC + EE	Indoor test 3 – ACC + EE (% shows angle)	Outdoor test – ACC
Walking (3)	Sitting on chair (3)	Walking (3 km/h) (3)	Cycling (14)
Running (3)	Lying on bed (4)	Walking (5 km/h) (3)	Cycling uphill (4)
Walking upstairs (3)	Typing on computer (3)	Walking (3 km/h, 10%) (3)	Cycling downhill (1)
Walking downstairs (3)	Folding clothes (3)	Walking (5 km/h, 10%) (3)	
	Cleaning surface (3)	Running (6 km/h) (3)	
		Running (10 km/h) (3)	
		Running (12 km/h) (3)	
		Running (6 km/h, 10%) (3)	

3.3.1 Physical Fatigue Assessment

The aim of the second experiment was to gather information about cardiovascular (CV) and reference parameters during different physical fatigue states for the physical fatigue assessment study (**Publication III**). The experiment consisted of three main activities: rested-state (RS) measurements in the morning, physically-fatigued-state (PFS) measurements in the afternoon and a workout session in-between (Figure 5). The workout session lasted for about an hour and consisted of multiple sets of various exercises such as squats, burpees, sit-ups, push-ups, planks and jumping jacks to induce physical fatigue. The analyzed parameters were divided into reference parameters, that usually need administered tests and cannot be obtained in real-time, and cardiovascular (CV) parameters, that could be continuously monitored and measured. The selected reference parameters were the score of a fatigue questionnaire, reaction time (RT), hand grip strength and countermovement jump (CMJ) height. Evaluated CV parameters were heart rate (HR), measures of heart rate variability (HRV), and blood pressure normalized pulse arrival time (PAT).

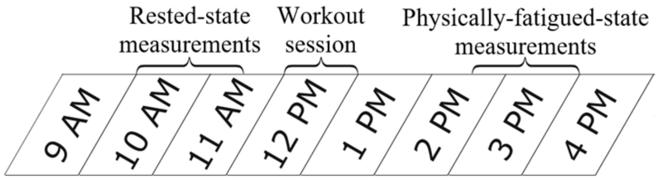

Figure 5. Overview of one experiment day. Cardiovascular and reference parameters were measured similarly in both measurement sets. The workout session consisted of multiple full-body exercises.

During the test both reference parameters and CV parameters were measured with multiple different systems and methods. At the start of the experiment the subjects were asked to complete a questionnaire to evaluate their current subjective fatigue level. The questionnaire adopted for the experiment was the Swedish Occupational Fatigue Inventory (SOFI), developed for the measurement of after-work fatigue (Åhsberg et al., 1998). The scale items were scored based on a 7-point Likert scale to assess fatigue from 0 (not at all) to 6 (to a very high degree). The scale items were as follows: (i) Physical Exertion, (ii) Physical Discomfort, (iii) Lack of Motivation, (iv) Sleepiness and (v) Lack of Energy.

Subject RT was measured using PC-PVT platform (Khitrov et al., 2014; Reifman et al., 2018) on a desktop computer (CPU: Intel Core i5-7500, GPU: Intel HD Graphics 630 (Intel, Santa Clara, California, USA), Mouse: Logitech G203 (Logitech, Lausanne, Switzerland)) with an external monitor (HP E233, Hewlett-Packard, Palo Alto, California, USA). The duration of the test was 5 minutes during which each participant performed about 75 simple RT measurements. The Inter-Stimulus Interval was selected between three to five seconds.

The hand grip strength was measured using the Grip Force Transducer (MLT004/ST, ADInstruments, Sydney, Australia) with PowerLab 4/25T (ADInstruments, Sydney, Australia) data acquisition device and LabChart software (v. 8.1.13, ADInstruments). The participants performed five maximal voluntary contractions with the dominant arm while seated. Hand grip strength was analyzed as the average of the maximums of the five repetitions.

Countermovement jump (CMJ) height was found based on the recording as the difference between standing position and highest point during the jump. The performance was filmed at 60 frames per second with a camera (OnePlus 6, OnePlus Technology, Shenzhen, China), which was statically mounted at a fixed distance. Each participant performed five maximal effort CMJ according to the recommended method (Petrigna et al., 2019) and the performance was assessed as the average jump height.

During the CV parameter measurements subjects performed a veloergometer test with alternating work (3 minutes) and recovery (5 minutes) phases. There were three work phases during which the subjects were asked to cycle respectively at three different power levels (60 W, 90 W and 120 W) while keeping the pace at 60 rotations per minute. The ECG signals were recorded at sampling rate of 1 kHz using PowerLab 4/25T (ADInstruments, Sydney, Australia) data acquisition device and LabChart software (v. 8.1.13, ADInstruments). HR and HRV parameters were calculated based on the R-peaks of the ECG signal, which were detected using the Hamilton-Tompkins algorithm (Hamilton et al., 1986). The assessed HRV parameters were SDNN (the standard deviation of all NN intervals) and RMSSD (the square root of the root mean square of the sum of all differences between successive NN intervals). PAT was found as the time difference between ECG R-peak and pulse wave signal rising front, which was registered using the same sensing unit with an external piezoelectric transducer attached to the fingertip (MLT 1010 pulse transducer, ADInstruments). Calculated PAT values were normalized based on blood pressure measurements to 120 mmHg using the relation of every 1 mmHg difference causing 1 ms discrepancy in PAT (Muehlsteff et al., 2008).

3.4 Chapter Summary

In this chapter the study designs and experimental setups are introduced and presented. The HAR and EE study were based on one experimental setup, where 25 test subjects performed various physical activities during which the accelerometer signals and indirect calorimetry values were measured. The accelerometer signals were preprocessed and used to train a decision tree based classifier to assess the effect of the classification window length, acceleration sampling frequency and different feature sets on the HAR classifier performance. Indirect calorimetry values were compared with multiple BMR and RMR predictive equations to choose the most suitable one for the EE models.

PFA was done based on a separate experimental setup, where multiple tests were conducted to measure the reference parameters (questionnaire score, reaction time, hand grip strength and countermovement jump) and cardiovascular parameters (heart rate, heart rate variability, blood pressure normalized pulse arrival time). Same measurements were conducted in the rested-state and physically-fatigued-state to propose a method for real-time PFA.

4 Results

In this chapter the results of the human activity recognition (HAR), energy expenditure (EE) and physical fatigue assessment (PFA) are presented. In HAR study the results with different classification window lengths, acceleration sampling frequencies, different feature sets and their combined effect are shown. In EE study the results of the multiple BMR and RMR predictive equations are compared to each other and with the indirect calorimetry results. For the PFA study the values for the reference parameters and the cardiovascular parameters in both the rested-state and the physically-fatigued-state are presented. In addition, a model for classifying between mildly fatigued and significantly fatigued states is proposed based on the two best performing cardiovascular measures.

4.1 Human Activity Recognition

The aim of the HAR study (**Publication I**) was to create an optimized physical activity classifier that would be suitable for implementation on real-time wearable systems. The focus was on testing various sampling frequencies, window lengths and number of features in order to reduce the power consumption, and to decrease the required memory buffer without compromising classification performance.

The classifier performance was evaluated using a leave-one-subject-out cross-validation scheme where each test subject's signals were classified with a classifier that was trained using the signals from all the other test subjects, which has also been previously used by other researchers (Moncada-Torres et al., 2014; Altini et al., 2012). The confusion matrix attained for one of the subjects is shown in Table 7.

Table 7. Confusion matrix of conducted activities vs classified activities based on all the segments from all the subjects (using all 110 features, 25 Hz sampling frequency and 3 s window length), where the results for each subject was found individually using a leave-one-subject-out cross-validation scheme. The activity types are Static (1), Low Intensity (2), Moderate Intensity (3), Rhythmical Intensity (4), Walking (5), Running (6) and Outdoor Cycling (7).

		Classified activity type							
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	Total
Conducted activity type	(1)	2897	68	4	1	21	6	3	3000
	(2)	43	835	9	11	1	0	1	900
	(3)	2	5	809	40	32	6	6	900
	(4)	3	2	47	797	29	10	12	900
	(5)	22	6	34	181	5943	111	3	6300
	(6)	1	0	21	16	171	4002	1	4212
	(7)	3	6	23	10	14	1	1083	1140
	Total	2971	922	947	1056	6211	4136	1109	17352

Sensitivity was chosen as a statistical measure to evaluate classification performance during feature selection. Sensitivity shows the proportion of true positives classified in relation to correct or real ones, i.e., true positives that are correctly identified. Classification results were evaluated using F1-score (also called F-score or F-measure), which is calculated as a harmonic mean of precision and sensitivity. While evaluating the results with different window lengths, sampling frequencies and number of features, F1-scores were calculated separately for each activity type. Additionally, an average F1-score for different parameter combinations was found as a mean of the activity type

F1-scores. A paired t-test ($p < 0.05$) was used to find statistical differences between the classification F1-scores of different activity types and averages while using different window lengths and sampling frequencies.

Overall average classification F1-score achieved was about 0.90 and depended on the used window length, sampling frequency, feature set and classified activity type. To best evaluate how each of these parameters affected the classifier, a method was employed where the parameter under focus was evaluated using different values while the classifier F1-scores were averaged over the other parameters. This allowed to individually evaluate the effect of the window length, sampling frequency and number of features on the classifier performance. Performance with different window lengths is shown in Figure 6, with different sampling frequencies in Figure 7 and with different feature numbers of Figure 8.

In a system with a physical activity classifier working in real time, the window length determines the delay of the system, since each classification is done after signals have been collected for a whole window. Window lengths of 5 s, 3 s, and 1 s were chosen to evaluate how different window lengths affect the classifier performance. Window lengths of 5 s and 3 s had similar results with the average F1-scores of 0.92 ±0.02 and 0.91 ±0.02, while the result with 1 s was 0.87 ±0.02 (Figure 6).

To test different sampling frequencies, the signals that were initially recorded with a sampling frequency of 512 Hz were later resampled using a MATLAB function resample. The classifier had similar average F1-score with 50 Hz (0.92 ±0.02) and 25 Hz (0.91 ±0.02), while the average F1-score with 13 Hz was lower (0.87 ±0.02) (Figure 7).

When using machine learning methods for HAR, the classifier training is done based on features that are extracted from signal fragments. The feature set has to capture specific and diverse information of posture and human motion to allow precise activity classification. Two different feature selection schemes were used to analyze how different number of features affects HAR and what is the minimal number of features to use without compromising classification performance. The feature sets of 110 features, 43 features, 28 features and 13 features, achieved with the first feature selection scheme, had similar average F1-scores between 0.89 and 0.90. The sequential forward selection (SFS) feature set had a slightly higher average F1-score of 0.92 ±0.03 (Figure 8).

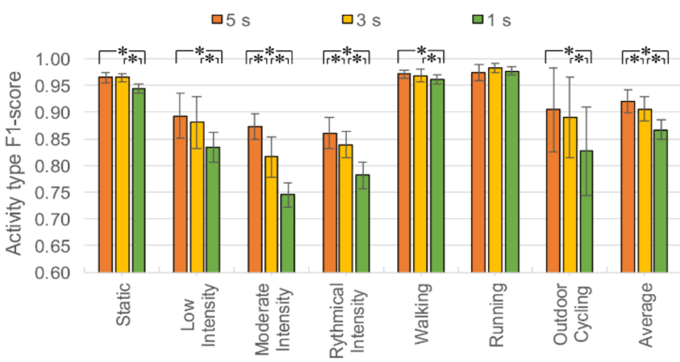

Figure 6. F1-scores of different activity types (mean ± SD) averaged over sampling frequencies and feature sets using different window lengths. Asterisks show significant statistical difference between different values of the window length (p < 0.05).

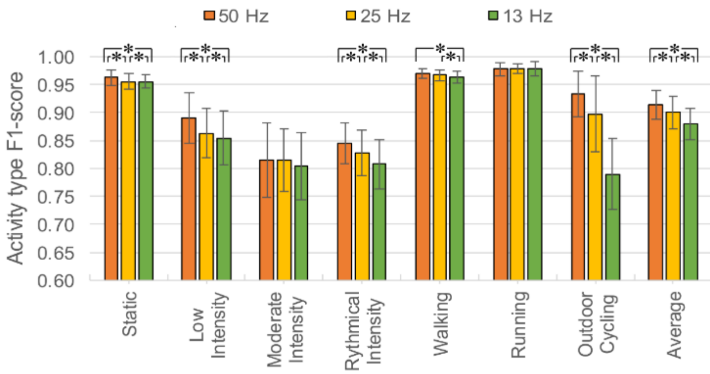

Figure 7. F1-scores of different activity types (mean ± SD) averaged over window lengths and feature sets using different sampling frequencies. Asterisk shows significant statistical difference between different values of the sampling frequency ($p < 0.05$).

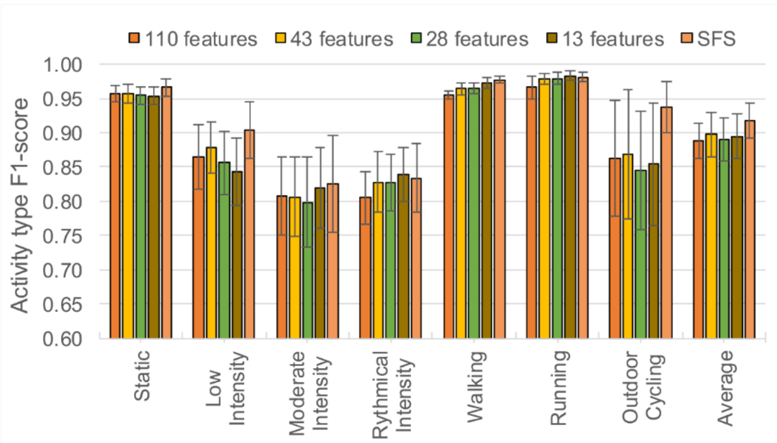

Figure 8. F1-scores of different activity types (mean ± SD) averaged over window lengths and sampling frequencies using different feature sets, where the SFS is the feature combination obtained with sequential forward selection.

Since both classification window length and sampling frequency of the acceleration signal affect the number of samples in classification fragment, it was also deemed important to evaluate their combined effect on the classification performance. Figure 9 shows the average classification F1-scores with different feature sets using different combinations of sampling frequencies and window lengths. The classification performance was better with combinations that had more samples per classification fragment, with the highest average of 0.93 ±0.05 achieved with the combination of 50 Hz and 5 s. The results with the combinations that had either 1 s window length or sampling frequency of 13 Hz were lower compared to other combinations with most feature sets. The SD values were large, since the results were averaged over different activity types with different F1-scores.

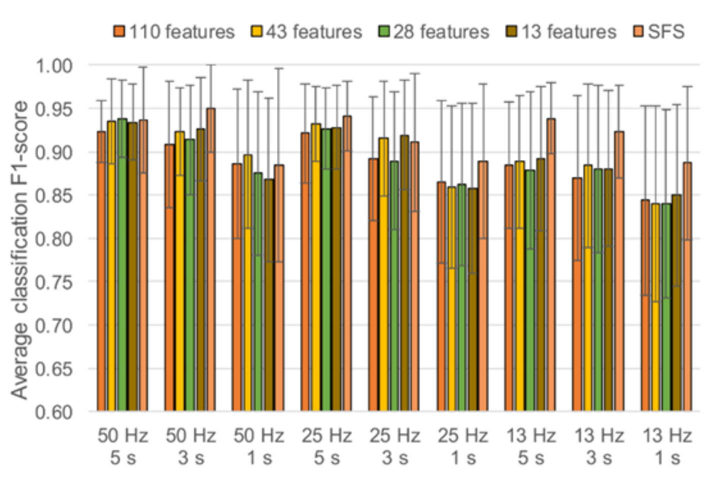

Figure 9. F1-scores (mean ± SD) averaged over all activities using different feature sets, window lengths and sampling frequencies, where the SFS is the feature combination obtained with sequential forward selection.

4.2 Physical Activity Intensity and Energy Expenditure Estimation

The aim of the EE study (**Publication II**) was to assess the basal metabolic rate (BMR) and resting metabolic rate (RMR) predictive equations (Table 2) by comparing the results of the different equations and validating them with indirect calorimetry (IC) values in order to choose the most suitable one for energy expenditure (EE) models.

The predictive equations explored in this study for BMR were Harris-Benedict, Schofield, FAO/WHO/UNU, Henry-Rees, and Kleiber; and for RMR were Mifflin-St Jeor, Owen and Livingston-Kohlstadt. The values achieved with RMR equations were divided by 1.1 to compare the results with BMR equations. Regression analysis was done in order to compare the different equations, assessed by the coefficient of determination (R^2). Based on the anthropometric data, BMR was calculated for each test subject using all equations. Using these results, R^2 was calculated for each pair of equations.

From the eight different BMR predictive equations explored in this study Mifflin-St Jeor formula had the best performance when compared to the results of the IC (Figure 10). Based on regression analysis most equations had similar results, with Owen and Kleiber formulas being the outliers (Table 8), which respectively had the lowest and highest average BMR results. The average BMR values with Mifflin-St Jeor formula (1447 ±204 kcal/day) were the closest with IC results (1485 ±255 kcal/day) and also had the lowest RMSE of 175 kcal/day compared to IC.

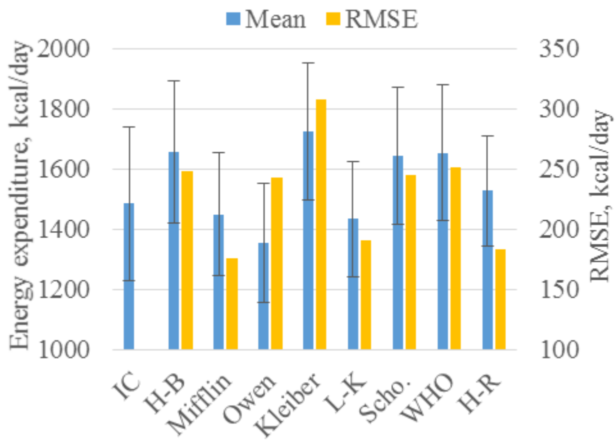

Figure 10. Average BMR (mean ± SD) with indirect calorimetry (IC) and predictive equations; RMSE of BMR between predictive equations and IC. H-B – Harris-Benedict; L-K – Livingston-Kohlstadt; Scho. – Schofield; WHO – FAO/WHO/UNU; H-R – Henry-Rees.

Table 8. Coefficient of determination R^2 between different predictive equations. H-B – Harris-Benedict; L-K – Livingston-Kohlstadt; Scho. – Schofield; WHO – FAO/WHO/UNU; H-R – Henry-Rees.

	Mifflin	Owen	Kleiber	L-K	Scho.	WHO	H-R
H-B	0.98	0.88	0.92	0.98	0.93	0.94	0.95
Mifflin		0.88	0.92	0.97	0.95	0.95	0.96
Owen			0.85	0.92	0.91	0.92	0.90
Kleiber				0.90	0.90	0.91	0.93
L-K					0.97	0.97	0.97
Scho.						1.00	0.98
WHO							0.99

4.3 Physical Fatigue Assessment

The aim of the physical fatigue assessment study (**Publication III**) was to propose a method for physical fatigue assessment employing a set of real-time measurable cardiovascular (CV) parameters. Measurements were conducted on 16 healthy participants (8 female) and consisted of a morning test set, physical exercise during the day and an afternoon test set. The analyzed cardiovascular parameters were heart rate (Mohanvelu et al., 2017; Thomson et al., 2016), heart rate variability measures SDNN (the standard deviation of all NN intervals) and RMSSD (the square root of the root mean square of the sum of all differences between successive NN intervals) (Mohanvelu et al., 2017; Shortz et al., 2017; Pecanha et al., 2017), and blood-pressure normalized pulse arrival time (PAT) (Mukkamala et al., 2015; Muehlsteff et al., 2008). The parameters were selected with the aim to keep the complexity of the overall measurement process and computational power requirements as low as possible for suitable use in wearable systems, and thus, only time-domain measures were considered.

For every test battery parameter the percentual change between the rested-state (RS) and the physically-fatigued-state (PFS) was found individually for each participant. The average reference and CV parameter values are shown in Table 9. A paired t-test ($p < 0.05$) was used to find statistical differences between the parameters. Statistically significant changes were found in the CMJ height and questionnaire score for the reference parameters and in the HR, SDNN and RMSSD for the CV parameters.

A linear correlation coefficient was calculated separately for each parameter pair to detect any linear relationship. Relatively strong linear correlation ($0.5 < r < -0.5$) was noted between several test battery measures. The correlation values are shown in Table 10.

Table 9. Average (mean±SD) reference and CV parameter values in the rested-state (RS), physically-fatigued-state (PFS) and their difference in percentage (DIF). Results are shown for all participants (A), female participants (F) and male participants (M). Reference parameters: Q – questionnaire, RT – reaction time, DYN – dynamometer hand grip force, CMJ – countermovement jump height; CV parameters: HR – average heart rate, SDNN – HRV parameter SDNN value, RMSSD – HRV parameter RMSSD, PAT – pulse arrival time. Values marked with asterisk (*) in bold showed statistical difference (paired t-test, P < 0.05).

			Q (%)	RT (ms)	DYN (N)	CMJ (cm)
Reference parameters	RS	A	14.0±7.6	208.7±11.3	360.3±99.1	38.2±8.7
		F	12.1±9.4	206.8±13.4	294.2±47.4	33.1±3.3
		M	15.8±5.3	210.6±9.4	426.4±93.8	43.3±9.7
	PFS	A	29.2±13.0	211.4±16.9	349.7±105.7	37.0±9.0
		F	30.0±17.7	211.7±17.2	286.4±48.7	31.6±3.3
		M	28.3±6.9	211.0±17.9	413.0±111.3	42.5±9.8
	DIF (%)	A	15.2%*	1.3%	-2.9%	-3.1%*
		F	17.9%*	2.4%	-2.7%	-4.5%*
		M	12.5%*	0.2%	-3.1%	-1.9%
			HR (bpm)	SDNN (ms)	RMSSD (ms)	PAT (ms)
CV parameters	RS	A	98.5±10.9	58.0±19.7	35.4±18.9	273.4±21.6
		F	100.6±9.7	52.8±13.3	31.7±12.7	267.5±15.7
		M	96.4±12.3	63.2±24.3	39.0±24.0	279.4±25.9
	PFS	A	107.9±12.2	45.7±15.9	25.0±13.8	268.1±23.8
		F	110.1±11.4	40.5±16.0	23.5±17.6	254.8±11.8
		M	105.6±13.4	50.8±15.1	26.5±9.5	281.3±26.0
	DIF (%)	A	9.5%*	-21.2%*	-29.3%*	-2.0%
		F	9.4%*	-23.2%*	-25.9%*	-4.7%
		M	9.6%*	-19.6%	-32.0%	0.7%

Table 10. Linear correlation coefficient values between different parameters based on all participants (A), male participants (M) and female participants (F). Parameter values are taken as the difference in % between the rested-state and physically-fatigued-state. Q – questionnaire, RT – reaction time, DYN – dynamometer hand grip force, CMJ – countermovement jump height, SDNN – HRV parameter SDNN, RMSSD – HRV parameter RMSSD, PAT – pulse arrival time, HR – average heart rate (between resting heart rate and average veloergometer cycling heart rate). Values above 0.5 or below -0.5 are marked in bold.

