André Costa Nunes Campos

Número de estudante 44957 (MIEI)

MLFatigueDetection
Machine Learning Based Walking Fatigue Detection

Dissertação para obtenção do Grau de Mestre em

Engenharia Informática

Orientador: Nuno Cardoso, Researcher,
Fraunhofer Portugal

Co-orientadora: Susana Maria S. Nascimento M. Almeida,
Associate Professor,
Faculdade de Ciências e Tecnologia
da Universidade Nova de Lisboa

Júri

Presidente: Doutor Carlos Augusto Isaac Piló Viegas Damásio
Vogais: Doutor Ricardo João Rodrigues Gonçalves
Mestre Nuno Filipe de Oliveira Cardoso

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

December, 2019

MLFatigueDetection
Machine Learning Based Walking Fatigue Detection

Este documento foi gerado utilizando o processador (pdf)LaTeX, com base no template "novathesis" [1] desenvolvido no Dep. Informática da FCT-NOVA [2].
[1] https://github.com/joaomlourenco/novathesis [2] http://www.di.fct.unl.pt

To my mother who always supported me and encouraged me to give my best in everything I do.

Acknowledgements

I would like to thank Associação Fraunhofer Portugal Research for giving me this opportunity to enrich myself both professionally and personally. A special mention to my advisor Nuno Cardoso for guiding me through the end of my student life and the start of my professional career. I'm incredibly grateful for his advice, his patience and his availability to help me through whatever I needed.

I would also like to thank professor Susana Nascimento, who was also always available to help me and clarify any questions I had, especially during the writing and revision of this dissertation.

Lastly, I would like to give a big thank you to my family, to whom I was not the most available person the past few months, and my friends who dealt with my disappearances weeks at a time.

Abstract

Leg fatigue can influence the gait patterns, therefore declining the postural stability and the motor performance, increasing the risk of falls. In order to improve the earlier detection of risks and the application of fall prevention strategies, automated solutions based on gait analysis must be developed. A sector of the population at risk is the workforce where a majority of workers admits to be fatigued and where falls can lead to serious workplace injuries or even deaths. In these cases, having the ability to detect if the user is fatigued in real time by simply using the motion sensors on the smartphone and processing it with machine learning can lead to the prevention of falls and the consequences these bring.

Phones and wearable devices were studied for their ability to be used to extract inertial sensor's data to provide enough information for the fatigue detection. Supervised machine learning algorithms, such as Support Vector Machines (SVM) and Neural Networks, will be used to process this information for fatigue level classification. Their performance will then be compared to find the best algorithm for fatigue detection. In addition to this comparative work, different conditions for the data collection and processing were tested in an effort to discover the optimal conditions for the implementation of the algorithms.

Keywords: Gait Patterns, Fall risk, Fall prevention, Fatigue, Inertial sensors, Supervised learning.

Resumo

Situações de fadiga nas pernas podem influenciar os padrões de marcha e, como tal, reduzir a estabilidade postural e a performance motora, aumentando assim o risco de quedas. Um sector da população em risco são os trabalhadores, onde a maioria admite estar fatigado e onde quedas podem levar a ferimentos graves ou mesmo mortes em ambiente de trabalho. Nestes casos, a habilidade de detetar se um utilizador se encontra fatigado em tempo real, usando apenas os sensores inerciais do smartphone e processá-lo com aprendizagem automática, pode levar à prevenção das quedas e das consequências que estas trazem.

Telemóveis e dispositivos *wearable* foram estudados como ferramentas para extrair dados de sensores inerciais, de forma a providenciar informação suficiente para a deteção de fadiga. Algoritmos de aprendizagem automática supervisionada, tais como Máquina de Vetores de Suporte (SVM) e redes neuronais, foram usados para processar esta informação para classificação da presença de fadiga. A performance destes algoritmos foi então comparada para descobrir o melhor algoritmo para deteção de fadiga. Em conjunto com este trabalho comparativo, diferentes condições de recolha e processamento de dados foram testadas num esforço para descobrir as melhores condições para a implementação dos algoritmos.

Palavras-chave: Padrões de marcha, Risco de queda, Prevenção de queda, Fadiga, Sensores inerciais, Aprendizagem supervisionada.

Contents

List of Figures

Acronyms

ANN	Artificial Neural Network.
HCV	Heel Contact Velocity.
IMC	Inertial Motion Capture.
IMU	Inertial Measurement Unit.
IoTiP	Internet of things in Package.
LOGO	Leave One Group Out.
MMH	Manual Material Handling.
MVE	Maximum Voluntary isokinetic Exertions.
OMC	Optical Motion Capture.
OSH	optimal Separating Hyperplane.
PCA	Principle Component Analysis.
RBF	Radial Basis Function.
RGC	Representative Gait Cycle.
ROM	Range of Motion.
RPE	Rating of Perceived Exertion.
SFL	Subjective Fatigue Level.
SL	Step Length.
SOM	Self-organising Maps.
SVM	Support Vector Machines.
TSFEL	Time Series Feature Extraction Library.

Chapter 1

Introduction

1.1 Motivation

Gait is defined in the dictionary as

> *a manner of walking or moving on foot*
>
> *Gait [Def. 1]. (n.d.). Merriam-Webster Online. In Merriam-Webster. Retrieved March 13, 2019, from https://www.merriam-webster.com/dictionary/gait.*

in other words, it is the most basic human locomotion tool. It is easily understandable how important it is in our everyday lives and the importance of studies done around it and on the effects of alterations to it.

If we think of a working environment, there are many physically demanding occupations, in which walking is still a primary activity, such as manufacturing, construction, agriculture and others. The high physical demands that accompany these occupations are bound to be the cause of fatigue in its workers. As such, when surveyed, approximately 45% of US manufacturing workers reported to be fatigued, in consequence of high amount of walking required in their occupation. The same survey showed, per shift, an average of 5.7h spent walking in the workforce [14].

Past studies indicate the relation between the presence of fatigue and specific alteration in gait patterns. It's important to understand these alterations to be able to identify them on the data that will be collected. In situations of fatigue, previous studies found an increase in step width, higher than double jerk cost and greater resulting acceleration [26]. In accordance with these results, other studies [9] found increases in step width and mediolateral trunk acceleration, paired with an increase in step length variability in fatigue situations. However, there are also contradictory finds; results have been obtained

[17] that indicate that step length variability appears less sensitive to fatigue, when compared to step width variability. Even with some minor inconsistencies in the findings, generally these studies are able to find multiple changes in gait patterns when fatigue is induced, which is a good basis for monitoring fatigue and creating fatigue detection platforms [17].

The utilization of machine learning algorithms to assist in the classification of altered gait patterns has been a topic studied by various authors [4, 9, 26] . Most studies have found success in identifying the alterations in gait patterns they were searching for, using said algorithms. One of the classification algorithms used in the area of fatigue detection is SVM, which was found to hold considerable potential to identify at-risk gait due to muscle fatigue [26].

1.2 Technology

In the area of gait pattern identification and studying, Optical Motion Capture (OMC) systems have been widely used for the acquisition of the necessary data, while in recent years developments in wearable technology propelled an increase in the number of studies using Inertial Motion Capture (IMC) systems, both as just a means of collecting data or as the focus of the study, comparing different methods of data gathering. The highest advantage of IMC systems is the possibility of studies on daily life conditions, without the burden of being stuck to the space where the equipment is set up. This can lead to solutions that can be implemented in the area of diagnosis and prevention.

The past studies that utilised OMC methods have resorted to equipment varying from simple cameras, pointed towards a treadmill [4], to full motion capture systems [13]. In all cases, this limits the environment where the data can be collected and, also due to the high costs of the needed equipment, heavily restricts the possibilities of real world solutions. IMC methods bring portability to the table and even though some of the more high-end equipment used in some studies, such as the MVN Link [15], still bring impeditive costs for broad-wide solutions, other studies have been made with a single wearable [2] or even a smartphone [19] for data collection and still obtaining positive results.

This widespread availability of wearable technology and usage of smartphones paired with the development of the technology present inside these devices shows a promising opening in new possibilities in terms of monitoring and prevention. Bearing this is mind, the data necessary for this dissertation was acquired from the sensors present in smartphones and wearables. This decision guarantees the possibility of extending the usefulness of the work developed on this dissertation to a real world scenario and allows it to be integrated in already existent solutions as future work.

1.3 Objectives

The main objective of the dissertation is the comparison of the performance of supervised learning algorithms when applied to the identification of fatigue from gait data. Most previous works tend to focus on whether the algorithm is capable of identifying fatigued gait; this work tries to go further by focusing instead on which the best conditions for this identification are.

Two separate comparative studies were made in an effort to reach the best possible results. The first, and most important, of the two is the comparative study of the supervised learning algorithms. For this, two algorithms, Neural Networks and SVM, and three of SVM's kernels, were tested. The performance of the algorithms was measured by their accuracy, sensitivity and specificity.

The second comparative study is on what the best locations to place the sensors are when only two maximum sensors can be used. For this, multiple sensors will be placed on various body parts of the participants, allowing for the analysis of which locations are more sensitive to changes in gait patterns and also for better classification results. This second comparative study allows to differentiate from previous studies where the placement of the devices was decided before the data collection and its influence was not questioned.

1.3.1 Main Contributions

In this dissertation, a comparison study was made on the classification of fatigue in gait using different machine learning algorithms and kernels, obtaining high results on the ability to distinguish fatigued and non-fatigued gait data. Good results were achieved with all algorithms and kernels, with the Radial Basis Function (RBF) kernel of the SVM classifiers being considered as the best performing algorithm/kernel combination. Along with the polynomial kernel, both achieved results above 96.5% in all metrics, except in the test where only the accelerometer data was used in which the performance of the polynomial kernel dropped to around 90%.

Not only that, but the best conditions to collect and process the data to achieve the best possible classification were also studied. This permitted to reach conclusions, such as the possibility to successfully collect the necessary data with just a single device. Another conclusion taken is on the best window duration to divide the data on. It was found that a 5 second window for data splitting obtains the best results in avoiding false positives, an important factor in any possible future use of the work developed in this dissertation.

1.4 Organization of the document

In the next chapters the concepts introduced in this introduction will be further explained and substantiated, culminating in a presentation of the practical work done, results obtained and conclusions gathered. In this section I want to take the time to briefly explain the points to address in each of the following chapters.

The next chapter is dedicated to the review of the state of the art, particularly an in depth analysis of four papers, the objectives proposed, the methods used to collect and process the data and the results obtained.

The third chapter will serve to explore theoretical concepts used through the elaboration of the dissertation. This includes an analysis of the importance of gait patterns and the expected changes in these patterns with fatigue. Also including a few important definitions on the field and an explanation of the possible causes and effects of the changes in gait patterns observed in previous studies and expected in this dissertation. After the introduction of necessary information about gait, the following section will be dedicated to the technology that was used, containing an explanation of the function of the necessary sensors on the collection of the data. A brief discussion of the use of the smartphone or wearable sensors to collect the data will also be present in this section. In the last section of this chapter there will be an introduction to Machine learning, followed by an in-depth analysis of the Machine learning algorithms that were used in the dissertation. As the goal of this dissertation was to compare the results achieved by multiple algorithms, a good understanding of the algorithms is essential as a basis for the work to follow.

In the fourth chapter, the practical work done in this dissertation is explained in multiple steps, starting with the elaboration of the experimental protocol and the decisions that were made during this process. This is followed by the demographic information of the participants, the steps taken to process the data and ending with a section explaining approaches attempted during the dissertation that didn't lead to positive results.

The fifth chapter shows the results obtained with the different tests during this dissertation. It is split into a section with all the variations of the duration of windows into which the data is split, a section with variations to other variables and finally a discussion of the results obtained.

In the last chapter some final conclusions are presented followed by a description of possible future work based on limitations found throughout the elaboration of the dissertation.

Chapter 2

State of the art

In this chapter, as the name implies, an analysis of previous studies will be done, in regards to the instrumentation and methods used and the results obtained.

Four studies were chosen for analysis, three of which used machine learning algorithms to identify fatigue in gait, the first two using SVM, while the fourth paired the SVM classification with a neural network in its implementation. Although the third study only targets the effects of fatigue on gait patterns, it was studied nevertheless as it was considered a relevant study in this area. At the end of the chapter, after a detailed description of each study, a short summary is presented with the methods used in each one and the results achieved.

2.1 Multiple inertial sensors and classification with SVM

In this study by Jian Zhang, Thurmon E. Lockhart and Rahul Soangra [26], wearable sensors, paired with other methods, were used to collect the gait and gait pattern data, processing it using an SVM classifier, testing different kernels, and obtained an accuracy of 96%. The objective of this study was to explore the potential of SVM to recognise and classify gait patterns associated with lower extremity muscular fatigue, using an inertial measurement unit.

Participants in the study were screened to avoid participants with existing factors that could influence gait patterns, such as the use of medication, the existence of neuromuscular diseases and balance or vision disorders. To collect the data multiple methods were used. Firstly, the trials were performed on a linear walkway (15.5 x 1.5m) with two force plates installed in the middle of it. Participants were equipped with 5 reflective markers, attached on heels and toes of both legs with one more at the sacrum. This was used to collect three-dimensional movement data using a six-camera ProReflex system

(Qualysis). Lastly, two IMU were attached to the participants, the first in the right shank with the objective of normalizing the gait data and the second at the sternum level (Figure 2.1). The data for this study was acquired with sampling frequency of 120 Hz, which was considered sufficient for human movement analysis in daily activities.

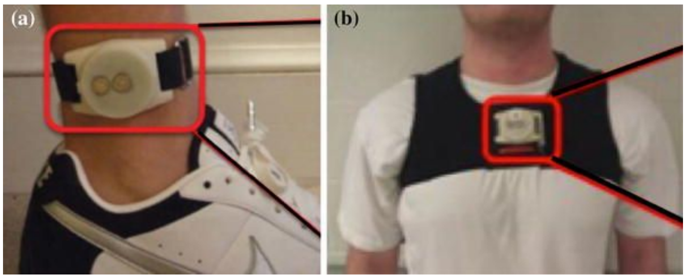

Figure 2.1: IMUs attached to participants in this first study

The reflective markers on both feet allowed the detection of the Step Length (SL), step width, Heel Contact Velocity (HCV) and single stance time, while the function of the one on the sacrum was to determine walking velocity. SL was calculated in this study using the points of the heel contacting the floor. Step width refers to the distance between the rear-end center lines of the heels. While HCV was calculated by using velocities of the heel in the horizontal direction at the foot dislocation of 1/60 s before and after the heel contact phase of the gait cycle.

