Master Thesis

DEVELOPMENT OF A FATIGUE ESTIMATION MODEL FOR PHYSICAL REHABILITATION EXERCISES

Maria José Pinto Bernal

Supervisor:
Prof.Dr. Marcela Múnera

Co-supervisor:
Prof. Dr. Carlos Andrés Cifuentes García

A thesis submitted in fulfillment of the requirements for the degree of
Master in Electronic Engineering

NOVEMBER 2021

"It doesn't matter how beautiful your theory is

it doesn't matter how smart you are

if it doesn't agree with experiment, its wrong

in that simple statement

is the key of science"

Richard Feynman−Gyorgyi

"Research is to see what everybody

else has seen, and to think

what nobody else has thought."

Albert Szent−Gyorgyi

Acknowledgements

I would like to thank my thesis advisors, who have closely followed the development of my work and were always available to provide all the guidance and advice through this arduous but pleasant process of becoming a researcher. To Marcela for her support in each of the ups and downs of my career, she taught me to move on, to make our voice count, and that it does not matter who or how we are because each person has a great value. To Carlos for believing in me, opportunities, and teaching me to always look for my best version and always give my best. Likewise, I would like to thank Oscar Perdomo for his support, teachings and encouragement during the development of this master's degree.

It is essential to emphasize that this work would not have been possible without many people, volunteers, relatives, and friends. In particular, I would like to thank my mom Liliana for her constant support, unconditional love, and teaching me always to keep going no matter how hard the situation becomes. Thank you to my brother Juan Pablo for being my companion in my adventures and constantly reminding me how valuable and ephemeral life is. Finally, and in a very special way, I would like to express my immense gratitude and love to my mentor, my inspiration, the center of my world, and the love of my life, my dad, who was always believed in me, taught me to always give the best of me without ever losing my identity, and with his unconditional support, wisdom and love have made me the person that I am. I love you, daddy, and this would not have been possible without you. Forever and ever.

Finally, this thesis contributes to the development and scope of the SORCAR research project. I would like to acknowledge the funding institutions of this project, such as the Minister of Science, Technology and Innovation Colciencias with the SORCAR project (Grant 813-2017) and the internal projects of the Colombian School of Engineering Julio Garavito.

Quiero agradecer a mis directores de tesis, quienes han seguido de cerca el desarrollo de mi trabajo y siempre estuvieron dispuestos a brindarme toda la orientación y consejos en este arduo pero placentero proceso de convertirme en investigadora. A Marcela por su apoyo *en cada uno de los altibajos de mi carrera, por enseñarme a seguir adelante, a hacer valer nuestra voz y que no importa quién o cómo seamos porque cada persona tiene un gran valor. A Carlos, por creer en mí, por las oportunidades, por enseñarme a buscar siempre mi mejor versión y enseñarme a dar siempre lo mejor de mí. Del mismo modo, quiero agradecer a Oscar Perdomo, por su apoyo, enseñanzas y motivación durante el desarrollo de ests tesís de maestría.*

Por otra parte, es importante destacar que este trabajo no hubiese sido posible sin el apoyo de un gran grupo de personas, familiares y amigos. En particular, quiero agradecer a mi mamá Liliana por tener siempre su apoyo, por su ejemplo, por su amor incondicional y por enseñarme a seguir siempre adelante sin importar lo difícil que sea la situación. Gracias a mi hermano Juan Pablo por ser mi compañero de aventuras y recordarme constantemente lo valiosa y efímera que es la vida. Finalmente, y de manera muy especial, quiero expresar mi inmensa gratitud y amor a mi mentor, mi inspiración, al centro de mi mundo y al amor de mi vida, mi papá, quien siempre creyó en mí, me enseñó a dar siempre lo mejor de mí sin perder nunca mi esencía, y gracias a su apoyo incondicional, sabiduría y amor ha hecho de mí la persona que soy. Te amo con mi vida entera, papá, y esto no hubiera sido posible sin ti. Por siempre y para siempre.

Finalmente, la realización de este proyecto de tesis contribuye con el desarrollo y alcance del proyecto SORCAR. De esta manera, quiero reconocer a las instituciones financiadoras de este proyecto, como el Ministerio de Ciencia, Tecnología e Innovación Colciencias (Grant 813-2017) y los proyectos internos de la Escuela Colombiana de Ingeniería Julio Garavito.

Abstract

Physical exercise (PE) contributes to achieving a successful rehabilitation program and rehabilitation processes assisted through social robots. However, the amount and intensity of exercise needed to obtain positive results are unknown. Several considerations must be kept in mind for PE implementation in rehabilitation as monitoring of patients' intensity, which is essential to avoid extreme fatigue conditions, which may cause physical and physiological complications. Machine learning models have been implemented to fatigue management but limited in practice due to the lack of understanding of how an individual's performance deteriorates with fatigue accumulation; that can vary based on the physical exercise, environment, and individual's characteristics. As a first step toward realizing the human-centered approach to artificial intelligence and expert systems, this master thesis lays the foundation for a data analytic approach to managing fatigue in walking tasks. The proposed framework capitalizes on continuously collected human performance data from wearable sensor technologies. It establishes criteria for a feature and machine learning algorithm selection for fatigue management, classifying four fatigue diagnoses state. Based on the proposed framework and a large number of test sets used during the evaluation of the classifiers, we have shown that (i) the random forest model presented the best performance with an average accuracy of $\geq 98\%$ and an F-score of $\geq 93\%$, this model was comprised of ≤ 16 features; and (ii) the prediction performance was analyzed by limiting the sensors used from four IMUs to two or even one IMU with an overall performance of $\geq 88\%$; hence, only one wearable sensor is needed for fatigue detection. This research presents an initial approach to a promising tool for physical rehabilitation, and regarding classification accuracy, it presents remarkable results according to the literature. We provide links to our data and code as supplementary materials to encourage future work in this crucial area.

Keywords: Fatigue diagnosis; classification models; inertial measurement units; Physical rehabilitation, walking rehabilitation; physical exercise.

List of Tables

List of Figures

Contents

Chapter 1

Introduction

The work presented in this document focuses on the development and preliminary validation of a fatigue identification model based on the individuals' exercise performance assessment. The main aim of the model is to classify four fatigue diagnosis stages (low, medium, high, and very high) for walking tasks, widely used in physical rehabilitation. Additionally, this thesis presents the framework for developing a fatigue identification classifier based on exercise performance and, a preliminary study aimed at validating the performance of the developed model in 25 healthy subjects.

This chapter introduces the main motivations and research objectives that lead to the development of this work. The main contributions, publications and the organization of this document are presented.

1.1 Motivation

Exercise rehabilitation during or after medical treatment is considered adequatee to restore physical and psychological function [2,3]. Recent reviews highlight some benefits provided by physical exercise used as a therapeutic measure. Physical exercise contributes to achieving and maintaining therapeutic goals and improving quality of life, physical functioning, functional capacities, muscle strength, emotional well-being, and even reducing depression and anxiety and increase self-esteem. Additionally, it can lower the risk of heart disease, diabetes, cancer, stroke, reduce the risk of orthopedic problems, recover mobility of limbs, strengthen the immune defense (influenza), among others [4, 5]. In this context, physical exercise can affect people's health conditions in many ways; hence, the American Health Association established physical exercise as one of the main components for improving people's health and decreasing morbidity and mortality levels [6].

Although these positive health-related outcomes of regular physical exercise are well documented, the amount and intensity of exercise needed to obtain these positive results are unknown [7, 8]. Several considerations must be kept in mind for its implementation in rehabilitation. For instance, elderly patients have more severe impairments and comorbidities than younger patients; therefore, the rehabilitation needs of the older population are different from those of younger patients [2]. The same occurs with individuals with different diseases. Consequently, an essential question when prescribing exercise is the optimal therapeutic dose required to produce a specific health benefit according to the individuals' needs?. Typically, when considering exercise dose about health outcomes, exercise is characterized by type, intensity, and volume (session duration and frequency) according to each patient's age, weight, fitness level and pathologies [9]. Recent studies have shown that intensity is the most relevant feature in prescribing physical exercise [10] because it determines the amount of energy expenditure and can be seen as the "dose" of the prescription [11]. Controlling exercise intensity avoids overtraining patients, affecting their rehabilitation and even causing health consequences (i.e., physical, or physiological complications) [12, 13].

Some successful rehabilitation programs have been explored to supervise the exercise intensity during therapies by monitoring the patient's fatigue state [2, 14–17]. Fatigue has generally been defined as a subjective state of tiredness or exhaustion and the reduction of capacity for regular activity [18]. Additionally, it is defined as the inability of the muscles to maintain the required level of strength during exercises. It can also result in the deterioration of health in the long term, including work-related musculoskeletal disorders [19], chronic fatigue syndrome [20] and compromised immune function [21]. Therefore, fatigue is a common concern among clinicians and individuals who participate in physical activities based on training or rehabilitation [18]. An essential first step in managing fatigue is the rapid and accurate detection of its occurrences. However, there is no scientifically accepted method to identify it because of the wide range of factors that can produce fatigue.

Therefore, nowadays, artificial intelligence systems appear as a complementary and further alternative to monitoring and identifying fatigue [22,23]. These models presented a significant potential for clinical scenarios because they provide an objective indicator of fatigue. Because these methods only considered two fatigue states, i.e., fatigued; or non-fatigued state, which limits the accurate monitoring of the user's exhaustion during therapy, these models restricts the possibility to determine the adequate "dose" (i.e., intensity) of the individuals to produce a specific health benefit according to their individuals' needs. Thus, the use of these models limits the improvement of the user's performance during therapy. Within this context, as a first step, this project proposes a framework for developing a fatigue identification model based on the individuals' exercise performance assessment to classify four fatigue diagnosis stages (low, medium, high, and very high). The fatigue diagnosis stages allow clinicians to pinpoint the hazard directly. They can then prescribe interventions from a large number of options, including assigning rest breaks (which can reduce the level of fatigue before it reaches potentially dangerous levels); or redesigning the activity (which can eliminate the development of fatigue) [22]. Likewise, four fatigue stages accurately monitor patients' fatigue conditions during exercise to avoid any injuries or affect rehabilitation. In this manner, patients, researchers, and clinical staff could benefit from the outcomes of this thesis and the overall SORCAR project (explained below): as they seek to understand better and provide a more efficient rehabilitation process.

1.2 Background

This thesis is developed in the context of the research project "*Human-Robot Interaction Strategies for Rehabilitation based on Socially Assistive Robotics*" (SORCAR) supported by the Ministry of Science, Technology, and Innovation *MinCiencias* (grant 801-2017), as well as, internal funding from the Colombian School of Engineering Julio Garavito (ECIJG). The SORCAR project is primarily led by Prof. Dr. Marcela C. Múnera and, Prof. Dr. Carlos A. Cifuentes (professors at the Department of Biomedical Engineering and head of the *Center for Biomechatronics* at ECIJG). The research team of this project is formed by a cooperation network comprising both national and international research groups and institutions.

The main goal of the SORCAR project is to extend the functionalities and capabilities of Socially Assistive Robotics systems (SAR), such as the development of advanced sensing strategies that allow the implementation of robust and reliable measuring devices useful in clinical scenarios. Likewise, this project focuses on the development and validation of interaction strategies applied in clinical scenarios. In this regard, the establishment of appropriate interaction between the user, the robotic platform, and the environment is required. To this end, it is essential to explore the implementation of the Human-Robot interface (HRi) to determine the user's condition and provide assistance through social interaction. The HRi should have a social interaction module, which are strategies implemented to provide motivation, feedback, or another type of assistance based on the user's condition. The above can be done by measuring several essential parameters implemented in physical rehabilitation, such as physiological and exercise intensity parameters. An example of the SORCAR project's HRi is presented in Figure 1.1. It is possible to appreciate the specific user's parameters used in the monitoring and control module, which correspond to cardiovascular parameters and exercise intensity parameters, as well as, the social interaction which is related to the patient's condition, during the execution of physical training.

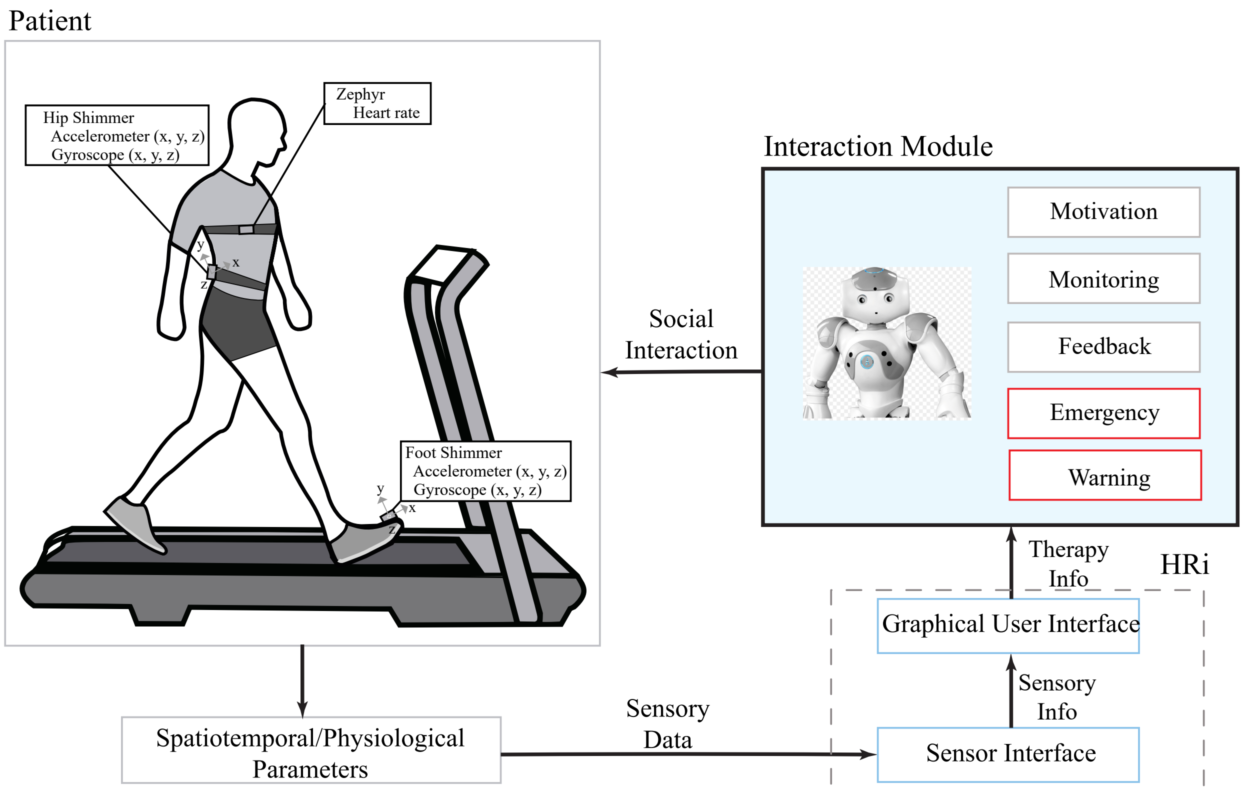

Figure 1.1: General model of the human-robot interface SORCAR project. This model considers two main components: (i) a Human-Robot interface (HRi) designed to retrieve all relevant information from the therapy, and (ii) an interaction module that will be able to provide feedback to the users regarding their performance and the therapy conditions, based on the information obtained from the patient.