		RT	DYN	CMJ	SDNN	RMSSD	PAT	HR
Q	A	0.36	-0.18	-0.43	0.13	0.02	-0.05	-0.13
	F	0.36	**-0.55**	-0.35	0.10	-0.07	0.18	0.02
	M	0.36	**0.74**	**-0.59**	0.36	0.26	-0.36	-0.49
RT	A		-0.24	-0.25	0.11	-0.03	0.03	-0.33
	F		**-0.8**	-0.04	**0.56**	0.03	0.41	**-0.63**
	M		0.40	-0.38	-0.13	0.01	-0.22	-0.04
DYN	A			-0.10	-0.24	0.11	-0.39	-0.10
	F			0.30	**-0.74**	0.19	-0.18	0.22
	M			**-0.59**	0.23	0.12	**-0.80**	**-0.53**
CMJ	A				-0.09	0.03	0.35	-0.09
	F				-0.29	0.23	0.26	-0.35
	M				-0.08	-0.18	0.25	0.18
SDNN	A					**0.71**	0.30	**-0.61**
	F					0.10	0.24	-0.33
	M					**0.93**	0.23	**-0.88**
RMSSD	A						0.24	**-0.57**
	F						-0.12	-0.31
	M						0.42	**-0.80**
PAT	A							-0.35
	F							**-0.79**
	M							0.13

While all participants followed the same study protocol, they experienced different levels of physical fatigue based on their physiological and physical background. To distinguish between mildly fatigued and significantly fatigued, the participants were grouped into two groups based on their relative change in CMJ height between the rest-ed-state and the physically-fatigued-state. All CV parameters (and multiple sub-parameters) were analyzed one-by-one to test selectivity against the mildly fatigued and the significantly fatigued groups based on F-score and accuracy found using a leave-one-subject-out cross-validation scheme with a decision stump. The two best performing parameters were named (i) *SDNN_DIF_N_AVG* (relative change of the resting SDNN value normalized with the average recovery phase value between the RS and the PFS) and (ii) *PAT_PFS_N_MIN* (resting PAT value normalized with the lowest recovery phase value during the physically-fatigued-state). These two parameters were used to train a linear SVM (Figure 11), with a decision boundary formula:

$$-0.0261x_1 0.3366x_2 - 1.6558 = 0, \qquad (2)$$

where x_1 is the parameter *SDNN_DIF_N_AVG* and x_2 is the parameter *PAT_PFS_N_MIN*.

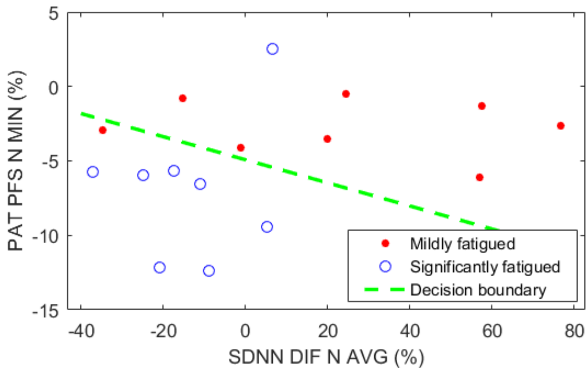

Figure 11. Linear SVM model for binary classification between the mildly fatigued and the significantly fatigued groups.

4.4 Chapter Summary

In this chapter the results of the studies are presented. For the HAR study, the classification performance with the different classification window lengths, acceleration sampling frequencies, different feature sets and their combined effect are shown respectively on Figures 6–8. Classification F1-scores with window lengths of 5 s and 3 s were similar, while results with 1 s were lower. All tested sampling frequencies performed similarly for most activity types, but the results with 13 Hz were considerably worse for the cycling activity. Initial set of 110 features were successfully decreased to 9–14 features without decreasing the classification performance.

For the EE study, the Figure 10 shows the energy expenditure values for the eight predictive equations explored in this study and their difference with the indirect calorimetry results. Based on the findings the Mifflin-St Jeor formula had the best performance – the average BMR values with Mifflin-St Jeor formula (1447 ±204 kcal/day) were the closest with IC results (1485 ±255 kcal/day) and also had the lowest RMSE of 175 kcal/day compared to IC.

In PFA study, the average values for the reference parameters and the CV parameters in the rested-state and the physically-fatigued-state are shown in Table 9. From the assessed cardiovascular parameters, the statistically significant change between the two states was noted in the average heart rate and heart rate variability measures SDNN and RMSSD. The correlations between different measures are shown in Table 10, strongest linear correlation was found between the reference parameter hand grip strength and CV parameter pulse arrival time. The two best performing CV parameters, which were based on heart rate variability and pulse arrival time, were used to create a linear SVM model presented on Figure 11 for classifying between the mildly fatigued and the significantly fatigued groups. There is one noticeable outlier shown in the figure. While this was not more thoroughly analyzed as part of this study, one potential reasoning could be considerably higher BMI of that subject (30.1 kg/m^2 vs an average of 23.0 ±2.9 kg/m^2).

5 Discussion

This thesis aims to provide results for creating or improving novel methods for physical fatigue assessment and human activity monitoring, which could be monitored in real-time and unobtrusively using wearable sensors and systems. This was done for three different topics: (i) optimization of the window length, sampling frequency and feature selection aimed to improve real-time human activity recognition (HAR) (**Publication I**), (ii) assessment of different basal metabolic rate (BMR) and resting metabolic rate (RMR) predictive equations to improve energy expenditure (EE) estimation (**Publication II**), (iii) assessing and proposing a method for teal-time physical fatigue assessment for wearable systems by utilizing a set of real-time and easily measurable cardiovascular parameters (**Publication III**). The main findings of these studies are shown in Table 11.

Table 11. The main findings of the human activity recognition (HAR), energy expenditure (EE) and physical fatigue assessment (PFA) studies.

Study	Major findings of the study
Optimization of parameters for real-time HAR (**Publication I**)	Classification performance with window lengths of 5 s and 3 s were similar, while results with 1 s were lower.
	Analyzed sampling frequencies performed similarly for most activity types, except for outdoor cycling, where 13 Hz was considerably worse.
	Similar or better results were achieved with the feature sets with 9 to 14 features, compared to the initial full feature set of 110 features.
EE with BMR ja RMR predictive equations (**Publication II**)	From the eight different BMR predictive equations explored Mifflin-St Jeor formula had the best performance.
	Most equations had similar results, with Owen and Kleiber formulas being the outliers, which respectively had the lowest and highest average BMR results.
	The average BMR values with Mifflin-St Jeor formula were the closest with IC results and had the lowest RMSE of 175 kcal/day compared to IC.
PFA with cardiovascular parameters (**Publication III**)	Statistically significant change between the rested state and physically-fatigued state was noted in the average heart rate and heart rate variability measures.
	The strongest linear correlation was found between the reference parameter hand grip strength and CV parameter pulse arrival time.
	The finest CV parameters for separating the mildly fatigued and significantly fatigued groups were based on heart rate variability and pulse arrival time.

In the HAR study (**Publication I**) it was analyzed for the first time how different window length, sampling frequency and feature set combinations affect the performance of physical recognition based on decision tree classifier in order to optimize the classifier for real-time wearable systems. The results of this study have been implemented into a smart work-wear prototype (Leier et al., 2018). The main findings were: 1) classification F1-scores with window lengths of 5 s and 3 s were similar, while results with 1 s were lower; 2) all sampling frequencies performed similarly for most activity types, with an exception of outdoor cycling; 3) Similar or better results were achieved with the feature sets with 9 to 14 features, achieved with either feature reduction scheme, compared to the initial full feature set of 110 features.

Window lengths of 5 s, 3 s and 1 s were used to analyze how different window lengths affect the performance of physical activity classifier. F1-scores of walking, running and low intensity activity types were similar with all window lengths, while the differences with moderate intensity, rhythmical intensity and outdoor cycling were larger. Even though window lengths between 3 s and 1 s have been found suitable for other studies (2.56 s in (Moncada-Torres et al., 2014), 2 s in (Loh et al., 2015), 1.5 s in (Aktaruzzaman et al., 2015), 1 s in (Bulling at al., 2014)), in this study the classifier performance had a larger drop when decreasing the classifier window down to 1 s while window lengths of 5 s and 3 s had similar results, The window length of 1 s had statistically significant differences with both 3 s and 5 s window lengths while classifying static, moderate intensity rhythmical intensity and outdoor cycling activity types. This could be caused by 1 s window length not being long enough to capture the movement of the body during activities where one period of movement exceeds the window length.

Different sampling frequencies of 50 Hz, 25 Hz and 13 Hz were used to investigate how sampling frequency affects classification performance. For most classified activity types no statistical differences were found between tested sampling frequencies, but there were large differences while classifying outdoor cycling. Previously it has been found that frequencies above 20 Hz cannot be expected to arise from voluntary human movement, where the accelerometer is not in contact with vibrating external sources (Bouten et al., 1997). It is likely that the 13 Hz sampling frequency was not high enough to capture the vibration during outdoor cycling.

A total of 110 features were extracted from acceleration signals for HAR. To reduce and optimize the number of features, two different feature selection schemes were used in this study. While the first scheme used different consecutive methods to reduce the number of features, the second scheme used forward SFS where features were added one-by-one. The first feature selection scheme enabled to reduce the feature set from 110 features to 13 features without decreasing the classifier performance. It is possible that the feature set with 13 features was overfit for the conditions used in this study and would perform worse in other conditions.

Compared to the feature sets of the first feature selection scheme, the SFS method used in the second scheme had higher performance with most window length and sampling frequency combinations. This difference was very noticeable when using the sampling frequency of 13 Hz. The number of features used in SFS feature sets was between 9 and 14. The large differences in average F1-scores shown in Figure 9 between SFS feature set and other feature sets while using sampling rates of 25 Hz and 13 Hz were mostly affected by outdoor cycling. Unlike other feature sets, SFS feature set had high F1-score while classifying outdoor cycling with all sampling frequency and window length combinations. The highest average classification F1 score was achieved with a parameter combination with SFS feature set (3 s window length, 50 Hz sampling frequency, 12 features), which also had the best performance while classifying static, low intensity, walking and outdoor cycling activity types. It was predictable that the SFS method would provide better results, since the SFS method chose the best features to maximize the classification sensitivity separately for each window length and sampling frequency combination, while with the first scheme features were selected based on one sampling frequency and window length combination. SFS method proved to be a simple comparison method for more comprehensive feature selection and showed that the effect of features depends on different classifier parameters, of which sampling frequency and window length were tested in this study.

Despite the recent advances in deep learning based activity recognition, which reduces the dependency on hand-crafted feature sets and thus could outperform more traditional machine learning methods, it is still far from being used in online mobile systems due to excessive computational power it requires (Wang et al., 2018). Thus the methods and results of this study provide useful information to other researchers for designing and implementing state-of-the-art physical activity recognition for real-time wearable systems.

From the eight different BMR predictive equations explored in the EE study (**Publication II**) Mifflin-St Jeor formula had the best performance. Based on regression analysis most equations had similar results, with Owen and Kleiber formulas being the outliers, which respectively had the lowest and highest average BMR results. The average BMR values with Mifflin-St Jeor formula (1447 ±204 kcal/day) were the closest with IC results (1485 ±255 kcal/day) and also had the lowest RMSE of 175 kcal/day compared to IC.

Mifflin-St Jeor equation has also been found as the most accurate and suitable for metabolic rate calculation in a comprehensive review study for both non-obese healthy adults and obese but otherwise healthy adults (Livingstone et al., 2005). In other experiments it has significantly overestimated RMR for underweight females (Aliasgharzadeh et al., 2015) or underestimated it for obese and overweight adults (Oliveira et al., 2011).

While the results of this study can be used to compare different predictive equations and for developing different EE models, there are some limitations which can be improved on in future studies. First, the test group in this study was very homogeneous since only healthy adults of same race were included, of which none were either obese or professional athletes. The findings of the study are suitable for calculating EE of BMR for a similar group, but might not be expandable for other groups. Secondly, the IC calorimetry tests conducted in this study were part of a larger experiment, which is why each position was held for a minimal amount of time needed to reach an EE plateau. Although lying down should have a higher EE than sitting or standing (Ainsworth et al., 2011), all positions were found to have similar EE levels in this study. The difference in EE between these positions could be more noticeable with a longer experiment time.

The physical fatigue assessment study (**Publication III**) study evaluated how exercise induced physical fatigue affects various test battery measures and whether real-time measurable cardiovascular (CV) parameters could provide sufficient information to classify between the mildly fatigued and significantly fatigued groups, aiming to provide information for real-time physical fatigue assessment. The main findings were: 1) from the assessed cardiovascular parameters, the statistically significant change between the rested-state and physically-fatigued-state was noted in the average heart rate and heart rate variability measures SDNN and RMSSD; 2) the strongest linear correlation was found between the reference parameter hand grip strength and CV parameter pulse arrival time (PAT); 3) the best CV parameters for separating the mildly fatigued and significantly fatigued groups were based on heart rate variability (HRV) parameter SDNN between the rested-state and the physically-fatigued-state and PAT changes during the physically-fatigued-state.

While most parameters were selected based on the findings of other studies, not all of them were found significant based on the results of this study. From the reference parameters, the score of fatigue questionnaire showed a statistically significant increase (of about 15.2%) between the rested-state and the physically-fatigue-state, which is consistent with previous findings (Dawson et al., 2011). Countermovement jump (CMJ) height decrease was statistically different for the whole group (average decrease of 3.1%) and the female subgroup, being in the same range as found in previous studies (Dawson et al., 2011; Thompson, 2019; Thompson et al, 2011). In accordance with the previous studies, the average value of reaction time (RT) increased (1.3%) (Thompson, 2019; Sabzi, 2012; Sant'Ana et al., 2017) and hand grip strength decreased (-2.9%) (Thompson, 2019; Thompson et al., 2017); however, the changes were not found statistically significant. The results of the hand grip strength test could be explained by the full-body workout regime that did not involve enough exercises for the specific arm muscles. It was expected that RT decreases due to physical fatigue (Allik et al, 2019); however, the present study did not reach such a result that can be explained by the different influence of the workout regime.

From the evaluated CV parameters, average heart rate had a statistically significant increase of 9.5%, which is in accordance with the previous studies (Mohanvelu et al., 2017; Thomson et al., 2016). The HRV parameters SDNN and RMSSD decreased respectively 21.2% and 29.3% between the rested-state and the physically-fatigued-state, which has also been noted by other researchers (Mohanvelu et al., 2017; Shortz et al., 2017; Pecanha et al., 2017). It was interesting to note, that PAT, which is a novel parameter for physical fatigue assessment studies, had a decrease of 2.0% for the whole group and 4.7% for the female subgroup, but for the male subgroup the value increased by 0.7%.

The linear correlation coefficient was found based on the relative individual changes between all measures. Strongest correlation between CV and reference parameter for the whole group was found between hand grip strength and PAT (linear correlation coefficient of -0.39). This finding was consistent with the male subgroup, where the linear correlation coefficient was -0.80. However, for the female subgroup the strongest correlation was found between HRV measure SDNN and hand grip strength (-0.79).

In total, 74 different sub-parameters were evaluated based on how well they classify between the mildly fatigued and the significantly fatigued groups that were compiled using the relative change in CMJ height value. These parameters were found using SDNN, RMSSD, PAT, and HR values from different veloergometer test phases and participant fatigue states. The best parameters for separating these groups were the relative change of the resting SDNN value normalized with average recovery phase value be-tween rested-state and physically-fatigued-state "*SDNN_DIF_N_AVG*" (F-score 0.842, accuracy 0.813) and the resting PAT value normalized with the lowest recovery phase value during the physically-fatigued-state "*PAT_PFS_N_MIN*" (F-score 0.875, accuracy 0.875). Simple linear support vector machine model was trained based on these two parameters to give an example of a possible use of these results. This model has a potential to reveal whether the user is "mildly fatigued" or "significantly fatigued" after a physically demanding day when implemented into a real-time monitoring system.

Based on the findings of this study, it was concluded that the test battery has added value for the assessment of physical fatigue. The evaluated CV parameters showed promising results compared to the reference parameters and thus could be used for real-time physical fatigue monitoring in workplace settings and for the general population. Furthermore, the novel parameters based on PAT were found to provide additional information and thus have the ability improve the overall quality of physical fatigue assessment.

Additional research is required to fully evaluate the utility of the test battery to determine the sensitivity of the variables with accumulation of fatigue. The authors also wish to point out several limitations that should be improved in the proceeding studies:

(i) the study was conducted on a relatively small number of participants, the results should be verified on a larger study group; (ii) induced physical fatigue could have been specific to the used workout and study protocol, it should be explored how different workout regimes affect the results; (iii) the experiments were conducted in lab settings and should be verified in real-life situations.

Conclusions

This thesis provides novel methods and knowledge for physical fatigue assessment and human activity monitoring, which could be applied in real-time and unobtrusively using wearable sensors and systems. In accordance with the aims of the thesis, the main efforts and the conclusions made were:

1. It was analyzed for the first time how different window length, sampling frequency and feature set combinations affect the performance of physical recognition based on decision tree classifier in order to optimize the classifier for real-time wearable systems (**Publication I**).
   Conclusion: Sampling frequency of 3 s and sampling frequency of 25 Hz were shown to be appropriate for activity recognition. Furthermore, using multiple feature selection methods, only 13 features from the initial 110 features were kept without decreasing the classifier performance.

2. Different basal metabolic rate predictive equations were compared and validated with indirect calorimetry data (**Publication II**).
   Conclusion: The most accurate predictive equation from the eight assessed formulas was the Mifflin-St Jeor equation, which had the most similar average energy expenditure and lowest root-mean-square error compared to the indirect calorimetry results.

3. It was evaluated how exercise induced physical fatigue affects various test battery measures and whether real-time measurable cardiovascular parameters could provide sufficient information to classify between different fatigue states (**Publication III**).
   Conclusion: Average heart rate and heart rate variability measures were best for classifying between fatigued and non-fatigued states. Additionally, the best parameters for separating the mildly fatigued and significantly fatigued groups were based on heart rate variability and pulse arrival time.

These results improve the current methods and provide important knowledge for real-time physical fatigue estimation. Improvements in human activity recognition (**Publication I**) can be used indirectly by improving the energy expenditure estimation by employing activity-specific models and directly by automatically detecting active, resting and recovery states to be used in fatigue estimation methods. More precise energy expenditure (**Publication II**) can be used as an input variable or to validate physical fatigue assessment methods. This information in conjunction with the validated fatigue assessment test battery measures (**Publication III**) is a strong basis for a physical fatigue assessment and human monitoring system, which works unobtrusively and continuously in real-time by using wearable sensors and systems.

References

Aaronson, L. S., Teel, C. S., Cassmeyer, V., Neuberger, G. B., Pallikkathayil, L., Pierce, J., Press, A. N., Williams, P.D., & Wingate, A. (1999). Defining and measuring fatigue. *Image J Nurs Sch, 31*(1), 45–50. doi: 10.1111/j.1547-5069.1999.tb00420.x

Ackerman, P. L., & Kanfer, R. (2009). Test length and cognitive fatigue: An empirical examination of effects on performance and test-taker reactions. *J Exp Psychol Appl, 15*(2), 163–181. doi: 10.1037/a0015719

Åhsberg, E., & Gamberale, F. (1998). Perceived fatigue during physical work: an experimental evaluation of a fatigue inventory. *Int J Ind Ergon, 21*(2), 117–131. doi: 10.1016/S0169-8141(96)00071-6

Ainsworth, B.E., Haskell, W.L., Herrmann, S.D., Meckes, N., Bassett, D.R. Jr., Tudor-Locke, C., Greer, J. L., Vezina, J., Whitt-Glover, M. C., & Leon, A. S. (2011) Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc, 43*(8), 1575–1581. doi: 10.1249/MSS.0b013e31821ece12

Aktaruzzaman, M., Scarabottolo, N., & Sassi, R. (2015). Parametric estimation of sample entropy for physical activity recognition. *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, 470–473. doi: 10.1109/embc.2015.7318401

Aliasgharzadeh, S., Mahdavi, R., Asghari Jafarabadi, M., & Namazi, N. (2015) Comparison of Indirect Calorimetry and Predictive Equations in Estimating Resting Metabolic Rate in Underweight Females. *Iran J Public Health, 44*(6), 822–829.

Allen, D. G., Lamb, G. D., & Westerblad, H. (2008). Skeletal Muscle Fatigue: Cellular Mechanisms. *Physiol Rev, 88*(1), 287–332. doi: 10.1152/physrev.00015.2007

Allik, A., Pilt, K., Viigimäe, M., & Fridolin, I. (2019). Pilot Study for Estimating Physical Fatigue Based on Heart Rate Variability and Reaction Time. Proceedings of the 2019 15th Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON 2019), Coimbra, Portugal, 193–200. doi: https://doi.org/10.1007/978-3-030-31635-8_23

Altini, M., Penders, J., & Amft, O. (2012). Energy expenditure estimation using wearable sensors. *Proceedings of the Conference on Wireless Health (WH '12),* New York, NY, United States. doi.org/10.1145/2448096.2448097

Altini, M., Penders, J., Vullers, R., & Amft, O. (2015). Estimating Energy Expenditure Using Body-Worn Accelerometers: A Comparison of Methods, Sensors Number and Positioning. *IEEE J Biomed Health Inform, 19*(1), 219–226. doi: 10.1109/jbhi.2014.2313039

Altun, K., Barshan, B., & Tunçel, O. (2010). Comparative study on classifying human activities with miniature inertial and magnetic sensors. *Pattern Recognit, 43*(10), 3605–3620. doi: 10.1016/j.patcog.2010.04.019

Amaro-Gahete, F., Jurado-Fasoli, L., De-la-O, A., Gutierrez, N., Castillo, M., & Ruiz, J. (2018). Accuracy and Validity of Resting Energy Expenditure Predictive Equations in Middle-Aged Adults. *Nutrients, 10*(11), 1635. doi: 10.3390/nu10111635

Amor, J. D., Hattersley, J. G., Barber, T. M., & James, C. J. (2015). Characterization of wrist-wearable activity measurement using whole body calorimetry in semi-free living conditions. *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* Milan, Italy, 3735–3738. doi: 10.1109/embc.2015.7319205

Anwer, S., Li, H., Antwi-Afari, M. F., Umer, W., & Wong, A. Y. L. (2021). Evaluation of Physiological Metrics as Real-Time Measurement of Physical Fatigue in Construction Workers: State-of-the-Art Review. *J Constr Eng Manag, 147*(5). doi: 10.1061/(asce)co.1943-7862.0002038

Arshad, M. H., Bilal, M., & Gani, A. (2022). Human Activity Recognition: Review, Taxonomy and Open Challenges. *Sens, 22*(17), 6463. doi: 10.3390/s22176463

Awais, M., Mellone, S., & Chiari, L. (2015). Physical activity classification meets daily life: Review on existing methodologies and open challenges. *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* Milan, Italy, 5050–5053. doi: 10.1109/embc.2015.7319526

Behm, D. G. (2004). Force Maintenance With Submaximal Fatiguing Contractions. *Can J Appl Physiol, 29*(3), 274–290. doi: 10.1139/h04-019

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Med Sci Sports Exerc, 14*(5), 377–381.