The heel contact and toe-off time events necessary for these calculations were confirmed with the ground reaction forces measured using the forceplates afore mentioned. To guarantee that the measures could be confirmed, participants were asked to redo trials where the foot placement was not accurate with the center of the forceplates. In this study five accepted walking trials were required per participant per state (non-fatigued, fatigued), with 6 to 7 complete gait cycles per trial.

The fatiguing task used in this study was squatting; the participants were asked to perform squats, while holding a weight equivalent to 5% of their body weight in front of themselves, repeatedly, at a rhythm of 22 repetitions per minute. This task was divided in sets of 5 minutes, and after each set three Maximum Voluntary isokinetic Exertions (MVE) measurements were made. This exercise cycle kept repeating until participants reached 60% of their baseline MVE, at which point participants were considered to be fatigued. After this point participants were asked to repeat the walking trials to collect the fatigued data, in a process equal to its non-fatigued counterpart.

The SVM classifier input used was the data from the IMU located at the sternum, the Representative Gait Cycle (RGC). RGC was seen as the period between two consecutive contacts of one foot to the ground, depicting the duration of a stride. This duration was identified by the angular velocity profiles of the IMU placed in the shank. This way RGC was considered to start when the right shank angular velocity reached a peak and to end when a consecutive peak was reached. The IMU signals from the sternum were then cut between the RGC and normalized, 0% being the start of the RGC and 100% the end.

Classification wise, both data sets, training and testing, included fatigued and non-fatigued RGC data. As mentioned, each trial consisted of 6-7 gait cycle, of which two middle RGCs data were extracted from each trial, totalling twenty RGCs extracted (ten RGCs from five trials of each state). For both intra and inter-subject classifications the training and testing data split was kept at 70/30(%). For intra-subject classification 7 RGCs of each state (fatigued, non-fatigued) were used for training and 3 of each for testing. Inter-subject classification used 238 RGCs for training and 102 for testing data sets.

The features selected for this study can be seen in figure 2.2. The justification behind this selection was the inclusion of all possible spatial and temporal information from the signal, all features used were extracted from raw signals.

	General features	**Domain knowledge based selected features**
Data input for feature extraction	Accelerometer (A_x, A_y, A_z) and gyroscope (G_x, G_y, G_z) signals in all three directions of normalized RGC	Resultant acceleration $\left(R=\sqrt{A_x^2+A_y^2+A_z^2}\right)$ Resultant Jerk $\left(J=\frac{dR}{dt}\right)$
	Mean	Resultant acceleration features
	Standard deviation	Skewness (temporal shift)
	Maximum	Energy
	Minimum	Dominant frequency
	Mean absolute value $\bar{x}=\frac{1}{N}\sum_{k=1}^{N}\lvert X_k\rvert$	Maximum acceleration
		Minimum acceleration
		Range of acceleration
	Skewness	Resultant jerk features
	Kurtosis	Skewness (temporal shift)
	Energy	Mean jerk at heel contact
	Number of slope sign changes	Absolute maximum jerk
	Number of zero crossings	Absolute minimum jerk
	Length of waveform	Range of jerk produced abs (max–min)
	Dominant frequency using low-pass filter and FFT	Jerk cost
		$JC=\int_0^T \left\lvert\frac{d^3r}{dt^3}\right\rvert^2 dt$

(1) General features and (2) selected features

Figure 2.2: The three feature sets used as inputs to SVM

In this study two steps were taken to preprocess the features before using the SVM classifier, the input features were normalized and the dimension of the feature space was reduced. To normalize the feature values, training and testing feature space were combined and divided by the maximum value of the particular feature. This way, input data was kept in a range between 0 and 1, 1 being the maximum value of the feature. To reduce the dimension of the features space Principle Component Analysis (PCA) was used. To evaluate the ability of the SVM classifier to be generalized five-fold cross-validation was adopted.

The SVM models were trained using linear, polynomial and RBF kernel, over the range of the cost parameter C (2^{-10}, 2^{10}). To evaluate the performance of the classifier the criteria used was accuracy, sensitivity and specificity.

The full schematic diagram of the SVM classification algorithm used in this study is illustrated in Figure 2.3.

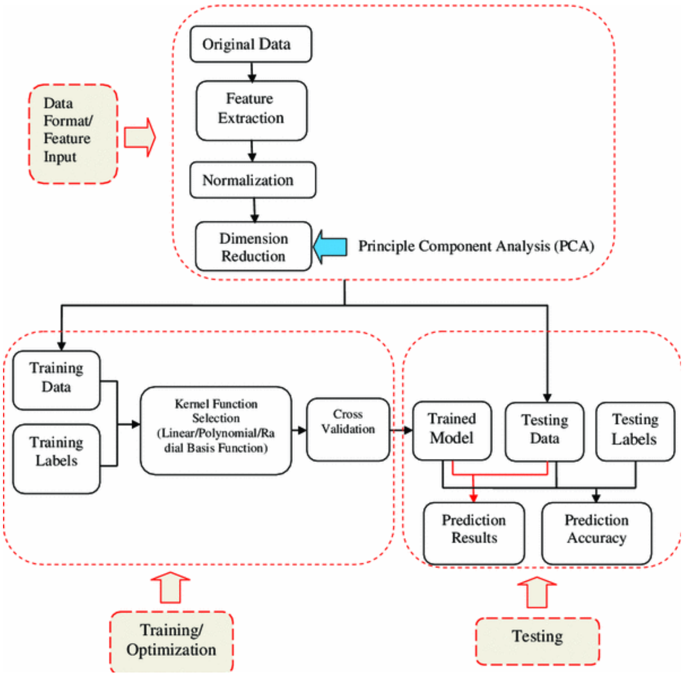

Figure 2.3: Schematic diagram of procedure of SVM classification

In this study, there were not found significant changes in SL after the fatiguing task, yet a wider base of support (12% wider) was observed. Even though walking velocity showed no statistical difference, HCV was considerably faster (p = 0.01) after the inducement of fatigue.

The results regarding the machine learning classification showed high intra-individual classification rates across the three kernels, with linear and RBF with similiarly high results (97 and 96% respectively) with polynomial presenting the lowest accuracy (about 88%).

In inter-subject classification SVM achieved about 90% accuracy with general features and 88% with selected features from the trunk kinematics.

With these results, especially the ones relating to intra-subject classification, this study concluded that IMUs can help in the identification of localized muscle fatigue and body sensors can be used for personalized monitoring to identify risk patterns in gait. It was also concluded that SVM is applicable to the classification of gait patterns after fatiguing task, by using features relevant to walking trunk kinematics.

2.2 Single inertial sensor with SVM classification

This second paper [2] studied the usage of a single IMU, placed at the right ankle, for purposes of fatigue symptoms monitoring using a template matching pattern recognition technique. This paired with the usage of machine learning algorithms for the classification of fatigue states. A single IMU at the ankle was used in an effort to keep the data collection as minimally intrusive and inexpensive as possible for a possible usage in a workplace environment. With this study two questions were posed: can fatigue induced changes in gait be detected by a single IMU, located at the right ankle, and can fatigue and non-fatigued states be distinguished by a computationally efficient classifier, in this case the $1 Recognizer was used. This classifier is an instance-based nearest-neighbour classifier with a 2-D Euclidean distance function. This was chosen due to its computationally efficiency and to the fact that it has been developed for motion recognition (finger gesture).

In this paper the work plan was divided into four steps: gathering of the data, pre-processing it, classification, and evaluation of the model as seen in figure 2.4.

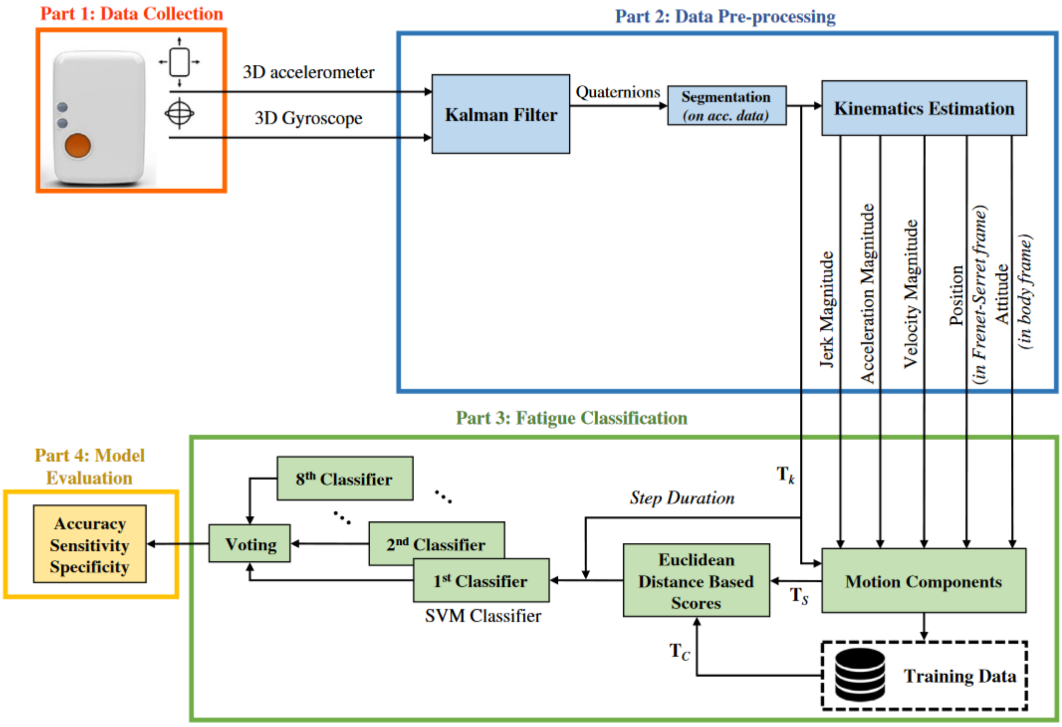

Figure 2.4: Block diagram of the proposed fatigue classification model

Due to the focus of this study in a healthy worker population, criteria for exclusion of participants included reported cardiovascular diseases, musculoskeletal disorders, and an history of injury or pain that could interfere the completion of the experiment.

The experimental procedures involved participants completing a three-hour Manual Material Handling (MMH) session. In this session participants were asked to continuously place weighted containers in pallets and transport them. This method was picked for representing a physically demanding task that usually is performed in warehouses and shipping operations. The advised set of tasks to be followed was picking up one of the eighteen cartons, placing it on the dolly, walking the path while pushing the dolly and, finally, delivering the carton in a prespecified location to simulate a warehouse. The steps of this task can be seen in figure 2.5.

The requested task consisted of three sets of these deliveries, the delivery path length for each box was eighty meters and participants were asked to achieve a rhythm of one set of eighteen boxes per thirty minutes.

Every ten minutes participants were asked to provide a Rating of Perceived Exertion (RPE) on a scale of 6 to 20 and, every thirty minutes, a Subjective Fatigue Level (SFL). For a participant to be included, subjective ratings of RPE and SFL were required to be higher than ten and five, respectively, at the end of the task.

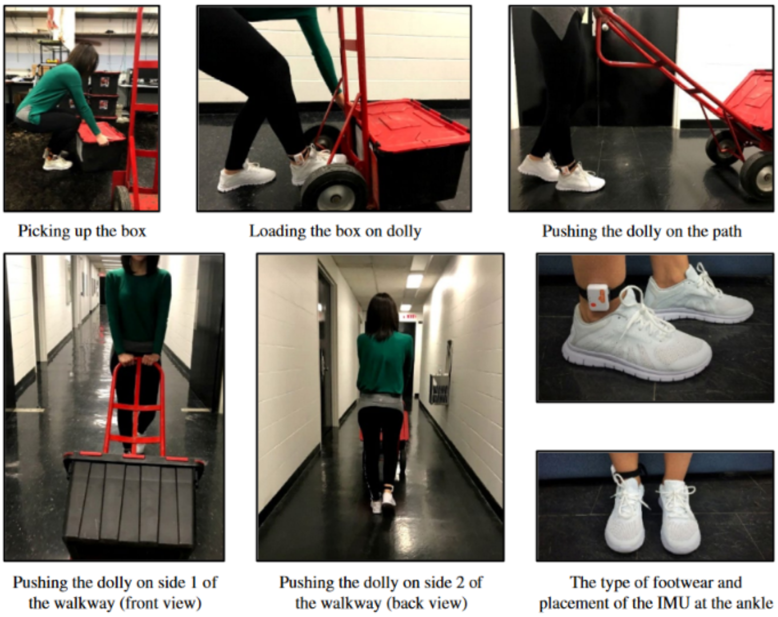

Figure 2.5: Detailed sequence of the task

The sensor used in this study was a Shimmer3, placed at the right ankle as previously mentioned, which contains a low-noise analogue accelerometer, a digital wide range accelerometer, a magnetometer and a digital gyroscope. The data for this study was recorded at a sampling rate of 51.2 Hz which was found to be enough for the purpose of the study and avoided having the data sets unnecessary large due to the high duration of the experiment. Matlab R2015b was the tool selected for post-processing and analysis of the signals.

To pre-process the data the first step adopted was passing the calibrated data through a Kalman filter. This was done with the assumption of uncorrelated white Gaussian process and measurement noises to assess, spatially wise, the orientation of the body frame in relation to the global frame of reference. The final purpose of this step was the estimation of the kinematics of motion, in this case, jerk, acceleration, velocity, position and posture. The next step was the introduction of a robust segmentation algorithm to allow the identification of the strides, the period between consecutive toe off and heel strikes. Afterwards, individual kinematics were estimated using different methods in an attempt to avoid bias.

Out of the initial data a set of 2000 sample point (equivalent to about fourty seconds) of pure walking were extracted from the first ten minutes, non-fatigued, and the last ten, fatigued. After the segmentation of these two sets, batches of twenty five strides from each set were selected as fatigued and non-fatigued sets to obtain the Euclidean distance-based scores as one feature for the final classification. This by comparing each testing stride with both training batches in $1 Recognizer.

The testing sets were obtained in equal fashion to the training ones, similarly ending up with two batches of 25 strides, guaranteeing only no overlap between training and testing sets. The testing batches were then concatenated into an even pool of data. Afterwards, one stride was randomly selected and passed through the modified $1 Recognizer classifier together with the two fatigued and non-fatigued training batches. In this classifier, to assign a score the testing stride was compared to each stride in the training batches based on the point-wise Euclidean distance between the testing segment and one by one of the fatigued and non-fatigued training segment classes.

After this, the results calculated in $1 Recognizer and the test stride duration, for each test stride, were used as feature data points for distinguishing between fatigue and non-fatigued states. The classifier used was a SVM using RBF, which was applied to the score and step duration feature data points of fifty random test strides. In this study the kernel parameter was optimised on the training data set and were selected values that maximize the classification rate. For this step 20% of the feature points were held as testing data points, using the rest for training. To validate the classifier 5-fold cross-validation was again used, for each subject data the fifty feature data points were randomly partitioned into five subgroups of ten each. In addition, a simple ensemble classification was provided using the majority voting of prediction, as seen in figure 2.4, such that fatigued was considered more than four votes. As was the case in the last study, the evaluation of the model was made using the accuracy, sensitivity and specificity.