According to the above and within the scope of the SORCAR project, this master thesis seeks to develop and validate a fatigue identification model based on the individuals' exercise performance to classify four fatigue diagnosis stages (low, medium, high, and very high). The purpose is to control and monitor the patients' fatigue conditions through the measurement of the patient's physiological and exercise intensity parameters during exercise to avoid any injuries or affect the rehabilitation process. A great application of this model could be its implementation in the SAR systems using HRi, because this could provide continuous monitoring and feedback on the patient's performance, allowing to control the intensity and the correct execution of the exercise. Similarly, alerts could be generated when the patient presents high fatigue levels or variation in exercise performance parameters. This is very promising for clinical scenarios and would help to avoid overtraining patients, affecting their rehabilitation and even causing physical or physiological complications.

1.3 Objectives

Considering the main motivations of this project, it is proposed to develop a machine learning model to estimate the patient's fatigue condition, based on physiological and exercise performance parameters obtained from healthy subjects. This model could allow clinicians to pinpoint the hazard directly; and then prescribe interventions to reduce or avoid fatigue. Thus, this project seeks to provide a robust fatigue classifier; and understand how an individual's performance deteriorates with fatigue accumulation in a walking task. In order to achieve the project proposal the following objectives are defined.

1.3.1 General Objective

The primary objective of this thesis is to develop and evaluate a system to estimate the patient's fatigue condition in walking exercises which is one of the most implemented exercises in physical rehabilitation.

1.3.2 Specific Objectives

- To perform a systematic review of literature to understand the estimation of associated fatigue in aerobic exercise during rehabilitation scenarios.
- To design a system to estimate the user's fatigue condition through monitoring exercise performance features and physiological parameters, for a aerobic exercise according to the systematic review.
- To perform an evaluation study, to validate the system's performance.

1.4 Contributions

The key contributions of this work are framed in the main activities of the SORCAR project. Specifically in the development and validation of strategies applied in clinical scenarios. The fulfillment of this master thesis is accompanied by a series of technical and scientific contributions presented as follows:

1. The design and implementation of a framework for developing a fatigue identification model based on the individuals' exercise performance assessment to classify four fatigue diagnosis stages.
2. The design and development of an experimental protocol for the validation of the machine learning techniques.
3. A public repository of the code and data implemented in this work to encourage adoption in practice and further investigations by researchers.

1.5 Publications

The work presented in this thesis has been reported to the scientific community employing the following publications:

1. (Journal Article) **Pinto-Bernal, Maria J.**, Cifuentes, C.A.; Perdomo, O.; Rincón-Roncancio, M.; Múnera. (2021). "A Data-Driven Approach to Physical Fatigue Management Using Wearable Sensors to Classify Four Diagnostic Fatigue States" Sensors 21, no. 19: 6401.`https://doi.org/10.3390/s21196401`
2. (Journal Article) Aguirre, A., **Pinto, M.J**, Cifuentes, C.A., Perdomo, O., Díaz, C.A.R.; Múnera, M. (2021). "Machine Learning Approach for Fatigue Estimation in Sit-to-Stand Exercise". Sensors 21, no. 15: 5006.`https://doi.org/10.3390/s21155006`
3. (Journal Article - Under review) Alvarez, P., Tello, A., **Pinto, M.J**, Cifuentes, C.A., Perdomo, O. Díaz; Rincón-Roncancio, M.; Múnera, M. (2021). "Fatigue estimation using wearable sensors in physical exercise in cycle ergometer". Sensors Journal.

4. (Book Chapter) **Pinto, M.J**, Aguirre, A., Cifuentes, C.A., Múnera, M. (2020). Wearable Sensors for Monitoring Exercise and Fatigue Estimation in Rehabilitation. Internet of Medical Things: Paradigm of Wearable Devices, CRC Press.

1.6 Document Organization

This master thesis document is structured as follows:

- **Chapter 1** presents the main motivations and research goals of this work. Additionally, this chapter describes the research project in which this thesis is framed, defining the critical contributions of this thesis.
- **Chapter 2** introduces the context of physical exercises and their modulation in rehabilitation programs and, the different strategies or alternatives for fatigue estimation.
- **Chapter 3** describes the current state of the machine learning techniques used to estimate fatigue. This chapter also addresses the literature review concerning the most representative systems implemented on fatigue estimation.
- **Chapter 4** presents a detailed structure of the framework designed for the fatigue classification model. This chapter also describes the dataset construction focused on the most relevant feature related to walking exercises.
- **Chapter 5** addresses the validation study focused on assessing the performance of the proposed fatigue classifier in healthy subjects.
- **Chapter 6** summarizes and highlights the main conclusions and remarks of this works. This chapter also proposes a set of future works to address with the fatigue classifier, regarding the integration with the HRi in SAR systems.

Chapter 2

Physical Exercise: Rehabilitation Scenarios

Historically, patients with chronic diseases were recommended to rest and avoid physical activity [24]. However, excessive rest and lack of physical activity decondition and, reduce functional capacity and quality of life [25]. Current medical opinion has changed to the belief that patients should be encouraged to engage in physical activity during rehabilitation [24], considering that most chronic diseases affect the patients' physical function. In this way, exercise rehabilitation during or after medical treatment is considered an effective means of restoring physical and psychological function. This chapter concisely describes the main components of physical exercise and the main conditions that affect it. This chapter, it is also presented a summary of the conventional techniques for fatigue estimation.

2.1 Physical Exercise

Before analyzing the benefits of Physical exercise (PE), it is essential to define PE precisely. Indeed, PE is a term often incorrectly used interchangeably with physical activity PA that is "any bodily movement produced by skeletal muscles that require energy expenditure" [26].

Then, PA includes any motor behavior such as daily and leisure activities, and it is considered a determinant lifestyle for general health status [27]. Instead, PE is "a sub-classification of PA that is planned, structured, repetitive, and has as a final or an intermediate objective the improvement or maintenance of one or more components of physical fitness" [26].

Cardiology was the first medical specialty in which physical exercise rehabilitation was implemented and evaluated [28]. PE is now commonly prescribed in cardiac patients and is integral to the rehabilitation program. Psychological, social, and physical benefits of physical exercise after myocardial infarction, coronary artery bypass grafting, heart transplantation, and stable congestive heart disease are well-documented [29, 30]. Based on the above, nowadays, rehabilitation treatment includes exercise to enable patients to decrease limitations and facilitate post-treatment exercise performance. Regular physical exercise has the potential to affect people's health conditions in many ways. Physical exercise can mitigate the effects of pathological fatigue (oncology rehabilitation) [31, 32], improve the cardiac system capability (cardiac rehabilitation) [33, 34], improve the respiratory system capability (pulmonary rehabilitation) [35, 36], recover joint strength after a surgery (musculoskeletal rehabilitation) [37, 38], lower the risk of stroke, diabetes, cancer, and reduces the risk of orthopedic problems (osteoporosis) [39, 40]. Therefore, exercise-based rehabilitation is an essential for cardiac, oncology, and orthopedic patients toward improved health and a physically active lifestyle.

At present, physical exercise therapy is being considered for approval as a prescribed medication by the Food and Drug Administration in the USA [41, 42]. The key features of such approval would necessarily include efficacy for the specific condition, effectiveness in the target population, recommended dosing for the designated outcome, mechanisms of action, and the safety or adverse event profile as illustrated in Table 2.1. Just as pharmacological therapies and dietary modifications are individualized for the patient, similarly, a tailored physical activity program could be prescribed for treatment, once approved.

Efficacy	Does it cause a specific health benefit as demonstrated by adequately designed randomized controlled trials?
Effectiveness	Is the specified benefit obtained by a reasonable percentage of the persons who undertaking the prescribed exercise regimen? Who will be the responders and the non-responders?
Dose	What dose of exercise provides a meaningful benefit for this specific condition? The prescribed dose needs to be defined in terms of type, intensity, frequency, and duration
Mechanism of action	What changes in structure or function caused by the exercise are responsible for the specified health benefit? In a therapy such as an exercise, there may be multiple mechanisms for a single health benefit.
Potential adverse events	What are the medical risks associated with the prescribed dose of exercise? What are the medical contraindications to the prescribed exercise, and what adjustments in dosing must be made for specific populations to reduce adverse events?

Table 2.1: Criteria required by the Food and Drug Administration of the USA to approve physical exercise as a prescribed medication [41, 42]

In consequence, several considerations must be kept in mind for PE implementation in rehabilitation. PE is characterized by type, intensity, and volume (session duration and frequency) according to the age, weight, fitness level, and pathologies of each patient [9]. Recent studies have shown that intensity is the most relevant feature in prescribing PE in rehabilitation [10, 43] because it determines the amount of energy expenditure. It can be seen as the "dose" of the prescription [11]. Controlling exercise intensity avoids over-training patients, affecting their rehabilitation and even suffer health consequences (i.e., physical or physiological complications) [12, 13, 44]. In this context, it is essential to clarify the meaning of the various terms associated with physical activity and exercise for consistent interpretation of exercise intensity in the context of dose-response issues.

2.2 Exercise therapy: characterizing the dose

An important question when prescribing exercise is the optimal therapeutic dose required to produce a specific health benefit is? Typically, when considering exercise dose concerning health outcomes, intensity is a critical factor in response to exercise to achieve these outcomes [17]. Intensity is the magnitude of the increase in energy expenditure necessary to perform the activity (aerobic or endurance exercise) or the force produced by the muscle contractions (resistance or strength exercise) [17, 45].

- *Aerobic exercise* (prevalence of oxidative metabolic pathway) involves large muscle groups in dynamic activities that substantially increase in heart rate and energy expenditure. Regular participation in aerobic exercise improves the function of the cardiovascular system and skeletal muscles, leading to central and peripheral adaptations that increase endurance performance.
- *Anaerobic exercise* (glycolytic and phosphagens pathways) involves very high-intensity exercise that uses glycogen and phosphocreatine stores for the most of the energy provided.
- *Resistance exercise* is an anaerobic training explicitly designed to increase muscular strength, power, and endurance by varying the resistance; the number of times the resistance is moved in a single set of exercises; the number of sets performed; and the rest interval between sets.

Considering the above, the intensity has been used to classify the physical exercise in three groups: low-intensity exercises that are composed of soft activities which demand low energy cost (50% of the maximal heart rate (HR_{max})), and are usually used for patients with extreme risk conditions [11]; moderate-intensity exercises that contemplate non-stopped activities with a long duration that require a low or moderate effort around 50% to 75% of the patient's HR_{max} [11]; and high-intensity exercises that are workouts that alternate hard-charging intervals with short duration (15s to 5 minutes) that increase the HR_{max} up to 85% to 100% with a recovery period of equal or longer duration than the work interval [11, 46]. Regarding prescribing exercise, moderate-intensity training is the most implemented in the rehabilitation process because it involves large muscle groups in dynamic activities that result in substantial increases in heart rate and energy expenditure [47,48]; allowing improvements in the cardiovascular system and skeletal muscle function [46,49,50]. Likewise, high-intensity exercises have become more attractive since providing more significant changes in specific metabolic pathways associated with aerobic metabolism, and improved functional capacities [51, 52]. The main challenge for the implementing these exercises is the difficulties in managing the intensity that makes their prescription a complex task [13, 46, 53].

The intensity then is a crucial factor in the responsiveness to exercise for achieving health outcomes. Not only does intensity play a significant role in producing favorable adaptations to exercise, but it also has a significant role in the various health risks produced by increases in exercise [17]. Therefore, to achieve a successful rehabilitation program, measuring parameters that can facilitate data analysis and interpretation of the intensity performed [54, 55]. Additionally, providing experts with relevant data about the patient allows them to guide a safe and optimized rehabilitation process. Several methods have been explored to supervise the exercise intensity during therapies by monitoring of the patient's fatigue state [2, 14–17].

Fatigue is generally defined as a subjective state of tiredness or exhaustion, and the reduction of capacity for regular activity [18]. Fatigue is a common concern among clinicians and individuals who participate in physical activities based on training or rehabilitation [18]. In the same way, fatigue is defined as the inability of the muscles to maintain the required level of strength during exercise. Fatigue also results in the deterioration of health in the long term, including work-related musculoskeletal disorders [19], chronic fatigue syndrome [20] and compromises immune function [21]. An essential first step in managing fatigue is the rapid and accurate detection of its occurrences. However, there is no scientifically accepted method to identify it because of the wide range of factors that can produce fatigue.

2.3 Fatigue detection techniques

Diverse fatigue detection techniques have been studied and used in rehabilitation that can be divided into two main categories: qualitative and quantitative. The first category, qualitative methods, are centered around the use of subjective scales of fatigue perception or fatigue surveys [56–58]. Whereas the second category, quantitative approaches, is based on the use of one or more sensor technologies to model changes in human performance.

2.3.1 Qualitative methods

As highlighted above, fatigue is a subjective experience, and it can be presented in different ways. Some questionnaires ask patients about their perceived tiredness level according to a pre-established ordinal numeric scale [59, 60], where the lower number represents a state of absence of fatigue. The higher number represents a state of extreme fatigue, i.e., the person does not feel able to continue with the activity [57]. Therefore, several perception scales have been developed and even modified according to their application [61]; divided into two main categories: one-dimensional and multidimensional scales [62].

One-dimensional fatigue scales are easier to use and contemplate one fatigue type, typically, patient's fatigue severity. In this regard, these scales are widely implemented during physical exercise rehabilitation [63], composed of different items to assess fatigue in a different time or social conditions. Several one-dimensional scales have been proposed considering the different rehabilitation scenarios as illustrated in Table 2.2 [62]. These questionnaires have an ordinal point scale to determine the level of fatigue according to the patient's answer. On the other hand, the multidimensional fatigue scale differs from the one-dimensional scale.

Mainly because these scales seek to analyze different fatigue factors and experiences such as duration, daily pattern, cognitive, behavioral, social, and the effect on daily activities, instead of only considering the intensity. The multidimensional scales are composed of more than one item and implemented to evaluate fatigue before or after the rehabilitation procedure [62,63]. The most common multidimensional scales are also illustrated in Table 2.2.

Perception scale	**Scale**	**Factors**	**Program**	**Item number**	**Point values**
One-dimensional	Fatigue Severity Scale [64]	Physical	Physical rehabilitation	9	from 1 to 7
	Borg CR10 [65]	Physical	Physical rehabilitation	10	from 1 to 10
	Fatigue Assessment Scale [66]	Physical	Oncology rehabilitation, Parkinson's disease, and post-stroke recovery	10	from 0 to 4
	Brief Fatigue Inventory [67]	Physical	Oncology rehabilitation	9	from 0 to 10
Multidimensional	Fatigue Scale [68]	Physical, Mental	Oncology rehabilitation, Multiple Sclerosis, Parkinson's disease, Neurologic rehabilitation, Stroke recovery	7	from 1 to 4
	Multidimensional Fatigue Inventory [68]	Physical, Mental, General, Reduce activity and motivation	Oncology rehabilitation	20	from 1 to 5
	Modified Fatigue Impact Scale [69]	Physical, Cognitive, Psychosocial	General physical rehabilitation	21	from 0 to 4

Table 2.2: One-dimensional and multidimensional perception scales to evaluate fatigue in rehabilitation procedure

In general terms, the most common subjective scale implemented considering these two categories is ten points Borg Rating of Perceived Exertion scale (Borg CR10), a 10-point scale composed of only 1 item; where the lower number represents a state of absence of fatigue, and the higher number represents a state of extreme fatigue [70]. The Borg scale is easy to use and is related to physiological parameters (e.g., heart rate or blood lactate) used for fatigue identification [57]. However, several studies have illustrated that perception scales present subjectivity since specific fatigue symptoms may vary depending on existing pathologies, environmental factors, and physical conditions. However, the understanding of how an individual's performance changes throughout rehabilitation is limited; the qualitative methods do not always represent the actual intensity has led to a decrease in reliability [57, 71]. Likewise, it is essential to highlight that these questionnaires are not suitable for real-time since they are not scalable and are potentially disruptive. For instance, consider a situation where there are 20 patients in the rehabilitation center, and their fatigue ratings are measured every five minutes. The administration of surveys in this situation would require a large number of surveyors and would disrupt the patient's rehabilitation [22].