Bouten, C., Koekkoek, K., Verduin, M., Kodde, R., & Janssen, J. (1997). A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Trans Biomed, 44*(3), 136–147. doi: 10.1109/10.554760

Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. *ACM Comput Surv, 46*(3), 1–33. doi: 10.1145/2499621

Castro-García, J. A., Molina-Cantero, A. J., Gómez-González, I. M., Lafuente-Arroyo, S., & Merino-Monge, M. (2022). Towards Human Stress and Activity Recognition: A Review and a First Approach Based on Low-Cost Wearables. *Electronics, 11*(1), 155. doi: 10.3390/electronics11010155

Chowdhury, A. K., Tjondronegoro, D., Chandran, V., & Trost, S. G. (2018). Physical Activity Recognition Using Posterior-Adapted Class-Based Fusion of Multiaccelerometer Data. *IEEE J Biomed Health Inform, 22*(3), 678–685. doi: 10.1109/jbhi.2017.2705036

Cunningham, J. J. (1991). Body composition as a determinant of energy expenditure: a synthetic review and a proposed general prediction equation. *Am J Clin Nutr, 54*(6), 963–969. doi: 10.1093/ajcn/54.6.963

Daly, J. M., Heymsfield, S. B., Head, C. A., Harvey, L. P., Nixon, D. W., Katzeff, H., & Grossman, G. D. (1985). Human energy requirements: overestimation by widely used prediction equation. *Am J Clin Nutr, 42*(6), 1170–1174. doi: 10.1093/ajcn/42.6.1170

Dawson, D., Ian Noy, Y., Härmä, M., Åkerstedt, T., & Belenky, G. (2011). Modelling fatigue and the use of fatigue models in work settings. *Accid Anal Prevention, 43*(2), 549–564. doi: 10.1016/j.aap.2009.12.030

Demerouti, E., Bakker, A. B., Vardakou, I., & Kantas, A. (2003). The convergent validity of two burnout instruments: A multitrait-multimethod analysis. *Eur J Psychol Assess*, 19(1), 12–23. doi: 10.1027/1015-5759.19.1.12

Eaton, S., & Eaton, S. B. (2003). An evolutionary perspective on human physical activity: implications for health. *Comp Biochem Physiol A Mol Integr Physiol, 136*(1), 153–159. doi: 10.1016/s1095-6433(03)00208-3

Farrahi, V., Niemelä, M., Kangas, M., Korpelainen, R., & Jämsä, T. (2019). Calibration and validation of accelerometer-based activity monitors: A systematic review of machine-learning approaches. *Gait Posture, 68*, 285–299. doi: 10.1016/j.gaitpost.2018.12.003

Fridriksdottir, E., & Bonomi, A. G. (2020). Accelerometer-Based Human Activity Recognition for Patient Monitoring Using a Deep Neural Network. *Sens, 20*(22), 6424. doi: 10.3390/s20226424

Frone, M. R., & Tidwell, M. C. O. (2015). The meaning and measurement of work fatigue: Development and evaluation of the Three-Dimensional Work Fatigue Inventory (3D-WFI). *J Occup Health Psychol, 20*(3), 273–288. doi: 10.1037/a0038700

Feng, Z., Mo, Lingfei., & Li, Meng. (2015). A Random Forest-based ensemble method for activity recognition. *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* Milan, Italy, 5074–5077. doi: 10.1109/embc.2015.7319532

Finsterer, J., & Mahjoub, S. Z. (2013). Fatigue in Healthy and Diseased Individuals. Am J *Hosp Palliat Care, 31*(5), 562–575. doi: 10.1177/1049909113494748

Frankenfield, D., Roth-Yousey, L., & Compher, C. (2005). Comparison of Predictive Equations for Resting Metabolic Rate in Healthy Nonobese and Obese Adults: A Systematic Review. *J Am Diet Assoc, 105*(5), 775–789. doi: 10.1016/j.jada.2005.02.005

Hamilton, P. S., & Tompkins, W. J. (1986). Quantitative Investigation of QRS Detection Rules Using the MIT/BIH Arrhythmia Database. *IEEE Trans Biomed, BME-33*(12), 1157–1165. doi: 10.1109/tbme.1986.325695

Harris, J. A., & Benedict, F. G. (1918). A Biometric Study of Human Basal Metabolism. *PNAS, 4*(12), 370–373. doi: 10.1073/pnas.4.12.370

Henry, C. J., & Rees, D. G. (1991). New predictive equations for the estimation of basal metabolic rate in tropical peoples. Eur J Clin Nutr, 45(4), 177–185.

Hiilloskorpi, H., Fogelholm, M., Laukkanen, R., Pasanen, M., Oja, P., Mänttäri, A., & Natri, A. (1999). Factors Affecting the Relation Between Heart Rate and Energy Expenditure During Exercise. *Int J Sports Med, 20*(7), 438–443. doi: 10.1055/s-1999-8829

Hills, A. P., Mokhtar, N., & Byrne, N. M. (2014). Assessment of Physical Activity and Energy Expenditure: An Overview of Objective Measures. *Front Nutr, 1*. doi: 10.3389/fnut.2014.00005

Hilty, D. M., Armstrong, C. M., Luxton, D. D., Gentry, M. T., & Krupinski, E. A. (2021). A Scoping Review of Sensors, Wearables, and Remote Monitoring For Behavioral Health: Uses, Outcomes, Clinical Competencies, and Research Directions. *J Behav Sci, 6*(2), 278–313. doi: 10.1007/s41347-021-00199-2

Hughes, S., Chapman, D. W., Haff, G. G., & Nimphius, S. (2019). The use of a functional test battery as a non-invasive method of fatigue assessment. *PLOS ONE, 14*(2). doi: 10.1371/journal.pone.0212870

Kannus, P. (1994). Isokinetic Evaluation of Muscular Performance. *Int J Sports Med, 15*(S1), S11–S18. doi: 10.1055/s-2007-1021104

Kennedy, R. A., & Drake, D. (2017). The effect of acute fatigue on countermovement jump performance in rugby union players during preseason. *J Sports Med Phys Fitness, 57*(10). doi: 10.23736/s0022-4707.17.06848-7

Kent-Braun, J. A., Fitts, R. H., & Christie, A. (2012). Skeletal Muscle Fatigue. Compr Physiol, *2*(2), 997–1044. doi: 10.1002/cphy.c110029

Khakurel, J., Melkas, H., & Porras, J. (2018). Tapping into the wearable device revolution in the work environment: a systematic review. *Inf Technol People, 31*(3), 791–818. doi: 10.1108/itp-03-2017-0076

Khitrov, M. Y., Laxminarayan, S., Thorsley, D., Ramakrishnan, S., Rajaraman, S., Wesensten, N. J., & Reifman, J. (2013). PC-PVT: A platform for psychomotor vigilance task testing, analysis, and prediction. *Behav Res Methods, 46*(1), 140–147. doi: 10.3758/s13428-013-0339-9

Khusainov, R., Azzi, D., Achumba, I., & Bersch, S. (2013). Real-Time Human Ambulation, Activity, and Physiological Monitoring: Taxonomy of Issues, Techniques, Applications, Challenges and Limitations. *Sens, 13*(10), 12852–12902. doi: 10.3390/s131012852

Kleiber, M. (1932). Body size and metabolism. *Hilgardia, 6*(11), 315–353. doi: 10.3733/hilg.v06n11p315

Kumari, P., Mathew, L., & Syal, P. (2017). Increasing trend of wearables and multimodal interface for human activity monitoring: A review. *Biosens Bioelectron, 90*, 298–307. doi: 10.1016/j.bios.2016.12.001

Layzer, R. B. (1990). Muscle metabolism during fatigue and work. *Baillieres Clin Endocrinol Metab, 4*(3), 441–459. doi: 10.1016/s0950-351x(05)80064-3

Larsen, R. T., Christensen, J., Juhl, C. B., Andersen, H. B., & Langberg, H. (2019). Physical activity monitors to enhance amount of physical activity in older adults – a systematic review and meta-analysis. *Eur Rev Aging Phys Act, 16*(1). https://doi.org/10.1186/s11556-019-0213-6

Lee, I.-M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. *Lancet, 380*(9838), 219–229. doi: 10.1016/s0140-6736(12)61031-9

Lee, J., & Kim, J. (2016). Energy-Efficient Real-Time Human Activity Recognition on Smart Mobile Devices. *Mob Inf Syst,* 1–12. doi: 10.1155/2016/2316757

Lee, K., & Kwan, M. P. (2018). Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results. *Comput Environ Urban Syst, 67*, 124–131. doi: 10.1016/j.compenvurbsys.2017.09.012

Leier, M., Jervan, G., Allik, A., Pilt, K., Karai, D., & Fridolin, I. (2018). Fall detection and activity recognition system for usage in smart work-wear. *Proceedings of the 2018 16th Biennial Baltic Electronics Conference (BEC),* Tallinn, Estonia. doi: 10.1109/bec.2018.8600959

Lerman, S. E., Eskin, E., Flower, D. J., George, E. C., Gerson, B., Hartenbaum, N., Hursh, S. R., & Moore-Ede, M. (2012). Fatigue Risk Management in the Workplace. *J Occup Environ, 54*(2), 231–258. doi: doi.org/10.1097/jom.0b013e318247a3b0

Lin, Y., Becker, E., Park, K., Le, Z., & Makedon, F. (2009). Decision making in assistive environments using multimodal observations. *Proceedings of the 2nd International Conference on Pervasive Technologies Related to Assistive Environments (PETRA 2009),* Corfu, Greece. doi: 10.1145/1579114.1579120

Liu, Q., Poon, C., & Zhang, Y. (2011). Time–frequency analysis of variabilities of heart rate, systolic blood pressure and pulse transit time before and after exercise using the recursive autoregressive model. *Biomed Signal Process Control, 6*(4), 364–369. doi: 10.1016/j.bspc.2011.03.009

Liu, S., Gao, R. X., & Freedson, P. S. (2012). Computational Methods for Estimating Energy Expenditure in Human Physical Activities. *Med Sci Sports Exerc, 44*(11), 2138–2146. doi: 10.1249/mss.0b013e31825e825a

Livingstone, M. B. E. (1997). Heart-rate monitoring: the answer for assessing energy expenditure and physical activity in population studies? *Br J Nutr, 78*(6), 869–871. doi: 10.1079/bjn19970205

Livingston, E. H., & Kohlstadt, I. (2005). Simplified Resting Metabolic Rate-Predicting Formulas for Normal-Sized and Obese Individuals. *Obes Res, 13*(7), 1255–1262. doi: 10.1038/oby.2005.149

Loh, D., Lee, T. J., Zihajehzadeh, S., Hoskinson, R., & Park, E. J. (2015). Fitness activity classification by using multiclass support vector machines on head-worn sensors. Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 502–505. doi: 10.1109/embc.2015.7318409

Lu, K., Yang, L., Seoane, F., Abtahi, F., Forsman, M., & Lindecrantz, K. (2018). Fusion of Heart Rate, Respiration and Motion Measurements from a Wearable Sensor System to Enhance Energy Expenditure Estimation. *Sensors, 18*(9). doi: 10.3390/s18093092

Luke, A., Maki, K. C., Barkey, N., Cooper, R., & McGee, D. (1997). Simultaneous monitoring of heart rate and motion to assess energy expenditure. *Med Sci Sports Exerc, 29*(1), 144–148. doi: 10.1097/00005768-199701000-00021

Mayhew, J. L., Brechue, W. F., Smith, A. E., Kemmler, W., Lauber, D., & Koch, A. J. (2011). Impact of Testing Strategy on Expression of Upper-Body Work Capacity and One-Repetition Maximum Prediction after Resistance Training in College-Aged Men and Women. *J Strength Cond Res, 25*(10), 2796–2807. doi: 10.1519/jsc.0b013e31822dcea0

Marra, M., Cioffi, I., Morlino, D., Vincenzo, O. D., Pagano, M. C., Imperatore, N., Alfonsi, L., Santarpia, L., Castiglione, F., Scalfi, L., & Pasanisi, F. (2020). New Predictive Equations for Estimating Resting Energy Expenditure in Adults With Crohn's Disease. *JPEN J Parenter Enteral Nutr, 44*(6), 1021–1028. doi: 10.1002/jpen.1790

Megaw, E. D. (1995) The definition and measurement of visual fatigue. *In J. R. Wilson & E. N. Corlett (Eds.), Evaluation of human work: A practical ergonomics methodology, 2nd Edition*, 840–863. Taylor & Francis.

McNair, D. M., Lorr, M., & Droppleman, L. F (1971) Profile of mood states.

Mifflin, M. D., St Jeor, S. T., Hill, L. A., Scott, B. J., Daugherty, S. A., & Koh, Y. O. (1990). A new predictive equation for resting energy expenditure in healthy individuals. *Am J Clin Nutr, 51*(2), 241–247. doi: 10.1093/ajcn/51.2.241

Mohanavelu, K., Lamshe, R., Poonguzhali, S., Adalarasu, K., & Jagannath, M. (2017). Assessment of Human Fatigue during Physical Performance using Physiological Signals: A Review. *Biomed Pharmacol J, 10*(4), 1887–1896. doi: 10.13005/bpj/1308

Moncada-Torres, A., Leuenberger, K., Gonzenbach, R., Luft, A., & Gassert, R. (2014). Activity classification based on inertial and barometric pressure sensors at different anatomical locations. *Physiol Meas, 35*(7), 1245–1263. doi: 10.1088/0967-3334/35/7/1245

Muehlsteff, J., Aubert, X. A., & Morren, G. (2008). Continuous cuff-less blood pressure monitoring based on the pulse arrival time approach: The impact of posture. *Proceedings of the 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, Vancouver, BC, Canada, 1691–1694. doi: 10.1109/iembs.2008.4649501

Mukkamala, R., Hahn, J. O., Inan, O. T., Mestha, L. K., Kim, C. S., Toreyin, H., & Kyal, S. (2015). Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice. *IEEE Trans Biomed Eng, 62*(8), 1879–1901. doi: 10.1109/tbme.2015.2441951

Noakes, T. D. (2012). Fatigue is a Brain-Derived Emotion that Regulates the Exercise Behavior to Ensure the Protection of Whole Body Homeostasis. *Front Physiol, 3*. doi: 10.3389/fphys.2012.00082

Nweke, H. F., Teh, Y. W., Mujtaba, G., & Al-garadi, M. A. (2019). Data Fusion and multiple classifier systems for human activity detection and Health Monitoring: Review and Open Research Directions. *Inf Fusion, 46*, 147–170. Doi: 10.1016/j.inffus.2018.06.002

de Oliveira, E. P., Orsatti, F. L., Teixeira, O., Maestá, N., & Burini, R. C. (2011). Comparison of Predictive Equations for Resting Energy Expenditure in Overweight and Obese Adults. *J Obes, 2011*, 1–5. doi: 10.1155/2011/534714

Overton, A. J. (2013) Neuromuscular Fatigue and Biomechanical Alterations during High-Intensity, Constant-Load Cycling [Doctoral dissertation, Edith Cowan University].

Owen, O. E., Kavle, E., Owen, R. S., Polansky, M., Caprio, S., Mozzoli, M. A., Kendrick, Z. V., Bushman, M. C., & Boden, G. (1986). A reappraisal of caloric requirements in healthy women. *Am J Clin Nutr, 44*(1), 1–19. doi: 10.1093/ajcn/44.1.1

Owen, O. E., Holup, J. L., D'Alessio, D. A., Craig, E. S., Polansky, M., Smalley, K. J., Kavle, E. C., Bushman, M. C., Owen, L. R., & Mozzoli, M. A. (1987). A reappraisal of the caloric requirements of men. *Am J Clin Nutr, 46*(6), 875–885. doi: 10.1093/ajcn/46.6.875

Peçanha, T., Bartels, R., Brito, L. C., Paula-Ribeiro, M., Oliveira, R. S., & Goldberger, J. J. (2017). Methods of assessment of the post-exercise cardiac autonomic recovery: A methodological review. *Int J Cardiol, 227*, 795–802. doi: 10.1016/j.ijcard.2016.10.057

Petrigna, L., Karsten, B., Marcolin, G., Paoli, A., D'Antona, G., Palma, A., & Bianco, A. (2019). A Review of Countermovement and Squat Jump Testing Methods in the Context of Public Health Examination in Adolescence: Reliability and Feasibility of Current Testing Procedures. *Front Physiol, 10*. doi: 10.3389/fphys.2019.01384

Pillsbury, W. B. (2009). The Fundamentals of Psychology [1917]. Cornell University Library.

Pinheiro Volp, A. C., Esteves de Oliveira, F. C., Duarte Moreira Alves, R., Esteves, E. A., & Bressan, J. (2011). Energy expenditure: components and evaluation methods. *Nutr Hosp, 26*(3), 430–440. doi: 10.1590/S0212-16112011000300002

Poehlman, E. T. (1989) A review: exercise and its influence on resting energy metabolism in man. *Med Sci Sports Exerc, 21*(5), 515–525.

Powers, D. (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *J Mach Learn Technol, 2*(1), 37–63. doi: 10.48550/arXiv.2010.16061

Preece, S. J., Goulermas, J. Y., Kenney, L. P. J., & Howard, D. (2009). A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data. *IEEE Trans Biomed, 56*(3), 871–879. doi: 10.1109/tbme.2008.2006190

Qi, J., Yang, P., Waraich, A., Deng, Z., Zhao, Y., & Yang, Y. (2018). Examining sensor-based physical activity recognition and monitoring for healthcare using Internet of Things: A systematic review. *J Bimed Inform, 87*, 138–153. doi: 10.1016/j.jbi.2018.09.002

Racette, S. B., Das, S. K., Bhapkar, M., Hadley, E. C., Roberts, S. B., Ravussin, E., Pieper, C., DeLany, J. P., Kraus, W. E., Rochon, J., & Redman, L. M. (2012). Approaches for quantifying energy intake and %calorie restriction during calorie restriction interventions in humans: the multicenter CALERIE study. *Am J Physiol Endocrinol Metab, 302*(4), E441–E448. doi: 10.1152/ajpendo.00290.2011

Reifman, J., Kumar, K., Khitrov, M. Y., Liu, J., & Ramakrishnan, S. (2018). PC-PVT 2.0: An updated platform for psychomotor vigilance task testing, analysis, prediction, and visualization. *J Neurosci Methods, 304*, 39–45. doi: 10.1016/j.jneumeth.2018.04.007

Rushall, B. S. (1990). A tool for measuring stress tolerance in elite athletes. *J Appl Sport Psychol, 2*(1), 51–66. doi: 10.1080/10413209008406420

Sabzi, A. H. (2012). The effect of different fatigue protocols on choice reaction time. Middle East J Sci Res, 12(8), 1092–1096. doi: 10.5829/idosi.mejsr.2012.12.8.7113

SAITO, K. (1999). Measurement of Fatigue in Industries. *Ind Health, 37*(2), 134–142. doi: 10.2486/indhealth.37.134

Sant'Ana, J., Franchini, E., da Silva, V., & Diefenthaeler, F. (2016). Effect of fatigue on reaction time, response time, performance time, and kick impact in taekwondo roundhouse kick. *Sports Biomech, 16*(2), 201–209. doi: 10.1080/14763141.2016.1217347

Schofield, W. N. (2017). Predicting basal metabolic rate, new standards and review of previous work. Hum Nutr Clin Nutr, 39, 5–41.

Seneviratne, S., Hu, Y., Nguyen, T., Lan, G., Khalifa, S., Thilakarathna, K., Hassan, M., & Seneviratne, A. (2017). A Survey of Wearable Devices and Challenges. IEEE *Commun Surv Tutor, 19*(4), 2573–2620. doi: 10.1109/comst.2017.2731979

Shirom, A., Nirel, N., & Vinokur, A. D. (2006). Overload, autonomy, and burnout as predictors of physicians' quality of care. *J Occup Health Psychol, 11*(4), 328–342. doi: 10.1037/1076-8998.11.4.328

Shortz, A. E., Franke, M., Kilic, E. S., Peres, S. C., & Mehta, R. K. (2017). Evaluation of Offshore Shiftwork using Heart Rate Variability. *Proc Hum Factor Ergon, 61*(1), 1036–1039. doi: 10.1177/1541931213601742

Schutz, Y. (1995). The basis of direct and indirect calorimetry and their potentials. *Diabetes Metab Rev, 11*(4), 383–408. doi: 10.1002/dmr.5610110406

Slawinski, J. S., & Billat, V. L. (2005). Changes in Internal Mechanical Cost during Overground Running to Exhaustion. *Med Sci Sports Exerc, 37*(7), 1180–1186. doi: 10.1249/01.mss.0000170101.99529.a3

Stadje, R., Dornieden, K., Baum, E., Becker, A., Biroga, T., Bösner, S., Haasenritter, J., Keunecke, C., Viniol, A., & Donner-Banzhoff, N. (2016). The differential diagnosis of tiredness: a systematic review. *BMC Fam Pract, 17*(1). doi: 10.1186/s12875-016-0545-5

Straczkiewicz, M., James, P., & Onnela, J. P. (2021). A systematic review of smartphone-based human activity recognition methods for health research. *NPJ Digit. 4*(1). doi: 10.1038/s41746-021-00514-4

Svertoka, E., Saafi, S., Rusu-Casandra, A., Burget, R., Marghescu, I., Hosek, J., & Ometov, A. (2021). Wearables for Industrial Work Safety: A Survey. *Sensors, 21*(11), 3844. doi: 10.3390/s21113844

Tapia, E. M. (2008) Using Machine Learning for Real-time Activity Recognition and Estimation of Energy Expenditure [Doctoral dissertation, Massachusetts Institute of Technology].