The highest accuracy was obtained with the combination of all templates (90%) but only slightly higher than using only the acceleration template (89%). This is explained due to accurate direct segmentation results and the fact that this information is collected and not calculated, a process that can be the source of errors. This was followed by position trajectory and velocity magnitude both equal with an accuracy of (86%).

The authors then conclude that the inducement of fatigue leads to alterations in temporal and spatial characteristics of gait kinematics and these can be used for fatigue detection, although not all kinematics performed the same for the prediction of fatigue. It is also concluded that this work can be extended to real-time fatigue monitoring due to the simplicity of the template matching technique and the use of a single IMU at the ankle.

2.3 Effects of fatigue on gait characteristics

The third paper, of the authorship of Xingda Qu and Joo Chuan Yeo [17], studied the alterations on gait characteristics, not only under the effect of fatigue, but also while carrying different loads. Out of the three studies analised in this chapter, this is the one furthest from the work that is going to be developed in this dissertation. The focus of this study is purely on the effects on gait patterns and not in its identification, as such it does not contain a machine learning component, and the method for data collection is also different, using a motion capture system and a treadmill. Nevertheless, it presents the results on changes in gait patterns in an understandable way with some explanations for the events, and in this way it can be a good basis of understanding for the gait pattern changes that may be found later in this dissertation.

Participants in this study were all young males, between twenty and thirty years old, with no disorders that could affect gait patterns. The demographic chosen is due to the study focus on military like conditions. Participants were equipped with twenty six reflective spherical markers placed in anatomical landmarks, on both sides of the body, the placement of these markers can be seen in figure 2.6. This, paired with an eight-camera motion capture system allowed for the collection of whole-body kinematics in three dimensions at a sampling frequency of 100 Hz. All tests were performed on a medical treadmill and participants were asked to perform the task under varying levels of load, 0 kg (no load), 7.5 kg (low load comprised of a 5 kg field-pack and a 2.5 kg light bullet vest) and 15 kg (high load comprised of a 12.5 kg field-pack and a 2.5 kg light bullet vest).

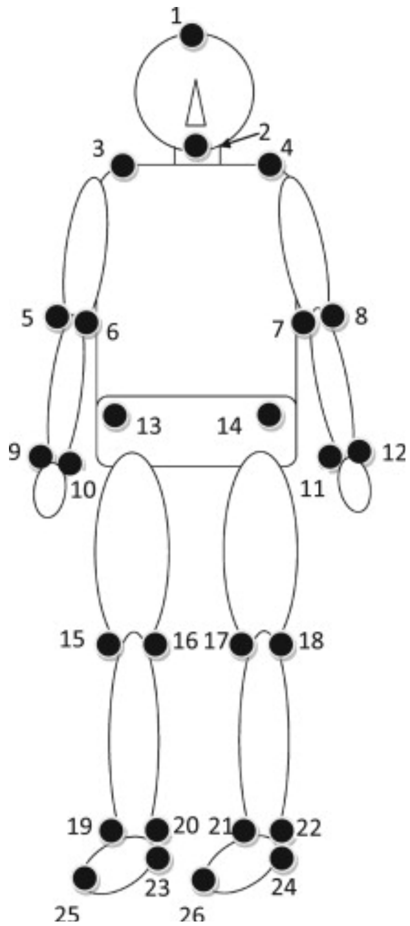

Figure 2.6: Placement of the reflective markers

Each participant performed three pre-fatigue walking trials, each with a different loading condition, with a break of, at least, three minutes between each other. Afterwards, the participants were subjected to the fatiguing exercise, for this they were instructed to run on the treadmill at a speed of 8 mph (near 13 km/h). During this exercise, at every thirty seconds, the participants were asked to rate their fatigue level using the RPE, the exercise was stopped when participants reported a rating of seventeen or above. After the fatiguing exercise, three post-fatigue walking trials were performed, each of them under the three loading conditions. To ensure that fatigue level would be maintained, the fatiguing exercise was repeated in between trials. The experimental procedure can be seen in figure 2.7.

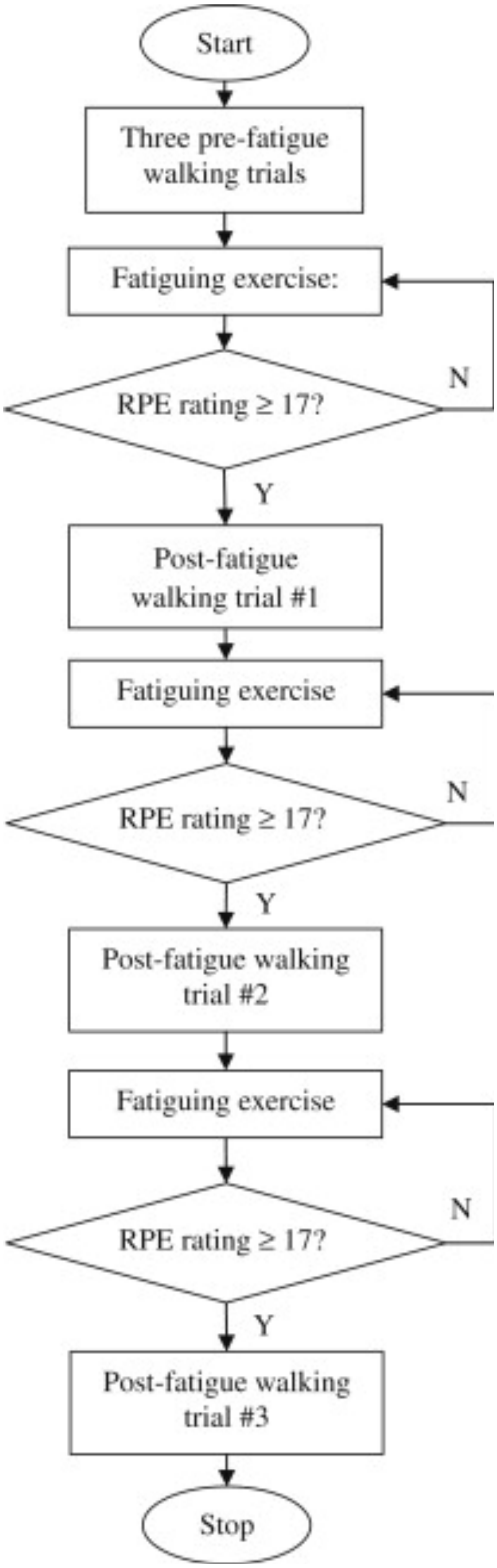

Figure 2.7: Flow of the experimental procedure

This study found that fatigue influenced step width variability, hip Range of Motion (ROM) and trunk ROM significantly, specifically all three factors suffered an increase after the induction of fatigue. Fatigue was also reported to be associated with a larger knee ROM but not as significantly. Effects caused by the back-carrying load were reported but as those will not be replicated they will not be discussed. The study also concluded that, due to the increased joint motions of the hip and trunk caused by fatigue, a higher energy expenditure is required. This indicates that people can adapt to the effects of fatigue by coordinating the joint motions as a way to maintain energy equilibrium. Still, higher muscle tensions can be a consequence of larger ROM, and as such fatigue can lead to injury, muscle strain, and joint problems during walking.

When compared to previous studies the findings related to step width and step length variability show an inconsistency. These inconsistencies were attributed to differences in the fatiguing protocol, as this study used a running test to induce fatigue while the one with contradictory findings used repeated sit-to-stand tasks. Running tests affect muscles controlling medial–lateral movements, on the other hand repeated sit-to-stand tasks can affect muscles responsible for the control of anterior–posterior movements and, as such, these fatiguing protocols can influence gait in different ways.

2.4 Classification using SVM and self-organizing maps

The fourth paper [10], authorship of Daniel Janssen, Wolfgang I.Schöllhorn, Karl M.Newell, Jörg M.Jäger, Franz Rost and Katrin Vehof, used both SVM and Self-organising Maps (SOM) to classify gait patterns before, during and after leg exhaustion. The data was collected with resource to a force plate. This study managed to obtained a 98.1% fatigue recognition success rate. Additionally, applied SOM allowed an alternative visualization of the development of fatigue in the gait patterns over the progressive fatiguing exercise regimen.

Nine participants were selected for this study, with an average age of 25.9 (± 3.14) years. All participants selected for this study were males with considerable experience in sport (in their majority track and field athletes). This selection was justified due to differences in the mechanisms of fatigue with age and gender, as such the participants were chosen to obtain an homogeneous group which would get fatigued with a comparable amount of exercise.

Participants were explained the experimental setup after which were required to become acquainted with it by simulating the test 3 to 5 times. The participants were asked to walk barefoot a distance of approximately 7 meters, starting at a point chosen by the participant in a way where their third foot contact would hit the placed force plate. The right foot was chosen, for all participants, to allow easier comparisons of the derived kinetic patterns, based on vertical ground reaction forces. The trial was repeated in the cases where the force plate was not hit centrally. To register the walking speed of the trials two pairs of double light-barriers were used.

The collections were split in multiple parts, firstly, to establish baseline values, participants were requested to perform 6 gait cycles. After these first trials started the fatiguing exercise in which the participants were required to completely exhaust their soleus and gastrocnemius muscles (both muscles in the back part of the lower leg). This was done by lifting and raising the rear foot while standing only on their toes, to accelerate the process participants were equipped with a barbell and additional weights on their shoulders, in average 44.4 ± 8.8 kg. The actual weight added to each participant was self-chosen, 7 participants chose to add 40 kg in weights while the remaining 2 added 60 kg. The exercise was stopped when the participants were unable to lift their rear foot anymore, the number of repetitions necessary to reach this point varied from 20 to 60 repetitions per subject and trial. Straight after the fatiguing exercise 6 trials were recorded, additionally in the third round, after a 3 minutes break, another 6 trials were recorded without any exercises.

In total there were recorded 18 gait cycles per participant from the three states, before, during and after fatigue. A recursive second order Butterworth low pass filter and a cutoff frequency of 100 Hz were used to filter the vertical ground reaction forces, which were then normalized in time to 100 sampling points, by means of a mathematical linear interpolation and by amplitude to the interval [0, 1], covering 0 to 100% of the gait cycle. This was made in an effort to minimize or remove any possible influence of speed and body weight to the recognition process, which allowed retracing intra-individual changes and inter-individual comparisons at the same time.

A traditional linear statistical analysis, by means of time discrete variables, was conducted first, to verify if the participants' gait patterns were affected by the introduction of fatigue, which allowed the comparison with the performance of nonlinear analysis methods. There were computed and analyzed statistically six commonly used time-discrete parameter, from the vertical ground reaction forces, which described the typical characteristics of the M-shaped curve. These include the forces and elapsed gait cycle time for both the first and second peaks and the valley in-between.

A nonlinear SVM with an RBF-kernel was chosen to test if the gait patterns contained information on the participant a pattern belongs to and the possibility of recognising the state of fatigue, both for just the states before fatigue and during fatigue and for all the fatigue states. To achieve this goal k-fold cross validation was used, with the training set being divided into k subsets, then trained with k-1 and tested with the remaining one. This was chosen with the purpose of guaranteeing the finding of optimal parameter for the SVM, avoiding over-fitting and delivering more reliable recognition rates. All 162 available gait patterns were considered in the classification process.

For inter-individual person recognition and recognition of fatigue only the kinetic data of the states before and during fatigue were used. A multiclass SVM with the "one-against-one" algorithm was used for person recognition, while for inter-individual recognition of fatigue a single SVM was chosen. In the latter, the data samples were allocated to the two classes before fatigue (+1) and during fatigue (-1) and presented to the algorithm. Two custom-developed data-preparation methods were utilised for both person recognition and recognition of fatigue:

In the first method, the time course of all gait patterns was used as time continuous input data. This approach was named the signal approach.

The second method, the deviations approach, involved calculating a synthetic model gait pattern for each participant. This was done by averaging all 18 of the participants' gait patterns data points. This synthetic model gait pattern was then subtracted from all participants' gait patterns to allow the exclusive extraction of intra-individual differences. In other words, to exclusively obtain the deviations from the average gait of the participant. Deviations were then used for further processing. The following figure (Figure 2.8) depicts the deviations approach under a schematic form. In it (a) represents the synthetic average model gait pattern calculated for each participant, in (b) the model gait pattern was subtracted from all participant's gait patterns, which is exemplified for one pattern, finally (c) shows the deviation from the 'average gait' of the participant which is used for further processing.

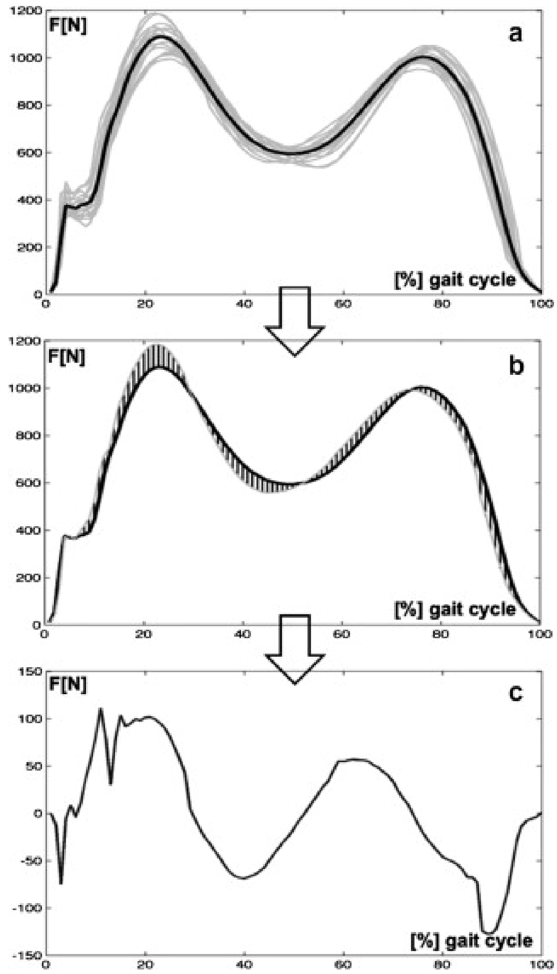

Figure 2.8: Schematic depiction of the deviations approach.

To check if fatigue would alter a participant's gait patters so much that individual characteristics would be affected, an additional test was implemented, which checks influence of fatigue on said characteristics. This could lead to a decrease of person recognition rates for the during fatigue condition, with a returning increase in the after fatigue condition due to regeneration.