2.3.2 Quantitative methods

To overcome the limitations presented in Section 2.3.1, and considering that continuous monitoring is essential for patients with chronic diseases during physical sessions, it is preferred to estimate the exercise intensity based on metrics obtained directly from the patients [11]. Therefore, quantitative approaches such as physiological parameters and exercise performance have been proposed [20, 72–75]. Regarding physiological parameters measurement, one of their applications is the indirect estimation of fatigue. The three parameters most related to fatigue are explained below [21, 76–78].

- **Oxygen uptake (VO_2)** represents the oxygen consumption that the body takes up and utilizes the exercising muscle [79]. This outcome is widely used in exercise physiology. It is considered as one of the best ways to quantify the patient's fatigue because it represents a linear relationship with the energy cost [80]. During exercise, the VO_2 increases exponentially until it reaches the point where oxygen supply matches oxygen demand, and then it stabilizes [79]. However, the VO_2 measurement requires complex instrumentation that makes it a technical problem, especially in rehabilitation therapies. Hence, it is generally used for research scenarios [81].

- **Heart rate** is one of the most used physiological parameters to control fatigue due to its measurement facility and the linear relationship with VO_2 [82]. This methodology generally consists of monitoring the heart rate reserve, which is the difference between the HR_{max} rate and the resting heart rate [11]. The HR_{max} can be estimated using a stress test or a clinical test that assesses the body's physiological behavior during different exercises [83]. Nevertheless, there are more practical ways that consider some subject's characteristics (e.g., the age and the gender) to get an HR_{max} approximation. One of the most implemented is Tanaka's formula, which uses the user's age (in years), as is shown in equation 5.1 [84].

$$HR_{max} = 206.9 - (0.7 * age) \tag{2.1}$$

- **Blood lactate** is one of the most often measured parameters during clinical exercise testing and, performance testing of athletes [85]. Clinicians need to understand the pathological response as well as the typical response to exertion. In response to progressive incremental exercise, lactate will increase exponentially [86]. An individual's endurance performance is well correlated with their blood lactate [85]; hence, lactate monitoring increases the confidence of healthcare personnel in assessing the patient's effort in physical therapies [86]. The blood lactate should be measured directly instead of estimated from other acid-base variables; therefore, it requires getting a blood sample, and a specialized instrument, which is not always easy to use during physical therapies [87]. Besides, it is required to monitor this parameter constantly for getting a good interpretation [87, 88].

Although the physiological parameters are considered accurate in measurement technique terms, they present difficulties monitoring in real-time due to their measurement process. In addition, they may present different behaviors depending on the exercise type performed (i.e., moderate or high-intensity exercise), which makes it difficult to relate them to the fatigue level. Regardless of the exercise type performed, (i) fatigue affects movement and gait characteristics as impaired motor control and postural instability [23,77]; and (ii) the exercise performance has a directly proportional relationship with fatigue [89]. Several methods for monitoring fatigue through exercise performance have been implemented using ambulatory sensors (e.g., electromyography and inertial sensors) or non-ambulatory sensors (e.g., motion analysis system) to identify when an event exists outside the typical pattern, which supports the rehabilitation process and the activities performance monitoring [90].

- **Electromyography (EMG)** is considered the gold standard to detect muscular fatigue considering that it directly asses the bio-electrical muscles function [91, 92]. This method consists measures the electrical activity generated during muscle contraction, and relaxation [93]. The methodology is to attach electrodes to the muscles to register electrical potentials [94]. These potentials are directly related to muscular strength, which allows estimating the effort and evaluating the performance of the exercise [95]. Nevertheless, these electrical signals are affected by the impedance of the skin and the electrodes' location. Therefore, an initial normalization process is required to avoid these problems leading to low adaptability [96, 97]. Consequently, the EMG processing is a complex task to execute in real-time since it requires power and frequency analysis [98] to identify fatigue progression. Therefore it inhibits their daily usage for real-time fatigue detection.

- **Inertial measurement units (IMUs)** are reliable and cheap sensors that are used to capture a person's acceleration and motion data in real scenarios without the use of external sources, or devices [99–101]. The IMU is essentially the combination of two components: accelerometers and gyroscopes; with them, this device can measure gravitational force, speed, and orientation. Likewise, with the combination of these components, it is possible to assess the activity performance through estimations of the kinematic and spatiotemporal parameters [102], and motion analysis [99] (e.g., gait features [103]). Although it is possible to identify the person's fatigue level using these sensors. Considering that the kinematic study in fatigue is still an early topic, the use of other physiological parameters like blood lactate [104], EMG [105] , or even perceived level of fatigue [106] are widely used to corroborate the results.

- **The motion analysis system** is widely used in fatigue estimation due to its high accuracy and robustness in the measurement of the kinematic parameters. The motion analysis system is based on infrared cameras to estimate the position of reflective markers to segment an object or an individual and measures variables such as position and orientation [51]. These systems can measure a many kinematic parameters, and according to the kinematic model, implemented estimate specific characteristics of each movement or exercise, which has led to its high applicability in these areas [107]. Motion analysis is widely used in fatigue estimation due to its high accuracy and robustness in measuring the kinematic parameters. Muscular fatigue affects movement and gait characteristics as impaired motor control, and postural instability [77]. The aim of using these devices is to quantify and measure these different parameters. Likewise, identify when an event exists outside the typical pattern to help in the rehabilitation process, and the performance of the activities [108]. However, motion capture systems often require unique setups, making them better suited for controlled environments; they also present certain limitations in terms of their sophisticated instrumentation and their high-cost [109].

Note that the use of these methodologies to evaluate the exercise performance depends heavily on the exercise type used [110]. For instance, if an exercise is performed on a treadmill, the commonly evaluated parameters are the cadence, width, length of step, and duration of each gait phase, among others [111]. In other words, the parameters and characteristics to be evaluated depend on each exercise and must be sure to avoid errors or uncertainties in the measurements.

As this chapter has already pointed out, humans' performance changes as a function of a person's individual characteristics (e.g., age, gender, fitness level, and injury history), time (which can be manifested through detrimental performance due to fatigue and improved performance due to learning effects) and degree of exercise difficulty. Therefore, to enhance fatigue estimation, artificial intelligence systems in optimizing and transforming human performance have been implemented as a further alternative to monitor and understand how an individual's performance deteriorates with fatigue accumulation [22, 23]. These models appear complementary to collected human performance data from the diverse detection techniques (i.e., qualitative or quantitative approaches) that classify fatigue levels. Thus, Chapter 3 describes the current state of the use of artificial intelligence systems to identify fatigue.

Chapter 3

Current State of Machine Learning models for Fatigue Estimation

3.1 Introduction

Managing fatigued patients is an important issue; it is a precursor to many detrimental short-term and long-term health outcomes. The short-term effects include discomfort, lowered strength, and a diminished motor control function [112]. In a clinical environment, those short-term effects lead to reduced performance, quality of session rehabilitation and increased incidence of injuries [113]. The problem of estimation of the actual physical load and fatigue appeared from older times [114], and nowadays, it has evolved to concrete challenges [115–117], especially in the context of human-machine and machine-human interactions [118]. In older times, the technical side of the problem was related to limitations of data monitoring, collecting, processing, and representing by the available tools. Nevertheless, the rapid development of information and communication technologies allows us to expand the scope of sensors and actuators, which have already become standard devices of ordinary devices [119].

Recent investigations have therefore shown that machine learning (ML) models appear as a complementary alternative for working with fatigue databases due to the facility that they give in the analysis of each feature; considering that ML models are a vital technique that has demonstrated the ability to translate large health data-sets into actionable knowledge. In general, the use of these models could improve patient safety [118, 119], improve quality of care [120], and reduce healthcare costs [121, 122]. Hence, these models enable the creation of tools that can classify fatigue levels based on the data extracted for each patient. ML models to classify fatigue have been demonstrated to be more efficient and give a wide range to explore the features within the data set.

Even though machine learning has a lot of demonstrated benefits, the successful utilization of machine learning requires a great effort from human experts, given that no algorithm can achieve good performance on all possible problems [123]. Even though healthcare researchers are familiar with clinical data, they still often lack the machine learning expertise necessary to apply these techniques to significant data sources. Healthcare researchers can and do work with expert data scientists [124], but the interactive process generally takes both parties a lot of time and effort. Given the above and considering that there is not a method established to measure the fatigue as mentioned in Chapter 2, it is challenging to devise and deploy machine learning solutions as the whole exercise begins with a lengthy data provisioning process, continues with finding the right collaborators, and involves a continuous back-and-forth between (ML) experts and domain experts. They lead to a slow assessment of the fatigue measures' reliability, validity, and usefulness for the physician and the researcher, given the few reviews to draw such information.

Considering the potential of automatic recognition, its challenges, and its contribution to the rehabilitation field, the present review presents recent studies that employed machine learning algorithms to identify fatigue. The review focuses on evaluating the pros and cons off each machine learning system to establish selection criteria for the most suitable solution based on the specific requirements. Specifically, the literature was reviewed to answer the main research questions: (i) Which is the most appropriate method to measure fatigue?; (ii) What are the most appropriate ML methods to recognize fatigue?; (iii) What are the methods that improve fatigue recognition?. To the best of the author's knowledge, no previous works on the state-of-the-art address this comparative analysis; this analysis highlights strategy capable of performing intelligent, accurate, rapid, and cost-effective clinical fatigue analysis.

3.2 Experimental Section

3.2.1 Search Strategy

The comprehensive electronic literature search was conducted in Web of Science Direct, IEEE, BMC, and PubMed on studies from 2000 onward. In this electronic search, the following keywords were included: [“fatigue” OR “exhaustion" OR "tiredness”] AND [“evaluation” OR “measure” OR "estimate"] AND [“exercise" OR "activity" OR "training” OR "rehabilitation"] AND [“feature selection methods”]. In addition, wildcard symbols, such as hyphens or inverted commas, were used to consider all possible variations of root words. To avoid missing critical studies, a cross-referencing was applied from each article found during the electronic search.

3.2.2 Data Extraction

The search results yielded titles of articles. Articles whose titles did not fit the research topic or did not respond to the study questions were excluded. The titles and abstracts of all included studies were checked for relevance concerning the research topic. Publications included in this systematic review were downloaded into Mendeley for screening. To make the review readable and focused on the author' intention, as claimed in the introduction section, a data extraction was conducted based on major themes: (i) fatigue measurement environment; (ii) fatigue measurement technique; and, (iii) performance of fatigue measurement. To this end, Table 3.1 extracted the data related to fatigue detection technique, description of the involved feature, applied machine learning algorithms, and main results.

All the information extracted from the selected studies served as the benchmark to broaden and discuss the concepts of the issues in question. A descriptive and comparative analysis was performed since the identified data were insufficient for a meta-analysis.

3.2.3 Inclusion Criteria

Articles obtained through these searches were evaluated using the title and abstract. The articles were included in this systematic review when they met the following criteria: (i) the study must be focused on physical fatigue measure; (ii) it must include at least an experiment, a pilot study or a trial with at least one group of participants constituted homogeneously; (iii) implemented machine learning approaches to identify fatigue or exhaustion; (iv) accomplished classification using normalized or non-normalized features extracted from biomechanical data, such as spatiotemporal parameters, kinematics, kinetics and physiological indexes; (v) applied feature selection methods only as a pre-processing technique for the classification stage; (vi) implemented at least one fatigue or intensity regulation; (vii) were written in English. Works that explored cross-validation methods and other strategies to improve machine learning performance were also included. In addition, we did not impose constraints regarding sample size (number of subjects, number of trials, or number of strides) or the dimension of the features dataset. Finally, we excluded conference proceedings when a journal article published by the same authors with the same contents was already included.

3.2.4 Search strategy field

The electronic search of the previously-mentioned database identified a total of 1094 published studies. Of these, 287 were included based on the first, and second inclusion criteria (See Figure 3.1) After reading the titles there were 147 potentially relevant articles were left, and after removing duplicates, 75 articles remained for abstract review. Regarding the application of inclusion/exclusion criteria 34 studies remained for full-text reading. Finally, 10 suitable adequate quality reviews were identified, and all of them are included and discussed in the following paragraphs.

3.2.5 Assessment of methodological quality

The methodological quality of the articles was assessed using the AMSTAR method (A MeaSurement Tool to Assess Systematic Reviews). This method was initially developed to evaluate the methodological quality of systematic reviews [125], but it is also used to evaluate the quality of individual clinical studies.

It comprises 11 concise criteria; each criterion is given a score of 1 if it is met or a score of 0 if it is not met, or if it is unclear or not applicable. The individual scores are then added to give a final score. An AMSTAR score of $8-11$ implies high quality, $4-7$ medium quality, and $0-3$ poor quality.

3.3 Results and Discussion

Table 3.1 synthesizes the physical activities in which fatigue is detected, fatigue detection technique, features dataset, classifier and recognition results of the ten collected studies that applied feature selection methodologies in fatigue recognition. These studies are the outcome

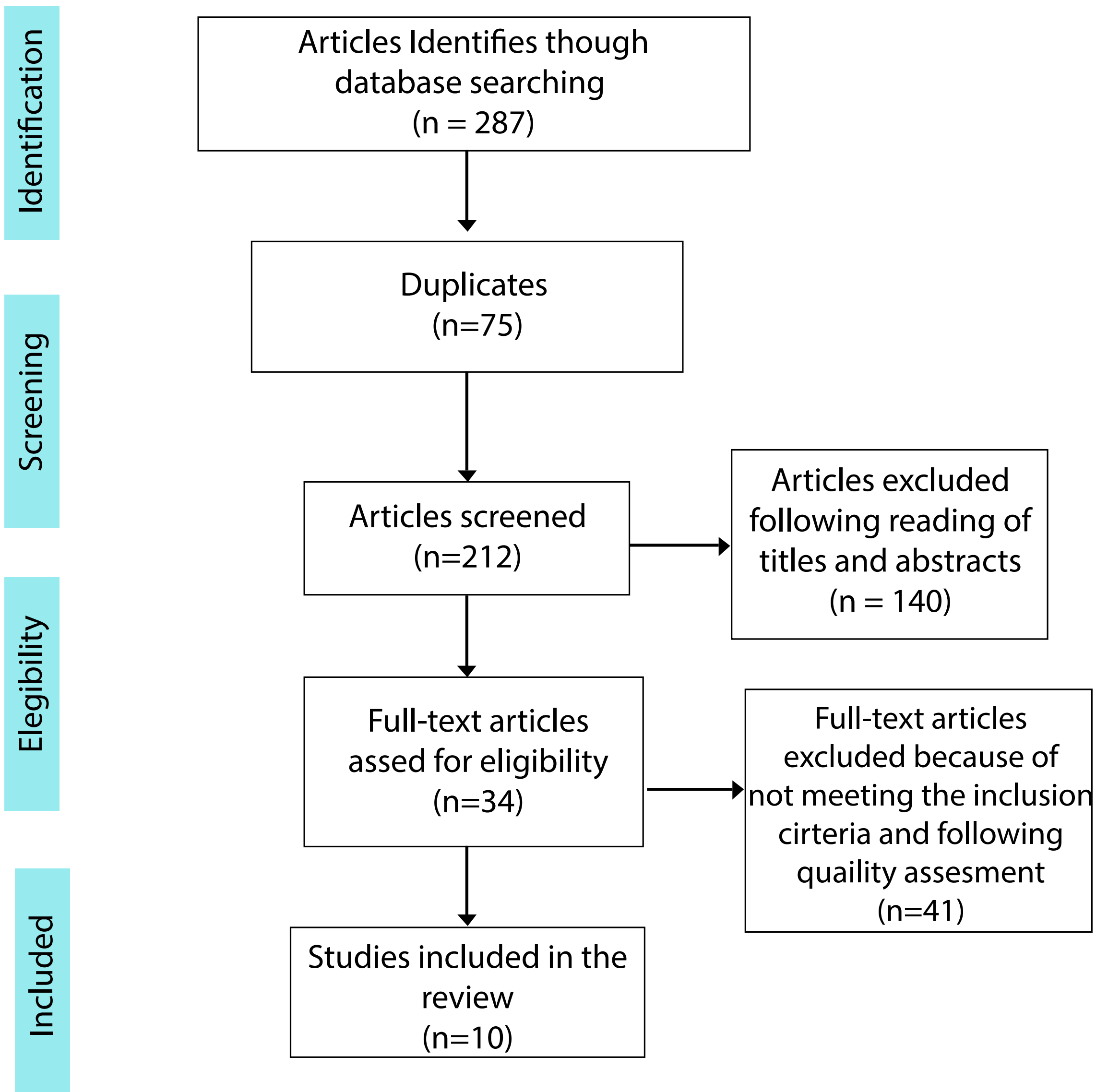

Figure 3.1: the PRISMA flow diagram for the systematic review detailing the database searching, the number of abstracts screened and the full texts retrieved

of the search strategy carried out in this literature review, which considered studies published since the year 2000 that exclusively involved feature selection methods as pre-processing in machine learning approaches for fatigue detection and analysis.