Terzis, G., Spengos, K., Manta, P., Sarris, N., & Georgiadis, G. (2008). Fiber Type Composition and Capillary Density in Relation to Submaximal Number of Repetitions in Resistance Exercise. *J Strength Cond Res, 22*(3), 845–850. doi: 10.1519/jsc.0b013e31816a5ee4

Thompson, B. J., Stock, M. S., & Banuelas, V. K. (2017). Effects of Accumulating Work Shifts on Performance-Based Fatigue Using Multiple Strength Measurements in Day and Night Shift Nurses and Aides. *Hum Factors, 59*(3), 346–356. doi: 10.1177/0018720816677814

Thompson, B. J. (2019). Does work-induced fatigue accumulate across three compressed 12 hour shifts in hospital nurses and aides? *PLOS ONE, 14*(2). doi: 10.1371/journal.pone.0211715

Thomson, R. L., Rogers, D. K., Howe, P. R., & Buckley, J. D. (2015). Effect of acute exercise-induced fatigue on maximal rate of heart rate increase during submaximal cycling. *Res Sports Med, 24*(1), 1–15. doi: 10.1080/15438627.2015.1076414

Vandewalle, H., Pérès, G., & Monod, H. (1987). Standard Anaerobic Exercise Tests. *Sports Med, 4*(4), 268–289. doi: 10.2165/00007256-198704040-00004

Wang, J., Redmond, S. J., Voleno, M., Narayanan, M. R., Wang, N., Cerutti, S., & Lovell, N. H. (2012). Energy expenditure estimation during normal ambulation using triaxial accelerometry and barometric pressure. *Physiol Meas, 33*(11), 1811–1830. doi: 10.1088/0967-3334/33/11/1811

Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern Recognit Lett,, 119*, 3–11. doi: 10.1016/j.patrec.2018.02.010

Wang, Y., Cang, S., & Yu, H. (2019). A survey on wearable sensor modality centred human activity recognition in health care. *Expert Syst. Appl. 137*, 167–190. doi: 10.1016/j.eswa.2019.04.057

Wannenburg, J., & Malekian, R. (2017). Physical Activity Recognition From Smartphone Accelerometer Data for User Context Awareness Sensing. IEEE Trans Syst Man Cybern, 47(12), 3142–3149. doi: 10.1109/tsmc.2016.2562509

Webster, J. D., Welsh, G., & Garrow, J. S. (1986). Description of a human direct calorimeter, with a note on the energy cost of clerical work. *Br J Nutr, 55*(1), 1–6. doi: 10.1079/bjn19860003

Weir, J. B. D. V. (1949). New methods for calculating metabolic rate with special reference to protein metabolism. *J Physiol, 109*(1–2), 1–9. doi: 10.1113/jphysiol.1949.sp004363

World Health Organisation (1985) Energy and protein requirements. Report of a joint FAO/WHO/UNU Expert Consultation. *World Health Organ Tech Rep Ser, 724*, 1–206.

Wu, P. P. Y., Sterkenburg, N., Everett, K., Chapman, D. W., White, N., & Mengersen, K. (2019). Predicting fatigue using countermovement jump force-time signatures: PCA can distinguish neuromuscular versus metabolic fatigue. *PLOS ONE, 14*(7). doi: 10.1371/journal.pone.0219295

Völker, I., Kirchner, C., & Bock, O. L. (2016). Relation between Multiple Markers of Work-Related Fatigue. Saf Health Work, 7(2), 124–129. doi: doi.org/10.1016/j.shaw.2015.11.003

Yan, Z., Subbaraju, V., Chakraborty, D., Misra, A., & Aberer, K. (2012). Energy-Efficient Continuous Activity Recognition on Mobile Phones: An Activity-Adaptive Approach. *Proceedings of the 2012 16th International Symposium on Wearable Computers (ISWC 2012),* Newcastle, UK, 17–24. doi: doi.org/10.1109/iswc.2012.23

Yu, Y., Li, H., Yang, X., Kong, L., Luo, X., & Wong, A. Y. (2019). An automatic and non-invasive physical fatigue assessment method for construction workers. *Autom Contr, 103*, 1–12. doi: 10.1016/j.autcon.2019.02.020

Acknowledgements

I am sincerely grateful to everyone who have helped me throughout my studies. My deepest gratitude goes to my supervisor Professor Ivo Fridolin and my co-supervisors Kristjan Pilt and Professor Gert Jervan for their support, guidance and time during my studies. I would also like to thank Moonika Viigimäe, for her contribution in publications and for the overall positive support.

Special thanks goes to my colleagues in the Department of Health Technologies for their willingness to help me with whatever issues I faced during my studies.

I would also like to acknowledge the organizations and programmes which have supported me throughout my PhD: Estonian Ministry of Education and Research under institutional research financing IUTs 19-1 and 19-2, Estonian Centre of Excellence in IT (EXCITE) funded by European Regional Development Fund, European Commission H2020 project 872614 - SMART4ALL, Study IT programme of HITSA and Ustus Agur scholarship. The research was funded partly by the Estonian Ministry of Education and Research under institutional research financing IUTs 19-1 and 19-2 and by Estonian Centre of Excellence in IT (EXCITE) funded by European Regional Development Fund and supported by Study IT programme of HITSA and Ustus Agur Scholarship by the Estonian Association of Information Technology and Telecommunications.

Last, but not least, I would like to thank my family and friends for all the support I received from them over the years.

Abstract

Advancing Novel Physical Fatigue Assessment and Human Activity Monitoring Methods towards Perzonalized Feedback with Wearable Sensors

The circumstances of current human existence are far different from remote past. Physical exertion is no longer a requirement for daily living and today's conditions allow an unprecedentedly sedentary lifestyle. This discordance between our contemporary lives and our genetic makeup has important health implications on skeletal density, cardiovascular diseases, obesity, body composition and insulin resistance. It is important to propagate active lifestyle, since research studies confirm that routine physical activity multiple benefits by lowering the risk of diabetes, cardiovascular disease and obesity, while increasing psychological well-being.

Advancement of technology has brought a surge of popularity for devices that help their users keep track of their physical activity, training schedule, exercises and lost calories. Since this makes training more interactive and allows users to have better overview of their progress, it often motivates the users to have a more active lifestyle. This is achieved by using wearable systems to conveniently measure, collect and analyze the user's physiological data. For convenient use wearables need to be small and unobtrusive, which in turn puts significant demand on optimizing different aspects of these system such as reducing power consumption. The general aim of the thesis is to advance novel physical fatigue assessment and human activity monitoring methods that could be applied in real-time by using wearable sensors and systems.

Firstly, the thesis aimed to improve real-time physical activity recognition by optimizing the window length, sampling frequency and feature selection **(Publication I)**. Physical activity recognition allows automatic recognition of physical activities. Real-time activity recognition provides valuable information for improving online feedback of the activity trackers or for providing extra safety by monitoring the status of the users working in high-risk environments. As a result of this thesis, both window length and sampling frequency were optimized and multiple feature selection methods were used to decrease the initial 110 features to 13 features without lowering classification performance.

Secondly, the thesis aimed to assess different basal metabolic rate predictive equations to improve energy expenditure estimation **(Publication II)**. Energy expenditure is an important parameter for the studies of physical activity and is often used as a correlate of its level. Energy expenditure is an important tool for adjusting the individuals' nutritional supply or to assess the health of a larger population. Based on the results of this thesis the Mifflin-St Jeor model performed the best by having the lower root-mean-square-error of 175 kcal/day.

Thirdly, the thesis aimed to propose a method for real-time physical fatigue assessment suitable for wearable systems by utilizing a set of real-time and easily measurable cardiovascular parameters **(Publication III)**. Fatigue is a term used to describe an altered physiological state that results in decreased mental or physical performance. The ability to effectively monitor fatigue is most desired since high prevalence of fatigue has been reported in many operational settings to induce safety problems by directly influencing the mental and physical ability of people to perform even light activities. This thesis demonstrated that the compiled test battery can

selectively assess the rested vs. physically-fatigue states and the obtained linear support-vector machine based model showed promising ability to classify between different fatigue states.

The current thesis shows multiple possibilities to further advance the current state-of-the-art physical fatigue assessment and human activity monitoring methods by improving their performance or optimizing them for suitable use in wearable system.

Lühikokkuvõte

Kantavatel seadmetel põhinevate füüsilise väsimuse hindamise ning inimese aktiivsuse monitoorimise meetodite adrendamine personaalseks tagasisideks

Tänapäeva inimeste eluviis erineb olulisel määral inimeste elust eelmisel sajandil või veel kaugemas minevikus. Füüsiline pingutus ei ole enam igapäevaelu eeldus ja tänapäeva tingimused võimaldavad inimesel hakkama saada enneolematult vähese füüsilise liikumisega. Selline lahknevus meie kaasaegse elu ja meie geneetilise ülesehituse vahel avaldab olulist mõju tervisele, näiteks luustiku tihedusele, südame-veresoonkonna haigustele, rasvumisele, keha koostisele ja insuliiniresistentsusele. Oluline on propageerida aktiivset elustiili, sest uuringud kinnitavad, et rutiinne füüsiline tegevus omab positiivset mõju, vähendades diabeedi-, südame-veresoonkonna haiguste ja rasvumise riski ning suurendades vaimset heaolu.

Tehnoloogia arenedes on populaarsust kogunud erinevad seadmed, mis aitavad nende kasutajatel jälgida oma füüsilist aktiivsust, treeninggraafikut, -harjutusi ja põletatud kaloreid. Kuna selline monitooring muudab treenimise interaktiivsemaks ning võimaldab kasutajatel saada paremat ülevaadet oma edusammudest, motiveerib see sageli kasutajaid harrastama aktiivsemat elustiili. Füüsilise aktiivsuse ja treeningu jälgimine põhineb kantavatel seadmetel, mis mõõdavad, koguvad ja analüüsivad kasutaja füsioloogilisi andmeid. Mugavaks kasutamiseks peavad kantavad süsteemid olema väiksed ja märkamatud, mis omakorda seab märkimisväärse nõude nende süsteemide erinevate aspektide optimeerimisele, näiteks energiatarbimise vähendamisele. Antud doktoritöö üldeesmärk on edendada uudseid füüsilise väsimuse hindamise ja inimese aktiivsuse monitoorimise meetodeid, mida saaks reaalajas rakendada erinevate andurite abil kantavates süsteemides.

Lõputöö esimeseks eesmärgiks oli täiustada füüsilise tegevuse reaalajas tuvastamist, optimeerides akna pikkust, diskreetimissagedust ja tunnusjoonte valikut **(Publikatsioon I)**. Reaalajas töötav liikumisviiside tuvastus annab väärtuslikku teavet aktiivsusmonitoride tagasiside kvaliteedi tõstmiseks või lisaohutuse tagamiseks, jälgides kõrge riskiga keskkondades töötavate kasutajate olekut. Selle lõputöö tulemusena optimeeriti nii klassifitseerimise akna pikkust kui ka kiirendusanduri diskreetimissagedust ning erinvate tunnusjoonte valimismeetodite abil vähendati tunnusjoonte arvu 110-lt 13-le ilma klassifitseerimise tulemust langetamata.

Teiseks lõputöö eesmärgiks oli hinnata erinevaid baasainekulu hindamise valemeid, et edendada energiakulu hinnangut **(Publikatsioon II)**. Energiakulu on kehalise aktiivsuse uuringute seisukohast väga oluline parameeter, mida kasutatakse sageli ka sellega korreleeruva parameetrina. Energiakulu on ka oluline parameeter inimeste toitumisharjumuste kohandamiseks või suurema elanikkonna tervise hindamiseks. Selle lõputöö tulemuste põhjal toimis kõige paremini Mifflin-St Jeori mudel, millel ruutkeskmine hälve 175 kcal oli väikseim.

Kolmandaks lõputöö eesmärgiks oli luua kantavatele seadmetele sobiv füüsilise väsimuse reaalajas hindamise meetod, kasutades reaalajas ja kergesti mõõdetavaid kardiovaskulaarseid parameetreid **(Publikatsioon III)**. Väsimus tähendab muutunud füsioloogilist seisundit, mille tagajärjeks on vaimsete ja füüsiliste võimete langus. Väsimuse kvaliteetne hindamine on väga oluline, kuna väsimuse tekkimine on levinud mitmetes erinevates töökohtades, mis omakorda suurendab tööõnnetuste riski,

mõjutades otseselt inimeste vaimset ja füüsilist võimet sooritada isegi lihtsamaid tegevusi. Antud doktoritöö tulemused näitavad, et koostatud parameetrite kogum suudab selektiivselt eraldada puhkeolekut ja füüsiliselt väsinud olekut ning loodud tugivektormasina põhine mudel näitas võimet eristada erinevaid väsimusseisundeid.

Käesolev lõputöö toob välja mitmeid võimalusi kaasaegse füüsilise väsimuse hindamise ja liikumisviiside tuvastamise meetodite edasiseks arendamiseks, parandades nende jõudlust või optimeerides neid kantavates süsteemides kasutamise eesmärgil.

Appendix 1 – Features of the Human Activity Recognition Study (Publication I)

X_D – Dynamic component of X axis; Y_D – Dynamic component of Y axis; Z_D – Dynamic component of Z axis; X_S – Static component of X axis; Y_S – Static component of Y axis; Z_S – Static component of Z axis; s_i – Measured signals in a fragment; N – number of signals in a fragment.

* - Features used in feature subset with 43 features.
** - Features used in feature subsets with 43 and 28 features.
*** - Features used in feature subsets with 43, 28 and 19 features.
**** - Features used in feature subsets 43, 28, 19 and 13 features.

Features adopted from Liu et al., 2012		
Nr.	Fragment used	Explanation
1	X_D	Accelerometer counts: Sum of the absolute values of the signals over a fragment $\lvert\bar{s}\rvert$
2	Y_D	
3	Z_D	
4	X_S	
5	Y_S	
6	Z_S	
7****	X_D	Mean of the absolute values
8****	Y_D	
9****	Z_D	
10***	X_S	
11***	Y_S	
12***	Z_S	
13	X_D	Standard deviation
14	Y_D	
15	Z_D	
16***	X_S	
17***	Y_S	
18***	Z_S	
19*	X_D	Coefficients of variation
20*	Y_D	
21*	Z_D	
22**	X_S	
23*	Y_S	
24*	Z_S	

25	X_D	Peak-to-peak amplitude
26	Y_D	
27	Z_D	
28	X_S	
29	Y_S	
30	Z_S	
31	X_D	Percentile (10^{th})
32	Y_D	
33	Z_D	
34	X_S	
35	Y_S	
36	Z_S	
37	X_D	Percentile (25^{th})
38	Y_D	
39	Z_D	
40	X_S	
41	Y_S	
42	Z_S	
43*	X_D	Percentile (50^{th})
44*	Y_D	
45*	Z_D	
46	X_S	
47	Y_S	
48	Z_S	
49	X_D	Percentile (75^{th})
50	Y_D	
51	Z_D	
52	X_S	
53	Y_S	
54	Z_S	
55	X_D	Percentile (90^{th})
56	Y_D	
57	Z_D	
58	X_S	
59	Y_S	
60	Z_S	
Body posture related features (adopted from Tapia 2008)		
61	X_S	Mean $\bar{s}$
62	Y_S	
63	Z_S	

64**	X_S Y_S Z_S	Mean over all axes
65	X_S	Area under signal (sum of the values)
66	Y_S	
67	Z_S	
68**	X_S Y_S	Mean distance between axes
69*	X_S Z_S	Mean distance between axes
70*	Y_S Z_S	Mean distance between axes
Motion shape related features (adopted from Tapia 2008)		
71*	X_D Y_D Z_D	Cumulative sum over absolute signal value
72	X_D Y_D Z_D	Mean of total signal vector magnitude
Motion periodicity related features (adopted from Tapia 2008)		
73****	X_D	Mean crossing rate (Number of times signal crosses its mean value over the fragment)
74****	Y_D	
75****	Z_D	
76***	X_S	
77***	Y_S	
78***	Z_S	
Features adopted from Moncada-Torres et al., 2014		
79	X_D	Percentile (3^{rd})
80	Y_D	
81	Z_D	
82****	X_S	
83****	Y_S	
84****	Z_S	
85	X_D	Percentile (20^{th})
86	Y_D	
87	Z_D	
88	X_S	
89	Y_S	
90	Z_S	

91	X_D	Percentile (97th)
92	Y_D	
93	Z_D	
94	X_S	
95	Y_S	
96	Z_S	
97****	X_D Y_D	Correlation coefficient between axes
98****	X_D Z_D	
99****	Y_D Z_D	
100	X_S	Root-mean-square
101	Y_S	
102	Z_S	
Additionally added features		
103****	X_S Y_S Z_S	Mean of total signal vector magnitude
104*	X_D Y_D Z_D	Mean of velocity modules
105*	X_D	Kurtosis
106*	Y_D	
107*	Z_D	
108*	X_D	Skewness
109**	Y_D	
110*	Z_D	

Appendix 2 – Publication I

Publication I

Allik, A., Pilt, K., Karai, D., Fridolin, I., Leier, M., & Jervan, G. (2019). Optimization of Physical Activity Recognition for Real-Time Wearable Systems: Effect of Window Length, Sampling Frequency and Number of Features. *Applied Sciences, 9*(22), 4833. DOI: 10.3390/app9224833

applied sciences

MDPI

Article

Optimization of Physical Activity Recognition for Real-Time Wearable Systems: Effect of Window Length, Sampling Frequency and Number of Features

Ardo Allik [1,*], Kristjan Pilt [1], Deniss Karai [1], Ivo Fridolin [1], Mairo Leier [2] and Gert Jervan [2]

1 Department of Health Technologies, Tallinn University of Technology, Ehitajate tee 5, 12616 Tallinn, Estonia; kristjan.pilt@taltech.ee (K.P.); deniss.karai@taltech.ee (D.K.); ivo.fridolin@taltech.ee (I.F.)

2 Department of Computer Systems, Tallinn University of Technology, Ehitajate tee 5, 12616 Tallinn, Estonia; mairo.leier@taltech.ee (M.L.); gert.jervan@taltech.ee (G.J.)

* Correspondence: ardo.allik@taltech.ee

Received: 20 October 2019; Accepted: 7 November 2019; Published: 12 November 2019

check for updates

Abstract: The aim of this study was to develop an optimized physical activity classifier for real-time wearable systems with the focus on reducing the requirements on device power consumption and memory buffer. Classification parameters evaluated in this study were the sampling frequency of the acceleration signal, window length of the classification fragment, and the number of classification features, found with different feature selection methods. For parameter evaluation, a decision tree classifier was created based on the acceleration signals recorded during tests, where 25 healthy test subjects performed various physical activities. Overall average F1-score achieved in this study was about 0.90. Similar F1-scores were achieved with the evaluated window lengths of 5 s (0.92 ± 0.02) and 3 s (0.91 ± 0.02), while classification performance with 1 s were lower (0.87 ± 0.02). Tested sampling frequencies of 50 Hz, 25 Hz, and 13 Hz had similar results with most classified activity types, with an exception of outdoor cycling, where differences were significant. Using forward sequential feature selection enabled the decreasing of the number of features from initial 110 features to about 12 features without lowering the classification performance. The results of this study have been used for developing more efficient real-time physical activity classifiers.

Keywords: accelerometer; activity classification; activity trackers; machine learning; wearable systems

1. Introduction

It is important to propagate active lifestyle, since routine physical activity has been found to have multiple benefits, such as preventing chronic diseases and increasing psychological well-being [1,2], while prolonged inactivity has been shown to lead to an increase of chronic diseases and obesity [1,3]. Advancement of technology has brought a surge of popularity for many activity trackers in the form of mobile phone apps or wearable systems. With these devices, users are able to keep track of their training schedule, exercises and lost calories [4]. Since this makes training more interactive and allows users to have better overview of their progress, then it often motivates the users to have a more active lifestyle and lose weight over sustained periods [5–7].

Wearable systems are used to conveniently measure, collect and analyze the user's psychological data. This requires wearables to be small and unobtrusive, which in turn puts significant demand on reducing power consumption of the system [8]. This is also significant for real-time physical activity classification, which can be used in wearables for online activity recognition by allowing automatic recognition of the activities the user is performing [9,10]. Real-time activity recognition provides valuable information for improving online feedback of the activity trackers or for providing extra safety by monitoring the status of the users working in high-risk environments [11].

Power consumption required for physical activity classification is determined by multiple different components. Some of these components are based on the processing of the acceleration values, such as sampling rate of the signal and filtering [12]. Other elements are based on classification mechanics, such as classification window length, feature calculation, and the used machine learning algorithm. While studies have explored classification mechanics such as training times of different physical activity classification algorithms [13,14], they do not provide valuable information for real-time classification, since classifier training can be done previously on a desktop computer and later implemented into the wearable system. For classification systems working in real time, it is important to focus on processing time of the calculations the system has to do online [13,15].

In an earlier study, our group explored how different accelerometer sampling frequencies, classification window lengths, and the number of correlating features affect the classifier performance [16]. Few studies before have evaluated how different window lengths (commonly chosen between 1.5 s [17] and 5 s [13]) affect physical activity classification performance [15,18], but the lack of gold standard in physical activity classification makes it difficult to compare these results [19]. It has been stated that frequencies above 20 Hz cannot be expected to arise from voluntary movement [20], but comparable performance has been reported while using lower sampling frequencies [12,21]. Various methods have been used for feature selection, such as the ReliefF algorithm [22], principal component analysis [13], or information gain [15], but not in connection with window length and sampling frequency.

The aim of this study was to create an optimized physical activity classifier that would be suitable for implementation on real-time wearable systems. The focus was on testing various sampling frequencies, window lengths and number of features in order to reduce the power consumption, and to decrease the required memory buffer without compromising classification performance. Other classification elements were chosen based on the results of other studies with emphasis on high classification performance and low power consumption.

2. Materials and Methods

Physical activity classification often uses machine learning methods, where the classification is usually based on acceleration signals. Overview of the steps taken to create and evaluate the classifier used in this study are shown in Figure 1.

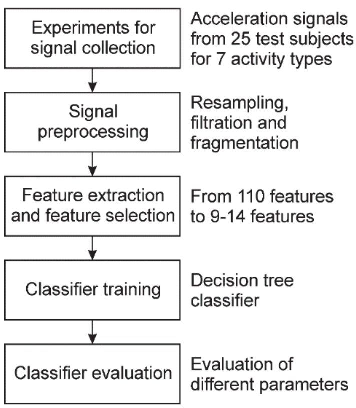

Figure 1. A summary of methods used in the study.

2.1. Instrumentation

Acceleration signals were measured with Shimmer3 (from here on Shimmer) sensor platform (Shimmer Research, Dublin, Ireland). While sensor fusion between accelerometers and gyroscopes has shown to increase classification performance in some studies [23], then others have found that gyroscope information does not contribute to activity recognition performance [22]. Due to the emphasis on designing physical activity classifier with low power consumption, gyroscope data were disregarded in this study.

The Shimmer sensor system has two built-in triaxial accelerometers: low noise accelerometer with the dynamic range of ±2 g and a wide range accelerometer with the dynamic range switchable between ±2 g to ±16 g (where 1 g equals to about 9.81 m/s^2). Since acceleration values during human motion surpass ±2 g [24], the data from wide range accelerometer was used with the dynamic range set to ±16 g. The wide range accelerometer uses STMicroelectronics LSM303AHTR sensor (Geneva, Switzerland), which has a numeric resolution of 16-bit. Acceleration was measured with a sampling rate of 512 Hz.