For the intra-individual testings was of interest the investigation of similarities and changes in individual gait patterns for the three states of fatigue. To this end PCA was used for dimension reduction in combination with classification done by pre-trained SOM, which offered a visualization of the similarities in the gait patterns in a two-dimensional space. The dataset acquired was not sufficiently large for this step and, as such, the SOM was pre-trained with a conjugation of data from the study and similar gait data from previous studies.

This study did not find significant alteration to walking speed under the experimental conditions, as such data analysis focused on the structure of the movement patterns and the alterations of the kinetic properties of the gait patterns caused by fatigue. In five of the six variables a statically significant difference between the before and during fatigue conditions was revealed by the linear statistical analysis of the time-discrete parameters.

From the conditions before and during fatigue the nonlinear person recognition, based of vertical ground reaction forces, achieved a rate of 100% with the SVM using the previously mentioned signal approach. In comparison, when using the deviations approach the rate dropped to 65.7%. Still on an inter-individual level, the recognition of fatigue in gait patterns achieved a rate of 96.3% using the signal approach and 98.1% when using the deviations approach.

When testing the person recognition for all participants with all three fatigue conditions separately, the rate of recognition of the SVM was of 100% for before fatigue, decreasing to 96.3% during fatigue and raising to 98.1% after fatigue.

The achieved rate of 100% for person recognition, using SVM with the signal approach, emphasizes the high individuality of properties in human movement. This findings were consistent with previous studies with similar rates of success. The person recognition rate using the deviations approach was expected to be much lower, which was attributed to different preconditions with the differences in the calculations to build the reference pattern. Even so, the contrasting with previous studies, in which the deviation approach improved person recognition rates, can open questions on the common strategy of data filtering.

The most surprising results were the recognition rates of fatigued gait patterns, including inter-individual recognition. The achieved rates of 96.3% (signal approach) and 98.1% (deviations approach) lead to the conclusion that the SVM was capable of distinguishing gait patterns, both before and during fatigue, by kinetic information over all participants. For this situation the deviations approach has the advantage of a more obvious extraction of intra-individual differences.

2.5 Summary

The studies analysed in this state of the art chapter were selected in an attempt to analyse at least one study related to each part of this study. As such, the first, second and fourth studies use SVM on their implementation, with different methods of data collection, sensor placements and fatiguing exercises between them. The last study also uses a neural network as part of its implementation. Differing from the previous, the third study was selected only due to its finds in alterations to gait characteristics and does not include any machine learning element. With this in mind, any comparison of results related to machine learning algorithms can never include the third study.

The first study collected data using two IMUs, one at the sternum level and one on the right shank, in partnership with other methods, with squatting being the fatiguing exercise chosen. Different gait related features were used and the classification was done using a SVM classifier with the linear, polynomial and RBF kernels being tested. This study achieved an accuracy of 97% with the linear kernel, 96% with the RBF and a lower 88% with the polynomial kernel. There were still inter-subject classification, in which were achieved 90% accuracy with general features and 88% with selected features.

The second study also includes SVM classification, although its pre-processing of data and steps taken to reach the classification are different. In this study data was collected with a single IMU, placed at the right ankle, contrary to last study the fatiguing exercise was a part of the collection itself with the participants being asked to push a dolly with weighted packages in between points. The highest accuracy achieved by this study was of 90%, when using a combination of all templates calculated, followed by 89% using only the accelerations template. Finally, both the position trajectory template and velocity magnitude obtained an accuracy of 86%.

As previously mentioned, the third study does not include any machine learning segment, focusing only on the alterations induced by fatigue on the gait patterns. Data was collected using both motion capture methods and a treadmill, with the fatiguing exercise chosen involving asking the participants to run on said treadmill. This study found a significant increase of step width variability, hip ROM and trunk ROM, when under the effects of fatigue, with an increase in knee ROM also being reported but not as significantly. The study also raises awareness to inconsistencies between its finds and the finds of previous studies, especially related to step width and step length variability. These inconsistencies were attributed to differences in the fatiguing protocol.

The last study used both SVM and a SOM to classify gait patterns. The data was collected using a force plate and the fatiguing exercise chosen asked the participants to lift and raise the rear foot while standing on their toes, carrying a barbell with weights on their shoulders. In this study the highest rate of fatigue recognition achieved was 98.1%, with the deviations approach, and 96.3%, with the signal approach. For person recognition a rate of 100% before fatigue was achieved, along with a rate of 96.3% during fatigue and 98.1% after fatigue.

Chapter 3

Theoretical work

3.1 Gait and fatigue

This chapter is dedicated to a more in-depth explanation of gait, the gait cycle, gait patterns, and the effects of fatigue on said gait patterns. There will be a review of previous findings in an attempt to predict the alterations expected to be found with testing.

3.1.1 Gait

As previously explained in the motivation section of chapter 1, gait refers to human locomotion on foot, a movement which has specific characteristics known as gait patterns or gait kinematics. These patterns are influenced by fatigue levels; as this factor increases, there is a decrease in postural stability and motor performance. This leads to an increase in the risk of falls and accidents. Fatigue is a major issue in a working environment, in the US workforce the estimations for fatigue prevalence is of 37.9% [21]. Fatigue in this context can lead to accidents where, depending on the industry, the outcome might be workplace injuries or even deaths.

To understand the way other studies identified changes in gait patterns it is important to understand the different parts of the gait cycle. The gait cycle is split into two main phases, the stance phase and the swing phase, where the stance phase attends for 60% of the cycle with the swing phase occupying the remaining 40% [12]. More detailed classifications vary in the way they split the gait cycle varying between 6 or 8 phases [12, 16]. The possible classifications are:

6 phase classification (as seen in figure 3.1):

1. Heel Strike
2. Foot Flat
3. Mid-Stance
4. Heel-Off
5. Toe-Off
6. Mid-Swing

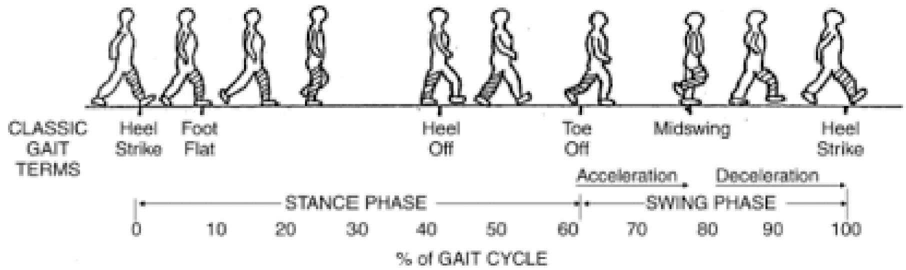

Figure 3.1: 6 phase gait cycle [6]

8 phase classification (as seen in figure 3.2):

1. Initial Contact
2. Loading Response
3. Midstance
4. Terminal Stance
5. Pre swing
6. Initial Swing
7. Mid Swing
8. Late Swing

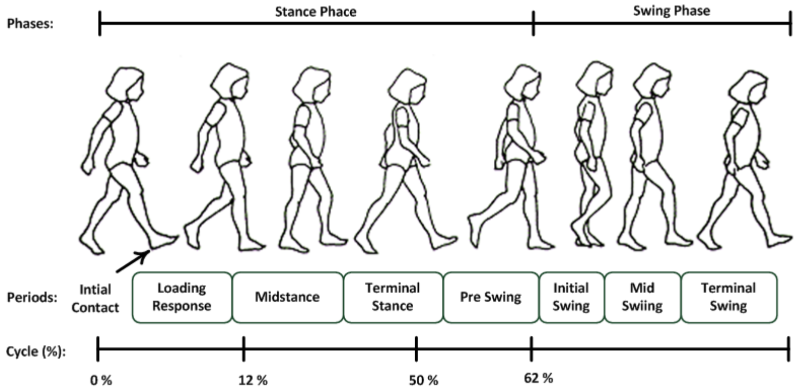

Figure 3.2: 8 phase gait cycle [20]

3.1.2 Fatigue influence on gait patterns

After understanding the fundamental parts of the gait cycle, this can be used to understand the findings of previous studies and their testing methods. Even with discrepancies in some of the findings, it is reported in several studies that fatigue affects the human gait kinematics and, as such, it is a factor in destabilizing the normal gait. The discrepancies in the finding can be attributed to the fatiguing protocols used and how these can affect different muscles [2], as an example running protocols may affect more medial-lateral muscles where anterior-posterior muscles may respond more to sit-to-stand tasks. These differences can lead to different motor responses, which can explain inconsistencies between findings [17].

It is important to mention that previous studies searched and found relations between fatigue and increases in knee ROM, hip ROM, and larger trunk ROM. Findings like these, although significant and deserving of mention, will not be discussed further due to the requirement of different methods of acquiring data, such as full body tracking [17].

Previous studies found increases in step width [26] and step width variability [17], with a reason for this increase hypothesised as a strategy to compensate for the results of fatigue and increase gait stability. There have been reports of decreased step length [3], meanwhile step length variability is one of the characteristics on which there were found contradictory results; sit-to-stand tasks led to an increase in step length variability [9], while running fatiguing methods did not influence this characteristic [17].

Jerk cost was reported to greatly increase, more than 2-fold, with fatigue, paired with a higher result acceleration [26], while significant increases to mediolateral trunk acceleration and trunk acceleration variability in the vertical direction, in between strides occur[9]. Another of the characteristics where the effects of fatigue are contradictory was gait speed as procedures that allowed participants to choose a comfortable walking speed found different results. Some studies reported no significant differences before and after the fatiguing procedure [9], whereas others found a decrease in stride speed, which was attributed to an anticipatory strategy to guarantee dynamic stability during obstacle crossing [3].

The results reported by previous studies indicate that the effects of fatigue on gait parameters can serve as a basis for continuous monitoring for fatigue detection [2]. Also being reported that these effects, even though some are too subtle to be identified by the human eye, can still provide information for an SVM classifier to predict a situation of fatigue [26].

3.2 Technology and sensors

This chapter is dedicated to exploring the possibilities available when it comes to obtaining the necessary data. For that, first there will be an explanation of the most useful for this study, sensors that exist inside most devices, such as accelerometers and gyroscopes, followed by a brief exploration of technology available in the market, mostly in the wearable department, finishing with an analysis and comparison of the use of smartphones and wearables to collect the data.

3.2.1 Inertial sensors

3.2.1.1 Accelerometer

An accelerometer, as the name implies, is a device used to measure acceleration forces. Despite having the appearance of a simple circuit, consists of multiple parts and can work in different ways. Two of these different ways of operating are the piezoelectric effect and the capacitance sensor. Out of these two, the most common form of accelerometer is the piezoelectric effect. This method utilises microscopic crystal structures that, when subject to the influence of accelerative forces, become stressed and create a voltage. The accelerometer uses this voltage to calculate the velocity and orientation.

The other form, the capacitance accelerometer, uses micro-structures placed next to the device to detect differences in electric capacity, capacitance. The influence of an accelerative force displaces one of the structures, causing the capacitance to be altered which the accelerometer will translate into voltage, so it can be interpreted.

Accelerometers can be made up of multiple axes, depending on the type of movement to be analysed. Smartphones usually utilise a three-axis model, allowing it to detect positioning in three dimensions (as seen in figure 3.3), while, for example, cars only use a 2D model when determining the moment of an impact. Regardless, these devices are very sensitive, due to the requirement of registering minimal shifts in acceleration. For this reason, smartphone accelerometers opt for having a lower value range, so they can achieve a higher precision (iPhone 4 range: ±2g, precision 0.018g). Even though different smartphones achieve different precisions, as there are no regulations imposed on the manufacturers in this area, higher-end smartphones usually have better results, but nowadays all smartphones have good precision.

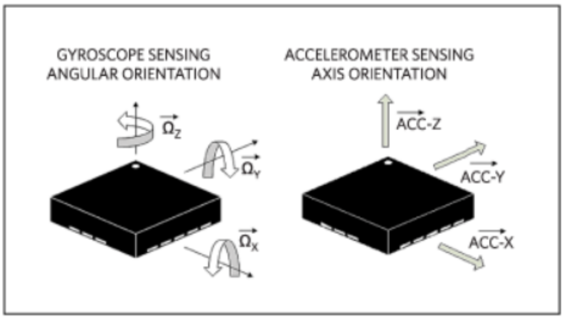

Figure 3.3: An accelerometer and a gyroscope and the forces these measure

3.2.1.2 Gyroscope

The gyroscope can act as a support to the accelerometer in a sense that it can help in situations where the accelerometer alone would fail. Situations like running can cause a similar acceleration to a fall but the use of the gyroscope can differentiate the two. This happens because the gyroscope allows the addition of the data related to the angle at which something happens, allowing for better accuracy rates, when compared to the use of only the accelerometer [18].

A gyroscope reads angular velocity. The gyroscopes found in smartphones and wearables measure the Coriolis force, this force is proportional to the angular rate of rotation in a rotating reference frame. To obtain the angular rate, micromachined gyroscopes then integrate the gyroscopic signal with the detected linear motion from the Coriolis effort. Other types of gyroscopes do exist, based on different principles, ranging from electronic gyroscopes, fiber optic gyroscopes, to extremely sensitive quantum gyroscope. In the analysis of human gait, gyroscopes can be used to determine the angular velocity, angle of feet or legs, depending on the placement of the sensor, and this can be used to identify the different gait phases [24].

3.2.2 Available Technology

In recent years there have been major developments in the availability of sensors in our every day lives, from the sensors contained in the smartphones we carry to small wearable devices that show up more and more on the market.

While current smartphones have all the necessary sensors included, wearables found in the market tend to be more specific with their function. Market-wise most wearables that contain the necessary accelerometer and gyroscope seem to fall into two separate categories. The first are small circuits or modules that are meant to be used for building your own device, such as the WitMotion BWT61 (figure 3.4) or the Adafruit wearable section, these can be very cheap but require the acquiring of the necessary parts and the construction of the full device, making them not viable on a future solution.

Figure 3.4: WitMotion BWT61

On the other hand, there are fully built solutions such as the Xsens MVN Awinda or the Shimmer3, this last one was used on a previous study with positive results [2], which can be seen in figure 3.5. These solutions are built for research purposes or high performance training. As such, these are quite expensive solutions, some even requiring pricing request, and their availability is very limited, this again limits the possibilities of future implementations on real world scenarios.

The outliers to these two categories are smart bands, which are increasing in popularity. Although not all bands have the required sensors, newer ones such as the Huawei

Figure 3.5: Shimmer3 sensor used on a previous study [2]

Band 3e are starting to add them, and still maintain a very affordable price. This technology shows the most promise, outside of smartphones, in possible solutions for future integrations.