Table 3.1: A summary of fatigue modeling research

Begin of Table					
Research	**Physical activity**	**Fatigue detection technique**	**Features description**	**Classifiers**	**Results**
Aguirre et al. (2021) [23]	Sit-to-stand	3D optical tracking, HR, Borg Scale.	32 kinematic and temporal parameters for each stand-to-stand cycle	RF, ANN, SVM, LR, KNN.	The best classification (83.2% of accuracy) was achieved using all features in the RF classifier, meanwhile the worst classification (66.6% of accuracy) was performed using KNN classifier.
Mamman et al. (2020) [22]	Manufacturing task	IMUs, HR, Borg Scale.	53 statistical and biomechanical parameters	SVM, RF, LR, PLR.	Maximum accuracy (85.0%) was reached using 7 to 53 features with RF classifier. The worst accuracy (62.4 %) was achieved using LR classifier.

Continuation of Table 3.1					
Research	**Physical activity**	**Fatigue detection technique**	**Features description**	**Classifiers**	**Results**
Ulinskas et al. (2018) [126]	Office work	Keystroke dynamics template	Statistical characteristics of keystroke data. Linear discriminant analysis select the most relevant features.	SVM	SVM achieved an average daytime fatigue recognition accuracy of 98.11.%
Gordienko et al. (2017) [122]	Phsyical load	HR, BCI, accelerometer, muscle movements monitor.	12 Spatial, temporal and combined metrics	LRM, DNN	DNN correctly classified the fatigue state in training section.
Mamman et al. (2017) [127]	Manufacturing task	IMUs, HR, Borg Scale.	217 descriptive statistics and kinematic parameters.	PLR.	Maximum accuracy (79.0%) was accomplished using 16 features.

Continuation of Table 3.1					
Research	**Physical activity**	**Fatigue detection technique**	**Features description**	**Classifiers**	**Results**
Zhang et al. (2014) [128]	Walking	3D optical tracking, IMUS.	The general features include all possible spatial and temporal information from the raw signals.	SVM	Accuracy of 96% was reached in distinguishing the two gait patterns (fatigue and no-fatigue).
Karg et al. (2014) [129]	Squats	3D optical tracking, Subjective scale.	99 spatial and temporal parameters	HMM, LR.	HMM was more accurate than linear regression with 79% the accuracy.
Lee et al. (2009) [130]	Walking	3D optical tracking.	6 lower body kinematic parameters including ranges of motion and path lengths from the phase portraits were used.	LDA, Statistical test	LDA provides a 92.5% accuracy of classification

Continuation of Table 3.1

Research	Physical activity	Fatigue detection technique	Features description	Classifiers	Results
Karg et al. (2008) [131]	Walking	3D optical tracking.	19 PCA and FT were used to classify the gait of normal and exhausted walkers	LDA, SVM, kNN, NB	The NB classifier achieved a recognition rate of 68%.
Yoshino et al. (2004) [132]	Walking	Subjective scale, EMG, Accelerometers.	Seven gait and physiological parameters	LR	The correlation coefficient between the fatigue level predicted and the subjective fatigue level obtained was around 77%.

End of Table

SVM = support vector machines, RF = random forest, LR = logistic regression, DNN = deep learning neural network, LMR = linear regression model, PLR = penalized logistic regression, HMM = hidden markov models, LDA = linear discriminant analysis, kNN = k-nearest neighbors, NB = naive bayes.

Studies developed to classify fatigue using machine learning models differ in several aspects. They can change based on fatigue, the type of movement performed, and the sensors used to analyze fatigue. In the same way, the literature on physical fatigue detection in clinical environments can be classified into (i) exhaustion detection and (ii) partial fatigue detection.

In the first group, studies attempted to identify high/extreme fatigue that results in an inability to generate muscle forces and, consequently, an individual's performance decrease and inability to perform the physical activity [102, 128, 129, 131]. The other studies focused on detecting fatigue without reaching exhaustion, where individuals can still perform their physical activity at a diminished level [22,122,126,127,130,132]. Since exhaustion in the physical activity is often on localized fatigue, the associated literature [102,122,128–133] were characterized as focusing on one physical activity only (e.g., walking, sit-to-stand, physical load or squats), were primarily utilizing 3D motion capture system, EMG/EEG, accelerometers or IMUs, were used to model the individual's performance. Only two studies developed models to focus on a more complex task [22, 127], which utilized IMUs and heart rate monitors.

Overall, these systems implemented quantitative approaches that used one or more sensor technologies to model changes in human performance. The utilized sensor technologies include: (i) heart rate (HR) sensors to measure heart-rates, which are indicative of whole-body fatigue [22,102,122,127]; (ii) inertial measurement units (IMUs), which are cheap and reliable sensors that are used to capture a person's acceleration and motion data [22,122,127,132,134]; (iii) electroencephalography (EEG), used to measure brain activity, which is vital in detecting mental fatigue [122]; (iv) electromyography (EMG), used to assess muscle activity and localized fatigue [132]; and (v) 3D optical tracking, which is utilized for motion capture and human modeling. Note that some of these technologies are not suitable for daily field implementation. Specifically, EEG and EMG are invasive, which inhibits their daily usage for real-time fatigue detection. Moreover, motion capture systems often require unique setups, which make them better suited for controlled environments. Meanwhile, Heart rate or IMUs sensors are easy to adapt and use in rehabilitation environments, providing a practical and valuable tool for the health staff [100, 101]; note that the above answers to the first research question raised in this review.

On the other hand, six studies used a subjective fatigue scale to reference of the individuals' fatigue perception in contrast to their implemented fatigue detection method [22,102,126,127, 129,132]. Ulinskas et al. [122] stated that considering a subjective scale as a reference makes it challenging to control the subjective fatigue factors and thus, not always have entirely reliable results. In addition, Gordienko et al. [122] showed that using only subjective scales to classify fatigue does not have promising results, whereas using objective parameters or the combination of these two methods improves the accuracy and reliability of the classifiers significantly.

Finally, even if the fatigue models presented have been proposed to classify fatigue, none of them except for [102] has been able to classify fatigue in more than two groups. In other words, they only evaluated two fatigue conditions," non-fatigue" and" fatigue," which were generally obtained in two separate steps; in the first step, the participants have no fatigue, while in the second step, fatigue is induced to the participants.

These models presented a significant potential for clinical scenarios because they provide an objective indicator of the fatigue. However, they considered only two fatigue states, i.e., fatigued or non-fatigued state limits the accurate monitoring of the user's exhaustion during therapy, restricts the possibility to determine the adequate "dose" (i.e., intensity) of the individuals to produce a specific health benefit according to their individuals' needs, and thus limits improve the user's performance during therapy. On the other hand, from a detailed literature review, we could not answer the second reach question because we could not identify any article that discussed or established the best classifier for identifying and diagnosing fatigue. The above may be attributed to the lack of understanding of how an individual's performance deteriorates with fatigue accumulation, which can vary based on user conditions and physical activity.

Considering the above and as a first step, in this master thesis, we focus on occupational fatigue since it is: (i) a precedent to exhaustion and (ii) more aligned to the clinical environment in advanced physical rehabilitation. As the second step, this work proposes a framework for developing a fatigue identification model based on the individuals' exercise performance assessment to classify four fatigue diagnosis stages (low, medium, high, and very high) as presented in Chapter 4. The fatigue diagnosis stages allow clinicians to pinpoint the hazard directly. They can then prescribe interventions from a large number of options, such as assigning rest breaks (which can reduce the level of fatigue, and thus, avoiding reaching potentially dangerous levels of fatigue); or activity redesign (eliminating the development of fatigue) [22].

Four fatigue stages accurately monitor patients' fatigue conditions during exercise to avoid injuries or affect the rehabilitation process. To this end, our framework benefits from the advances and widespread use of wearable sensors for data collection. These sensors offer an individualized insight into the individuals' performance and present a unified performance benchmark that does not depend on process cycle time (essential for real-time scenarios). Furthermore, our proposed framework is evaluated monitoring fatigue in walking tasks since (i) localized muscle fatigue is a potential risk factor for injury or falls as muscle fatigue adversely affects proprioception, movement coordination, and muscle reaction times, leading to postural instability and gait alterations [106, 132–135]; therefore, gait patterns associated with fatigue may help in the assessment of fatigue-related fall risks or injuries in various environments; and (ii) moderate-intensity training as walking exercises are one of the most used in physical rehabilitation process due to improving cardiovascular system, and skeletal muscle function [46, 48–50].

3.4 Conclusions

Physical fatigue is a significant safety concern in rehabilitation environments, and monitoring physical fatigue is essential to prevent an accident, injury occurrences, and a successful program rehabilitation. The utilization of predictive models for physical fatigue modeling can better understand the physiology and psychology of fatigue. This literature review covers the state-of-the-art on machine learning approaches and their respective fatigue detection methodologies. The reviewed methods may vary based on fatigue detection techniques and the features used. Likewise, the ML implementation varies based on supervised and unsupervised learning; linearity and non-linearity models; the possibility of leading to over-fitting or not; division of classification in training and test phases or not; and the necessity of defining split criteria or not.

Fatigue stated detection is a powerful automatic tool that may provide an objective analysis and constant monitoring of the patient's rehabilitation, thus avoiding injuries and affectations to their process. Recent studies have evidenced that wearable sensory systems provide these data sets given their potential for long-term and free setting applications and time and cost-effectiveness.

From this literature analysis, we verify that proper and reliable fatigue recognition should involve several phases. The first phase is feature extraction to characterize the fatigue state. Second, methods of feature normalization may be applied to achieve a more robust classification. Then, feature selection methods are implemented to select the most significant features to distinguish the classes based on the dependence of classifier performance on the number and type of features. The next stage before the classification algorithm is to form the training and testing data sets through cross-validation procedures. Cross-validation methods also prevent over-fitting and generalize the classifier performance. The implementation of these three methodologies answers the third search question raised in this review since these are reliable tools that improve the performance of fatigue recognition.

In summary, automatic recognition of fatigue through machine learning algorithms is likely to offer an objective and prompt assessment of the subject's clinical status. Hence, it provides a potential patient's monitoring and improves their physical rehabilitation program. However, the classifiers models were implemented only to determine whether the user is fatigued or not fatigued, limiting the accurate monitoring of the user's exhaustion during therapy and restricting the possibility of prescribing interventions according to the patient's needs. In this context, the classification model presented in this work determines four states of fatigue (low, medium, high, very high) to improve monitoring and intervention prescription, allowing clinicians to assign adequate rests and training intervals for each patient to avoid injury.

Chapter 4

Proposed framework for physical fatigue management using wearable sensors[1]

Figure 4.1 presents an overview of the proposed framework for managing physical fatigue. The first phase is comprised of fourth main steps: (i) sensor selection, where practitioners should identify appropriate sensors for fatigue detection; (ii) data preprocessing and feature generation, where the sensors' data are prepared for analysis and generated the dataset; (iii) model construction and validation, where statistical and data analytic models are trained for distinguishing between the four fatigue states (non-fatigued, low-fatigue, moderate-fatigue, high-fatigue); (iv) data analysis, where the classifier models are evaluated based on accuracy, precision, recall and F-score. The outcome from phase 1 is the selection of an appropriate model for prospective analysis. The subset of features/predictors most frequently used in predicting the fatigue state is identified in the second phase. Likewise, in this phase, the classifier is evaluated by constraining the number of sensors used.

[1]This chapter is mostly based on the following journal article:
Pinto-Bernal, Maria J., Cifuentes, C.A.; Perdomo, O.; Rincón-Roncancio, M.; Múnera. (2021). "A Data-Driven Approach to Physical Fatigue Management Using Wearable Sensors to Classify Four Diagnostic Fatigue States" Sensors 21, no. 19: 6401.https://doi.org/10.3390/s21196401

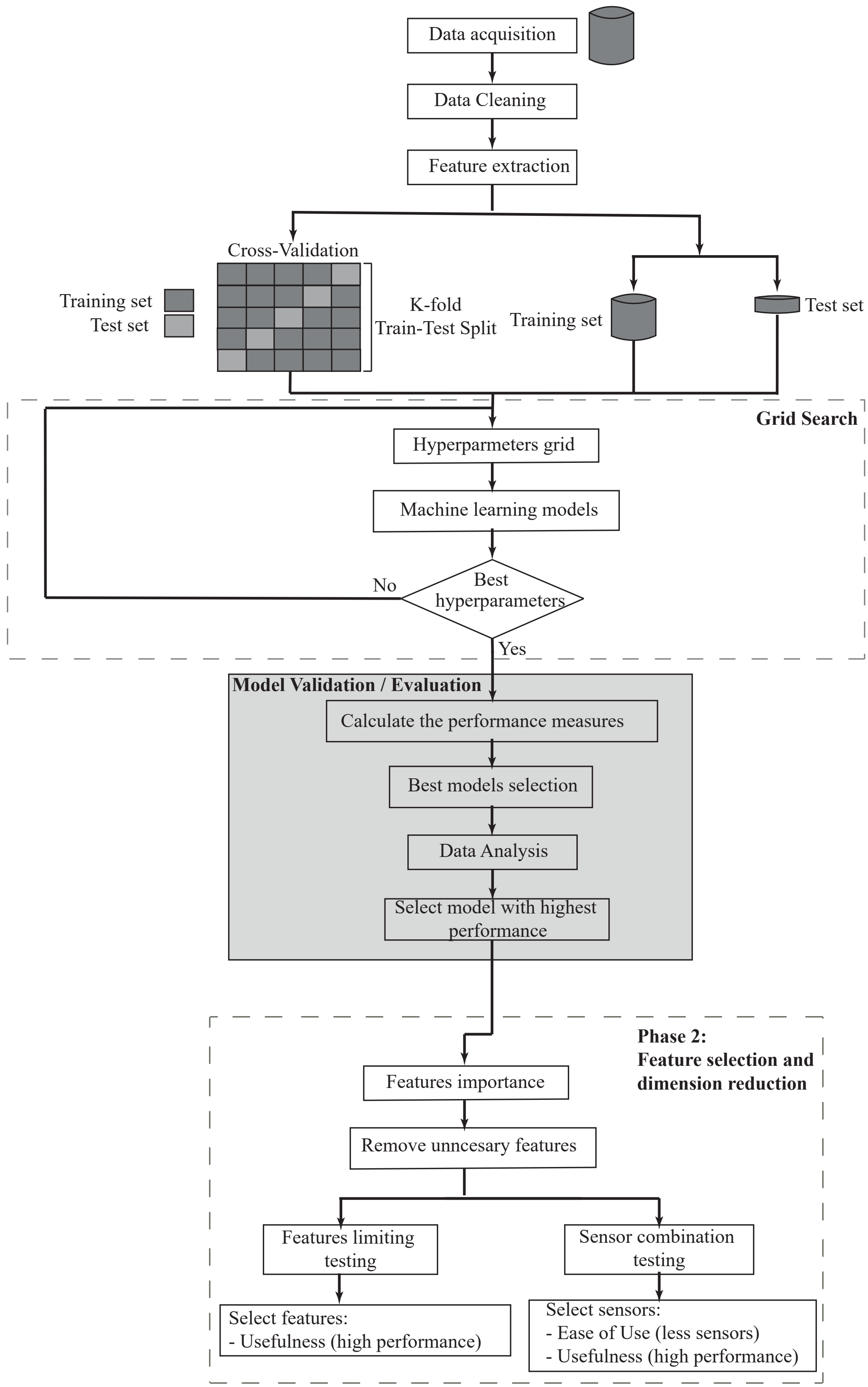

Figure 4.1: A flowchart that illustrates an overview of the proposed method [1].