2.2. Study Group

The study was approved by the Tallinn Medical Research Ethics Committee. The main study group consisted of 25 healthy 21–45 year old test subjects (with an exception of one 57-year-old male), of whom 13 were male and 12 female. Average age was 32.0 ± 8.8 years (median 30.0) for the whole group, 32.8 ± 10.0 years (median 30.0) for males, and 31.0 ± 7.7 years (median 30.0) for females. A separate study group was used to measure the signals of outdoor cycling. This group consisted of 5 males with an average age of 38.4 ± 5.3 years (median 37.0).

2.3. Test Overview and Recorded Signals

Test subjects performed various physical activities during which acceleration signals were measured and recorded using the Shimmer sensor system. The sensor was located on the left wrist for feasibility of implementing the results in an activity tracker worn on the wrist. Even though using multiple sensors has been shown to increase the classification performance [25,26], having a wearable system with only one sensor is more comfortable and convenient for the user.

Each test subject conducted activities based on a precise schedule, where each activity was carried out for a fixed amount of time, shown in Table 1. For classification, these activities were grouped into different activity types, shown in Table 2. Indoor activities were divided into three different parts, during which each activity was performed for 3 min, with the exception of lying down, which lasted 4 min. There were short pauses between each activity, which were later discarded from the signals.

Table 1. Conducted activities and their duration in minutes.

Indoor Test 1	Indoor Test 2	Indoor Test 3 (% Shows Angle)	Outdoor Test
Walking (3) Running (3) Walking upstairs (3) Walking downstairs (3)	Sitting on chair (3) Lying on bed (4) Typing on computer (3) Folding clothes (3) Cleaning surface (3)	Walking (3 km/h) (3) Walking (5 km/h) (3) Walking (3 km/h, 10%) (3) Walking (5 km/h, 10%) (3) Running (6 km/h) (3) Running (10 km/h) (3) Running (12 km/h) (3) Running (6 km/h, 10%) (3)	Cycling (14) Cycling uphill (4) Cycling downhill (1)

Table 2. Classified activity types.

Activity Type	Activities Concluded
Static	Lying, sitting, standing
Low Intensity	Typing on computer
Moderate Intensity	Folding clothes
Rhythmical Intensity	Cleaning a surface with a towel
Walking	Walking in a corridor, walking on a treadmill, walking upstairs, walking downstairs
Running	Running in a corridor, running on a treadmill
Outdoor Cycling	Cycling outdoors on different terrains

In the first part, test subjects walked in a corridor, ran in the corridor, walked upstairs, and walked downstairs. Altogether, a total of 12 min of acceleration signals were used from this part.

The second part consisted of sitting on a chair, lying on a bed, typing on a computer while sitting, standing, folding clothes while standing, and cleaning a surface while standing. A total of 19 min of signals were used from the second part.

The third indoor part consisted of walking on a treadmill at different speeds and angles (3 km/h, 5 km/h, 3 km/h with uphill angle 10%, 5 km/h with uphill angle 10%) and running on treadmill at different speeds and angles (6 km/h, 10 km/h, 12 km/h, 6 km/h with 10% uphill angle). A total of 24 min of signals were used from this part.

Outdoor cycling signals were recorded separately with a different study group. These signals consist of 14 min of cycling on a plain road, 4 min of cycling uphill, and 1 min of cycling downhill.

2.4. Resampling and Sampling Frequency

As an aim of this study, it was tested how different sampling frequencies affect the classification results. Lowering the sampling frequency, f_s, decreases the number of samples in the classification fragment, s_f, which is calculated as follows:

$$s_f = f_s \cdot w_f, \tag{1}$$

where w_f is the window length of a fragment given in seconds.

To test different sampling frequencies, the signals that were initially recorded with a sampling frequency of 512 Hz were later resampled using a MATLAB function resample (R2016b, MathWorks, Natick, MA, USA). This function applies interpolation and decimation in order to achieve the desired sampling rate. In case of interpolation, the function inserts points with 0-values between each of the original samples of the signal, after which the signal is low-pass filtered at half of the desired sampling rate. To obtain the final result, decimation is applied by selecting samples from the filtered output [27]. The sampling frequencies of 50 Hz, 25 Hz, and 13 Hz were chosen for evaluating the effects of different sampling frequencies on classifier performance.

2.5. Filtering

Following resampling, filtering was applied to separate the recorded acceleration signals into static and dynamic components for physical activity classification. The static component in the acceleration signal is mostly affected by gravity and captures the posture information, while the dynamic component is based on motion and captures the human movement information.

In this study, the static component was found using a third order low-pass Butterworth infinite impulse response (IIR) filter. The passband and stopband edge frequencies and ripples were 0.1 Hz and 0.5 Hz, and 1 dB and 20 dB, respectively. The dynamic component was found by subtracting the static component from the original signal by taking into account the group delay of the low pass filter.

2.6. Fragmentation and Window Length

For classifier training, acceleration signals were fragmented into shorter consecutive fragments. Before fragmentation, the short pauses in the signals between different conducted activities were removed and only signals recorded during activities listed in Table 2 were kept. While some studies opt for an overlap between windows to increase the classification performance, in this study, no overlap was used to keep the computational power minimal.

In a system with a physical activity classifier working in real time, the window length determines the delay of the system, since each classification is done after signals have been collected for a whole window. The number of samples in the fragment is determined by both the sampling frequency and the window length according to Equation (1).

To evaluate how different window lengths affect the classifier performance, the window lengths of 5 s, 3 s, and 1 s were chosen, which are near the values usually used for physical activity classification in previous studies [13,17].

2.7. Feature Extraction

When using machine learning methods for physical activity classification, the classifier training is done based on features that are extracted from signal fragments. The feature set has to capture specific and diverse information of posture and human motion to allow precise activity classification. The initial set of 110 features used in this study were mostly adopted from previous studies by other researchers: (1) 60 various time-domain features from [28]; (2) 10 body posture related, 6 motion shape related features and 6 motion periodicity related features from [15]; (3) 24 various time-domain features from [22]; and (4) 9 separately added additional features.

Only time-domain features were chosen in this study in order to keep computing power minimal. While activity recognition studies have also used frequency-domain and wavelet transform features, the transforms needed to calculate these features would require extra resources. Additionally, it has been found that time-domain features give comparable results to other feature types [29].

2.8. Feature Selection

Another major aim of this study was to analyze how different number of features affects physical activity classification and what is the minimal number of features to use without compromising classification performance. For that, two different feature selection schemes were used to optimize the feature set.

One scheme was based on various methods that were used successively (Figure 2). This scheme used the features extracted with sampling frequency of 50 Hz and window length of 3 s and the achieved optimized feature set was later used with other frequency and window length combinations.

First, correlating features were removed based on a large correlation matrix that showed each feature's correlation coefficient with other features. From feature pairs or groups with a very high correlation (correlation coefficient larger than 0.9 or lower than -0.9), only the simpler features in terms of computational power requirements and complexity were kept. By using this method, 67 features were removed from the initial set, and a new subset of 43 features was formed. This method and the results have also been described in the previous study done by the authors [28].

Further feature optimization was done with one-way analysis of variance (ANOVA). The purpose of one-way ANOVA is to determine whether data from several groups of a factor have a common mean. ANOVA was used in this work to find out which features did not differentiate between any of the activities and thus did not provide any useful information for activity classification. Based on ANOVA results, 15 features were removed that were found not to affect classifier performance, and a new subset of 28 features was formed.

Finally, a sequential backward selection (SBS) procedure was repeated, where each feature was again removed one-by-one (those calculated similarly over all axes were removed together), and the feature that decreased the classifier performance the least was removed. After removing features this way, the classifier performance was still persistent with 13 features used. Further removal of features showed a decrease in activity classification sensitivities.

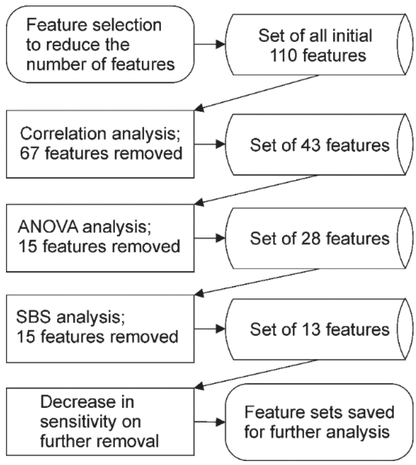

Figure 2. First feature selection scheme using correlation analysis, ANOVA, and backwards sequential feature selection with the number of features removed in each step.

The second feature selection scheme used in this study was a sequential forward selection (SFS) method similar to the last steps used in the first scheme (Figure 3). In this method, features were added one-by-one by conducting physical activity classification with each feature and, for every iteration, the best feature was kept. Features were added until the overall average classification sensitivity did not improve by more than 0.001. This method was completed for every sampling frequency and window length combination, and was used to compare the results of the first method.

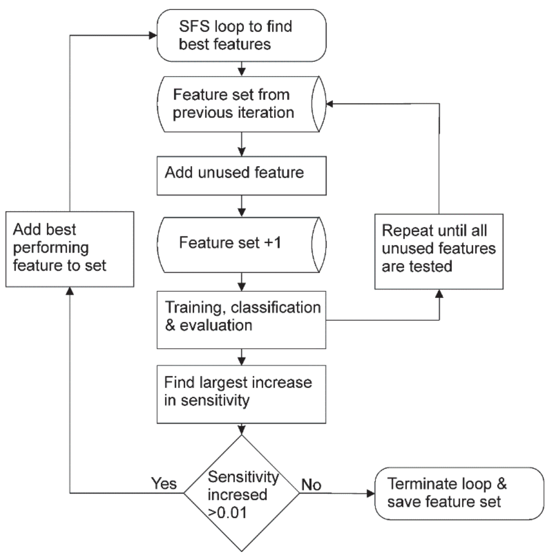

Figure 3. Forward sequential feature selection (SFS) method used in the second feature selection scheme.

2.9. Classifier Training

A machine learning based decision tree classification algorithm was chosen, which has been previously used in real-time physical activity classification and proposed as the most suitable in terms of performance and computational power needed for real-time classification [15,30]. The classifier was trained based on training data using MATLAB's function fitctree, which returns a fitted binary classification decision tree based on the input variables.

2.10. Classifier Evaluation

The classifier performance was evaluated using a leave-one-out cross-validation scheme where each test subject's signals were classified with a classifier that was trained using the signals from all the other test subjects. This method has been previously used in other physical activity classification studies to reduce overfitting errors [29,31].

Sensitivity (also called recall or true positive rate) was chosen as a statistical measure to evaluate classification performance during feature selection. Sensitivity shows the proportion of true positives classified (True_positives) in relation to correct or real ones (Real_positives), i.e., true positives that are correctly identified [32], and it is calculated as follows:

$$\begin{gathered}\text{Sensitivity} = \text{True\_positives}/\text{Real\_positives} = \\ \text{True\_positives}/(\text{True\_positives} + \text{False\_Negatives}).\end{gathered} \tag{2}$$

Classification results were evaluated using F1-score (also called F-score or F-measure), which is calculated as harmonic mean of precision and sensitivity [27], using the following formulas:

$$\begin{gathered}\text{Precision} = \text{True\_positives}/\text{Predicted\_positives} = \\ \text{True\_positives}/(\text{True\_positives} + \text{False\_positives}),\end{gathered} \tag{3}$$

$$\text{F1-score} = (2 \cdot \text{Sensitivity} \cdot \text{Precision})/(\text{Sensitivity} + \text{Precision}). \tag{4}$$

While evaluating the results with different window lengths, sampling frequencies and number of features, F1-scores were calculated separately for each activity type. Additionally, an average F1-score for different parameter combinations was found as a means of the activity type F1-scores.

A paired *t*-test ($p < 0.05$) was used to find statistical differences between the classification F1-scores of different activity types and averages while using different window lengths and sampling frequencies.

3. Results

3.1. Classifier Performance with Different Window Lengths

An overall average classification F1-score of about 0.90 was achieved for the physical activity classifier in this study, depending on the used window length, sampling frequency, feature set, and classified activity type. To evaluate how each of these parameters affected the classifier individually, classifier F1-scores were averaged over other parameters.

Figure 4 shows the classification F1-score of activity types for the different window lengths when averaged over different sampling frequencies (50 Hz, 25 Hz, 13 Hz) and feature sets (110 features, 43 features, 28 features, 13 features, and SFS feature set). The classifier had better performance with the average F1-score over 0.9 classifying static, walking and running activity types. Window lengths of 5 s and 3 s had similar results with the average F1-scores of 0.92 ± 0.02 and 0.91 ± 0.02, while the result with 1 s was 0.87 ± 0.02.

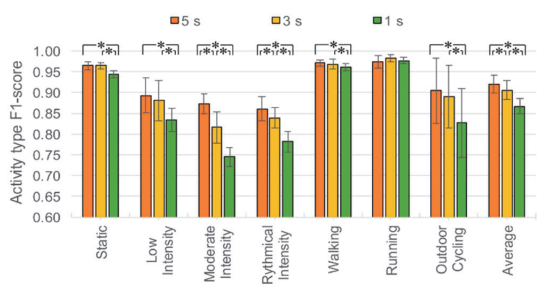

Figure 4. F1-scores of different activity types (mean ± SD (Standard deviation)) averaged over sampling frequencies and feature sets using different window lengths. Asterisks show significant statistical difference between different values of the window length ($p < 0.05$).

Statistically significant differences (marked with an asterisk in Figure 4) were found in moderate intensity and rhythmical intensity activity types between window lengths of 5 s and 3 s. Window length of 1 s had a statistical difference classifying every activity type other than running compared to both 5 s and 3 s window length.

3.2. Classifier Performance with Different Sampling Frequencies

To compare the results with different sampling frequencies, F1-scores were averaged over different window lengths and feature sets (Figure 5). Overall, the classifier had similar average F1-score with 50 Hz (0.92 ± 0.02) and 25 Hz (0.91 ± 0.02), while the average F1-score with 13 Hz was lower (0.87 ± 0.02).

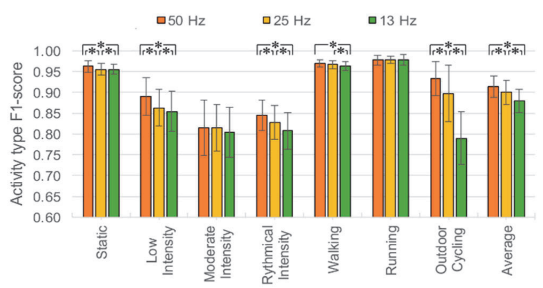

Figure 5. F1-score of different activity types (mean ± SD) averaged over window lengths and feature sets using different sampling frequencies. Asterisks show a significant statistical difference between different values of the sampling frequency ($p < 0.05$).

Statistically significant differences between different sampling frequencies (marked with an asterisk in Figure 5) were found for most activity types with the exceptions of moderate intensity and running.

Very large differences in classification performance were noted while classifying outdoor cycling, where the F1-score was 0.93 ± 0.04 with 50 Hz, 0.90 ± 0.07 with 25 Hz and 0.79 ± 0.06 with 13 Hz.

3.3. Classifier Performance with Different Feature Sets

To evaluate how the feature selection methods and the number of features used for classification affect the classifier performance, the results were averaged over different sampling frequencies and window lengths while using different feature sets (Figure 6). The feature sets of 110 features, 43 features, 28 features and 13 features, achieved with the first feature selection scheme, had similar average F1-scores between 0.89 and 0.90. The SFS feature set had a slightly higher average F1-score of 0.92 ± 0.03.

The SFS feature set had a major increase in performance compared to other feature sets classifying outdoor cycling (0.94 ± 0.04 compared to an average of 0.86 ± 0.09 with other sets) and a slight increase in classifying low intensity activity type (0.90 ± 0.04 compared to an average of 0.86 ± 0.04).

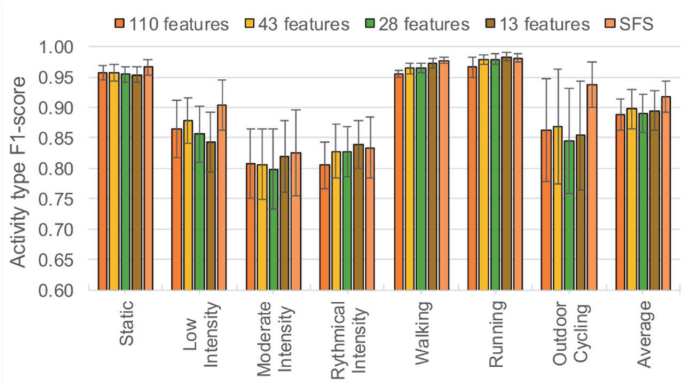

Figure 6. F1-scores of different activity types (mean ± SD) averaged over window lengths and sampling frequencies using different feature sets.

Since both classification window length and sampling frequency of the acceleration signal affect the number of samples in classification fragments, it is important to evaluate their combined effect on classification performance. Figure 7 shows the average classification F1-scores with different feature sets using different combinations of sampling frequencies and window lengths. The SD values were large, since the results were averaged over different activity types with different F1-scores.

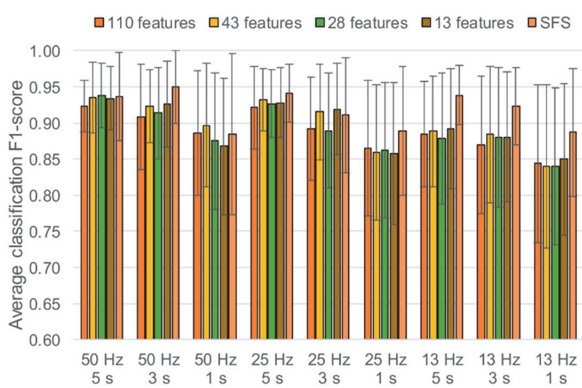

Figure 7. F1-scores (mean ± SD) averaged over all activities using different feature sets, window lengths and sampling frequencies.

The average F1-scores of all the combinations of sampling frequencies and window lengths were similar to all of the feature sets of the first feature selection scheme. The classification performance was better with combinations that had more samples per classification fragment, with the highest average of 0.93 ± 0.05 achieved with the combination of 50 Hz and 5 s. The results with the combinations that had either 1 s window length or sampling frequency of 13 Hz were lower compared to other combinations with most feature sets.

Compared to the feature sets of the first feature selection scheme, the SFS method used in the second scheme had higher performance with most window length and sampling frequency combinations. This difference was very noticeable with 13 Hz sampling frequency. The number of features used in

SFS feature sets was between 9 and 14 (Table 3), being remarkably lower than the number of features in most of the feature sets achieved with the first feature selection scheme.

Table 3. Number of features in sequential feature selection (SFS) feature sets with different sampling frequencies and window lengths.

Sampling Frequency and Window Length Combination	Number of Features in the SFS Feature Set
50 Hz, 5 s	12
50 Hz, 3 s	12
50 Hz, 1 s	9
25 Hz, 5 s	11
25 Hz, 3 s	12
25 Hz, 1 s	12
13 Hz, 5 s	11
13 Hz, 3 s	11
13 Hz, 1 s	14

3.4. Best Parameter Combination for Different Activity Types

While the results of this study generalized the effect of different sampling frequencies, window lengths, and number of features over various activity types, then it might also be useful to know the best combination for each activity type separately. Table 4 shows the parameter combination the highest F1 score for each classified activity type. The values are shown separately for both feature reduction schemes in order to compare the differences.

Table 4. Parameter combination with highest F1-score for different activity types and the average for both feature reduction schemes.

Activity Type	Window Length (s)	Sampling Frequency (Hz)	Number of Features	F1 Score
Static	5	25	110	0.97
	3	50	12 (SFS)	0.98
Low Intensity	5	13	110	0.93
	3	50	12 (SFS)	0.97
Moderate Intensity	5	50	110	0.90
	5	13	11 (SFS)	0.91
Rhythmical intensity	5	50	13	0.90
	5	25	12 (SFS)	0.89
Walking	3	50	43	0.98
	3	50	12 (SFS)	0.98
Running	3	25	13	0.99
	3	50	12 (SFS)	0.99
Outdoor Cycling	5	50	43	0.97
	3	50	12 (SFS)	0.98
Average	5	50	28	0.94
	3	50	12 (SFS)	0.95

4. Discussion

In this study it was analyzed for the first time how different window length, sampling frequency, and feature set combinations affect the performance of physical recognition based on decision tree classifiers in order to optimize the classifier for real-time wearable systems. The results of this study have been implemented into a smart work-wear prototype [11]. The main findings were: (1) classification F1-scores with window lengths of 5 s and 3 s were similar, while results with 1 s were lower; (2) all sampling frequencies performed similarly for most activity types, with an exception of outdoor cycling; (3) similar or better results were achieved with the feature sets with 9 to 14 features, achieved with either feature reduction scheme, compared to the initial full feature set of 110 features.

The window lengths of 5 s, 3 s and 1 s were used in this study to analyze how different window lengths affect the performance of physical activity classifier. F1-scores of walking, running and low intensity activity types were similar to all window lengths, while the differences with moderate intensity, rhythmical intensity, and outdoor cycling were larger. Even though window lengths between 3 s and 1 s have been found to be suitable for other studies (2.56 s in [22], 2 s in [26], 1.5 s in [17], 1 s in [18]), in this study, the classifier performance had a larger drop when decreasing the classifier window down to 1 s, while window lengths of 5 s and 3 s had similar results. The window length of 1 s had statistically significant differences with both 3 s and 5 s window lengths while classifying static, moderate intensity rhythmical intensity and outdoor cycling activity types. This could be caused by 1 s window length not being long enough to capture the movement of the body during activities where one period of movement exceeds the window length.

Different sampling frequencies of 50 Hz, 25 Hz, and 13 Hz were used to investigate how sampling frequency affects classification performance. For most classified activity types, no statistical differences were found between tested sampling frequencies, but there were large differences while classifying outdoor cycling. Previously, it had been found that frequencies above 20 Hz cannot be expected to arise from voluntary human movement, where the accelerometer is not in contact with vibrating external sources [20]. It is likely that the 13 Hz sampling frequency was not high enough to capture the vibration during outdoor cycling.

A total of 110 features were extracted from acceleration signals for physical activity classification. To reduce and optimize the number of features, two different feature selection schemes were used in this study. While the first scheme used different consecutive methods to reduce the number of features, the second scheme used forward SFS where features were added one-by-one. The first feature selection scheme enabled the reduction of the feature set from 110 features to 13 features without decreasing the classifier performance. It is possible that the feature set with 13 features was overfit for the conditions used in this study and would perform worse in other conditions.