3.2.3 Outlook

Both types of devices explored in this chapter fit the requirement of adaptability to real world (out of the lab) conditions, which guarantees a chance to expand the work developed on this dissertation to complement existing solutions, in case of positive results. That said, the greatest advantage of using only the inbuilt sensors of smartphones is the widespread availability and popularity of them, nowadays a large part of the population owns a smartphone and the tendency is for the number to increase. This way, there is no need to acquire devices which limits real world applications.

Previous studies found that phones can achieve a high level of precision, enough for gait analysis, with results comparable with other methods of data collection [23], and as such can be viable data collection tools for a real world solution. The two major foreseeable disadvantages of using a smartphone to collect the data is the fact that the smartphone does not have a static placement, that is while in a testing environment we can guarantee that the smartphone is always in the front pocket, in the real world people carry their phones in different places, anywhere from back pockets or shirt pockets to hand bags or backpacks. This difference can cause a distinction between the test results and a real life application. The second disadvantage is the fact that people may not fully understand what the collection of data entails and mistrust the use of their smartphone for it. In this case, there is an argument to be made that a small wearable device, to be attached to the ankle for example, with a single purpose and that can be removed when not needed, could be preferable.

Even though wearable devices have the advantage of having a fixed placement and can create less complications, they are not without disadvantages either. As previously mentioned, wearable devices always require the acquisition of the device itself and, depending on the sensor used, the cost might be considerably high. That said, the availability of these devices has been increasing in the last years and, with it, it is expected to see a decrease in prices. Another pertinent question is how intrusive these devices are; the development of the technology allows the devices to be extremely small and lightweight but they may still cause a felling of being monitored, as their presence is more noticeable and foreign than the everyday tool of the smartphone.

Due to the availability of both smartphones and wearable sensors at the facilities where the dissertation will be developed, both were used to collect data, allowing a comparison of both methods of collection based on results, instead of only foreseeable advantages and disadvantages.

3.3 Machine Learning

Machine learning is the study of algorithms that computers use to perform a task without requiring explicit instructions, using instead patterns and inference. It is a subset of artificial intelligence. Machine learning algorithms create a mathematical model from an initial sample data to make predictions or decisions without being explicitly programmed to perform the task.

There are four basic types of machine learning problems, supervised, unsupervised, semi-supervised and reinforced. Supervised machine learning require the data to be labelled and uses it to come up with a rule to predict the output of future inputs; these problems can be divided into regression problems for continuous values or classification problems for discrete classes. In unsupervised machine learning the data is not labelled with the objective usually being to find some structure in the data; this can be used as a first step in broader learning task, as a preprocessing step, for example. Semi-structured learning mixes the two previous approaches and in it some data is labelled but most is not. Lastly, reinforced learning is best one when we want to optimize some output but don't have direct feedback in every case; a good example is playing a game, every move has a cost and a benefit but only at the end of the game it is visible the final result, the conjugation of all the moves.

3.3.1 Support Vector Machines

SVM is a classification algorithm that is specialized in solving binary classification problems. SVMs define the learning problem as a quadratic optimization problem, whose error surface is free of local minima and has global optimum, originating from the statistical learning theory proposed by Vapnik [25]. For binary classification problems, such as the distinction between fatigue and non-fatigued gait patterns, the objective is to discover an OSH between the two data sets, as can be seen in figure 3.6. To quickly define an hyperplane, assuming an Euclidean space with n dimension, an hyperplane is an n-1 dimensional subset of the space which divides it into two detached parts. To find the

OSH the SVM tries to maximize the margin between data points of both classes. By maximizing this margin there is a better chance future data points will be correctly classified. To do this SVM first utilizes a kernel function to transform the input data into a higher dimensional space. In this transformed larger version of the feature space, it creates an OSH that linearly divides the two classes, the nearest data vectors to the constructed line, in the transformed space, are named support vectors and hold useful information regarding the OSH [4][8].

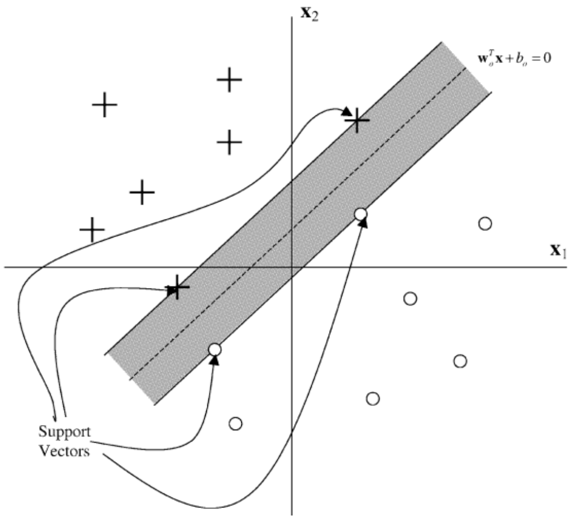

Figure 3.6: Example of a two-class problem showing OSH that SVM uses to divide the data and associated Support Vectors

The problem of pattern recognition may be stated as follows: Given a Θ training data set, having input features (x_i) and classification output (d_i), with the following form

$$\begin{aligned} &\Theta = \{(x_1, d_1), (x_2, d_2), \ldots, (x_N, d_N)\} \\ where&: x_i \in \Re^m \\ &d_i \in \{+1, -1\} \\ &N \; is \; the \; number \; of \; samples \end{aligned} \tag{3.1}$$

In this dissertation , d_i can be considered +1 for fatigued and −1 for non-fatigued gait. We assume g(x) is some unknown function to classify the feature vector x

$$\begin{aligned} &g(x) : \Re^m \rightarrow \{+1, -1\} \\ &where \; m \; is \; dimension \; of \; the \; feature \; vector. \end{aligned} \tag{3.2}$$

As previously mentioned, SVM method wants to find a hyperplane in m dimensional space that linearly separate the two classes +1,−1. The equation of the hyperplane is then

$$w^T x + b = 0 \tag{3.3}$$

with w being the adjustable weight vector and b the bias of the hyperplane. The linearly separable case can be represented mathematically as

$$
\begin{aligned}
w^T x + b &\leq 0 \quad for\ d_i = -1 \\
w^T x + b &> 0 \quad for\ d_i = +1.
\end{aligned} \tag{3.4}
$$

When implementing the SVM classifier, there are two tuning parameters that need to be discussed. The first, usually represented by C, controls the trade off between a smother decision boundary and the correct classification of points. While a higher value of C leads to more points being classified correctly, due to the fact that any algorithm should be trained with only a portion of the data, a larger value for C can lead to problems of over-fitting and lead to worse results when testing the algorithm with new data. The second parameter, usually called gamma, regulates reach of the influence of each point. Low values of gamma lead to a bigger reach for every point and vice-versa. In other words, an higher value for gamma give points close to the decision boundary a bigger influence to over it, even ignoring points that are far away from it. This results in a line that curves when necessary to adapt to the points close by. On the other hand, lower gamma values give more weight to far away points and lead to a more linear curve.

The majority of real life problems, the problem discussed in this dissertation included, are not linearly separable. To be able to solve this issue it is possible to apply nonlinear transformation on the data. While this is easy for modest values of m it is easy to see how it can become unmanageable with higher values of m. A way to surpass this is using a kernel trick, basically using basis expansions such as polynomials or splines to map the data from the input space into an enlarged feature space. Linear boundaries in this substantially higher dimension space obtain better separation between classes, and translate back into the original feature space as nonlinear boundaries[8]. This allows to avoid the curse of dimensionality by hiding the potentially high dimension of that feature space. The kernel function K(x,y) is related to the nonlinear feature mapping function φ(x) by

$$
\begin{aligned}
K(x,y) &= \varphi(x)^T \varphi(y) \\
where&: x \in \Re^m \\
\varphi(x) &\in \Re^h.
\end{aligned} \tag{3.5}
$$

3.3.1.1 Kernels

A kernel is a function that grabs a given input data and transforms it into an high-dimensional space where the classification of such data is possible. As previously mentioned, a better separation between classes is obtained by the linear boundaries in this higher dimension space, which translate back as nonlinear boundaries on the original feature space. Kernel function can fall into one of two categories, linear or nonlinear. Selecting the adequate kernel for the problem directly influences the accuracy of the SVM classification; a well selected kernel may minimize generalization error, and increase classification accuracy. Overall the choice of a kernel/regularisation parameters can be automated by cross-validation, a method used in past studies[4] and chosen as adequate for this study.

The simplest kernel function is the linear kernel; this is suitable when there is a high number of features in the training data. This kernel can also be called dot-product, it is given by the inner product plus an optional constant c: $K(x, x_i) = x^T.x_i + c$.

Polynomial kernels (of degree “d”: $K(x, x_i) = ((x.x_i) + c)^d$) are non-stationary kernels and work well for normalized training data. The optimization parameters for this kernel are the slope, alpha, the degree of the polynomial, d, and the constant, c, which controls the trade off in influence of higher and lower order terms.

RBF kernel (with width “g”: $K(x, x_i) = \exp\{-|x - x_i|^2/g^2\}$.) is normally a reasonable first choice due to its ability to non-linearly map the data into a higher dimensional space. [26] The RBF kernel creates a non-linear combination of features to transform the data to an high-dimensional feature space.

To briefly explain the Radial Basis Function, in it as the data point x moves away from the center x_i the RBF function decreases in a swift and monotonic way. The rate at which the Gaussian RBF decreases is controlled by the width g as the higher the value of g the slower the rate of decline.

3.3.2 Neural Networks

Out of the multiple types of neural networks, the multilayer feed-forward neural network, also called multilayer perceptron or back-propagation network, has been standard in a wide range of applications, including gait analysis [10][11]. As such, part of this explanation will focus on this type, although it is still applicable to other types of networks.

Artificial Neural Network (ANN) are simply nonlinear statistical models, an ANN is a two-stage regression or classification model, usually represented by a network diagram of inputs and outputs, with the middle being occupied by a processing section also called hidden layer/s.[8] Statistically wise, the tradition is the inputs being the independent variable with the outputs being dependent. The hidden layers receive their name due to them not being directly observed, and there can be more than one hidden layer.

ANN can be compared to a flexible mathematical function with multiple configurable internal parameters. To be able to accurately represent the intricate relationships between gait kinematics, the internal parameters need to be optimized. In supervised learning, examples of inputs and desired outputs are presented to the network, in an effort to accurately represent as many examples as it can, the network iteratively self-adjusts to the presented data. This learning process is considered to be complete after a chosen criterion falls bellow a preselected threshold. After the ANN finishes training it is able to receive new, previously unseen, inputs and attempt to accurately predict the output.

There is one assumption when deploying a multilayer feedforward neural network (with one hidden layer), in that the input and output data are related by a continuous functional relationship, although this is rarely mentioned due to it being so general and nonrestrictive [11].

3.3.3 Performance assessment

To follow the steps of previous studies[26][2], a possibility when assessing the performance obtained by the classifiers is using the accuracy, sensitivity and specificity obtained. These three metrics have been used as means to classify the quality of predictions in multiple studies and articles, from areas such as gait analysis[26][2] to general medical diagnostics [27]

In the setting of this study, sensitivity represents the ability to correctly identify fatigued gait from only the fatigued data present in the overall data, in other words, it portrays the ability of the classifier to find the presence of fatigue. A classifier with high sensitivity scores is able to identify all situations where a user is fatigued without missing.

$$Sensitivity = \frac{TP}{TP + FN} x100\%$$

Meanwhile, specificity indicates the ability of the classifier to avert situation of false detection, showing its ability to correctly classify situation without fatigue. The higher the specificity, the more non-fatigued data is classified as such, and the better the classifier is at avoiding classifying a situation without fatigue as fatigued. It is possible to obtain highly specific results but not very sensitive or vice versa. A good prediction is one that achieves high results for both sensitivity and specificity, although depending on the requirements asked from the prediction one might be more important than the other.

$$Specificity = \frac{TN}{TN + FP} x100\%$$

Accuracy indicates the overall correct predictions, in other words, the ability of the algorithm to correctly identify both fatigued and non-fatigued gait. Since accuracy requires correct prediction, both positive and negative, it is possible to see how this metric is much related to the results of sensitivity and specificity.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} x100\%$$

On the formulas of these metrics multiple acronyms were used; in this scenario, TP represents the number of true positives, a situation where the classifier identified a fatigued gait that was labelled as fatigued; TN being the number of true negatives, a situation where the the classifier identified a non-fatigued gait that was labelled as non-fatigued; FP and FN are the number of false positives and negatives or the number of cases incorrectly classified as fatigued and non-fatigued, respectively.

It is important to mention that for the conditions required for this study an high relevance will be given to the specificity results. Overall, all metrics are important and should ideally achieve high results but any future use of the work developed during this dissertation would require a high capability to avoid false positives.

Chapter 4

Experimental study

In this chapter the different steps taken during the experimental part of this study are explained, starting with the experimental protocol and data collection process. This is followed by a brief insight into the participants' demographics that volunteered to be a part of this study. Afterwards, the methods implemented to process the data collected and train the classifiers are explained and, lastly, a few alternative approaches that were attempted during development are presented.

4.1 Experimental protocol

An experimental protocol was written to regulate the collection of data, necessary to this study. This section explains the steps present in this protocol and followed for the data collection process.

Before starting the collection, all participants were asked to fill-in a form with a few personal questions for statistical analysis and preferences related to the collection. This form also served to obtain the consent of the participants for the collection and utilization of their data. This form will be included as an annex (Annex I) but the information collected from it will be presented in the Demographic information section.

After filling-in the form, the participants were informed of the process necessary to the collection of data. First, the participants were equipped with 4 Internet of things in Package (IoTiP) devices placed in specific locations, one in the dominant side ankle, two at belt height, one of them at the center of the back and the other facing forward above the pocket area, and finally one centered at the top of the chest. These IoTiP devices were developed in-house, at Fraunhofer, and include both accelerometer and gyroscope sensor used during this collection, among others.

Figure 4.1: IoTIP device used to collect data

In addition, the smartphone used to connect the sensors was also collecting data itself and participants were asked to keep it on their right pocket during the data collection. These placements were chosen according to previous studies, that obtained good results, and also due to the diversity in the placements, as an attempt to facilitate the observation of the influence that the placement of said devices has on the performance of the supervised learning algorithms.

After a few moments to acclimatise to the wearing of the sensors, the participants are given the smartphone and are requested to put it in their pockets in order start the first collection.

The first collection is considered the non-fatigued collection; each participants is requested to walk five minutes on a treadmill, at a self-selected comfortable pace. The treadmill's display is covered as not to influence the participant's speed selection, especially important for the second collection.