4.1 Phase 1: Fatigue detection

4.1.1 Sensor selection

The first step is to identify the metrics and the characteristic used based on the selected physical exercise critical to the quality of the process/product. Note that the methodologies presented in Chapter 2 to evaluate the exercise performance depend heavily on the exercise type used [110]. In other words, the parameters and characteristics to be evaluated depend on each exercise and must be sure to avoid errors or uncertainties in the measurements. For instance, if an exercise is performed on a treadmill, the commonly evaluated parameters are the cadence, width, length of step, and duration of each gait phase, among others [111].

The above means identifying (i) the exercise, (ii) the clinical scenario, (iii) the parameters that change their behavior upon fatigue while exercising. An essential part of this is talking with clinicians or patients to understand the likelihood of fatigue and its associated symptoms. This helps isolate the type of fatigue being experienced. Along with interviews, literature review, historical data on injuries should be evaluated to determine the body parts that experience higher rates of injury. The result of this stage is an understanding of how an individual's performance changes throughout rehabilitation and which body parts are essential to be monitored during their physical rehabilitation.

Based on these metrics and the information gathered in the first step, the measurement means can be identified. It can include (i) the use of wearable or camera-based sensors, (ii) the use of physiological parameters, and (iii) the use of subjective scales of fatigue perception. The choice of approach would depend on the budget, intensity of exercise, and clinical environment. Considering the advantages and disadvantages of these systems presented in Chapter 2, Cavuoto and Megahed [136] discussed several fatigue indicators and reported that these indicators could be monitored using pervasive wearable sensors. Likewise, in a follow-up works, Maman et al. [22, 127] showed that four IMU sensors (located at the ankle, hip, torso, and wrist) coupled with a heart rate sensor could be used to detect fatigue in different tasks. We suggest using these wearable sensors for fatigue detection. More importantly, our framework presents a systematic approach to answer the question: "what are the gains associated with wearing an extra sensor?" In essence, this question attempts to quantify whether the hassle and cost associated with wearing an extra sensor can be justified with a significant/practical improvement in fatigue detection.

4.1.2 Data preprocessing

The first step in analyzing data is to ensure that the data is correct and cleaned. Therefore, the fourth primary cleaning step was proposed. First, the gyroscope and accelerometer outputs were treated with a second-order low-pass filter Butterworth for noise removal [137, 138]. Second, the collected data was visualized to check any additional erroneous data (e.g., faulty sensor values (too high and too low)), i.e., data that were not corrected through the automated filtering in the previous step. Third, the data from the different sensors were synchronized and any observations captured outside of the experimental window were eliminated. Fourth, the data was partitioned using a non-overlapping time window, where the selection of the length of the time window depended on the identification of each participant's gait event.

Considering that gait presents a repetitive behavior, an automated procedure was implemented to detect each participant's gait cycle. Specifically, the process consisted of determining each heel-strike event by detecting the second minimum value in the angular velocity signals of each stride time. This process outcome is represented in Figure 4.2, where a sample of each heel strike of the corresponding signal from a participant's test record can be observed. In addition, a zoom of a signal part is presented in Figure 4.2, where gait events detection throughout the angular velocity of an inertial sensing system over two gait cycles can be appreciated [139]. The selection of the correct threshold value was carried out as in [140], whose study validated all possible thresholds of the gait cycle within a range, and whose limits were visually established from signals acquired in a preliminary analysis.

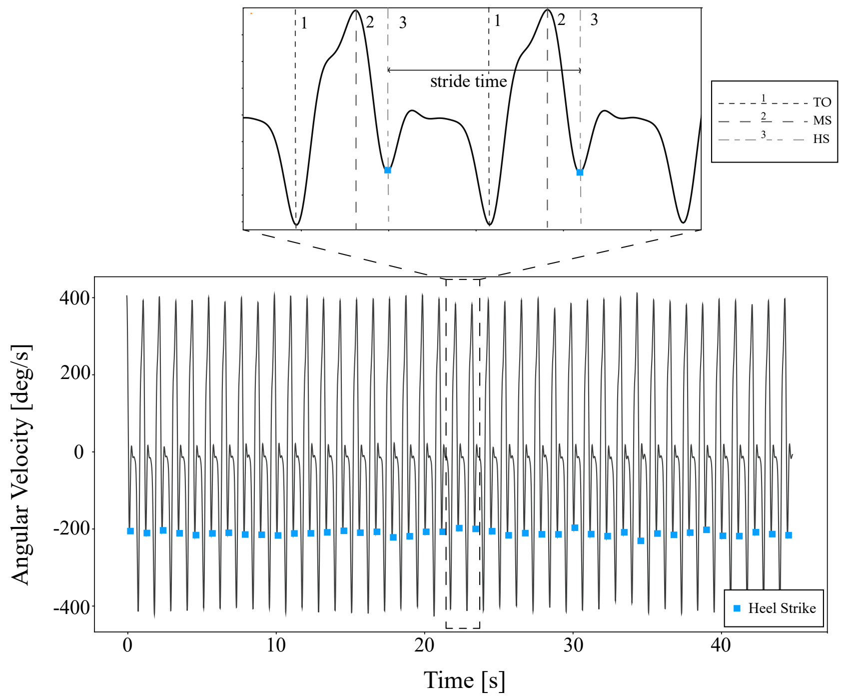

Figure 4.2: Heel strike detection using an inertial detection system on a participant's test record. The zoom part represents two gait cycles and the identification of each gait phase: TO = Toe-Off (first dashed line), MS = Mid-Swing (second dashed line) and HS = Heel Strike (third dashed line) [1].

Regarding feature generation two IMUs were attached at the participants' foot instep and around L5-S1 to measure the acceleration associated with a person's dynamic motion (spatiotemporal and kinematic parameters). These features capture the intensity and spread, commonly used in the fatigue detection literature [141–143]. Likewise, two more Shimmer3 were used and configured to obtain EMG signals in four muscles (tibialis anterior, rectus femoris, biceps femoris, and gastrocnemius) to register electrical potentials [94]. These potentials are directly related to muscular strength, which allows estimation the effort and evaluating the participants' exercise performance [95]. The description of the proposed features is provided in Table 4.1. Note that these features are calculated for each time window, i.e., for each gait cycle.

It is essential to highlight that feature variability caused by each participant's physical condition makes it difficult to directly compare the volunteer registers, which requires a normalization of the data according to each initial subject performance [150–152]. All features extracted were normalized by dividing them with the corresponding initial value (see equation 4.1). Note that the test number zero of each volunteer was taken as a reference since it was considered that the volunteer did not have fatigue, which was corroborated with the blood lactate, the Borg CR10 scale, and the multidimensional fatigue inventory questionnaire.

$$F_i = \left(\frac{F_i}{F_{reference}}\right) \tag{4.1}$$

where F_i corresponded to each feature extracted in each time window and $F_{reference}$ is the average of the first ten-time window values of the test number zero where volunteers were not fatigued.

4.1.3 Model construction and validation

In machine learning, there are several ways of data partitioning for experimentation. The most popular methods are typically referred to as training/test partitioning or cross-validation [153,154]; both were implemented in this research to evaluate the best classifier performance.

The training/test partitioning typically involves partitioning of the data into a training set and a test set in a specific ratio, e.g., 70% of the data are used as the training set and 30% of the data are used as the test set. This data partitioning can be done randomly or fixed. The fixed way is typically avoided (except when order matters) as it may introduce systematic differences between the training set and the test set, which leads to sampling representativeness issues. To avoid such systematic differences, the random assignment of instances into training and test sets is typically used [154, 155].

Cross-validation is conducted by partitioning a data set into n folds (or subsets), followed by an iterative combination of the folds into different training and test sets. In other words, each of the n folds is, in turn, used as the test set at one of the n iterations, while the rest of the folds are combined as the training set [156]. A typical approach to cross-validation is dividing the dataset into ten folds, where the models are selected based on the average/median prediction performance across ten non-overlapping test datasets. The literature suggests that 10-fold cross-validation may reduce the variation between the train and test performance [157]. However, cross-validation is generally more expensive in terms of computational cost than training/test partitioning.

Regarding the selection of the machine learning models, several classification methods are viable candidates for utilization in fatigue prediction. However, our framework's perspective, makes it, impossible to predetermine which methods will work best for fatigue prediction in walking tasks. Because these methods are data-driven and thus, are application-dependent, i.e., dependent on the exercise, extracted features, sensors, scenarios, etc. Several methods were applied during our preliminary analysis of the data to develop the fatigue prediction model. The models evaluated included: logistic regression (LR), decision trees (DT), k-nearest neighbors (KNN), support vector machine (SVM), naive Bayes (NB), linear discriminant (LDA), artificial neuronal network (ANN), and random forest (RF). The open-source python library "scikit-learn" [158] was used to execute a quick general training for these classifiers. Afterward, according to the accuracy metric, and due to their relatively poor performance LDA, and NB were eliminated. Hence, our case study focused on using the best six classification models (LR, KNN, SVM, RF, ANN, and DT), adjusted and retrained, by modifying their training hyperparameter automatically through computational iterations. For a detailed introduction on the classifier mentioned above, the following paragraphs provide a brief explanation of each classification model.

A statistical model such as LR attempts to build a relationship among the input variables and response employing parametric methods. In other words, it uses a logistic function to model conditional probability. Hence, LR is a supervised learning algorithm technique where the probability of a dichotomous outcome is a function of the predictors/features [127, 159]. Although LR is a simple yet effective classification algorithm, its performance can vary significantly with sparse data [127]. Moreover, non-parametric approaches such as KNN, SVM, and ANN, are commonly used in human performance modeling applications [160–163]. KNN is a simple classifier, an easy-to-implement supervised machine learning algorithm that can solve classification and regression problems. The algorithm assumes that similar things are near each other; therefore, it requires the computation of the distance of the unlabeled object to all the labeled objects in the training set [164]. Regarding the SVM classifier, which is a supervised learning method that uses kernel functions for data classification and regression analysis, its methodology consists of using a hyperplane to separate one-dimensional data to a high-dimensional space from a given labeled data set [165, 166] to identify the optimal hyperplane to classify the given data with minimum error [166].

The ANN classifier, is a supervised machine learning classifier that seeks to classify an observation as belonging to some discrete class based on inputs. This classifier is a set of connected input-output networks in which weight is associated with each connection. It consists of one input layer, one or more intermediate layers, and one output layer. Learning of neural networks is performed by adjusting the weight of the connection. By updating the weight iterative, the performance of the network is improved [167]. Finally, concerning DT model is a kind of a flowchart where each node represents a test, and the successive children represent the outcomes of the test. Each of the leaf nodes depicts the labels of the resulting class. The ensemble classification algorithm; utilizes trees as base classifiers in the RF model to generate many classifiers and aggregate their results via voting. It means that each tree in the random forest spits out a class prediction and the class with the most votes becomes our model's prediction. The premise of this method is that combining a large number of single classifiers allows for a more diverse representation of the data and consequently a more accurate prediction [22, 168, 169].

4.1.4 Data Analysis

Developments in machine learning classifiers from imbalanced data have been mainly motivated by numerous real-life applications since it faces the problem of the unequal representation of the data [170]. Most of the machine learning models used for classification have been designed around the assumption of an equal examples distribution for each class [171]. This means that an incorrect model application may focus on learning the characteristics of the abundant observations only, neglecting the examples from the minority class increasing false positives. However, whether is a slight imbalance, i.e., the distribution of examples is uneven by a small amount in the training dataset (e.g., 4:6) is often not a concern [172, 173]. In this context, to evaluate the performance of the proposed fatigue detection models it was essential to consider the following performance measures.

Recall or true positive rate (TPR) that captures the ability to detect the fatigued cases, i.e., it quantifies the number of positive class predictions made of all positive examples in the dataset; which is computed as follows

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{4.2}$$

where a true positive (TP) is considered if the classifier prediction and the reference value match; otherwise, such classification is regarded as a false positive (FP), i.e., the number of registers that belong to other groups and were wrongly estimated. Likewise, accuracy, presents the percentage of correct classifications made by a given model (Equation 4.3, where n represents the entire amount of data).

$$Accuracy = \frac{1}{n}(\frac{TP + TN}{TP + FN + FP + FN}) \tag{4.3}$$

where the non-fatigue state is similarly detected by classifier and reference signal correspond to true negative (TN); otherwise, they have been accounted for by false negatives (FN). Precision, which quantifies the number of positive class predictions that belong to the positive class, was measured as follows

$$Precision = \frac{TP}{TP + FP} \tag{4.4}$$

The F-score (Equation 4.5) was the last metric, which provides a single score that balances precision and recalls concerns in one number.

$$F1_{micro}score = \frac{Precision * Recall}{Precision + Recall} * 2 \tag{4.5}$$

4.2 Phase 2: Feature selection and dimension reduction

Once the best prediction model is identified, it is essential to consider that an critical aspect for technology adoption is usability. When the number of potential features/predictors is large, the computational complexity for model training increases. Feature reduction is typically applied to reduce the computational burden. In general, models are more interpretable if the number of features is smaller, which could (i) equalize or even increase the performance metrics of the classifier by removing unnecessary features from the data; and (b) an increased generalization capability [174]. In the context of our framework, usability can be measured using a total number of features selected; therefore we hypothesize that the chosen prediction model will have a relatively low number of features. Considering that the proposed framework enables the diagnosis of fatigue and the recommendation of an appropriate intervention. The feature selection/reduction was performed through univariate statistical approaches where the features selection was based on their relationship to the response and their prediction performance. From this step, practitioners can understand which features affect fatigue and how they are associated with changes in the potential classifier. Therefore, any unchanged features in the fatigued and non-fatigued states should be removed. That suggests the result would be a more interpretable fatigue classifier with relatively large prediction performance, i.e., an optimal fatigue detection classifier with a low false alarm rate.