Compared to the feature sets of the first feature selection scheme, the SFS method used in the second scheme had higher performance with most window length and sampling frequency combinations. This difference was very noticeable when using the sampling frequency of 13 Hz. The number of features used in SFS feature sets were between 9 and 14 (Table 3). The large differences in average F1-scores shown in Figure 7 between SFS feature set and other feature sets while using sampling rates of 25 Hz and 13 Hz were mostly affected by outdoor cycling. Unlike other feature sets, the SFS feature set had a high F1-score while classifying outdoor cycling with all sampling frequency and window length combinations. The highest average classification F1 score was achieved with a parameter combination with SFS feature set (3 s window length, 50 Hz sampling frequency, 12 features), which also had the best performance while classifying static, low intensity, walking and outdoor cycling activity types (Table 4).

It was predictable that the SFS method would provide better results, since the SFS method chose the best features to maximize the classification sensitivity separately for each window length and sampling frequency combination, while, with the first scheme, features were selected based on one sampling frequency and window length combination. The SFS method proved to be a simple comparison method for more comprehensive feature selection and showed that the effect of features depends on different classifier parameters, of which sampling frequency and window length were tested in this study.

Despite the recent advances in deep learning based activity recognition, which reduces the dependency on hand-crafted feature sets and thus could outperform more traditional machine learning methods, it is still far from being used in online mobile systems due to excessive computational power it requires [33]. Thus, the methods and results of this study provide useful information to other researchers for designing and implementing state-of-the-art physical activity recognition for real-time wearable systems.

5. Conclusions

This study evaluates the effects of sampling frequency of the acceleration signal, window length of the classification fragment, and number of features on classifier performance. The methods were chosen in order to reduce the requirements on computational power and available memory and are suitable for implementing physical activity classification in real-time systems.

We acknowledge some limitations in our approach that could be improved on in the future studies. First, sampling frequency and window length values evaluated in this study were chosen as a representative of the values used in other studies (low value, mid-range value, high value), but the optimum value could be somewhere between or even out of the explored range. It would be possible to classify larger numbers of different activity types and the acceleration signals should be measured under normal daily living conditions, which would allow for better physical activity classification during everyday life. The results could be evaluated with other machine learning algorithms that are used for physical activity classification, such as support-vector machines, Bayesian networks, and k-nearest neighbor algorithms, in order to see if there are any differences in the effects of the explored parameters.

Author Contributions: Conceptualization, A.A., K.P., and I.F.; formal analysis, A.A. and K.P.; methodology, A.A., K.P., and I.F.; investigation, A.A., K.P., D.K., and M.L.; data curation, D.K. and M.L.; writing—original draft preparation, A.A.; writing—review and editing, K.P., I.F., and G.J.; visualization, A.A.; validation, M.L.; supervision, I.F. and G.J.

Funding: The research was funded partly by the Estonian Ministry of Education and Research under institutional research financing IUTs 19-1 and 19-2, and by Estonian Centre of Excellence in IT (EXCITE) funded by European Regional Development Fund.

Acknowledgments: The authors wish to thank Siiri Mägi and Karl Erlenheim for assistance during the experiments, medical doctors Mari Meren and Ave Nagelmann from the Department of Pulmonology, North Estonia Medical Centre Foundations, Tallinn, Estonia for providing the environment for the study, and also those subjects who so kindly participated in the experiments.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyzes, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Dunstan, D.W.; Howard, B.; Healy, G.N.; Owen, N. Too much sitting—A health hazard. *Diabetes Res. Clin. Pract.* **2012**, *97*, 368–376. [CrossRef] [PubMed]
2. World Health Organization (WHO). *Obesity: Preventing and Managing the Global Epidemic*; WHO Technical Report Series 894; World Health Organization (WHO): Geneva, Switzerland, 2000.
3. Warburton, D.E.R.; Nicol, C.W.; Bredin, S.S.D. Health benefits of physical activity: The evidence. *Can. Med. Assoc. J.* **2014**, *174*, 801–809. [CrossRef] [PubMed]
4. Evenson, K.R.; Goto, M.M.; Furberg, R.D. Systematic review of the validity and reliability of consumer-wearable activity trackers. *Int. J. Behav. Nutr. Phys. Act.* **2015**, *12*, 1–22. [CrossRef] [PubMed]
5. Coughlin, S.; Stewart, J. Use of Consumer Wearable Devices to Promote Physical Activity: A Review of Health Intervention Studies. *J. Environ. Health Sci.* **2016**, *2*, 1–6. [CrossRef] [PubMed]
6. Maher, C.; Ryan, J.; Ambrosi, C.; Edney, S. Users' experiences of wearable activity trackers: A cross-sectional study. *BMC Public Health* **2017**, *17*, 880. [CrossRef] [PubMed]
7. Middelweerd, A.; Mollee, J.S.; Wal, C.N.; Brug, J.; Velde, S.J. Apps to promote physical activity among adults: A review and content analysis. *Int. J. Behav. Nutr. Phys. Act.* **2014**, *11*, 1–9. [CrossRef] [PubMed]
8. Senevirante, S.; Hu, Y.; Nguyen, T.; Lan, G.; Khalifa, S.; Thilakarathna, K.; Hassan, M.; Senevirante, A. A Survey of Wearable Devices and Challenges. *IEEE Commun. Surv. Tut.* **2017**, *4*, 2573–2620. [CrossRef]
9. Lee, K.; Kwan, M.P. Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results. *Comput. Environ. Urban Syst.* **2018**, *67*, 124–131. [CrossRef]

10. Wannenburg, J.; Malekian, R. Physical Activity Recognition from Smartphone Accelerometer Data for User Context Awareness Sensing. *IEEE Trans. Syst. Man. Cybern. Syst.* **2017**, *47*, 3142–3149. [CrossRef]
11. Leier, M.; Pilt, K.; Allik, A.; Karai, D.; Jervan, G.; Fridolin, I. Fall detection and activity recognition system for usage in smart work-wear. In Proceedings of the 16th Biennial Baltic Electronics Conference, Tallinn, Estonia, 8–10 October 2018.
12. Yan, Z.; Subbaraju, V.; Chakraborty, D.; Misra, A.; Aberer, K. Energy-Efficient Continuous Activity Recognition on Mobile Phones: An Activity-Adaptive Approach. In Proceedings of the 16th International Symposium on Wearable Computers, Newcastle, UK, 18–22 June 2012; pp. 17–24.
13. Altun, K.; Barshan, B.; Tuncel, O. Comparative study on classifying human activities with miniature inertial and magnetic sensors. *Pattern Recognit.* **2010**, *43*, 3605–3620. [CrossRef]
14. Feng, Z.; Mo, L.; Li, M. A Random Forest-based ensemble method for activity recognition. In Proceedings of the 37th Annual International Conference of the IEEE-EMBC, Milan, Italy, 25–29 August 2015; pp. 5074–5077.
15. Tapia, E.M. Using Machine Learning for Real-time Activity Recognition and Estimation of Energy Expenditure. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, June 2008.
16. Allik, A.; Pilt, K.; Karai, D.; Fridolin, I.; Leier, M.; Jervan, G. Activity classification for real-time wearable systems: Effect of window length, sampling frequency and number of features on classifier performance. In Proceedings of the IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, Malaysia, 4–8 December 2016; pp. 460–464.
17. Aktaruzzaman, M.; Scarabottolo, N.; Sassi, R. Parametric estimation of sample entropy for physical activity recognition. In Proceedings of the 37th Annual International Conference of the IEEE-EMBC, Milan, Italy, 25–29 August 2015; pp. 470–473.
18. Bulling, A.; Blanke, U.; Schiele, B. A Tutorial on Human Activity Recognition Using Body-Worn Inertial Sensors. *ACM Comput. Surv.* **2014**, *46*, 33. [CrossRef]
19. Awais, M.; Mellone, S.; Chiari, L. Physical activity classification meets daily life: Review on existing methodologies and open challenges. In Proceedings of the 37th Annual International Conference of the IEEE-EMBC, Milan, Italy, 25–29 August 2015; pp. 5050–5053.
20. Bouten, C.V.C.; Koekkoek, K.; Verduin, M.; Kodde, R.; Janssen, J.D. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Trans. Biomed. Eng.* **1997**, *44*, 136–147. [CrossRef] [PubMed]
21. Lee, J.; Kim, J. Energy-Efficient Real-Time Human Activity Recognition on Smart Mobile Devices. *Mob. Inf. Syst.* **2016**, *2016*, 2316757. [CrossRef]
22. Moncada-Torres, A.; Leuenberger, K.; Gonzenbach, R.; Luft, A.; Gassert, R. Activity classification based on inertial and barometric pressure sensors at different anatomical locations. *Physiol. Meas.* **2014**, *35*, 1245–1263. [CrossRef] [PubMed]
23. Wang, Z.; Yang, Z.; Dong, T. A Review of Wearable Technologies for Elderly Care that Can Accurately Track Indoor Position, Recognize Physical Activities and Monitor Vital Signs in Real Time. *Sensors* **2017**, *17*, 341. [CrossRef] [PubMed]
24. Chuang, F.C.; Yang, Y.T.C.; Wang, J.S. Accelerometer-based Energy Expenditure Estimation Methods and Performance Comparison. In Proceedings of the 2nd International Conference on Advances in Computer Science and Engineering (CSE 2013), Los Angeles, CA, USA, 1–2 July 2013; pp. 99–103.
25. Chowdhury, A.K.; Tjondronegoro, D.; Chandran, V.; Trost, S.G. Physical Activity Recognition Using Posterior-Adapted Class-Based Fusion of Multiaccelerometer Data. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 678–685. [CrossRef] [PubMed]
26. Loh, D.; Lee, T.J.; Zihajehzadeh, S.; Hoskinson, R.; Park, E.J. Fitness activity classification by using multiclass support vector machines on head-worn sensors. In Proceedings of the 37th Annual International Conference of the IEEE-EMBC, Milan, Italy, 25–29 August 2015; pp. 502–505.
27. Rajamani, K.; Lai, Y.-S.; Furrow, C.W. An efficient algorithm for sample rate conversion from CD to DAT. *IEEE Signal Process. Lett.* **2000**, *7*, 288–290. [CrossRef]
28. Liu, S.; Gao, R.X.; Freedson, P.S. Computational methods for estimating energy expenditure in human physical activities. *Med. Sci. Sports Exerc.* **2012**, *44*, 2138–2146. [CrossRef] [PubMed]
29. Preece, S.J.; Goulermas, J.Y.; Kenney, L.P.J.; Howard, D. A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. *IEEE Trans. Biomed. Eng.* **2009**, *56*, 871–879. [CrossRef] [PubMed]

30. Altini, M.; Penders, J.; Amft, O. Energy Expenditure Estimation Using Wearable Sensors: A New Methodology for Activity-Specific Models. In Proceedings of the Wireless Health, San Diego, CA, USA, 23–25 October 2012.
31. Bao, L.; Intille, S.S. Activity Recognition from User-Annotated Acceleration Data. *Pervasive Comput.* **2004**, *3001*, 1–17.
32. Powers, D.M.W. Evaluation: From precision, recall and F-factor to ROC, informedness, markedness and correlation. *J. Mach. Learn. Tech.* **2011**, *2*, 37–63.
33. Wang, J.; Chen, Y.; Hao, S.; Peng, X.; Hu, L. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognit. Lett.* **2019**, *119*, 3–11. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Appendix 3 – Publication II

Publication II

Allik, A.; Mägi, S.; Pilt, K.; Karai, D.; Fridolin I.; Leier, M.; Jervan, G. (2018). Comparison of Predictive Equations for Basal Metabolic Rate. *Proceedings of the 7th International Conference on Wireless Mobile Communication and Healthcare (MobiHealth 2017)*, Vienna, Austria, 261−264. DOI: 10.1007/978-3-319-98551-0

Comparison of Predictive Equations for Basal Metabolic Rate

Ardo Allik(✉), Siiri Mägi, Kristjan Pilt, Deniss Karai, Ivo Fridolin, Mairo Leier, and Gert Jervan

Tallinn University of Technology, Akadeemia Tee 1, 12616 Tallinn, Estonia
ardo.allik@ttu.ee

Abstract. The aim of this study was to compare and evaluate multiple predictive equations for basal metabolic rate in order to choose the most suitable one for energy expenditure models. Eight different predictive equations were compared to each other using regression analysis and with the results of indirect calorimetry tests with 25 participants. Mifflin-St Jeor, Livingston-Kohlstadt and Henry-Rees predictive equations performed better than other formulas with Mifflin-St Jeor having the lowest RMSE of 175 kcal/day compared to the results of indirect calorimetry. The results of this study can be used to develop more accurate energy expenditure models.

Keywords: Energy expenditure · Predictive equations · Basal metabolic rate Resting metabolic rate · Activity trackers · Physical activity

1 Introduction

Monitoring the physical activity (PA) is moving towards activity specific energy expenditure (EE) models that first recognise the activity and then apply a suitable EE algorithm for the specific activity type [1], which relies on accurate assessment of basal metabolic rate (BMR). For dietetics purposes BMR is commonly estimated using predictive equations, that use simple anthropometric variables such as the weight, height, age and sex of the person [2]. The aim of this study was to assess the BMR predictive equations by comparing different equations and validating their results with IC in order to choose the most suitable one for EE models.

2 Methods

The predictive equations explored in this study for BMR were Harris-Benedict [3], Schofield [4], FAO/WHO/UNU [5], Henry-Rees [6], and Kleiber [7] and for RMR were Mifflin-St Jeor [8], Owen [9, 10], Livingston-Kohlstadt [11]. EE values achieved with different predictive equations were compared to each other and with indirect calorimetry (IC) measurements. IC measurements were done using open-circuit indirect spirometry device "CareFusion MasterScreen CPX", which calculates EE based on Weir equation [12]. System was calibrated before each test subject.

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2018
P. Perego et al. (Eds.): MobiHealth 2017, LNICST 247, pp. 261–264, 2018.
https://doi.org/10.1007/978-3-319-98551-0

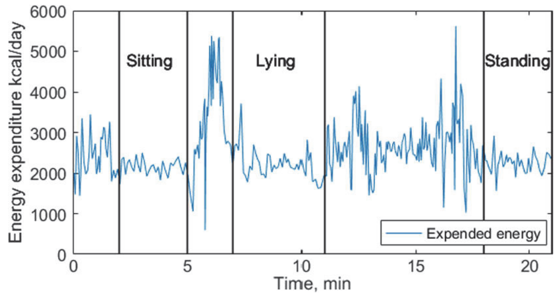

Fig. 1. Energy expenditure of one test subject during indirect calorimetry experiment.

The study group consisted of 25 healthy Caucasian adults, of whom 13 were male and 12 female. During IC measurements, EE was measured during three different positions – sitting on a chair, lying on a bed and standing up. The EE values during the experiment of one test subject are shown on Fig. 1. EE was calculated only based on the last minute of each activity. Even though standing and sitting should have approximately 1.3 times higher EE than lying [13], it was not possible to differentiate between these activities in this study (based on t-test results, $p < 0.05$). EE values from IC were divided by 1.3 and the values achieved with RMR equations were divided by 1.1 in order to compare the results with BMR equations.

3 Results

(1)

	1	2	3	4	5	6	7	8
1	1.00	0.98	0.88	0.92	0.98	0.93	0.94	0.95
2	0.98	1.00	0.88	0.92	0.97	0.95	0.95	0.96
3	0.88	0.88	1.00	0.85	0.92	0.91	0.92	0.90
4	0.92	0.92	0.85	1.00	0.90	0.90	0.91	0.93
5	0.98	0.97	0.92	0.90	1.00	0.97	0.97	0.97
6	0.93	0.95	0.91	0.90	0.97	1.00	1.00	0.98
7	0.94	0.95	0.92	0.91	0.97	1.00	1.00	0.99
8	0.95	0.96	0.90	0.93	0.97	0.98	0.99	1.00

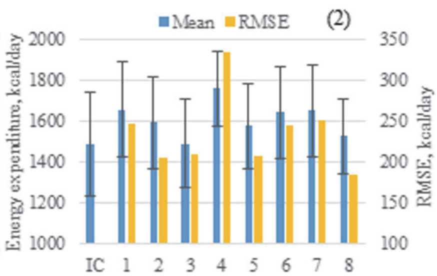

Fig. 2. (1) Coefficient of determination R^2 between different predictive equations. (2) Mean and SD of average BMR with indirect calorimetry (IC) and predictive equations; RMSE of BMR between predictive equations and IC. 1 – Harris-Benedict, 2 – Mifflin-St Jeor, 3 – Owen, 4 – Kleiber, 5 – Livingston-Kohlstadt, 6 – Schofield, 7 – FAO/WHO/UNO, 8 – Henry-Rees

4 Discussion

Based on regression analysis most equations had similar results, with Owen and Kleiber formulas being the outliers, which respectively had the lowest and highest average BMR results (Fig. 2). The average BMR values with Mifflin-St Jeor formula (1447 ± 204 kcal/day) were the closest with IC results (1485 ± 255 kcal/day) and also had the lowest RMSE of 175 kcal/day compared to IC. Based on paired t-test ($p < 0.05$), the results with Mifflin, Livingston-Kohlstadt and Henry-Rees equations were not statistically distinguishable from IC results.

While the results of this study can be used to compare different predictive equations and for developing different EE models, there are some limitations which can be improved on in future studies. First, the test group in this study was very homogeneous since only healthy adults of same race were included. Secondly, the IC calorimetry tests conducted in this study were part of a larger experiment, which is why each position was held for a minimal amount of time needed to reach an EE plateau.

The research was funded partly by the Estonian Ministry of Education and Research under institutional research financing IUTs 19-1 and 19-2, and by Estonian Centre of Excellence in IT (EXCITE) funded by European Regional Development Fund.

References

1. Altini, M., Penders, J., Amft, O.: Energy expenditure estimation using wearable sensors: a new methodology for activity-specific models. In: Wireless Health 2012, San Diego (2012)
2. Frankenfield, D., Roth-Yousey, L., Compher, C.: Comparison of predictive equations for resting metabolic rate in healthy nonobese and obese adults: a systematic review. J. Am. Diet. Assoc. **105**, 775–789 (2005)
3. Harris, J.A., Benedict, F.G.: A biometric study of human basal metabolism. Proc. Natl. Acad. Sci. USA **4**, 370–373 (1918)
4. Schofield, W.N.: Predicting basal metabolic rate, new standards and review of previous work. Hum. Nutr. Clin. Nutr. **39**, 5–41 (1985)
5. FAO/WHO/UNU: Energy and protein requirements: Report of a joint FAO/WHO/UNU expert consultation. World Health Organ. Technical report. Series 724 (1985)
6. Henry, C.J., Rees, D.G.: New predictive equations for the estimation of basal metabolic rate in tropical peoples. Eur. J. Clin. Nutr. **45**, 177–185 (1991)
7. Kleiber, M.: Body size and metabolism. Hilgardia **6**, 315–351 (1932)
8. Mifflin, M.D., St Jeor, S.T., Hill, L.A., Scott, B.J., Daugherty, S.A., Koh, Y.O.: A new predictive equation for resting energy expenditure in healthy individuals. Am. J. Clin. Nutr. **51**, 241–247 (1990)
9. Owen, O.E., et al.: A reappraisal of caloric requirements in healthy women. Am. J. Clin. Nutr. **44**, 1–19 (1986)
10. Owen, O.E., et al.: A reappraisal of the caloric requirements of men. Am. J. Clin. Nutr. **46**, 875–885 (1987)
11. Livingston, E.H., Kohlstadt, I.: Simplified resting metabolic rate-predicting formulas for normal-sized and obese individuals. Obes. Res. **13**, 1255–1262 (2005)

12. Weir, J. B. de V.: New methods for calculating metabolic rate with special reference to protein metabolism. J. Physiol. **109**, 1–9 (1949)
13. Ainsworth, B.E., et al.: 2011 Compendium of physical activities: a second update of codes and MET values. Med. Sci. Sports Exerc. **43**, 1575–1581 (2011)

Appendix 4 – Publication III

Publication III

Allik, A., Pilt, K., Viigimäe, M., Fridolin, I., & Jervan, G. (2022). A Novel Physical Fatigue Assessment Method Utilizing Heart Rate Variability and Pulse Arrival Time towards Personalized Feedback with Wearable Sensors. *Sensors, 22*(4), 1680. DOI: 10.3390/s22041680

sensors

MDPI

Article

A Novel Physical Fatigue Assessment Method Utilizing Heart Rate Variability and Pulse Arrival Time towards Personalized Feedback with Wearable Sensors

Ardo Allik [1,*], Kristjan Pilt [1], Moonika Viigimäe [1], Ivo Fridolin [1] and Gert Jervan [2]

[1] Department of Health Technologies, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia; kristjan.pilt@taltech.ee (K.P.); moonika.viigimae@taltech.ee (M.V.); ivo.fridolin@taltech.ee (I.F.)
[2] Department of Computer Systems, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia; gert.jervan@taltech.ee
* Correspondence: ardo.allik@taltech.ee

Abstract: This paper proposes a novel method for physical fatigue assessment that can be applied in wearable systems, by utilizing a set of real-time measurable cardiovascular parameters. Daylength measurements, including a morning test set, physical exercise during the day, and an afternoon test set were conducted on 16 healthy subjects (8 female and 8 male). To analyze cardiovascular parameters for physical fatigue assessment, electrocardiography, pulse wave and blood pressure were measured during the test sets. The fatigue assessment questionnaire score, reaction time, countermovement jump height and hand grip strength were also measured and used as reference parameters. This study demonstrates that (i) the compiled test battery can selectively assess the rested vs. physically-fatigued states; (ii) the obtained linear support-vector machine, trained using the heart rate variability based parameter (F-score 0.842, accuracy 0.813) and pulse arrival time based parameter (F-score 0.875, accuracy 0.875) shows a promising ability to classify between the physically mildly fatigued and significantly fatigued states. Despite the somewhat limited study group size, the results of the study are unique and provide a significant advancement on the existing physical fatigue assessment methods towards a personalized and continuous real-time fatigue monitoring system with wearable sensors.

Keywords: physical fatigue; fatigue assessment; heart rate variability; pulse arrival time; wearables

check for updates

Citation: Allik, A.; Pilt, K.; Viigimäe, M.; Fridolin, I.; Jervan, G. A Novel Physical Fatigue Assessment Method Utilizing Heart Rate Variability and Pulse Arrival Time towards Personalized Feedback with Wearable Sensors. *Sensors* **2022**, *22*, 1680. https://doi.org/10.3390/s22041680

Academic Editor: Leonardo Abbene

Received: 10 December 2021
Accepted: 18 February 2022
Published: 21 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Fatigue has been used as a term to describe an altered physiological state that results in decreased mental or physical performance, which may be caused by sleep loss, circadian changes or high workload [1,2]. The ability to effectively monitor fatigue is most desired due to multiple reasons since the complaint of fatigue is high in the general population, ranging from 18.3% to 27% [3]. As fatigue can directly influence the mental and physical ability of people to perform even light activities, workers' fatigue stemming from high demand jobs, long duty periods and accumulative sleep debt is a significant problem in modern industry. The high prevalence of fatigue has likewise been reported in many operational settings to induce safety problems [4]. Whereas previous studies of fatigue have mostly been focused on fatigue tests in different (work) settings, evaluation of muscular fatigue, subjective symptoms of fatigue, indicators of nervous strain, and the practical application of fatigue tests [5], further examination of the prior measures of fatigue addressed in these studies suggests that a practical need for a new multidimensional measure of fatigue exists.