After this first collection, it is necessary to induce a state of physical fatigue on the participant. For this goal one of two exercises were used, the participants were asked to either squat repeatedly or briskly walk up and down a flight of stairs. The exercise was chosen by the participant in the previously mentioned form and, in both cases, ending the exercise was dependent on the feedback provided by the participant. The fatiguing exercise was considered complete when the participant reported a rate of perceived exertion superior to 15 and a fatigue level superior to 5. The participants were informed of the scales used to measure these values before starting the exercise.

The scales used to assess the stopping point of the exercise were as follows:

How you might describe your exertion	Borg rating of your exertion	Examples (for most adults <65 years old)
None	6	Reading a book, watching television
Very, very light	7 to 8	Tying shoes
Very light	9 to 10	Chores like folding clothes that seem to take little effort
Fairly light	11 to 12	Walking through the grocery store or other activities that require some effort but not enough to speed up your breathing
Somewhat hard	13 to 14	Brisk walking or other activities that require moderate effort and speed your heart rate and breathing but don't make you out of breath
Hard	15 to 16	Bicycling, swimming, or other activities that take vigorous effort and get the heart pounding and make breathing very fast
Very hard	17 to 18	The highest level of activity you can sustain
Very, very hard	19 to 20	A finishing kick in a race or other burst of activity that you can't maintain for long

Table 4.1: Borg scale of perceived exertion[5]

0	Nothing at all	
0.3		
0.5	Extremely weak	Just noticeable
0.7		
1	Very weak	
1.5		
2	Weak	Light
2.5		
3	Moderate	
4		
5	Strong	Heavy
6		
7	Very strong	
8		
9		
10	Extremely strong	"Maximum"
11		
$\int$		
•	Absolute maximum	Highest possible

Table 4.2: Borg CR10 Scale used to measure the perception of intensity of any experience[7]

After the participants reported the level of exertion and fatigue required, the fatiguing exercise is considered finished and the second and final collection of data can be initiated. The placement of the sensors is quickly checked to guarantee that there were no shifts during the exercise and the participants are requested to once again walk on the treadmill for another five minutes. As previously, the participant could regulate the speed of the treadmill, without being able to see its display, which is particularly important in this collection to avoid visual influence when deciding the speed of the treadmill.

After this second collection, the overall trial is considered to be finished and the sensors are removed from the participant.

The existence of two possible fatiguing exercises was considered due to previous studies that found contradictory alterations in gait patterns after fatigue. In these studies, the explanation given for the contradictions found was the use of different exercises to fatigue the participants. These different exercises would fatigue different muscles and lead to different alteration in the gait patterns.

4.2 Demographic information

In total 12 participants volunteered to be a part of this study, 9 male and 3 female. The average age of the participants was 26.25 (±4.80) years. The age of the participants that volunteered to participate influenced the direction of this study, moving away from any conclusions possible about an, originally planned, older population. The participants average height was 174.50 (±6.70) cm, while the average weight was 70.92 (±13.14) kg.

In regards to the fatiguing exercise, the majority of participants (8 participants, 67%) chose the squatting exercise, while only 4 participants (33%) chose to walk up and down stairs. Those who chose squatting took in average 3 minutes and 38 seconds to report the previously mentioned stopping levels of fatigue and exertion, while participants that chose the stairs exercise took, a slightly lower, 3 minutes and 6 seconds in average.

Looking back at the experimental protocol, the display of the treadmill was covered during both collections in an attempt to discover if the participants would prefer an higher or lower pace while fatigued. While in average after the fatiguing exercise the walking velocity of the participants increased, from 4.06 to 4.22, this did not occur in every instance, as such it can not be concluded that the introduction of fatigue leads to an increase in walking velocity

Only one of the participants reported their left foot as their dominant one and only two reported health problems related to fatigue or gait but neither were considered impeditive enough that their data couldn't be used.

4.3 Data processing

In total, the collections generated 120 files of data, each of these files contained the data for both the accelerometer and gyroscope of one sensor and amounts to five minutes of gait.

To process the data the first step taken was the reading and parsing of the files. In an attempt to reduce writing errors from the devices, the first 40 seconds of each collection were skipped, as during this starting period the smartphone was prone to have problems communicating with the devices leading to larger gaps in between lines in the file and significantly lower frequencies of data. After that, the following two minutes of the collection are used. Only two minutes were used for two reasons, the first to avoid possible delays in between the stopping of the participant and the moment the collection is stopped on the smartphone. The second reason behind the two minutes was an attempt at averting a possible cooling down of the participants on the post-fatigue trial, hiding the effects of the fatigue caused by the exercise.

The necessary features for the learning algorithms are calculated using the (Time Series Feature Extraction Library (TSFEL)) which automates this process, aiding on exploratory feature extraction tasks on time series. TSFEL is optimized for time series and automatically extracts over 50 different features on the statistical, temporal and spectral domains. TSFEL calculates 50 features for each dimension of each sensor, as this study uses both the accelerometer and gyroscope sensors, each writing in three dimensions (Ax, Ay, Az, Gx, Gy, Gz), the total number of features returned is 300. A list of the 50 features calculated for each of the dimensions is included in annex (Annex II).

Afterwards, it is necessary to identify from which type of collection (fatigued or non-fatigued) each line of the resulting dataframe came from; to this effect an extra column is added identifying with a 1 if it is from a fatigued collection and a 0 otherwise. The resulting dataframe is split into 10 folds to allow for cross validation. At a time, one of the folds is saved for testing, while the other nine are used in training. The data is passed through PCA to reduce the number of features that do not add relevant information to the dataframe, reducing the number of features to under 100, the exact value depends on which device positioning is being used. For this process the training data was used to fit the PCA, which is then used to transform the test data. Before passing through PCA, it is necessary to standardize the data.

PCA is widely used to reduce the dimensionality of the features by condensing the information into a smaller group of composite dimension, all while retaining as much information as possible. The PCA analyses a data set and identifies the features with the highest variance using these to create a smaller data set with minimal losses to the original data set descriptive power. A smaller data set has the advantages of requiring less processing power and can have less noise in the data. This process does however transform the original data set, the principal components calculated by the PCA are influenced by the original features but are not a copy of any of them, with a linear combination of features being used to keep the descriptive power of the original data set. It is possible to know the contribution of a feature to a principal component but it is not correct to say that features are kept.[1]

After this processing section,the training data is used to optimise the necessary parameters of the different learning algorithms. For this optimisation process the training data is divided further into 5 folds, cross-validating the results using the possible parameters.

Lastly, the processed data is passed to the learning algorithms. Due to the requirements and to save time, the processed data was given to all the algorithms to be tested one by one. During this step the tested algorithm is trained with the training data and then given the testing data to identify. The results are stored and the next algorithm is given the already processed data saving processing time. Although 10 fold cross validation was used to validate the results of the algorithms, the splitting of the data only in the temporal field can bring problems on the way of over-fitting of the data. This problem will be further explained in the following section when explaining an alternative approach attempted during development.

4.4 Previous approaches

During development different approaches were attempted; the first, and most important, to mention is related to the way the features are calculated. Although in the end it was opted to maintain the automatic approach, during development there was an attempt at a different way to calculate features. This second attempt involved identifying individual steps from the gait data and calculate the necessary gait patterns. To this effect, the first step taken was to implement a step detection algorithm. This was done by analysing the temporal distribution of a collection; firstly the three axis of the acceleration were joined into the total acceleration on that moment[26]. On this total acceleration, steps were identified by three consecutive positive zero crossings, after an initial peak (to try to eliminate possible initial errors).[22]

After dividing all the data by steps, the next stage was calculating the wanted features; the first, jerk cost, was calculated from the total acceleration of the step using the following formula[22]:

$$SL = K\sqrt[4]{a_{max} - a_{min}}$$

The second feature calculated in this manner was the step length. Similarly to the previous feature, the formula is based on the total acceleration (where R represents the total acceleration)[26]:

$$J = \frac{dR}{dt}$$

Although these two features were calculated successfully, after that the problems with this development path out-weighted its value. The first problem showed itself even before these features were finalised, calculating features one by one was taking too long to implement; for every feature a formula to calculate it would have to be researched, implemented and tested, with the assumption that all these steps would succeed. The second and major issue is the fact that not all necessary features could be calculated due to limitations in the data. When the experimental protocol was written, in an attempt to limit battery usage, it was planned to only use the accelerometer and gyroscope of the devices. Unfortunately, at the time the real necessity of the usage of magnetometer was overlooked. The magnetometer is used to convert the data from the coordinates of the device to earth's coordinates. This transformation is necessary to use a lot of the formulas found for more features, such as step width. Although not completely impeditive, the high time investment necessary to continue this path caused the whole idea to be left for future work.

Another approach that unfortunately didn't lead to a successful state dealt with the way the results are validated. In an attempt to obtain better results for future uses of the trained algorithms, an alternative way of validating the results was tested. This alternative involved using a Leave One Group Out (LOGO) cross-validation, where each group contained all the windows of data from collections of one participant. This way, when the classifier tested with the data of one participant, it was guaranteed that no data windows from that participant were used during the training of the algorithm. This approach lowered significantly the results, possibly due to the low amount of participants in the study. As such this approach would only be revisited in the case of a new collection of data, with an increased number of participants.

Chapter 5

Results

Due to the comparative nature of this study, multiple tests were made while changing different conditions of the implementation. This variations included dividing the collection in different time windows, using different device placements or using only the accelerometer data. The conditions of each individual test will be explained before presenting its results. All results will be presented under the metrics mentioned in the theoretical work, as such three tables will be presented, one for each of the metrics. Each of these tables represent one of the metrics for one testing condition, on them, each column represents one of the learning algorithms and each line which sensor placement.

5.1 Window duration variations

These are the most basic variations of tests; in these only the duration of the windows in which the data is split is changed. For this study finding a balanced window size is important to allow us to achieve good results while maintaining a short processing time. While a longer window duration allows for faster processing time, some of the details from single steps can be lost leading to worse performances. Dividing the same data into smaller windows should lead to better performances but also takes longer to process.

5.1.1 5 second windows

In this first variation the duration chosen was 5 seconds. For all tests under this section data from both accelerometer and gyroscope was used, using only two minutes of each collection to skip both the start and end of the collections.

Accuracy:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	86.40	96.51	**97.38**	92.13
Chest	82.68	93.00	92.83	92.00
Belt back	85.00	95.24	97.07	93.77
Belt forward	85.58	97.26	93.98	88.49
Phone	81.66	94.27	94.82	92.02
All	66.32	92.51	92.18	89.23

In this first table the algorithm that obtained the best accuracy was the SVM with the RBF kernel, which means it was the best algorithm, and kernel, in correctly identifying both fatigued and non-fatigued data. Both the RBF and polynomial kernels obtained very high results, followed by the neural network and finally the SVM classifier with the linear kernel. The data from the ankle device obtained the best result for all the SVM kernels, while only the Neural Network obtained a better result using the data from the belt back device.

Sensitivity:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	87.62	**98.58**	96.88	90.67
Chest	85.06	93.75	94.82	92.28
Belt back	83.37	94.78	95.75	90.64
Belt forward	86.44	98.40	95.70	82.70
Phone	84.33	95.90	95.71	92.77
All	65.53	93.79	92.24	89.53

The algorithm that obtained the best result identifying true positives was the SVM with the polynomial kernel. In this setting this means the algorithm that was better able to identify windows with fatigued data out of all windows with fatigued data. Again, both the RBF and polynomial kernels obtained very high results, followed by the neural network and finally the SVM classifier with the linear kernel. Similar to the accuracy table, all the SVM kernels obtained their best result with the ankle placement with the only difference coming from the Neural Network obtaining its best result from the phone data.

Specificity:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	84.48	94.61	97.92	93.42
Chest	80.69	92.41	91.01	91.67
Belt back	86.54	95.47	**98.24**	96.76
Belt forward	84.81	96.18	92.07	94.24
Phone	79.17	92.56	93.94	91.22
All	67.10	91.22	92.10	88.94

In the identification of true negatives, which equals to correctly identifying windows of non-fatigued data, the algorithm which obtained the best result was the SVM with the RBF kernel. As previously, both the RBF and polynomial kernels obtained very high results, followed by the neural network and finally the SVM classifier with the linear kernel. This third metric is the one where most differences where found in terms of the device placement, where the belt region was superior. In this case, the best result for all algorithms was found with belt back placement, with the exception of the polynomial kernel where the belt forward obtained a better result.

Looking at the overview for these testing conditions, in both the accuracy and the specificity metrics the RBF kernel obtained the highest result of each of the tables, while the highest sensitivity was obtained with the Polynomial kernel. As previously, the first two performance assessment metrics obtained their best result for the same device placement (ankle), while the specificity obtained its best results with the devices at the belt level, mostly with the belt back device. The only outlier result observed was the lowest performance of the SVM algorithm with the linear kernel when processing the data from all device placements. In this situation performance in all metrics decreases significantly.

5.1.2 3 second windows

The duration chosen for the second variation was 3 seconds, and as explained previously the remaining variables are kept with the same values.

Accuracy:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	86.60	97.22	97.63	94.12
Chest	84.41	92.17	94.00	92.16
Belt back	87.00	96.21	97.51	95.35
Belt forward	88.38	95.93	**97.96**	94.60
Phone	74.79	89.53	93.06	89.69
All	65.38	91.66	92.31	87.48

As with the previous variation, the algorithm that obtained the best accuracy was the SVM with the RBF kernel. Again, the RBF kernel was accompanied by the polynomial kernel on obtaining the best results, followed by the neural network and finally the linear kernel of the SVM classifier. In this variation there was not a single placement that obtained most of the best results, with the belt forward being the best placement for both the linear and RBF kernel, while the polynomial kernel achieved its best result with the ankle placement and the Neural Network with the belt back one.

When comparing the results of this variation with those of the previous one, the 3 second window obtained the higher best result, although this did not happen for every algorithm, nor for every device placement. In the first four device placements (ankle, chest, belt back and belt forward) the 3 second window obtained better accuracies in 14 out of a possible 16 combinations, while for the last two placements (phone and all) this variation only obtained a better result once in 8 combinations. Still, the accuracy results achieved with both these conditions are similar, with most combinations obtaining under 5% difference between variations.

Sensitivity:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	87.60	97.55	97.65	94.71
Chest	84.12	92.23	93.94	93.22
Belt back	85.73	97.34	96.98	95.03
Belt forward	88.22	98.08	**98.40**	93.11
Phone	75.77	93.02	93.84	89.57
All	65.55	92.69	92.70	88.00

In terms of the best percentage of true positives, the algorithm that obtained the best results was the SVM classifier with RBF kernel. As with previous cases, the best results were obtained with the RBF and polynomial kernels, followed by the neural network and the linear kernel of the SVM classifier. In this metric the best placement for all the SVM kernels was the belt forward, while the Neural Network achieved a better result with the belt back placement.