Table 4.1: Features generated

N°	Feature	Description	Ref.
0	gait_mean_acce	Average gait acceleration	
1	gait_std_acce	Average gait acceleration std	
2	gait_max_acce	Average gait maximum acceleration	
3	gait_var_acce	Average gait acceleration variance	
4	gait_median_acce	Average median gait acceleration	
5	gait_energy_acce	Average gait acceleration energy	
6	gait_entropy_acce	Average gait acceleration entropy	
7	gait_kurtosis_acce	Average gait acceleration kurtosis	
8	gait_maxfreq_acce	Average gait acceleration maxfreq	
9	gait_stdfreq_acce	Average gait gyro stdfreq	
10	gait_mean_gyro	Average gait angular velocity mean	
11	gait_std_gyro	Average gait angular velocity std	
12	gait_max_gyro	Average gait maximum angular velocity	
13	gait_var_gyro	Average gait angular velocity variance	
14	gait_median_gyro	Average median gait angular velocity	
15	gait_energy_gyro	Average gait angular velocity energy	
16	gait_entropy_gyro	Average gait angular velocity entropy	
17	gait_curtosis_gyro	Average gait angular velocity kurtosis	
18	gait_maxfreq_gyro	Average gait angular velocity maxfreq	
19	l2_mean_acce	Average ts acceleration	[55], [142], [144], [145] [141], [146], [147], [148], [149], [133], [129].
20	l2_std_acce	Average ts acceleration std	
21	l2_max_acce	Average ts maximum acceleration	
22	l2_var_acce	Average ts acceleration variance	
23	l2_median_acce	Average ts acceleration velocity	
24	l2_energy_acce	Average median ts acceleration energy	
25	l2_entropy_acce	Average ts acceleration entropy	
26	l2_kurtosis_acce	Average ts acceleration kurtosis	
27	l2_maxfreq_acce	Average ts acceleration maxfreq	
28	l2_stdfreq_acce	Average ts acceleration maxfreq std	
29	l2_mean_gyro	Average ts angular velocity mean	
30	l2_std_gyro	Average ts angular velocity std	
31	l2_max_gyro	Average ts maximum angular velocity	
32	l2_var_gyro	Average ts angular velocity variance	
33	l2_median_gyro	Average median ts angular velocity	
34	l2_energy_gyro	Average ts angular velocity energy	
35	l2_entropy_gyro	Average ts angular velocity entropy	
36	l2_Kurtosis_gyro	Average ts angular velocity kurtosis	
37	l2_maxfreq_gyro	Average angular velocity maxfreq	
38	l2_stdfreq_gyro	Average ts angular velocity maxfreq std	
39	rms_gastro	RMS envelope of the gastrocnemius signal	
40	rms_tibilisAnterior	RMS envelope of the tibilis anterior signal	
41	rms_rectusFemoris	RMS envelope of the rectus femoris signal	
42	rms_bicepsFemoris	RMSenvelope of the biceps femoris signal	

Ref. = References; ts = torso swing; std = standard deviation; Maxfreq = maximun frequency; stdfreq = frequency standard deviation; RMS = root-mean-square.

Chapter 5

Experimental Trials and Validation [1]

The study has been divided into: a proposed framework for a fatigue classifier to develop and assess the fatigue estimation models as presented in Chapter 4, and an experimental setup to obtain the corresponding dataset. To this end, this chapter presents the experimental protocol implementation, as well as the validation of the six classification strategies to automatically identify four fatigue diagnosis levels (low-fatigue, moderate-fatigue, high-fatigue, and very high-fatigue) in walking exercise through several experiments with healthy patients; by monitoring 43 kinematic/temporal and biomechanical features.

5.1 Experimental setup

The experimental protocol was performed to quantify the detection success ratio of the fatigue classifiers. A total of 24 healthy subjects (14 males, 10 females, 21.75 $\pm$ 1.16 years old, 1.71 $\pm$ 0.09 m) performed the study. The subjects performed regular physical and

[1] This chapter is mostly based on the following journal article:
Pinto-Bernal, Maria J., Cifuentes, C.A.; Perdomo, O.; Rincón-Roncancio, M.; Múnera. (2021). "A Data-Driven Approach to Physical Fatigue Management Using Wearable Sensors to Classify Four Diagnostic Fatigue States" Sensors 21, no. 19: 6401.https://doi.org/10.3390/s21196401

had no known physical or cognitive disability, injuries, pain, or any impediment perming exercise (see Table 5.1 for further information). The participants enrolled in this study, must not present any fatigue state, i.e., they had to be in a non-fatigued state to avoid affecting their test performance. To this end, five different types of fatigue: general fatigue, physical fatigue, mental fatigue, reduced motivation, and reduced activity, were prior assessed using the "multidimensional fatigue inventory" [175] questionnaire. All subjects presented non-fatigued conditions, and were informed about the scope and purpose of the experiment. Written consent was obtained from each of them before the study. The Ethics Committee of the Colombian School of Engineering Julio Garavito, Bogota, Colombia approved the protocol (See Appendix A).

Table 5.1: Summary of participants' descriptive data (M ± SD). BMI, body mass index.

Gender	**Age [Years Old]**	**BMI [kg/m^2]**	**Walking Speed [m/s]**
Male	21.83 ± 1.40	22.84 ± 2.90	0.18 ± 0.37
Female	21.64 ± 0.74	22.25 ± 3.09	0.18 ± 0.35

Volunteers were first instructed to perform three $10m$ tests at a self-selected speed to determine their average overground speed, which was successively set on a treadmill (NIZA RX K153D-A-3, SportFitness, Bogota, Colombia). Participants were equipped with four Shimmer3 (Shimmer, Dublin, Ireland) IMUs units, one located on their foot dominant instep and one located around the center of L5-S1 with a sample rate of 128 Hz. The remaining two units were configured with a sampling rate of 512 Hz to obtain EMG signals in four muscles (tibialis anterior, rectus femoris, biceps femoris, and gastrocnemius) of the participant's dominant leg. One IMU was located on the outer lateral part of the thighs and one on the calves' outer side with elastic bands. The EMG signal was recorded from a pair of Ag–AgCl electrodes (interelectrode distance 3 cm) after cleansing the skin with alcohol. In addition, participants were equipped with Zephyr HxM BT (Medtronic, Ireland) on their chest with an elastic band, with a sample rate of 1 Hz. The selection of the Zephyr BT sensor is based on accuracy, reliability, cost, availability, and comfort [176, 177]. The experimental setup described is illustrated in Figure 5.1.

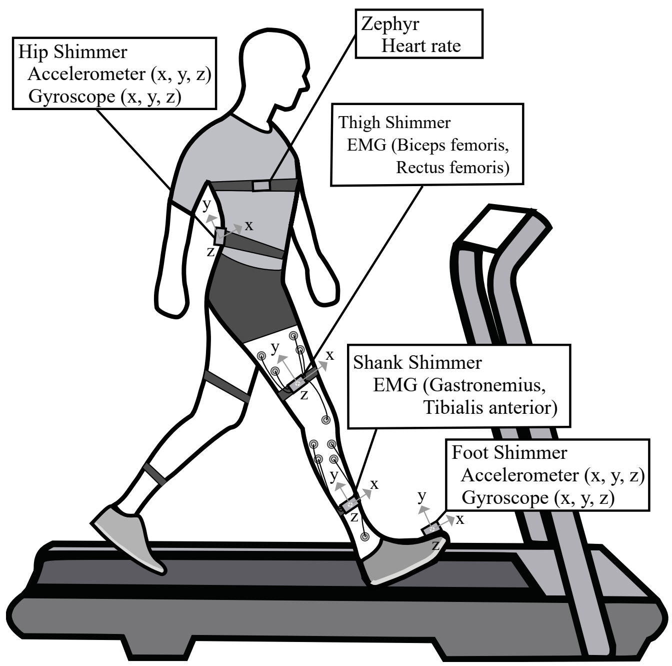

Figure 5.1: Experimental setup. Each participant was instrumented on their dominant side with an IMU placed on the dorsal side of their foot and with an IMU placed on their spine between L2 and S1. Two more IMUS are located on each thigh and shank with different electrodes in the tibialis anterior, rectus femoris, biceps femoris, and gastrocnemius. Participant's heart rate was captured using the Zephyr sensor located on their chest [1].

Participants were then asked to walk for at least 120s on the already configured treadmill at a zero-degree inclination, where different parameters were assessed. The first two parameters directly indicate the fatigue level: Blood Lactate concentration ($[La^{-}]$) and a perceptive fatigue scale using the Borg CR10. Therefore, they were used as reference values to diagnose and classify the participants' fatigue levels. At the same time, kinematic/spatiotemporal parameters, and EMG signals were recorded. Afterward, a participants' fatigue inducement was carried out. Volunteers had to perform as fast as possible a physical exercise circuit composed of four exercises: high knees, jumping jacks, squats, and short runs. At the end of the physical exercise circuit the participants returned to the treadmill, and the whole process was repeated. Note that the data acquisition, only started once the self-selected speed was reached, and the treadmill speed was only reduced after all data were acquired to prevent data capture during the transient state.

The above was repeated four times to increase the participants' fatigue level. The difference between each round was the execution time of the physical circuit that increased progressively. In other words, the time corresponding to the performance of each exercise increased by 15s each round; for the first time each exercise is performed for the 30s, the second round for 45s, the third round for the 60s, and the last round for 75s. If the participants' HR overcame 90% of the HR_{max}, or a 10 Borg value was notified, the test was immediately concluded. The entire experiment, including donning/doffing times related to instrumentation procedures and walking tasks, was completed within 60 min for all volunteers.

It is worth highlighting that to get an approximation of each volunteer's HR_{max}, the Tanaka's formula using the user's age (in years) was implemented as is shown in equation 5.1. Tanaka equation is recommended for healthy individuals such as those involved in this study because this equation significantly overpredicts maximal HR. Therefore, for people who present some diseases is recommendable to adapt this method using exercise testing [178].

$$HR_{max} = 206.9 - (0.7 * age) \tag{5.1}$$

Regarding the measured parameters, they were measured every time that the participant returned to the treadmill. The blood lactate sample was taken from the participant's earlobe with a new lancet. The blood was collected with a new test strip, and finally, the strip was inserted into the Lactate Pro2 (Arkray, Japan) to measure the blood lactate level. Likewise, the Borg CR10 scale was obtained by asking the volunteer how tired they felt according to the scale. Table 5.2 was used to explain the values meaning to each volunteer according to the four fatigue levels (low, moderate, high, very high). The participant was also instructed, only to focus on the total effort sensation and not on shortness of breath or muscle pain.

As described previously, blood lactate level and Borg CR10 scale were used as a reference

Table 5.2: Borg scale description and classification

Borg CR10 value	Description	Classification
0	No exertion at all	
1	Very easy	Low
2	Easy	
3	Somewhat moderate	
4	Moderate	Moderate
5	Somewhat hard	
6	Hard	
7	Very Hard	High
8	Very Very Hard	
9	Extremely hard	Very High
10	Maximum exertion	

and based on the hypothesis that the volunteers were taken from a rest or zero state to a maximum fatigue state. A linear correlation between these variables was carried out. To this end, the blood lactate level was normalized for each volunteer by dividing it with the corresponding initial value (see equation 4.1), considering that the $[La^{-}]$'s behavior increase exponentially during incremental exercise [179–181]. In contrast, Borg CR10 increment is linearly associated with an increase in exercise intensity [60, 181]. The relationship between these variables was approximated using a correlation. Specifically, a linear regression was conducted using least-squares methods as suggested in previous studies [182–186] (See Figure 5.2). From this correlation (p=0.9302), it was possible to corroborate the fatigue state of each participant dividing it into four states.

5.2 Results

Table 5.3 summarizes the samples numbers of the four fatigue states. The "Low" class presents most of the registers with 32.65%, followed by "Moderate" with and "High" classes with similar proportions (26.41% and 25.08%). "Very High" class has the lowest value, with a difference of 10% between "Moderate" and "High" classes; and a difference of 16.70%

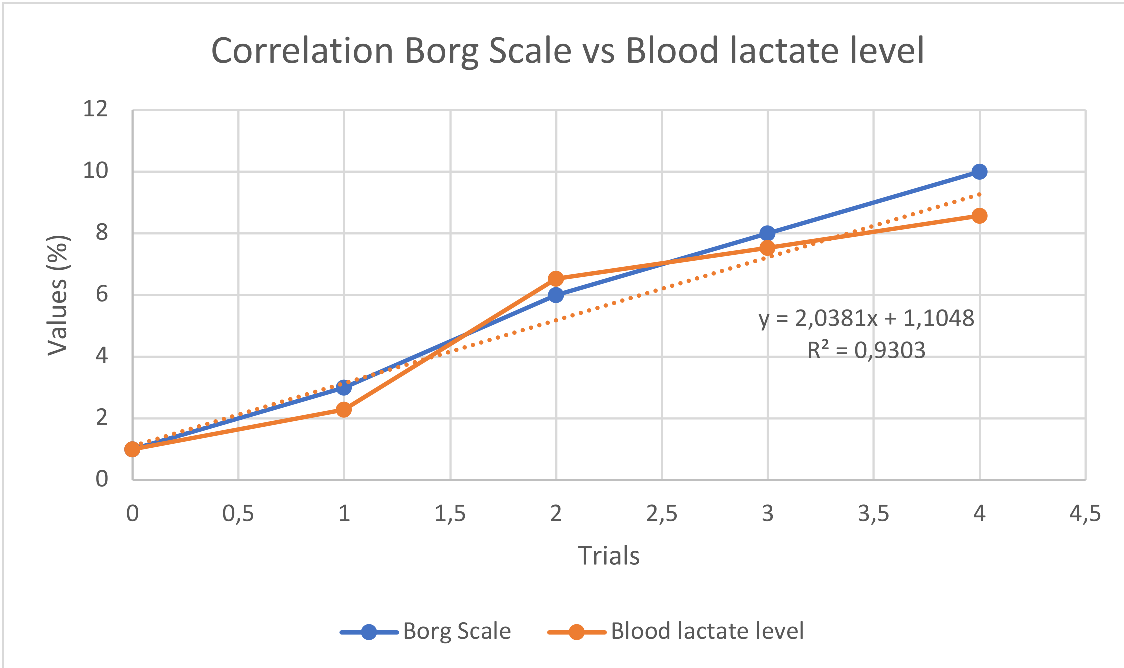

Figure 5.2: Borg scale and blood lactate values expressed as a function of the trials.

regarding "Low" class. The difference between the classes represents an imbalanced dataset.

Table 5.3: Data distribution in the dataset according to each fatigue state.

Class	**Number of samples**
Very High	463 (15.86%)
High	732 (25.08%)
Moderate	771 (26.41%)
Low	953 (32.65%)
Total	**2919 (100%)**

The predictive performance of the six models is summarized in Table 5.4. The table shows the mean for each of the four metrics. The reported results are based on 2919 dataset samples from the respective partitioning method: training/test (i.e., 80% of the data was used as the training set, and 20% of the data was used as a test set). A higher value is desired for the first three numeric columns, since it reflects a better prediction performance. Note that since the data set obtained is imbalanced, it could cause an increase of the false positives. F-score was considered as the most relevant feature for the performance classifier selection. Moreover, Table 5.4 reports the parameters and the performance of the six classifiers implemented. The logistic regression (LR) classifier implements the large-scale bound-constrained optimization as a penalty algorithm (solver = newton-cg), and a value of 1000000 for its inverse of regularization strength learning parameter (C = 1000000). Using Euclidean distance, the k-nearest neighbor (KNN) method classified the registers by a majority vote of its nearest elements with 3 neighbors (K = 3). Decision tree (DT) method using the function to ensure the quality of a split (criterion = entropy), tree depth to control the size of the tree to prevent overfitting (max depth = 12), can create arbitrarily small leaves (min samples split = 11) and guarantees that each leaf has a minimum size (min samples leaf = 4), avoiding low-variance, over-fit leaf nodes in regression problems. The support vector machine (SVM) with a radial basis function kernel (kernel = balanced) and a constrain value of 64 (C = 64). The artificial neuronal network with a stochastic gradient-based optimizer (solver=adam), 100, 100, and 100 as hidden layer sizes (HLS = (100, 100, 100)), activation function for the hidden layer (activation = tanh), learning rate schedule for weight updates (learning rate = constant), and regulation term parameter (alpha = 0.0001). Finally, the best model is a random forest classifier with 100 estimators (n estimators = 100), which means that the model integrates 100 decision tree models to merge their prediction. Note that all the hyperparameters used for the generation of the classifier are presented in Table 5.4, allowing their easy replication.