Due to the definitional difficulties and multiple causes of fatigue, no single instrument can be applied as a gold standard for fatigue measurement. Fatigue may also have several confounding factors such as medication, psychological and cognitive conditions, and deconditioning [6,7]. The multi-factorial nature of fatigue suggests that a single universal test to measure fatigue may not exist [8]. Fatigue assessment studies have usually compiled different test batteries of various measures [9,10], which may be classified into six different categories: (i) questionnaires on subjective feelings of fatigue, (ii) psychological tests, (iii) neuropsychological tests, (iv) biochemical indexes, (v) physiological tests and (vi) autonomic nervous function tests [8].

It is important to note that most of these measurement methods require special conditions and testing environment and are therefore not suitable for real-time assessment of physical fatigue. Yet, human activity monitoring has advanced towards utilizing new wearable technologies and devices, which are able to conveniently measure, collect and analyze the user's physiological data [11]. Although a considerable effort is being made in wearable sensors, for comfortable use, wearables would need to be small and unobtrusive, which in turn requires keeping the power consumption and computational power as low as possible [11,12]. For these reasons, in order to assess fatigue continuously or repeatedly by wearable sensors, a novel approach that utilizes real-time measurable physiological signals will be needed.

This study proposes a set of cardiovascular (CV) parameters for physical fatigue assessment, which could be measured continuously in real-time. Since only heart rate variability (HRV) and reaction time (RT) were explored in an earlier study on physical fatigue [13], it can be hypothesized that compiling an enhanced test-battery of real-time CV parameters could yield a more effective outcome regarding the reference parameters for physical fatigue assessment. Correlations in-between the test battery parameters were analyzed on the individual level, which to the best of the authors' knowledge, has not been performed in the previous physical fatigue studies. A strong correlation between various measures could be used to improve the overall assessment quality or decrease the required computational power and complexity of the measurement system by removing redundant parameters. Building a model based on the data for classifying between different fatigue states would provide a basis for further development towards a continuous real-time fatigue monitoring system.

This study introduces a method for real-time physical fatigue assessment that can be applied in wearable systems, by utilizing a set of real-time and easily measurable cardiovascular (CV) parameters.

2. Materials and Methods

2.1. Study Design

An experiment was conducted on all the subjects (n = 16) in three main activities, including measurements in the morning and afternoon with a workout session in-between (Figure 1). The calculated or measured reference parameters included the fatigue assessment questionnaire score, reaction time (RT), hand grip strength and countermovement jump (CMJ) height. The evaluated CV parameters included heart rate (HR), measures of heart rate variability (HRV) and blood pressure normalized pulse arrival time (PAT).

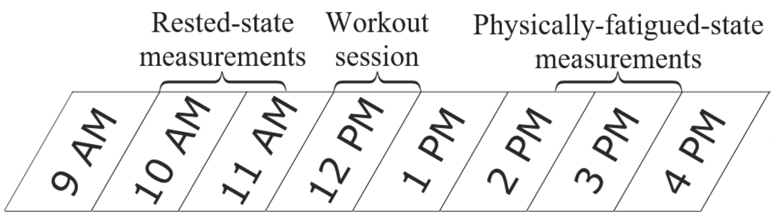

Figure 1. Overview of an experiment day. Cardiovascular and reference parameters were measured similarly in both test sets. The workout session involved multiple full-body exercises.

Rested-state (RS) measurements were obtained in the morning during 10:00–11:30. Following the RS measurements, all subjects performed the same exhausting full-body workout to induce physical fatigue. The 60-min workout was comprised of multiple sets of various exercises such as squats, burpees, sit-ups, push-ups, planks and jumping jacks. Physically-fatigued-state (PFS) measurements were obtained in the afternoon during 14:30–16:00. All measurements were conducted by the same person with the aim to avoid interoperator variability.

2.2. Study Group

The study group was comprised of 16 healthy 18 to 48 year-old subjects (8 female and 8 male) with the anthropometric parameters outlined in Table 1. The subjects signed a consent form prior to being enrolled in the experiment and the procedures performed were in accordance with the established ethical standards. The experiment was approved by the Tallinn Medical Research Ethics Committee (No. 1954).

Table 1. Subject anthropometric parameters. Age (years), height (cm), weight (kg) and body mass index (BMI) (kg/m2) mean values with standard deviations (SD).

Count	Age Mean ± SD; Range	Height Mean ± SD; Range	Weight Mean ± SD; Range	BMI Mean ± SD; Range
Total (16)	28.3 ± 7.9; 18–48	173.9 ± 8.1; 163–190	69.9 ± 12.3; 55–91	23.0 ± 2.9; 18.3–30.1
Female (8)	28.4 ± 7.0; 18–42	169.1 ± 5.9; 163–180	63.9 ± 10.5; 55–89	22.4 ± 3.5; 18.3–30.1
Male (8)	28.3 ± 9.2; 18–48	178.6 ± 7.3; 166–190	75.9 ± 11.4; 60–91	23.7 ± 2.2; 20.4–26.4

To reduce the effects of possible confounding factors, the following individuals were excluded from the study group: (i) individuals suffering from a disease or condition causing significant fatigue (congestive heart failure, respiratory failure, cancer, anemia, sleep disorders or a major psychiatric condition); (ii) individuals taking medicines that cause fatigue (beta blockers, diuretics or narcotics); (iii) pregnant or breastfeeding women; (iv) individuals that had worked night shifts within the past month; (v) individuals that consume alcoholic drinks daily, use of illicit drugs or smoking; (vi) individuals that had experienced a serious illness within two weeks before the test. The subjects of the study were asked to avoid any hard workout 24 h before the test and not to consume caffeinated drinks or foods.

2.3. Test Battery Design

Most of the parameters were chosen based on the literature overview of previous fatigue assessment studies. The analyzed parameters were divided into reference parameters that usually need administered tests and cannot be obtained in real-time, and CV parameters that could be continuously monitored and measured. These parameters were measured in the same order in both the rested and physically-fatigued states. The selected reference parameters included the score of a fatigue questionnaire [9], RT [10,14,15], hand grip strength [10,16] and CMJ height [9,10,16]. The evaluated CV parameters involved HR [1,17], HRV [1,2,18] and PAT. The parameters were selected with the aim to keep the complexity of the overall measurement process and computational power requirements as low as possible for suitable use in wearable systems, and thus, only time-domain measures were considered. While the effects of muscle fatigue on HRV have been additionally analyzed using frequency-domain and non-linear measures [19], these were not explored due to the following reasons: (i) information captured by frequency-domain parameters has been found to correlate with the measures analyzed in this study [20]; (ii) transforms needed for spectral analysis require extra resources.

2.4. Fatigue Questionnaire

At the start of the experiment the subjects were asked to complete a questionnaire to evaluate their current subjective fatigue level. The questionnaire adopted for the experiment was the Swedish Occupational Fatigue Inventory (SOFI), developed for the measurement of after-work fatigue [21] (Figure S1). The scale items were scored based on a 7-point Likert scale to assess fatigue from 0 (not at all) to 6 (to a very high degree). The scale items were as follows: (i) physical exertion (having palpitations, sweaty, out of breath and breathing heavily); (ii) physical discomfort (tense muscles, numbness, stiff joints and aching); (iii) lack of motivation (lacking concern, passive, indifferent and uninterested); (iv) sleepiness (falling asleep, drowsy, yawning and sleepy); (v) lack of energy (worn out, spent, drained and overworked). The score of the questionnaire was analyzed as the percentage of the maximum score.

2.5. Reaction Time Measurement

Subject RT was measured using the PC-PVT platform developed and validated for psychomotor vigilance testing [22,23]. The test was conducted on a desktop computer (CPU: Intel Core i5-7500, GPU: Intel HD Graphics 630 (Intel, Santa Clara, CA, USA)), Mouse: Logitech G203 (Logitech, Lausanne, Switzerland)) with an external monitor (HP E233, Hewlett-Packard, Palo Alto, CA, USA). The protocol was selected to be similar to the one applied in previous fatigue assessment studies [10]. During a 5-min test each subject performed about 75 simple RT measurements. Similar to multiple other RT measurement studies, in this study, the inter-stimulus interval was selected between 3 to 5 s [14].

2.6. Hand Grip Strength Measurement

Hand grip strength was measured using Grip Force Transducer dynamometer (MLT004/ST, ADInstruments, Sydney, Australia) with PowerLab 4/25T (ADInstruments, Sydney, Australia) data acquisition device and LabChart software (v. 8.1.13, ADInstruments). The subjects performed five maximal voluntary contractions with the dominant arm while seated. Hand grip strength was analyzed as the average of the maximums of the five repetitions.

2.7. Countermovement Jump Measurement

Each subject performed five maximal effort CMJ according to the recommended method [24]. The subjects were instructed to have their feet shoulder-width apart and hold their hands on their hips while performing the test. The performance was filmed and recorded at 60 frames per second with a camera (OnePlus 6, OnePlus Technology, Shenzhen, China), which was statically mounted at a fixed distance from the subject. Based on the recording, the height of each jump was calculated as the difference between the distance of a marked position (below the ribcage of the subject's torso) from the ground while standing and at the maximum jump height (Figure 2). The performance was estimated as the average of the jump heights of each CMJ repetition.

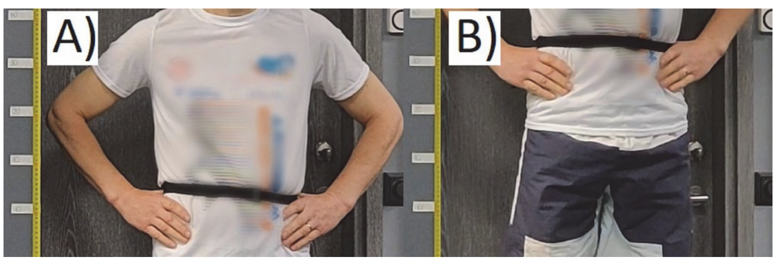

Figure 2. Countermovement jump (CMJ) height measurements. The subject is (**A**) standing and (**B**) at the maximum height. CMJ height was found as the vertical difference of the belt, which was firmly attached around the torso.

2.8. Veloergometer Test

Following CMJ measurements, a veloergometer test was performed on the Tunturi T6 Alpha 300 veloergometer (Tunturi, Turku, Finland). During the test, signals were measured to calculate different parameters based on heart electrical activity and pulse waveform in the resting and recovery phases between cycling. The test schedule is shown in Table 2. The subject sat on the veloergometer saddle calmly for at least 3 min prior to the heart rate measurement in a 5-min resting phase. Then the subject cycled at 60 rotations per minute on three different power levels (60 W, 90 W and 120 W) for three minutes with each level test followed by a 5-min recovery phase [18,25,26]. The power levels were manually changed during the recovery phases by the test supervisor.

Table 2. Schedule of the veloergometer test.

Activity	Duration in Minutes
Resting	5
Cycling @ 60 W	3
Recovery	5
Cycling @ 90 W	3
Recovery	5
Cycling @ 120 W	3
Recovery	5

2.9. Heart Electrical Activity Measurement

Subject ECG signals were recorded during the veloergometer test at the sampling rate of 1 kHz using the PowerLab 4/25T (ADInstruments, Sydney, Australia) data acquisition device and LabChart software (v. 8.1.13, ADInstruments). HR and HRV parameters were calculated based on the measured ECG signals to compare the heart electrical activity between the rested state in the morning and the physically-fatigued state in the afternoon. R-peaks of the ECG signals were detected by adopting the Hamilton–Tompkins algorithm [27], and manually verified in order to eliminate any errors. An example of subject HR in the veloergometer test is displayed in Figure 3.

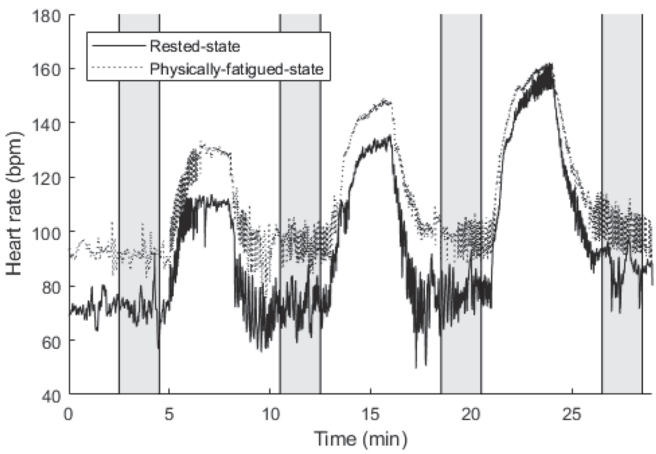

Figure 3. Heart rate of a subject during the veloergometer test in the rested state and the physically-fatigued state. The heart electrical activity parameters were calculated based on the grayed areas.

HRV parameters assessed in the test include time-domain measures SDNN (standard deviation of all NN intervals) and RMSSD (square root of the root mean square of the sum of all differences between successive NN intervals), which are the two commonly employed HRV parameters for the analysis of heart electrical activity [28]. The values of heart electrical activity parameters were calculated based on a 2-min signal section during the 'slow phase' of heart rate recovery [18], which was selected between 2.5 and 4.5 min during the resting phase and after cycling (grayed areas in Figure 3).

2.10. Pulse Arrival Time Measurement

While prior studies have explored PAT in exercise settings [29], this study is likely to be the first one to evaluate PAT for physical fatigue assessment. The pulse wave was registered synchronously with the heart electrical activity parameters employing the same sensing unit with an external piezoelectric transducer attached to the fingertip (MLT 1010 pulse transducer, ADInstruments, Sydney, Australia). PAT was found as the time difference between the ECG R-peak and the pulse wave signal rising front. PAT values were calculated based on a 1-min-signal section, selected between 3 and 4 min during the resting phase and after cycling. These sections and PAT of a sample subject are presented in Figure 4.

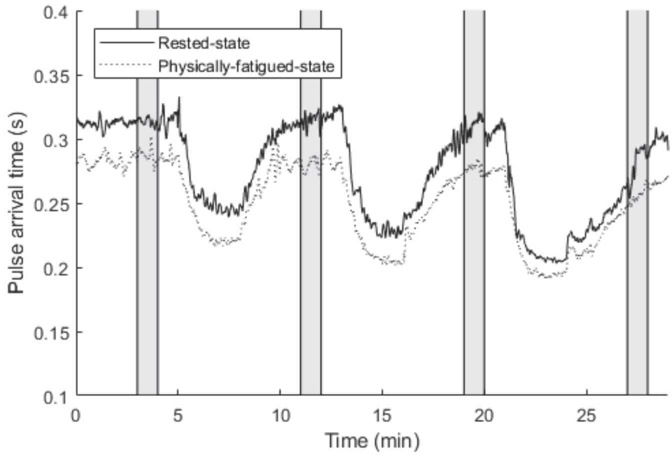

Figure 4. Pulse arrival time of a subject (before normalizing with blood pressure) during veloergometer test in the rested state and the physically fatigued state. Parameter values were calculated based on the grayed areas.

With the aim to reduce intersubject variability, PAT values were subsequently normalized based on blood pressure measurements, conducted in the same time sections [30]. The values were normalized to the systolic blood pressure of 120 mmHg, adopting the relation of every 1 mmHg difference causing 1 ms discrepancy in PAT [31].

2.11. Statistical Analysis

The test battery parameters were presented as mean +/− standard deviation (SD) over all subjects for the rested state and the physically-fatigued state. The paired *t*-test ($p < 0.05$) was used to find statistical difference between the measurements.

For every test battery parameter, the percentual change between the rested state and the physically-fatigued state was found individually for each subject. A linear correlation coefficient was calculated separately for each parameter pair to detect any linear relationship.

2.12. Grouping Based on Fatigue Levels

While all the subjects followed the same study protocol, they experienced different levels of physical fatigue based on their physiological and physical background. To distinguish between the mildly fatigued and significantly fatigued, the subjects were arranged into two groups in regard to their relative change in CMJ height between the rested state and the physically-fatigued state. From the measured reference parameters, CMJ height was selected since (i) two reference parameters, CMJ height and questionnaire score, had the ability to differentiate the rested state and the physically-fatigued state; (ii) compared to questionnaire score, CMJ height was seen as a more objective parameter.

All CV parameters (and multiple subparameters) were analyzed individually to test selectivity against the mildly fatigued and significantly fatigued groups. The total number of parameters analyzed was 74, including the values of SDNN, RMSSD, PAT and HR of different veloergometer test phases and subject fatigue states.

The parameters were evaluated by creating a decision stump using MATLAB's function *fitctree* that returns a fitted binary classification decision tree based on the input parameter. The leave-one-out cross-validation scheme was employed, where for each subject the decision stump was created based on the values measured from all other subjects. Applying this method, F-score and accuracy were found for the evaluation of each parameter.

Finally, a linear support vector machine (SVM) model was trained based on the two best performing parameters for binary classification between the mildly fatigued and the significantly fatigued groups. This classifier has been given as a valid example of a possible use of the findings of this study. The model was trained based on all the participants using MATLAB's function *fitcsvm* and the decision boundary was given using the formula:

$$w_1x_1 + w_2x_2 + b = 1, \quad (1)$$

where w_1 and w_2 are the coefficients for the parameters x_1 and x_2 and b is the bias term.

3. Results

3.1. Average Parameter Values

From the measured reference parameters, the questionnaire score had an increase of 15.2% for the whole study group, 17.9% for the female subgroup and 12.5% for the male subgroup for the rested vs. physically-fatigued states (see Table 3). This should confirm that the subjects felt more tired during the physically-fatigued state as compared to the rested state. CMJ height had a statistically significant decrease for the whole group −3.1% and for the female subgroup −2.7%, but interestingly not for the male subgroup (−1.9%). While the average RT increased by 1.1% and hand grip strength decreased −2.9% for the whole group, these changes were not found statistically significant ($p = 0.458$; $p = 0.113$).

Table 3. Reference parameter values. Average (mean ± SD) values for the reference parameters in the rested state (RS), physically-fatigued state (PFS) and their difference in percentage (DIF). Results are shown for the whole study group (A), female subgroup (F) and male subgroup (M). Q—questionnaire, RT—reaction time, DYN—dynamometer hand grip force and CMJ—countermovement jump height. Values marked with asterisk (*) indicate statistical difference (paired *t*-test, $p < 0.05$).

		Q (%)	RT (ms)	DYN (N)	CMJ (cm)
RS	A	14.0 ± 7.6	208.7 ± 11.3	360.3 ± 99.1	38.2 ± 8.7
	F	12.1 ± 9.4	206.8 ± 13.4	294.2 ± 47.4	33.1 ± 3.3
	M	15.8 ± 5.3	210.6 ± 9.4	426.4 ± 93.8	43.3 ± 9.7
PFS	A	29.2 ± 13.0	211.4 ± 16.9	349.7 ± 105.7	37.0 ± 9.0
	F	30.0 ± 17.7	211.7 ± 17.2	286.4 ± 48.7	31.6 ± 3.3
	M	28.3 ± 6.9	211.0 ± 17.9	413.0 ± 111.3	42.5 ± 9.8
DIF (%)	A	15.2% *	1.3%	−2.9%	−3.1% *
	F	17.9% *	2.4%	−2.7%	−4.5% *
	M	12.5% *	0.2%	−3.1%	−1.9%

For CV parameters (as in Table 4), only HR had a statistically significant increase for all study groups, including 9.5% for the whole group, 9.6% for the male subgroup and 9.4% for the female subgroup. Heart variability parameters SDNN (−21.2%, −19.6% and −23.2%, respectively) and RMSSD (−29.3%, −32.0% and −25.9%, respectively) had a statistically significant decrease for the whole group and the female subgroup, but not for the male subgroup ($p = 0.102$). The changes in average values of blood pressure normalized PAT (−2.0%, −4.7% and 0.7%) were not statistically significant ($p = 0.172$).

3.2. Correlation

Relatively strong linear correlation ($0.5 < R < -0.5$) was noted between several test battery measures (Table 5). For the whole group, these levels were found between HRV parameters SDNN and RMSSD (0.71); HR and SDNN (−0.61); HR and RMSSD (−0.57). For the male subgroup, the same levels were found between the questionnaire score and hand grip strength (0.74), CMJ height (−0.59); hand grip strength and CMJ height (−0.59), PAT (−0.80), HR (−0.53); SDNN and RMSSD (0.93), HR (−0.88); between RMSSD and HR (−0.80). For the female subgroup, linear correlation levels were found between questionnaire score and hand grip strength (−0.55); RT and hand grip strength (−0.80), SDNN (0.56), HR (−0.63); hand grip strength and SDNN (−0.74); PAT and HR (−0.79).

Table 4. Cardiovascular parameter values. Average (mean ± SD) values for the cardiovascular parameters in the rested state (RS), physically-fatigued state (PFS) and their difference (value, %) (DIF). Results are shown for the whole study group (A), female subgroup (F) and male subgroup (M). HR—average heart rate, SDNN—HRV parameter SDNN value, RMSSD—HRV parameter RMSSD and PAT—pulse arrival time. Values marked with asterisk (*) indicate statistical difference (paired *t*-test, $p < 0.05$).

		HR (bpm)	**SDNN (ms)**	**RMSSD (ms)**	**PAT (ms)**
RS	A	98.5 ± 10.9	58.0 ± 19.7	35.4 ± 18.9	273.4 ± 21.6
	F	100.6 ± 9.7	52.8 ± 13.3	31.7 ± 12.7	267.5 ± 15.7
	M	96.4 ± 12.3	63.2 ± 24.3	39.0 ± 24.0	279.4 ± 25.9
PFS	A	107.9 ± 12.2	45.7 ± 15.9	25.0 ± 13.8	268.1 ± 23.8
	F	110.1 ± 11.4	40.5 ± 16.0	23.5 ± 17.6	254.8 ± 11.8
	M	105.6 ± 13.4	50.8 ± 15.1	26.5 ± 9.5	281.3 ± 26.0
DIF (%)	A	9.5% *	−21.2% *	−29.3% *	−2.0%
	F	9.4% *	−23.2% *	−25.9% *	−4.7%
	M	9.6% *	−19.6%	−32.0%	0.7%

Table 5. Linear correlation coefficient values between different parameters based on the whole study group (A), male subgroup (M) and female subgroup (F). Parameter values are taken as the difference in % between those in the rested state and physically-fatigued state. Q—questionnaire, RT—reaction time, DYN—dynamometer hand grip force, CMJ—countermovement jump height, SDNN—HRV parameter SDNN, RMSSD—HRV parameter RMSSD, PAT—pulse arrival time and HR—average heart rate (between the resting heart rate and average veloergometer cycling heart rate). Values above 0.5 or below −0.5 are marked in bold.