Contrary to the previous metric, the best sensitivity was obtained with the 5 second window, but similarly the results were not constantly better with one of the variations, with the 3 second window having an higher sensitivity in 11 out of the 24 total combinations. While in most combinations the sensitivity achieved between the two variations is close for the linear kernel, with the phone data the 3 second window only achieves a 75.77%, which is significantly lower than the 84.33% from the 5 second variation.

Specificity:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	85.51	96.84	97.42	93.45
Chest	84.14	91.89	93.88	91.17
Belt back	88.44	95.46	**97.91**	95.45
Belt forward	88.36	93.60	97.49	96.17
Phone	73.80	85.97	92.27	89.84
All	65.27	90.64	91.93	86.95

In this variation the best result, in terms of true negatives, was obtained by the SVM with the RBF kernel. Continuing the trend, both the RBF and polynomial kernels obtained very high results, followed by the neural network and finally the SVM classifier with the linear kernel. Placement wise the results are again mixed, while with the linear and RBF kernels obtaining better results with the belt back placement, the polynomial performed better with the data from the ankle device, and the Neural Network got its best result from the belt forward device.

This last metric follows the sensitivity, with the highest better specificity coming from the 5 second window. This variation only obtained better specificities in 9 of the 24 combination, with the 5 second window being superior for two device placements (phone and all). Overall results between these 2 variations were once again close with the biggest difference coming from the polynomial kernel, using the phone data, where the 3 second variation obtained a worse specificity by over 6%.

Recapping, for the 3 second window variation, all three metrics obtained their best result using the RBF kernel, with the first two (accuracy and sensitivity) obtaining it using the data from the belt forward device placement, while the specificity best result came from the belt back one. This variation obtained results close to the ones from the 5 second window, with neither variation obtaining consistently better results.

5.1.3 10 second windows

In this third variation the duration chosen was 10 seconds. As with the previous cases, this was the only variable changed in this tests, all others were kept with the same value.

Accuracy:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	82.97	**96.79**	94.69	87.04
Chest	83.92	92.45	93.10	91.42
Belt back	81.77	96.28	96.30	93.69
Belt forward	80.87	94.74	81.33	89.61
Phone	83.57	95.91	95.99	91.59
All	67.35	84.76	93.41	86.16

The algorithm that obtained the best accuracy, in this third variation, was the SVM with the polynomial kernel. Both it and the RBF kernel obtained high results which were followed by the neural network and with lower results the linear kernel of the SVM classifier. Looking at the best accuracies from the device placement perspective reveals a diverse origin for the results, with the linear kernel achieving its best result with the chest placement, while the polynomial achieved it with the ankle device, and both the RBF kernel and Neural Network achieved it with the data from the belt back placement.

Comparing these results to those of the previous variations, the maximum accuracy achieved under these conditions is slightly inferior. Even though the 10 second windows achieve better results in 9 combinations, when compared to each previous variations, there are combinations where the result obtained is considerably lower. For example, the accuracy achieved with the RBF kernel, and the data from the belt forward device was only 81.33%, which is considerably inferior to the 93.98% obtained using 5 second windows and pales to the 97.96% from the 3 second variation.

Sensitivity:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	81.13	96.75	95.17	85.82
Chest	85.94	91.61	94.70	91.80
Belt back	83.95	95.61	96.50	94.39
Belt forward	77.79	95.36	87.77	88.70
Phone	85.28	96.54	**96.79**	91.09
All	71.56	81.12	94.12	83.88

In this variation, the best sensitivity was obtained by the SVM classifier with the RBF kernel. It was once again accompanied by the polynomial kernel in its high results with the neural network and the linear kernel of the SVM classifier obtaining slightly lower results. Device placement-wise the results are again varied, with each algorithm, and kernel, obtaining its best sensitivity from a different placement; chest, ankle, phone and belt back were the best placements for the linear kernel, polynomial kernel, RBF kernel and Neural Network, respectively.

As was the case with the accuracy, the maximum sensitivity achieved was slightly inferior to those of previous variations. Still, the 10 second windows managed to obtain better results in 11 out of the 24 combinations when compared to the 5 second windows, with that number dropping to 8 when compared to the 3 second variation. The results also show a bigger difference at points, with linear kernel having a worse performance, using the belt forward device, by around 10% but then an increase, when using the data from all devices, by over 5%.

Specificity:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	83.83	97.00	94.54	90.42
Chest	81.70	92.39	91.85	90.92
Belt back	80.06	**97.67**	96.42	93.25
Belt forward	82.49	94.61	77.86	91.43
Phone	81.93	95.37	95.24	92.09
All	63.29	88.44	92.57	88.25

Finally, in terms of specificity the best result was obtained by the SVM with the polynomial kernel. As with the previous variation, it was accompanied by the RBF kernel in its high results, followed by the neural network and the linear kernel of the SVM classifier. Only the linear kernel performed better with the ankle device placement, with the rest of the algorithms, and kernels, finding their best specificity when using the data from the belt back device.

Looking at the maximum specificity, for this variation, it is slightly inferior to the previous variations best results, as was the case with the other metrics. This variation can, however, outperform the previous two in 9 and 10 of the 24 possible combinations, respectively. There is a visible lower performance of the RBF kernel, with the belt forward device, that causes a massive difference to previous variations.

The 10 second windows brought a bigger variation of results. While some algorithms still achieved good results under these conditions, the best result achieved was always inferior to previous variations. On this variation both the accuracy and specificity found their best results with the polynomial kernel, while the best specificity was achieved with the RBF kernel. Each metric got their best result from a different device placement, respectively, ankle, phone and belt back for accuracy, sensitivity and specificity.

5.2 Other variations

5.2.1 Accelerometer data only

This test was performed with only the accelerometer data, the window duration chosen was 5 seconds, which is explained in the discussion of results, and only 2 minutes of data per collection. The reasoning behind this test was to discover if similarly positive results could be achieved while only using the accelerometer sensor of the devices. If such results could be achieved, it would open the possibility of saving battery on the used devices, or even using smaller, cheaper devices, possibly less intrusive for a potential user.

Accuracy:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	74.79	79.60	91.47	82.25
Chest	79.82	89.43	91.03	89.90
Belt back	78.77	84.61	**97.46**	93.46
Belt forward	72.74	90.47	93.20	93.87
Phone	71.11	88.16	90.57	81.90
All	60.45	85.54	87.48	81.96

In this test the previous trend was somewhat maintained, with the RBF kernel of the SVM classifier obtaining the best accuracy. Differing from previous results, however, the polynomial kernel dropped behind the neural network in accuracy with the lowest results still being obtained by the linear kernel of the SVM. The best placement was once again not unanimous, with the belt forward leading to the best performance for the polynomial kernel and Neural Network, while the linear kernel got its best result using the data from the chest device and the RBF kernel from the belt back placement.

Since 5 second windows were used to split the data, the comparisons will be made against the 5 second window variation in an attempt to isolate the effects of the use of only accelerometer data. The best accuracy was achieved by this variation but overall the results achieved were inferior, with this variation only outperforming the 5 second window variation on 2 combinations out of 24 possible. To worsen the situation, in multiple combinations this variation is outperformed, by the 5 second window variation, by over 10%.

Sensitivity:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	74.55	93.94	91.82	80.44
Chest	81.28	93.06	93.97	90.68
Belt back	74.56	73.54	**96.88**	91.25
Belt forward	76.04	90.34	96.21	94.48
Phone	71.98	87.27	90.57	84.72
All	61.22	85.74	87.98	82.03

For sensitivity the RBF kernel obtained the best result once again. Both the polynomial kernel and the neural network followed with slightly lower results and once again the linear kernel obtained the lowest results. In this variation, each algorithm obtained its best sensitivity with a different device placement. The linear kernel achieved it using the chest device, while the polynomial kernel best performance was when using the data from the ankle device, the belt back was the best device for the RBF kernel and, finally, the Neural Network performed better with the belt forward device.

The best sensitivity ended up being slightly inferior to the 5 second window variation. There is an increase from the last metric as this variation obtained a better sensitivity in

3 combinations, one more than for the accuracy. Even still, the overall results are still inferior, with situations where the difference between the sensitivity of the two variations is larger than 15%.

Specificity:

	SVM Linear	SVM Polynomial	SVM RBF	NN
Ankle	75.47	66.15	91.49	83.57
Chest	78.70	86.46	87.96	89.03
Belt back	82.23	95.03	**97.99**	94.91
Belt forward	69.66	91.59	90.38	93.33
Phone	70.39	89.06	90.51	78.62
All	59.70	85.35	86.96	81.89

In terms of specificity the trend continued with the RBF kernel of the SVM obtaining the best result. In results, it was followed by the neural network, with polynomial kernel of the SVM coming after, and finally the linear kernel. For these conditions, the device placement for the best specificity ended up being the belt back device for all algorithms.

Equally to the previous metric, the best specificity is slightly inferior to the 5 second variation counterpart but, unlike previously, for the specificity this variation never outperforms the 5 second window. In certain combinations, the difference between the performance of the two variations is once again substantial, reaching extremes of over 15% apart.

Overall, the best performing algorithm for this variation was the SVM algorithm, with the RBF kernel, achieving the best result in all metrics. The best placement was also the same for all metrics, with belt back device outperforming the other placements. This variation ended up with overall inferior results and when compared with the 5 second window variation, which used the same windows size to divide the data, these inferior results can be attributed to the main difference of this variation, the use of only accelerometer data.

5.3 Discussion

Summary of results:

Best accuracy:

Variations	SVM Linear	SVM Polynomial	SVM RBF	NN
5 sec window	86,40 (Ankle)	**97,26 (Belt for.)**	97,38 (Ankle)	93,77 (Belt back)
3 sec window	**88,38 (Belt for.)**	97,22 (Ankle)	**97,96 (Belt for.)**	**95,35 (Belt back)**
10 sec window	83,92 (Chest)	96,79 (Ankle)	96,30 (Belt back)	93,69 (Belt back)
Accelerometer only	79,82 (Chest)	90,47 (Belt for.)	97,46 (Belt back)	93,87 (Belt for.)

Best sensitivity:

Variations	SVM Linear	SVM Polynomial	SVM RBF	NN
5 sec window	87.62 (Ankle)	**98.58 (Ankle)**	96.88 (Ankle)	92.77 (Phone)
3 sec window	**88,22 (Belt for.)**	98,08 (Belt for.)	**98,40 (Belt for.)**	**95,03 (Belt back)**
10 sec window	85.94 (Chest)	96,75 (Ankle)	96,79 (Phone)	94,39 (Belt back)
Accelerometer only	81,28 (Chest)	93,94 (Ankle)	96,88 (Belt back)	94,48 (Belt for.)

Best specificity:

Variations	SVM Linear	SVM Polynomial	SVM RBF	NN
5 sec window	86,54 (Belt back)	96,18 (Belt for.)	**98,24 (Belt back)**	**96,76 (Belt back)**
3 sec window	**88,44 (Belt back)**	96,84 (Ankle)	97,91 (Belt back)	96,17 (Belt for.)
10 sec window	83,83 (Ankle)	**97,67 (Belt back)**	96,42 (Belt back)	93,25 (Belt back)
Accelerometer only	82,23 (Belt back)	95,03 (Belt back)	97,99 (Belt back)	94,91 (Belt back)

The previous tables present the best results in each of the metrics chosen, sorted by algorithm used and test variation. Each square identifies which sensor placement was used to obtain that result (the belt forward location was shortened to belt for. for ease of read) and the best result of each algorithm, column, is highlighted in bold.

According to the objectives of this dissertation, the first goal was the realization of a comparative study of learning algorithms for the given problematic. Looking at the results of the first section of tests, all algorithms had a positive performance on the conditions tested with a few minor exceptions. Comparing the results, the linear kernel ended up performing the worst out of the tested algorithms, and kernels, which was visible by the continuous trend of being the classifier with the worst results in all tables. The neural network achieved high results but slightly lower than the other two kernels of the SVM classifier in almost all tables. The two highest performing classifier/kernel combination were the polynomial and RBF kernels of the SVM classifier. Both achieved the highest results in all tables, with the RBF kernel usually achieving the highest result of all. Also to mention the reduction in performance of the polynomial classifier when only the accelerometer data was used to train and test the classifiers.

To reach a conclusion about the best result achieved, a discussion on the impact of each one of the metrics in a future application is necessary. Looking back at the section about these metrics, in the theoretical work chapter, each of their functions is clarified. For the problem studied in this dissertation, achieving a good accuracy was always an objective but due to any possible future uses of this work, it was decided that a very important factor would be avoiding false positives. In the event of a version of this work being integrated into a larger fall prevention algorithm, getting caught in false positives could influence the algorithm into taking wrong conclusions. To reach this goal of avoiding false positives the most important metric is the specificity, which measures the percentage of negative cases, for this dissertation this means, windows of data without fatigue that were successfully identified. The higher the specificity, the fewer the cases where a negative situation was classified as positive, i.e.false positives. With this in mind, the combination of conditions and algorithms that achieved the highest specificity was the SVM classifier with the RBF classifier, when using a window duration of 5 seconds and the data from the device placed on the back of the waist region, with a specificity of 98,24%.

In terms of device placement, all single device locations had similar results, although both the ankle and the belt back (device placed in the waist region on the back) achieved the best performances. Considering what was previously stated about the importance of the specificity results, due to the necessities of this study, there is a noticeable trend of the best specificity being obtained with the device at the back of the belt, which happened in 8 of the 12 results. It is also noticeable that the conjugation of all devices wielded lower performances, paired with the higher processing time necessary for these tests. It was concluded that a simpler approach with a single device could suffice in this fatigue detection problem.

The tested variations of duration of the window in which the data is divided did not seem to have a big impact on the overall performance. Considering that the 3 seconds test took the longest processing time and, as such, 5 seconds was chosen as an appropriate duration for the other tests.

Using only the data from the accelerometers, the results achieved were noticeably worse, especially when using the SVM classifier with the polynomial kernel, which apart from this test was one of the classifiers and kernels with the highest results. Even so, acceptable performances were still achieved, mainly with the RBF kernel, which was able to somewhat maintain the good performances achieved on the other tests.

Chapter 6

Conclusions

Overall, the learning algorithms obtained good performances when trying to identify fatigued gait, obtaining a maximum of 97.96% accuracy (overall performance), 98.58% sensitivity (identification of fatigued data) and 98.24% specificity (identification of non-fatigued data). Looking back at the studies presented on the "State of the Art"chapter, these results are on a par to those achieved in them, which achieved maximum accuracies of 97%, 90% and 98.1% (in the order they were presented).