The next step is to analyze how the prediction performance varies while limiting the number of features or the number of sensors used in the study. The random forest method was selected to analyze the effect of removing features according to the highest performance metric reported in Table 5.4. The results of this approach are presented in Table 5.5, where (i) feature reduction using all sensors, and (ii) when features are limited to those from one and two sensor combinations.

Figure 5.3 presents the box plot of each reliability metric for the six machine learning implemented methods. Hence, each method contains four box plots, where the middle horizontal line represents the median value, the four quartiles are contained by the vertical lines, and

Table 5.4: Mean performance of the classification methods for fatigue detection in walking task. Bold values show the best score for each performance metric

Model	Hyperparameters	Accuracy	Precision	Recall	F1-Score
RF	**estimators = 100**	**0.965**	**0.931**	**0.929**	**0.928**
ANN	activation = tanh solver = adam HLS = (100,100,100) alpha = 0.0001 learning_rate = 'constant' max_iter = 1000	0.949	0.896	0.898	0.894
SVM	kernel = rbf class_weight ='balanced' C=64	0.907	0.809	0.809	0.806
DT	criterion = entropy max depth = 12 min samples split = 11 min samples leaf = 4	0.907	0.806	0.805	0.804
KNN	neighbors = 3	0.908	0.807	0.805	0.804
LR	solver = newton-cg C = 1000000	0.822	0.626	0.624	0.620

Table 5.5: Mean performance of the random forest model for fatigue detection in walking task using feature reduction and different sensors combinations. The best-performing model is in **bold**.

Sensors	Estimators	Features	Accuracy	Precision	Recall	F1-Score
Thigh (EMG), Shank (EMG), L5-S1,Foot	60	25	0,965	0,934	0,928	0,930
	40	**16**	**0,965**	**0,932**	**0,927**	**0,929**
	80	13	0,960	0,921	0,916	0,917
	80	11	0,963	0,926	0,925	0,925
	100	8	0,946	0,895	0,883	0,888
L5-S1, Foot	80	17	0,940	0,883	0,876	0,879
L5-S1	80	16	0,839	0,678	0,658	0,664
Foot	80	19	0,921	0,856	0,827	0,838

the boxes and the black dots are atypical data. It can be seen that the RF method always presents the highest values, showing the lowest dispersion and, therefore, the lowest variance.

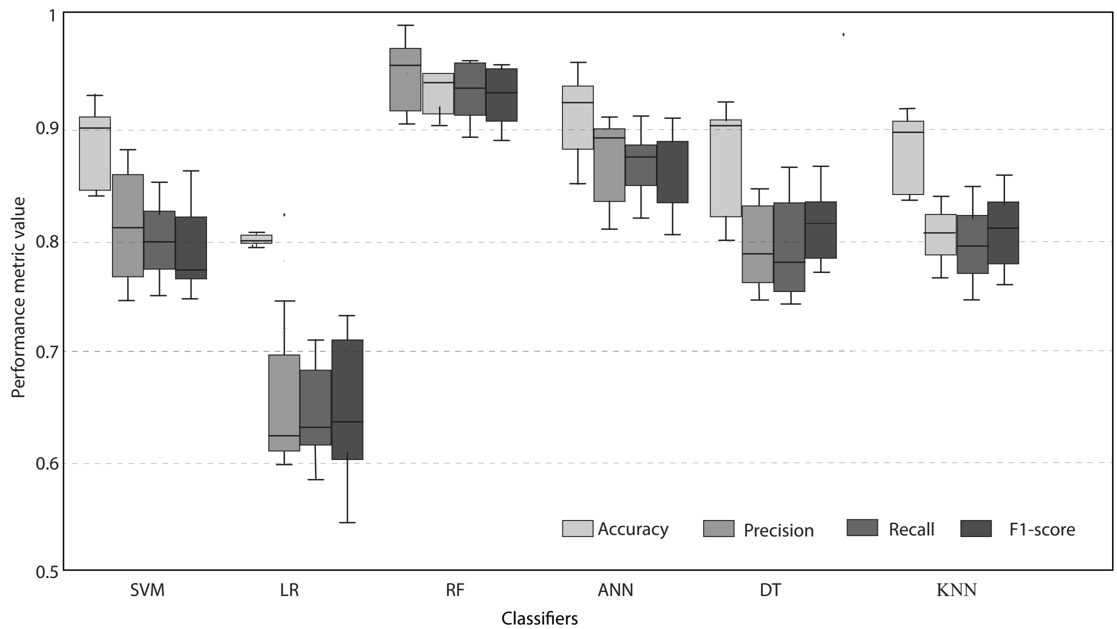

Figure 5.3: Box plot of the performance metric results for the six best machine learning methods. SVM = support vector machines, LR = logistic regression RF = random forest, ANN = artificial neuronal networl, DT = decisiton trees, KNN = k-nearest neighbors.

The confusion matrix obtained from the best three RF classifiers models with feature reduction implemented after exploring in a grid search manner is shown in Figure 5.4, where along the x-axis are listed the true class labels and along the y-axis are the class prediction. Along the first diagonal are the correct classifications, whereas all the other entries show misclassifications.

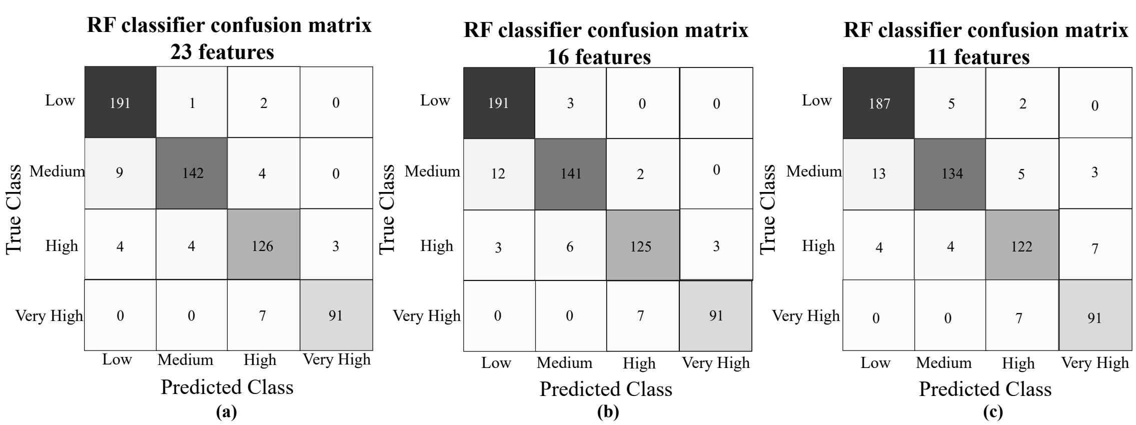

Figure 5.4: RF classifier confusion matrix for (**a**) 23 features, (**b**) 13 features (**c**) 11 features.

The features reduction used was essential to consider the feature importance properties for the initial RF model. The feature importance property measures a relative weight value to each feature, representing a direct relationship with the importance of the corresponding feature for this classical machine learning model. Figure 5.5 presents a cumulative graph representing the relative importance values obtained for each feature, sorted from the highest to the lowest values.

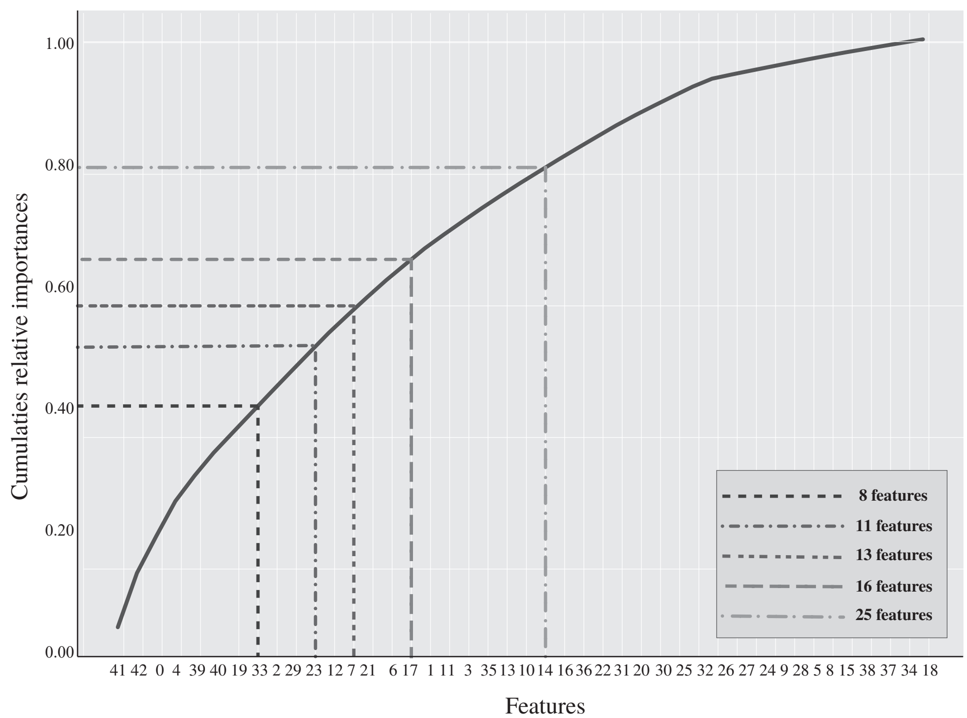

Figure 5.5: Features relative importance for random forest classifier using the original train data.

5.3 Discussion

The determination of the four fatigue states was achieved using qualitative methods such as the Borg scale and, as quantitative methods using blood lactate measurement. This last parameter was the most considered and used to measure the performance and fatigue of individuals. It is because, in response to progressively increasing exercise, lactate will increase exponentially. An individual's endurance performance is well correlated with their blood lactate [85]; therefore, lactate monitoring increases the confidence of healthcare personnel in assessing the patient's effort in physical therapies [86]. In other words, blood lactate is a direct indicator of fatigue. This, in turn, makes the fatigue classification model more accurate and reliable, since the data delivered by the sensors are directly related to what is happening physiologically with the user.

The heart rate was not considered a feature in the classification models considering that this variable is more related to controlling the intensity of the exercise rather than determining the individual's physical fatigue (i.e., it is not directly proportional to the individual's fatigue state) [82]. For instance, given a scenario where the individual is exercising at a constant speed or execution, the HR may remain stable and the individual may be experiencing a level of fatigue, which is not reflected in the measurement, and vice versa. The second reason for not considering HR was bearing in mind that it is determined by different variables such as age, the physical condition of the individual, comorbidities, and gender; therefore, when standardizing this variable it may cause a degree of uncertainty [187, 188].

According to the dataset distribution presented in Table 5.3, the main difference between classes was 16.79%, which is considered a slight imbalance, and it is acceptable for data analysis and training computational models [172, 173]. Considering the report by *Fernandez et al.* [171], that slight imbalance can often be treated like a common classification predictive modeling problem as long as true negatives are considered. The selection of the best model was performed mainly based on the F-score.

Fourth main observations from the predictive performance of our six models presented in Table 5.4 need to be highlighted. First, as expected from the preliminary analysis, the classifiers implemented with training/test presented a higher performance in all metrics than the classifiers implemented with cross-validation. Hence, a simple train-test split is sufficient for larger datasets. Second, the performance of the five models, except for the LR classifier, is relatively high with an overall average F-score greater than 80.4%. Third, according to the literature review, it was expected that the LR classifier presented a better performance given the positive results obtained with this classifier in previous studies [22, 127, 132]. However, this model presented the lowest performance with a 62% F-score, not representing an optimal or good classifier. This can be associated with two main factors: (i) these studies considered only two fatigue conditions: fatigue or non-fatigued, whereas this work contemplated four fatigue states. Therefore, our work increased the probability of failing in the estimation and suggested that this classifier performed better for binary classifications problems. (ii) To the lack of understanding of how an individual's performance deteriorates with fatigue accumulation, which can vary based on user conditions, physical activity, features extracted, and fatigue detection technique. Therefore, it is essential to have a general framework for fatigue estimation classifiers as presented in Chapter 4 to guide the implementation, evaluation, and continuous improvement of fatigue monitoring in rehabilitation scenarios regardless of physical activity or user conditions. Fourth, the RF model presented the best performance in all features with an overall of 92% considered as an optimum classifier for the fourth states estimation fatigue in walking tasks.

Once the best prediction model is identified, the next logical research question examines how the prediction performance varies while limiting the number of features and sensors used. To evaluate this question, we utilize the random forest model since Table 5.4 showed that it had the highest mean accuracy, precision, recall, and F-Score compared to the other classifiers. A list of predictive/important features was established from most important to least important (see Figure 5.5) and a feature reduction was applied to have a more interpretable classifier. Table 5.5 reports the prediction results when features are limited to 25, 16, 13, 11, and 8. From the results in Table 5.5, one can see that the prediction performance does not vary significantly as the number of features' is reduced. Some even match the RF performance utilizing all features (i.e., 43 features). Considering that the RF model presented a similar performance among 23, 16, and 11 features, their confusion matrix was analyzed to observe the prediction performance in two minority classes, i.e., very high and high fatigue states.

These predictions are more interesting and valuables since they are essential to avoid any risk in rehabilitation scenarios. The RF classifier presented fewer misclassifications between the low and minority classes with 16 features illustrated in Figure 5.4(**b**). Based on this observation, it is recommended to use the RF classifier with 16 features with an overall of 91.5% in all features. While the prediction performance is almost the same, the unnecessary features from the data were removed, optimizing computational costs and running time. Moreover, a smaller number of features facilitates the interpretation of the model, which is essential in fatigue identification and diagnosis phases.

Regarding the prediction performance variation, while limiting the sensor used, it is also reported in Table 5.4. Note that the sensors used to measure EMG are removed considering it that EMG is invasive, complicating its daily usage for real-time fatigue detection as highlighted in Chapter 2. Hence, the characteristics obtained by the EMG sensor were not considered. This means that the main features that detected fatigue were eliminated according to the sensor used. As expected without EMG measurements the classifier presented a performance reduction since these features presented great importance in fatigue detection (see Figure 5.5). However, the classifier performance remains positive using two sensors (IMUS placed on footstep and around L2-S1) with an F-score of 88% and an overall average of 94%. Similarly, using one IMU located in the footstep has a performance of 84% and an overall average of 92%. Based on this observation, it suggests using only the IMU located on foot. While the prediction performance is almost the same, the costs incurred by the clinics are much lower, and the usability of the system by using only one sensor is significantly improved. These results are comparable with the studies that proposed fatigue estimation models during walking tasks [130–132, 134], which have shown accuracy values among 77% and 92%; with only two fatigue conditions, fatigued and no-fatigued state. Besides, these studies [130, 131, 134] use motion capture systems as fatigue detection techniques that often require special setups, which make them better suited for controlled environments than real rehabilitation scenarios.

Considering these observations, one can indicate three main contributions. First, the framework proposed has shown higher detection performance (with fewer features), and detecting four fatigue diagnosis states in walking tasks; that allow clinicians to monitor the patient better, pinpoint the hazard; and prescribe and manage interventions according to each individual's needs. Second, the insights from the fatigue identification phase of our framework can be used to inform sensor placement and selection. Third, and more importantly, our framework presents a systematic approach that can answer the question: "what are the gains associated with wearing an extra sensor?" In essence, this question is left open to the researchers and practitioners to attempt to quantify whether the hassle and cost associated with wearing an extra sensor can be justified with a significant/practical improvement in fatigue detection when developing models for detecting/managing fatigue in other settings or their target application. The results and data used in this study can be accessed through the following link `https://figshare.com/projects/Fourth_fatigue_diagnosis_states/119580`. These codes can be used to develop predictive physical fatigue models.