		RT	**DYN**	**CMJ**	**SDNN**	**RMSSD**	**PAT**	**HR**
Q	A	0.36	−0.18	**−0.43**	0.13	0.02	−0.05	−0.13
	F	0.36	−0.55	−0.35	0.10	−0.07	0.18	0.02
	M	0.36	0.74	−0.59	0.36	0.26	−0.36	−0.49
RT	A		−0.24	−0.25	0.11	−0.03	0.03	−0.33
	F		−0.8	−0.04	0.56	0.03	0.41	−0.63
	M		0.40	−0.38	−0.13	0.01	−0.22	−0.04
DYN	A			**−0.10**	−0.24	0.11	**−0.39**	**−0.10**
	F			0.30	−0.74	0.19	−0.18	0.22
	M			−0.59	0.23	0.12	−0.80	−0.53
CMJ	A				−0.09	0.03	0.35	−0.09
	F				−0.29	0.23	0.26	−0.35
	M				−0.08	−0.18	0.25	0.18
SDNN	A					**0.71**	0.30	**−0.61**
	F					**0.10**	0.24	**−0.33**
	M					**0.93**	0.23	**−0.88**
RMSSD	A						0.24	**−0.57**
	F						−0.12	**−0.31**
	M						0.42	**−0.80**
PAT	A							−0.35
	F							−0.79
	M							0.13

3.3. Grouping Subjects Based on Fatigue States

In order to test how effectively the real-time measurable parameters differentiate between the mildly fatigued and significantly fatigued study groups, all these parameters (and multiple subparameters) were analyzed individually by conducting a classification task using only the chosen feature. Based on these findings, the two highest performing parameters were found as follows:

1. Relative change of the resting SDNN value normalized with the average recovery phase value between the rested-state and the physically-fatigued-state SDNN_DIF_N_AVG (F-score 0.842, accuracy 0.813) (Figure 5).
2. Resting PAT value normalized with the lowest recovery phase value during the physically-fatigued-state PAT_PFS_N_MIN (F-score 0.875, accuracy 0.875) (Figure 6).

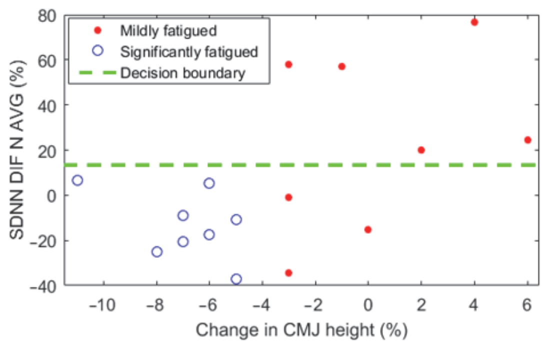

Figure 5. Change in CMJ height value compared to SDNN_DIF_N_AVG value. The subjects were arranged into the mildly fatigued and significantly fatigued groups. The shown threshold is found based on the data from all the subjects.

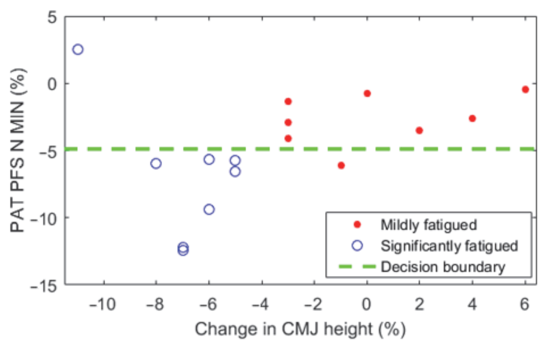

Figure 6. Change in CMJ height value compared to PAT_PFS_N_MIN value. The subjects were arranged into the mildly fatigued and significantly fatigued groups. The shown threshold is found based on the data from all the subjects.

These two parameters were used to train a linear SVM to classify new subjects or users into the mildly fatigued group or the significantly fatigued group (Figure 7). The linear SVM obtained demonstrates a promising ability to classify between the mildly fatigued or significantly fatigued physical states. The decision boundary can be described with the following formula according to (1):

$$-0.0261x_1 - 0.3366x_2 - 1.6558 = 0 \tag{2}$$

where x_1 is the parameter SDNN_DIF_N_AVG and x_2 is the parameter PAT_PFS_N_MIN.

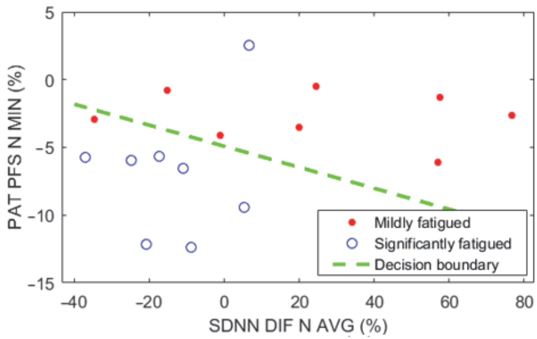

Figure 7. Linear SVM model for binary classification between the mildly fatigued and significantly fatigued groups.

4. Discussion

This study has evaluated how exercise induced physical fatigue affects various test battery measures, and whether real-time measurable cardiovascular (CV) parameters could provide sufficient data to classify between the mildly fatigued and significantly fatigued groups with the aim to provide information for real-time physical fatigue assessment. The main findings can be listed as follows: (i) from the assessed cardiovascular parameters, the statistically significant change between the rested state and physically-fatigued state was noted in the average heart rate and heart rate variability measures SDNN and RMSSD; (ii) the strongest linear correlation was found between the reference parameter hand grip strength and CV parameter pulse arrival time (PAT); (iii) the finest CV parameters for separating the mildly fatigued and significantly fatigued groups were based on heart rate variability (HRV) parameter SDNN between the rested state and the physically-fatigued state and PAT changes during the physically-fatigued state.

While most study parameters were selected drawing on the findings of previous studies, not all of the parameters were found significant based on the results of this study. From the reference parameters, the score of fatigue questionnaire showed a statistically significant increase (of about 15.2%) between the rested-state and the physically-fatigue-state data, which is consistent with the previous findings [3]. Countermovement jump (CMJ) height decrease was statistically different for the whole study group (average decrease of 3.1%) and the female subgroup, remaining in the same range as found in the previous studies [3,10,16]. In accordance with the previous studies, the average value of reaction time (RT) increased (1.3%) [10,14,15] and hand grip strength decreased (−2.9%) [10,16]; however, these changes were not found statistically significant. The results of the hand grip strength test could be explained by the full-body workout regime that did not involve a sufficient number of exercises for the specific arm muscles. It was expected that RT would decrease due to physical fatigue [13]; however, the present study did not reach such a result. A possible explanation for the difference in the data obtained in previous studies may be that different modes of physical work or exercise were used. It is also likely that the fatigue level during physically-fatigued-state measurements was not sufficiently high to limit the inhibitory control of some subjects [32].

From the evaluated CV parameters, the average heart rate had a statistically significant increase of 9.5%, which is in accordance with the previous studies [1,17]. The HRV parameters SDNN and RMSSD decreased by 21.2% and 29.3%, respectively, between the rested-state and the physically-fatigued-state measurements, which has also been noted by other studies [1,2,18]. The presence of some outliers in the results of the effects of fatigue on the cardiovascular parameters may be due to a faster and better recovery of these subjects compared to others [33]. It is interesting to note that PAT, which is a novel parameter for physical fatigue assessment studies, had a decrease of 2.0% for the whole group and 4.7% for the female subgroup, but for the male subgroup the value increased by 0.7%.

The linear correlation coefficient was found based on the relative individual changes between all measures. The strongest correlation between CV and reference parameter for the whole group was found between hand grip strength and PAT (linear correlation coefficient of -0.39). This finding was consistent with the date obtained in the male subgroup, where the linear correlation coefficient was -0.80; however, for the female subgroup the strongest correlation was found between HRV measure SDNN and hand grip strength (-0.79).

In total, 74 different subparameters were evaluated based on how effectively they classify between the mildly fatigued and significantly fatigued study groups compiled using the relative change in CMJ height value. These parameters were found using SDNN, RMSSD, PAT and HR values from different veloergometer test phases and subject fatigue states. The finest parameters for differentiating between these groups were the relative change of the resting SDNN value normalized with average recovery phase value between the rested-state and physically-fatigued-state measurements (SDNN_DIF_N_AVG; F-score 0.842, accuracy 0.813) and the resting PAT value normalized with the lowest recovery phase value during the physically-fatigued-state (PAT_PFS_N_MIN; F-score 0.875, accuracy 0.875). A simple linear support vector machine model was trained based on these two parameters to provide a valid example of a possible application of these results. This model, when implemented into a real-time monitoring system, has a potential to reveal whether the user is 'mildly fatigued' or 'significantly fatigued' after a physically demanding day.

Based on the findings of this study, it can be concluded that the test battery devised for the experiment has an added value for the assessment of physical fatigue. The evaluated CV parameters showed promising results compared to the reference parameters and thus could be used for real-time physical fatigue monitoring in workplace settings and for the general population. Novel parameters based on PAT were also found to provide additional information for the improvement of the overall quality of physical fatigue assessment.

This study has encountered several limitations that should be considered in the following studies, namely: (i) as the study was conducted on a relatively small number of subjects, the results should be verified on a larger study group; (ii) since the induced physical fatigue is likely to have been specific to the used workout and study protocol, it should be further explored how different workout regimes affect the results; (iii) the experiments were conducted in the lab settings and should thus be verified in real-life situations; (iv) it is possible that extra information for fatigue assessment could be obtained by additional exploration into frequency-domain parameters. While the above limitations should not reduce the impact of this novel study, further research is required to fully evaluate the effectiveness of the test battery compiled for the study in order to determine the sensitivity of variables with the accumulation of fatigue.

5. Conclusions

This study has proposed a novel method for real-time physical fatigue assessment employing a set of real-time and easily measurable cardiovascular (CV) parameters. Induced physical fatigue was found to cause statistically significant change in the score of the fatigue questionnaire and countermovement jump height. For the cardiovascular parameters assessed, the statistically significant change was noted in the average heart rate and heart rate variability measures SDNN and RMSSD. The subparameters based on heart rate variability and blood pressure normalized pulse wave arrival time showed the highest performance in classifying between the mildly fatigued and the significantly fatigued groups. The findings of the study can be used to enhance the existing physical fatigue assessment methods and provide a solid ground for further research in the development of a continuous real-time fatigue monitoring system. Further studies with a larger study group will be required to verify the obtained findings in multiple real-life situations.

Supplementary Materials: The following materials are available online at https://www.mdpi.com/article/10.3390/s22041680/s22041680/s1, Figure S1: The Swedish Occupational Fatigue Inventory questionnaire [21], used to evaluate current subjective fatigue levels of the subjects.

Author Contributions: Conceptualization, A.A., K.P. and I.F.; methodology, A.A, K.P., M.V., I.F. and G.J.; software, A.A. and K.P.; formal analysis, A.A., K.P., M.V., I.F. and G.J.; investigation, A.A., K.P. and I.F.; resources, I.F. and G.J.; data curation, A.A.; writing—original draft preparation, A.A.; writing—review and editing, A.A., K.P., M.V., I.F. and G.J.; visualization, A.A.; supervision, K.P., I.F. and G.J.; project administration, I.F. and G.J. All authors have read and agreed to the published version of the manuscript.

Funding: The research was funded partly by the Estonian Ministry of Education and Research under institutional research financing IUTs 19-1 and 19-2, by European Commission H2020 project 872614 - SMART4ALL and by Estonian Centre of Excellence in IT (EXCITE) funded by European Regional Development Fund.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Tallinn Medical Research Ethics Committee (protocol no. 1954; 15 June 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the subjects to publish this paper.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mohan, J.; Lamshe, R.; Poonguzhali, S.; Adalarasu, K.; Jagannath, M. Assessment of Human Fatigue during Physical Performance using Physiological Signals: A Review. *Biomed. Pharmacol. J.* **2017**, *10*, 1887–1896. [CrossRef]
2. Shortz, A.E.; Franke, M.; Kilic, E.S.; Peres, S.C.; Mehta, R.K. Evaluation of Offshore Shiftwork using Heart Rate Variability. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2017**, *61*, 1036–1039. [CrossRef]
3. Dawson, D.; Chapman, J.; Thomas, M.J. Fatigue-proofing: A new approach to reducing fatigue-related risk using the principles of error management. *Sleep Med. Rev.* **2012**, *16*, 167–175. [CrossRef] [PubMed]
4. Shahraki, S.; Bin Abu Bakar, N. Effects of nervous fatigue on workforce productivity. *Int. J. Acad. Res.* **2011**, *3*, 370–378.
5. Yu, Y.; Li, H.; Yang, X.; Kong, L.; Luo, X.; Wong, A.Y. An automatic and non-invasive physical fatigue assessment method for construction workers. *Autom. Constr.* **2019**, *103*, 1–12. [CrossRef]
6. Finsterer, J.; Mahjoub, S.Z. Fatigue in Healthy and Diseased Individuals. *Am. J. Hosp. Palliat. Med.* **2014**, *31*, 562–575. [CrossRef]
7. Stadje, R.; Dornieden, K.; Baum, E.; Becker, A.; Biroga, T.; Bösner, S.; Haasenritter, J.; Keunecke, C.; Viniol, A.; Donner-Banzhoff, N. The differential diagnosis of tiredness: A systematic review. *BMC Fam. Pr.* **2016**, *17*, 147. [CrossRef] [PubMed]
8. Saito, K. Measurement of Fatigue in Industries. *Ind. Health* **1999**, *37*, 134–142. [CrossRef]
9. Hughes, S.; Chapman, D.W.; Haff, G.G.; Nimphius, S. The use of a functional test battery as a non-invasive method of fatigue assessment. *PLoS ONE* **2019**, *14*, e0212870. [CrossRef]
10. Thompson, B.J. Does work-induced fatigue accumulate across three compressed 12 hour shifts in hospital nurses and aides? *PLoS ONE* **2019**, *14*, e0211715. [CrossRef]
11. Kumari, P.; Mathew, L.; Syal, P. Increasing trend of wearables and multimodal interface for human activity monitoring: A review. *Biosens. Bioelectron.* **2017**, *90*, 298–307. [CrossRef] [PubMed]
12. Seneviratne, S.; Hu, Y.; Nguyen, T.; Lan, G.; Khalifa, S.; Thilakarathna, K.; Hassan, M.; Seneviratne, A. A Survey of Wearable Devices and Challenges. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 2573–2620. [CrossRef]
13. Allik, A.; Pilt, K.; Viigimäe, M.; Fridolin, I. Pilot Study for Estimating Physical Fatigue Based on Heart Rate Variability and Reaction Time. In Proceedings of the 2019 15th Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON), Coimbra, Portugal, 26–28 September 2019; Volume 76, pp. 193–200.
14. Sabzi, A.H. The Effect of Different Fatigue Protocols on Choice Reaction Time. *MEJSR* **2012**, *12*, 1092–1096. [CrossRef]
15. Sant'Ana, J.; Franchini, E.; Da Silva, V.; Diefenthaeler, F. Effect of fatigue on reaction time, response time, performance time, and kick impact in taekwondo roundhouse kick. *Sports Biomech.* **2016**, *16*, 201–209. [CrossRef]
16. Thompson, B.J.; Stock, M.; Banuelas, V.K. Effects of Accumulating Work Shifts on Performance-Based Fatigue Using Multiple Strength Measurements in Day and Night Shift Nurses and Aides. *Hum. Factors J. Hum. Factors Ergon. Soc.* **2017**, *59*, 346–356. [CrossRef] [PubMed]
17. Thomson, R.L.; Rogers, D.K.; Howe, P.R.; Buckley, J.D. Effect of acute exercise-induced fatigue on maximal rate of heart rate increase during submaximal cycling. *Res. Sports Med.* **2015**, *24*, 1–15. [CrossRef]

18. Peçanha, T.; Bartels, R.; Brito, L.; Paula-Ribeiro, M.; Oliveira, R.S.; Goldberger, J.J. Methods of assessment of the post-exercise cardiac autonomic recovery: A methodological review. *Int. J. Cardiol.* **2017**, *227*, 795–802. [CrossRef]
19. Guidi, A.; Greco, A.; Felici, F.; Leo, A.; Ricciardi, E.; Bianchi, M.; Bicchi, A.; Valenza, G.; Scilingo, E.P. Heart rate variability analysis during muscle fatigue due to prolonged isometric contraction. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea, 11–15 July 2017; Volume 2017, pp. 1324–1327.
20. Shaffer, F.; Ginsberg, J.P. An Overview of Heart Rate Variability Metrics and Norms. *Front. Public Health* **2017**, *5*, 258. [CrossRef]
21. Ahsberg, E.; Gamberale, F.; Gustafsson, K. Perceived fatigue after mental work: An experimental evaluation of a fatigue inventory. *Ergonomics* **2000**, *43*, 252–268. [CrossRef]
22. Khitrov, M.Y.; Laxminarayan, S.; Thorsley, D.; Ramakrishnan, S.; Rajaraman, S.; Wesensten, N.J.; Reifman, J. PC-PVT: A platform for psychomotor vigilance task testing, analysis, and prediction. *Behav. Res. Methods* **2014**, *46*, 140–147. [CrossRef]
23. Reifman, J.; Kumar, K.; Khitrov, M.Y.; Liu, J.; Ramakrishnan, S. PC-PVT 2.0: An updated platform for psychomotor vigilance task testing, analysis, prediction, and visualization. *J. Neurosci. Methods* **2018**, *304*, 39–45. [CrossRef] [PubMed]
24. Petrigna, L.; Karsten, B.; Marcolin, G.; Paoli, A.; D'Antona, G.; Palma, A.; Bianco, A. A Review of Countermovement and Squat Jump Testing Methods in the Context of Public Health Examination in Adolescence: Reliability and Feasibility of Cur-rent Testing Procedures. *Front. Physiol.* **2019**, *10*, 1384. [CrossRef] [PubMed]
25. Ng, J.; Sundaram, S.; Kadish, A.H.; Goldberger, J.J. Autonomic effects on the spectral analysis of heart rate variability after exercise. *Am. J. Physiol. Circ. Physiol.* **2009**, *297*, H1421–H1428. [CrossRef] [PubMed]
26. Goldberger, J.J.; Le, F.K.; Lahiri, M.; Kannankeril, P.J.; Ng, J.; Kadish, A.H. Assessment of parasympathetic reactivation after exercise. *Am. J. Physiol. Circ. Physiol.* **2006**, *290*, H2446–H2452. [CrossRef]
27. Hamilton, P.S.; Tompkins, W.J. Quantitative Investigation of QRS Detection Rules Using the MIT/BIH Arrhythmia Database. *IEEE Trans. Biomed. Eng.* **1986**, *33*, 1157–1165. [CrossRef] [PubMed]
28. Achten, J.; Jeukendrup, A.E. Heart Rate Monitoring. *Sports Med.* **2003**, *33*, 517–538. [CrossRef]
29. Liu, Q.; Poon, C.; Zhang, Y. Time–frequency analysis of variabilities of heart rate, systolic blood pressure and pulse transit time before and after exercise using the recursive autoregressive model. *Biomed. Signal Process. Control* **2011**, *6*, 364–369. [CrossRef]
30. Mukkamala, R.; Hahn, J.-O.; Inan, O.T.; Mestha, L.K.; Kim, C.-S.; Toreyin, H.; Kyal, S. Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 1879–1901. [CrossRef]
31. Muehlsteff, J.; Aubert, X.A.; Morren, G. Continuous cuff-less blood pressure monitoring based on the pulse arrival time approach: The impact of posture. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2008**, *2008*, 1691–1694.
32. Mlynski, C.; Reza, A.; Whitted, M.; Cox, C.; Garsea, A.; Wright, R.A. Fatigue influence on inhibitory control: Cardiovascular and performance findings elucidate the role of restraint intensity. *Psychophysiology* **2021**, *58*, e13881. [CrossRef]
33. Wright, R.A.; Shim, J.J.; Hogan, B.K.; Duncan, J.; Thomas, C. Interactional influence of fatigue and task difficulty on cardiovascular response: Demonstrations involving an aerobic exercise challenge. *Psychophysiology* **2012**, *49*, 1049–1058. [CrossRef] [PubMed]

Curriculum vitae

Personal data

Name:	Ardo Allik
Date of birth:	3rd November 1990
Place of birth:	Tamsalu, Lääne-Viru County, Estonia
Citizenship:	Estonian

Contact data

E-mail:	ardo.allik@taltech.ee

Education

2016-2022	PhD, Information and Communication Technology, Tallinn University of Technology
2014–2016	MSC, Biomedical Engineering and Medical Physics, Tallinn University of Technology
2010–2014	BSC, Engineering Physics, Tallinn University of Technology
2007–2010	Highschool, Tallinn Secondary Science School

Language competence

Estonian	Native
English	Fluent
Russian	Basic

Professional employment

2021–...	Tallinn University of Technology, School of Information Technologies, Department of Health Technologies, Engineer
2016–2020	Tallinn University of Technology, School of Information Technologies, Department of Health Technologies, Junior Researcher
2014–2015	North Estonian Regional Hospital, Biomedical engineer
2012–2014	Tallinn University of Technology, Institute of Marine Systems, Engineer

Elulookirjeldus

Isikuandmed

Nimi:	Ardo Allik
Sünniaeg:	03.11.1990
Sünnikoht:	Tamsalu, Lääne-Virumaa, Eesti
Kodakondsus:	Eesti

Kontaktandmed

E-post:	ardo.allik@taltech.ee

Hariduskäik

2016–2022	Doktoriõpe, Info- ja kommunikatsioonitehnoloogia, Tallinna Tehnikaülikool
2014–2016	Magistriõpe, Biomeditsiinitehnika ja meditsiinifüüsika, Tallinna Tehnikaülikool
2010–2014	Bakalaureuseõpe, Tehniline füüsika, Tallinna Tehnikaülikool
2007–2010	Keskharidus, Tallinna Reaalkool

Keelteoskus

Eesti keel	Emakeel
Inglise keel	Kõrgtase
Vene keel	Algtase

Teenistuskäik

2021–...	Tallinna Tehnikaülikool, Infotehnoloogia teaduskond, Tervisetehnoloogiate instituut, Inserner
2016–2020	Tallinna Tehnikaülikool, Infotehnoloogia teaduskond, Tervisetehnoloogiate instituut, Nooremteadur
2014–2015	SA Põhja-Eesti Regionaalhaigla, Biomeditsiinitehnika insener
2012–2014	Tallinna Tehnikaülikool, TTÜ Meresüsteemide Instituut, Tehnik

ISSN 2585-6901 (PDF)
ISBN 978-9949-83-914-8 (PDF)