With the current conditions, it was found that a single device can suffice for the collection of data allowing for a less intrusive data collection. As previously explained, the considered best device positioning for this classification ended up being the lumbar region, represented as belt back in this study, although most device placements obtained satisfactory results. Looking at the three tables of the best results for each metric out of the 48 combinations of variations and algorithms the result was obtained with the belt back position in 20 of those combinations. In comparison the ankle and belt forward placements were the best in 11 combination each, with the chest placement trailing behind in only 4 combinations and finally the phone placement was the best in 2 of those combinations.

For the processing of said data, the SVM classifier, with both the RBF kernel and the polynomial kernel, obtained the best results in most situations. Not only the best result for each metric was obtained by one of these kernels, accuracy and specificity by the RBF and sensitivity by the polynomial, these kernels also achieved results above 96% in all metrics (except for the test conditions using only accelerometer data, where the polynomial kernel had a loss in performance). In comparison, the best results achieved by the neural network varied between 92% and 97% in all metrics, while the SVM classifier with the linear kernel only achieved results between 79% and 89%.

Different tests were done in an attempt to discover the best conditions to collect and process gait data. When trying to analyse the best conditions for each algorithm, an emphasis was given to the specificity metric. As such, the best conditions were considered to be the ones with the best specificity except in situations where there would be a significantly negative impact the other metrics. This way, it was observed that the best conditions for the linear kernel, of the SVM algorithm, were the 3 second window times with data from the belt forward device, obtaining an accuracy of 88,38%, sensitivity of 88,22% and specificity of 88,36%. For the polynomial kernel, the 10 second window was deemed the best, when using data from the belt back device. This combinations achieved an accuracy of 96,28%, sensitivity of 95,61% and specificity of 97,67%. The RBF kernel attained its best results when using data from the belt back device, splitting it into 5 second windows. This way, it achieved an accuracy of 97,07%, sensitivity of 95,75% and specificity of 98,24%. Finally, the best results for the Neural Network came from the belt forward device, with 3 second windows, achieving an accuracy of 94,60%, sensitivity of 93,11% and specificity of 96,17%.

Out of all the previously presented combinations of algorithm, collection and processing conditions, the best combination for the requirements of this study was considered to be the window duration of 5 seconds, while using data from the belt back device position and processing it with the SVM algorithm and RBF kernel.

6.1 Future work

During the development of the dissertation there were some paths that either posed some problems during implementation or simply did not lead to successful results at the current time. As such, these approaches were abandoned and left to explore as future work.

The most important issue to discuss in this section is the sample size of the data. Any possible follow up to this work would require a new data collection with an increased number of participants to be able to draw conclusions. Another possible development path still in the data collection realm is targeting a more specific demographic, such as the elderly. This way, the conclusions reached can be of more use for a demographic heavily affected by gait and mobility issues. A re-planned data collection could also allow the calculation of new features, specifically counteracting the problems faced due to the lack of magnetometer data.

Another point to add while talking about new features is the implementation of the necessary formulas to calculate the features themselves. Along with the previous mentioned extra information, that requires a new data collection, all the extra necessary work to implement new features would have to be done. As mentioned during the "Previous approaches"section (4.4) in chapter 4, every new feature would have to be researched, implemented and tested one by one.

Lastly, also possibly due to lack of data, the necessary steps to successfully implement, and use, a LOGO, cross-validation, scheme are encouraged to be taken. While the used 10-fold cross-validation is valid, the previously mentioned LOGO guarantees better results for future implementations of the algorithms.

Bibliography

[1] H. Abdi and L. J. Williams. "Principal component analysis." In: *Wiley Interdisciplinary Reviews: Computational Statistics* 2.4 (2010), pp. 433–459. issn: 19395108. doi: 10.1002/wics.101. url: http://doi.wiley.com/10.1002/wics.101.

[2] A. Baghdadi, F. M. Megahed, E. T. Esfahani, and L. A. Cavuoto. "A machine learning approach to detect changes in gait parameters following a fatiguing occupational task." In: *Ergonomics* 61.8 (2018), pp. 1116–1129. issn: 0014-0139. doi: 10.1080/00140139.2018.1442936. url: http://www.ncbi.nlm.nih.gov/pubmed/29452575https://www.tandfonline.com/doi/full/10.1080/00140139.2018.1442936.

[3] F. A. Barbieri, P. C. R. dos Santos, R. Vitório, J. H. van Dieën, and L. T. B. Gobbi. "Effect of muscle fatigue and physical activity level in motor control of the gait of young adults." In: *Gait & Posture* 38.4 (2013), pp. 702–707. issn: 0966-6362. doi: 10.1016/J.GAITPOST.2013.03.006. url: https://www.sciencedirect.com/science/article/pii/S0966636213001604?via{\%}3Dihub.

[4] R. Begg, M. Palaniswami, and B. Owen. "Support Vector Machines for Automated Gait Classification." In: *IEEE Transactions on Biomedical Engineering* 52.5 (2005), pp. 828–838. issn: 0018-9294. doi: 10.1109/TBME.2005.845241. url: http://ieeexplore.ieee.org/document/1420704/.

[5] G. A. BORG. "Psychophysical bases of perceived exertion." In: *Medicine & Science in Sports & Exercise* 14.5 (1982), 377???381. doi: 10.1249/00005768-198205000-00012. url: https://www.ncbi.nlm.nih.gov/pubmed/7154893?dopt=Citation.

[6] R. Evans and D. Arvind. "Detection of Gait Phases Using Orient Specks for Mobile Clinical Gait Analysis." In: *2014 11th International Conference on Wearable and Implantable Body Sensor Networks*. IEEE, 2014, pp. 149–154. isbn: 978-1-4799-4959-5. doi: 10.1109/BSN.2014.22. url: http://ieeexplore.ieee.org/document/6855633/.

[7] A. Hareendran, N. Leidy, B. U Monz, R. Winnette, K. Becker, and D. A Mahler. "Proposing a standardized method for evaluating patient report of the intensity of dyspnea during exercise testing in COPD." In: *International journal of chronic obstructive pulmonary disease* 7 (May 2012), pp. 345–55. doi: 10.2147/COPD.S29571.

[8] T. Hastie, R. Tibshirani, and J. Friedman. *Springer Series in Statistics The Elements of Statistical Learning Data Mining, Inference, and Prediction*. Tech. rep.

[9] J. L. Helbostad, S. Leirfall, R. Moe-Nilssen, and O. Sletvold. "Physical Fatigue Affects Gait Characteristics in Older Persons." In: *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 62.9 (2007), pp. 1010–1015. ISSN: 1079-5006. DOI: 10.1093/gerona/62.9.1010. URL: https://academic.oup.com/biomedgerontology/article-lookup/doi/10.1093/gerona/62.9.1010.

[10] D. Janssen, K. M. Newell, J. M. Jäger, F. Rost, and K. Vehof. "Diagnosing fatigue in gait patterns by support vector machines and self-organizing maps." In: *Human Movement Science* 30.5 (2011), pp. 966–975. ISSN: 0167-9457. DOI: 10.1016/J.HUMOV.2010.08.010. URL: https://www.sciencedirect.com/science/article/pii/S0167945710001478{\#}s0030.

[11] M. Karg, G. Venture, J. Hoey, and D. Kulic. "Human Movement Analysis as a Measure for Fatigue: A Hidden Markov-Based Approach." In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22.3 (2014), pp. 470–481. ISSN: 1534-4320. DOI: 10.1109/TNSRE.2013.2291327. URL: http://ieeexplore.ieee.org/document/6716986/.

[12] A. Kharb, D. V. Saini, Y. Jain, and S. Dhiman. *A REVIEW OF GAIT CYCLE AND ITS PARAMETERS*. 2011. URL: https://www.semanticscholar.org/paper/A-REVIEW-OF-GAIT-CYCLE-AND-ITS-PARAMETERS-Kharb-Saini/527ac5081f2b490414defe9a57c3e17

[13] S.-u. Ko, J. M. Hausdorff, and L. Ferrucci. "Age-associated differences in the gait pattern changes of older adults during fast-speed and fatigue conditions: results from the Baltimore longitudinal study of ageing." In: *Age and Ageing* 39.6 (2010), pp. 688–694. ISSN: 0002-0729. DOI: 10.1093/ageing/afq113. URL: https://academic.oup.com/ageing/article-lookup/doi/10.1093/ageing/afq113.

[14] L. Lu, F. M. Megahed, R. F. Sesek, and L. A. Cavuoto. "A survey of the prevalence of fatigue, its precursors and individual coping mechanisms among U.S. manufacturing workers." In: *Applied Ergonomics* 65 (2017), pp. 139–151. ISSN: 0003-6870. DOI: 10.1016/J.APERGO.2017.06.004. URL: https://www.sciencedirect.com/science/article/pii/S000368701730131X{\#}sec5.

[15] S. Marreiros, M. Schepers, G Bellusci, M. de Zee, M. Andersen, and A. Karatsidis. "Comparing Performance of Two Methods to Process Inertial Data in Gait Analysis." In: 2017.

[16] J. Perry, J. M. Burnfield, and L. M. Cabico. *Gait analysis : normal and pathological function*. SLACK, 2010, p. 551. ISBN: 9781556427664.

[17] X. Qu and J. C. Yeo. "Effects of load carriage and fatigue on gait characteristics." In: *Journal of Biomechanics* 44.7 (2011), pp. 1259–1263. ISSN: 0021-9290. DOI: 10.1016/J.JBIOMECH.2011.02.016. URL: https://www.sciencedirect.com/science/article/pii/S0021929011001084{\#}s0040.

[18] A. Z. Rakhman, L. E. Nugroho, Widyawan, and Kurnianingsih. "Fall detection system using accelerometer and gyroscope based on smartphone." In: *2014 The 1st International Conference on Information Technology, Computer, and Electrical Engineering*. IEEE, 2014, pp. 99–104. ISBN: 978-1-4799-6431-4. DOI: 10.1109/ICITACEE.2014.7065722. URL: http://ieeexplore.ieee.org/document/7065722/.

[19] P. Raknim and K.-c. Lan. "Gait Monitoring for Early Neurological Disorder Detection Using Sensors in a Smartphone: Validation and a Case Study of Parkinsonism." In: *Telemedicine and e-Health* 22.1 (2016), pp. 75–81. ISSN: 1530-5627. DOI: 10.1089/tmj.2015.0005. URL: https://www.liebertpub.com/doi/10.1089/tmj.2015.0005.

[20] S. Raychoudhury. "Three dimensional kinematics of human metatarsophalangeal joint during locomotion | Research Explorer | The University of Manchester." Doctoral dissertation. 2015. URL: https://www.research.manchester.ac.uk/portal/en/theses/three-dimensional-kinematics-of-human-metatarsophalangeal-joint-during-locomotion(547d151c-a29a-448c-a8fb-004739795c73).html.

[21] J. A. Ricci, E. Chee, A. L. Lorandeau, and J. Berger. "Fatigue in the U.S. Workforce: Prevalence and Implications for Lost Productive Work Time." In: *Journal of Occupational and Environmental Medicine* 49.1 (2007), pp. 1–10. ISSN: 1076-2752. DOI: 10.1097/01.jom.0000249782.60321.2a. URL: http://www.ncbi.nlm.nih.gov/pubmed/17215708https://insights.ovid.com/crossref?an=00043764-200701000-00001.

[22] T. Sayeed, A. Samà, A. Català, and J. Cabestany. "Comparison and adaptation of step length and gait speed estimators from single belt worn accelerometer positioned on lateral side of the body." In: *2013 IEEE 8th International Symposium on Intelligent Signal Processing*. 2013, pp. 14–20. DOI: 10.1109/WISP.2013.6657475.

[23] O. S. Schneider. "Gait sensing on smartphones to support mobile exercise applications and games." In: (2012). DOI: 10.14288/1.0052125. URL: https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0052125.

[24] W. Tao, T. Liu, R. Zheng, H. Feng, W. Tao, T. Liu, R. Zheng, and H. Feng. "Gait Analysis Using Wearable Sensors." In: *Sensors* 12.2 (2012), pp. 2255–2283. ISSN: 1424-8220. DOI: 10.3390/s120202255. URL: http://www.mdpi.com/1424-8220/12/2/2255.

[25] V. N. Vapnik. *The Nature of Statistical Learning Theory*. New York, NY: Springer New York, 2000. ISBN: 978-1-4419-3160-3. DOI: 10.1007/978-1-4757-3264-1. URL: http://link.springer.com/10.1007/978-1-4757-3264-1.

[26] J. Zhang, T. E. Lockhart, and R. Soangra. "Classifying Lower Extremity Muscle Fatigue During Walking Using Machine Learning and Inertial Sensors." In: *Annals of Biomedical Engineering* 42.3 (2014), pp. 600–612. ISSN: 0090-6964. DOI: 10.

1007/s10439-013-0917-0. URL: http://link.springer.com/10.1007/s10439-013-0917-0.

[27] W. Zhu, N. Zeng, and N. Wang. *Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS ® Implementations*. Tech. rep. 2010.

Annex I

Participation Form

This annex contains the form filled by the participants of the study before the data collection process. More information on this process is available in the experimental protocol section of the fourth chapter.

Personal Data

Sex: M ☐ F ☐

Age: _____________

Height (cm): ___________

Weight (kg): ___________

Dominant foot: Right ☐ Left ☐

Shoe size (EUR): __________

Health problems (that may affect gait): ____

____

____

Medication (that may affect gait): Yes ☐ No ☐

Preferred fatiguing exercise: Squatting ☐ Walk up and down stairs ☐

☐ I accept that this data may be used as statistics for this study

____ ________________________

Participant's Signature **Date**

Annex II

Features calculated

This annex contains a table with name of the features returned by the TSFEL tool. For the sake of legibility and in an effort for the table to fit in a single page the Histogram features were truncated as the full table would include an entry for each from 0 to 9. This table only includes one sixth of the number of features as there is one for each dimension and sensor combination (Ax, Ay, Az, Gx, Gy, Gz). This is explained in the data processing section of the fourth chapter.

Autocorrelation
Centroid
Fundamental frequency
Histogram 0
Histogram 9
Index of highest fft
Interquartile range
Kurtosis
Linear regression
Max
Max power spectrum
Maximum frequency
Maximum peaks
Mean
Mean absolute deviation
Mean absolute diff
Mean diff
Median
Median absolute deviation
Median absolute diff
Median diff
Median frequency
Min
Minimum peaks
Root mean square
Signal distance
Skewness
Spectral centroid
Spectral decrease
Spectral kurtosis
Spectral maximum peaks
Spectral roll-off
Spectral roll-on
Spectral skewness
Spectral slope
Spectral spread
Spectral variation
Standard Deviation
Sum absolute diff
Total energy
Variance
Zero crossing rate

Table II.1: Name of features calculated by the TSFEL tool