The implementation of our proposed model requires an understanding of the specific features selected in the features reduction process. Significant features contributing to the determination of physical fatigue in this model include:

- **EMG RMS signals (features 41,42,39,40)** represent the square root of the average power of the EMG signal for a given period. Decrease overtime of these signals led to the detection of muscle fatigue.

- **Gait Acceleration Mean (feature 0)** reflects the mean duration of each gait cycle. The fatigued musculoskeletal system is less able to attenuate heel strike-initiated shock waves, which could be observed as an increase in the amplitude of the acceleration measured at the foot. If the mean gait cycle time increased significantly with elapsed walking time indicates that the individual is fatigued.

- **Gait Acceleration Median (feature 4)** The median value for each gait cycle acceleration.

- **Spine Acceleration Mean (features 19 and 29)** represents the torso acceleration over each gait cycle. These features show that if participants kept consistent torso movement over gait cycles, it likely corresponds to their walking behavior and is less likely to report physical fatigue.

- **Spine Acceleration Median (features 33 and 23)** measures the central tendency of the torso acceleration distribution. Whereas the participant maintains a high level of spine acceleration, then they are more likely to feel physically fatigued.

- **Gait Maximum Acceleration (features 2 and 12)** as the gait cycle time increases significantly with increasing fatigue, gait acceleration decreases. If the participant reduces their walking speed generated a decrease in peak gait acceleration is generated indicating that the participant is fatigued.

The selected features for the best model for physical fatigue detection were shown in Figure 5.5. They are consistent with previous studies that have used IMUs for monitoring physical activity. Common features computed from the acceleration signal are the mean [23, 141, 142, 144, 146] ,and variance or standard deviation [23, 142, 147, 189] and the entropy and energy of the data [23, 102].

Concerning the results using only one IMU located around the center of L5-S1, according to the literature review, it was expected to have a better classifier performance considering that the torso is the body part commonly used for physical fatigue detection and development models as reported in the following studies [22,127,130–132]. Specifically, *Mamman et al.* [22] reported that one IMU on the torso was enough to detect fatigue in manual material handling environments. However, our results differ from this results. Although that the mean back rotational position was selected as a mainly important feature for the classifier, as illustrated in Figure 5.5 with features numbers 19, 33, 29, and 23. The results showed that the accuracy and F-score were reduced to 88% and 66%, respectively using only the IMU on the torso; which is not promising. The above may suggest that (i) the features extracted from this sensor differ from other related studies; hence, the feature extraction methods influenced the classification performance, and should be improved. (ii) The use of a single sensor placed at L5-S1 in walking tasks may not be sufficient and consequently, the classifier's performance decreases significantly. (iii) The fourth fatigue diagnoses states may increase the probability of failing in fatigue estimation using this sensor considering only two fatigue diagnostic conditions presented in the previous studies. Therefore, in future work, it is suggested to revise the obtained features to improve the performance of the classification model.

Similarly, it was obtained that the mean back rotational position was selected as a mainly important feature as illustrated in Figure 5.5 with features numbers 19, 33, 29, and 23. Moreover, *Mamman et al.* [22] reported that IMU located the torso was enough for detecting fatigue in manual material handling environments. However, the features utilized differ from the features extracted in this work. The above may suggest three things: (i) the feature extraction methods influenced the classification performance; hence, features extraction with this sensor needs to be improved; (ii) the use of only one sensor placed on L5-S1 in walking tasks the classifier performance decrease significantly; and (iii) fourth fatigue diagnoses states may increase the probability of failing in fatigue estimation concerning only two fatigue diagnoses conditions. Therefore, as future work, it is suggested to review the characteristics obtained to improve the performance of the classification model.

On the other hand, there are a few limitations that must be acknowledged for this study. First, the sample size is small due to the confinement caused by the COVID19, bio-safety protocols, and the time required for each participant. However, the sample size is consistent with other studies focused on lab-based modeling of physical fatigue presented in Table 3.1.

Second, the effect of different demographic variables (i.e., age, sex, physical condition) needs to be explored in future models of physical fatigue due to all volunteers were healthy people; and the features may show different behaviors and patterns with patients, or other groups with various physical conditions. Third, the evaluation of our framework's performance was limited to focused lab experiments. Future studies should evaluate how this framework performs in clinical scenarios.

Chapter 6

Conclusions and Future works

As discussed throughout this thesis, physical fatigue is a significant safety concern in rehabilitation environments, and monitoring patients' intensity is essential to avoid extreme fatigue conditions, which may cause physical and physiological complications. Therefore, it is required to monitor the patient's condition during their physical rehabilitation to prevent accident and injury occurrences and contribute to the success of their rehabilitation program.

An essential first step in managing fatigue is the rapid and accurate detection of its occurrences. However, considering fatigue is a subjective experience and the wide range of factors that can produce fatigue, there is no scientifically accepted method to identify it. Several studies have been carried out to propose indirect and direct methods for quantifying fatigue, such as qualitative methods centered around using subjective fatigue scales and quantitative approaches to monitor users' physiological parameters and exercise performance.

Three main aspects are considered (i) the understanding of how an individual's performance changes throughout rehabilitation are limited; hence, the qualitative methods do not always represent the actual intensity has led to a decrease in reliability; (ii) physiological parameters present difficulties to monitor in real-time due to their acquisition processes; and (iii) the novelty and potential of the fatigue estimation method based on exercise performance. In this master's thesis, patient exercise performance monitoring was selected and implemented for fatigue estimation using computational models.

Machine learning models have been implemented for fatigue management but are limited in practice due to the lack of understanding of how an individual's performance deteriorates with fatigue accumulation; that can vary based on the physical exercise, environment, and individual's characteristics. Hence, as the main contribution of this master thesis, an integrated framework was proposed with the main steps for developing a fatigue classification model regardless of the type of exercise performed. This classification model contemplates using a few features to facilitate its interpretation, which is essential in identifying and diagnosing fatigue, primarily through visual analytical approaches that allow practitioners to identify risks that need to be addressed through an appropriate intervention strategy.

The experimental studies made possible to implement and evaluate six classifiers models that determined four fatigue stages instead of only designating if the user is fatigued or not fatigued, as previous studies have presented. The above allows to improve the fatigue monitoring and enable clinicians to prescribe an optimal therapeutic dose according to the individual's needs avoiding injuries and affectations to the process. The best classifier performance obtained greater accuracy, even though identifying four fatigue states increases the error probability. In particular, an accuracy of 96% and an F-score of 93% were obtained with a random forest model, generating an optimal classifier. Note that the determination of the four fatigue states was achieved using blood lactate measurement. This parameter was the most used to measure the performance and fatigue of individuals, thus being a direct indicator of fatigue. This makes the method more accurate and reliable since the data delivered by the sensors are directly related to what happens physiologically with the user.

The classifier performance model was also analyzed to check the number of features and sensors used in future real-time applications. The dimensionality of data is vital to optimizing computational costs and running time. In addition, with fewer features, the total number of sensors required to estimate them could be reduced. The results showed that the classifier could use a single sensor, maintaining 94% and 92% accuracy and F-score, respectively. Therefore, clinicians will also be more inclined to adopt the framework if it requires fewer sensors since it will: (i) cheaper; for example, requiring one or two IMU instead of four, would reduce the cost; (ii) make the process less invasive to the patient; (iii) reduce the time needed for the patient to wear and strap all the sensors.

As a future work and considering the main goal of the SORCAR project, a great application would be to implement this physical fatigue detection model in SAR systems. In essence, the system is proposed to be composed of one sensor interface, aiming to measure all features relevant to the RF classifier. The interface architecture comprises different nodes (e.g., sensor and SAR nodes) to adequately handle the difference in sampling rates and the amount of information acquired by the system. In the case of the sensor interface, the output of each node is the data processed, and for the SAR node is the corresponding behavior and feedback. The data from the sensors would be stored in a database to be analyzed online to manage the interaction of the SAR with the user.

The SAR's behaviors should be designed considering three situations [190, 191] (e.g., motivation, warning, and emergency, which are triggered depending on the data provided by the interface) while monitoring the users' performance. The aim is to positively influence the constant monitoring of users' performance and provide feedback to motivate them while the physical exercise intensity and the correct exercise execution are controlled. Besides, SARs' should be designed to communicate with the therapists if an emergency event occurs during therapy (e.g., physical fatigue above the maximum allowable level, dizziness or any parameter is out of normal). The aforementioned allows for a successful rehabilitation program through continuous patient monitoring and providing feedback to clinicians about the patient's performance to avoid overtraining, which can affect their rehabilitation and even suffer physical, or physiological complications. It should be emphasized that this proposal is also applicable to other robot-assisted physical exercise scenarios since this model allows a close measurement of users' physiological events.

Appendix A

Ethics committee approval

In this appendix the original version of the Ethics Committee of the Colombian School of Engineering's approval of this master's thesis is attached.

AVAL 04-2019
COMITÉ DE ETICA DE INVESTIGACIÓN

Asistentes	**Cargo**
Martha Pimienta Giraldo	Subdirectora de fomento y desarrollo a la investigación
Ricardo Martinez Rozo	Médico y profesor de Ingeniería Biomédica
Gladys Rocio González Leal	Profesora Centro de Estudios Ambientales
Paola Andrea Mora Osma	Abogada de la Oficina Jurídica

EL COMITÉ DE ÉTICA DE INVESTIGACIÓN de la Escuela Colombiana de Ingeniería Julio Garavito, certifica mediante la presente acta del 20 de septiembre de 2019 que se revisó la propuesta **"Estimación de fatiga en ejercicio aeróbico"** cuyo investigador principal es la estudiante de maestria Maria Jose Pinto de la Escuela Colombiana de Ingeniería y cuya tutora es Marcela Múnera Ramírez profesora de la Escuela Cololmbiana de Ingeniería.

Se revisaron los siguientes documentos:

- Protocolo
- Consentimiento informado en donde se encuentra registrado: las estrategías para dar a conocer a los participantes la investigación, riesgos y beneficios, como se garantizará la privacidad y el anonimato de los mismos y confidencialidad de los datos de investigación, la cadena de custodia de la información obtenida y las restricciones para su uso por terceros
- Hoja de vida del investigador principal y coinvestigadores

Adicionalmente se revisaron los siguientes aspectos:

- Utilidad del protocolo para los participantes, la sociedad o el conocimiento
- Evaluación riesgos y beneficios
- Procedimientos, metodologías y procesos de investigación, el manejo divulgación y archivo de los datos obtenidos.

Adicionalmente se revisó que la investigación no vulnerará la dignidad de los sujetos, no constituye una amenaza bajo ninguna circunstancia, ni causa daño emocional ni moral a los investigados y se ajusta a estándares científicos y éticos propios

Concepto

EL COMITÉ DE ÉTICA DE INVESTIGACIÓN de la Escuela Colombiana de Ingeniería, **aprueba el proyecto ""Estimación de fatiga en ejercicio aeróbico"**. De conformidad con la legislación vigente, este proyecto se clasifica como:

INVESTIGACIÓN SIN RIESGO PARA USUARIOS Y PACIENTES.

Para constancia de lo anterior se firma en la ciudad de Bogotá D.C., el 23 de septiembre de dos mil diez y nueve (2019)

Martha Cecilia Pimienta Giraldo
Miembro del Comité de Etica de la Investigación

Ricardo Martínez Rozo
Miembro del Comité de Etica de la Investigación

Gladys Rocio González Leal
Miembro del Comité de Etica de la Investigación

Paola Andea Mora Osma
Miembro del Comité de Etica de la Investigación

Appendix B

Quality Assessment of the Included Reviews

In this appendix the original version of the evaluation table for the systematic review used in this master's thesis is attached.

Author and year of study		
Was the study question or objective clearly stated		
Was the study population clearly and fully described		
Were the subject comparable		
Was the procedure clearly described		
Was the data extraction fully described		
Were the outcomes measures clearly defined, valid, reliable and implemented consistently across all study participants		
Was the measure of follow-up adequate		
Were the statistical methods well described?		
Were the results were described		
Quality rating (Good, Fair, Poor)		
Score		

Table B.1: Assessment of studies using NIH Quality Assessment

Glossary

ECIJG In Spanish *Escuela Colombiana de Ingeniería Julio Garavito*.

EEG Electroencephalography.

EMG Electromyography.

HR Heart rate.

HR$_{max}$ Heart rate maximum.

HRi Human-Robot interface.

IMU Inertial measurement unit.

IMUs Inertial measurement units.

MinCiencias In Spanish *Ministerio de Ciencia, Tecnología e Innovación*.

ML Machine learning.

PA Physical Activity.

PE Physical Exercise.

SORCAR Human-Robot Interaction Strategies for Rehabilitation based on Socially Assistive Robotics.

USA United States of America.

Bibliography

[1] M. J. Pinto-Bernal, C. A. Cifuentes, O. Perdomo, M. Rincón-Roncancio, and M. Múnera, "A data-driven approach to physical fatigue management using wearable sensors to classify four diagnostic fatigue states," *Sensors*, vol. 21, no. 19, p. 6401, 2021.

[2] M. R. Salakari, T. Surakka, R. Nurminen, and L. Pylkkänen, "Effects of rehabilitation among patients with advances cancer: a systematic review," *Acta Oncologica*, vol. 54, no. 5, pp. 618–628, 2015.

[3] S. Zanuso, A. Jimenez, G. Pugliese, G. Corigliano, and S. Balducci, "Exercise for the management of type 2 diabetes: a review of the evidence," *Acta diabetologica*, vol. 47, no. 1, pp. 15–22, 2010.

[4] A. E. Bauman, "Updating the evidence that physical activity is good for health: an epidemiological review 2000–2003," *Journal of science and medicine in sport*, vol. 7, no. 1, pp. 6–19, 2004.

[5] D. E. Warburton, C. W. Nicol, and S. S. Bredin, "Health benefits of physical activity: the evidence," *Cmaj*, vol. 174, no. 6, pp. 801–809, 2006.

[6] W. L. Haskell, I.-M. Lee, R. R. Pate, K. E. Powell, S. N. Blair, B. A. Franklin, C. A. Macera, G. W. Heath, P. D. Thompson, and A. Bauman, "Physical Activity and Public Health," *Medicine & Science in Sports & Exercise*, vol. 39, pp. 1423–1434, aug 2007.

[7] N. Priest, R. Armstrong, J. Doyle, and E. Waters, "Interventions implemented through sporting organisations for increasing participation in sport," *Cochrane database of systematic reviews*, no. 3, 2008.

[8] M. Livingstone, P. Robson, J. Wallace, and M. McKinley, "How active are we? levels of routine physical activity in children and adults," *Proceedings of the Nutrition Society*, vol. 62, no. 3, pp. 681–701, 2003.

[9] M. L. Pollock, G. A. Gaesser, J. D. Butcher, J. P. Després, R. K. Dishman, B. A. Franklin, and C. E. Garber, "The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults," 1998.

[10] P. Schnohr, J. L. Marott, J. S. Jensen, and G. B. Jensen, "Intensity versus duration of cycling, impact on all-cause and coronary heart disease mortality: The Copenhagen City Heart Study," *European Journal of Preventive Cardiology*, vol. 19, pp. 73–80, feb 2012.

[11] D. E. Warburton, "Prescribing exercise as preventive therapy," *Canadian Medical Association Journal*, vol. 174, pp. 961–974, mar 2006.

[12] E. H. Cup, A. J. Pieterse, J. M. ten Broek-Pastoor, M. Munneke, B. G. van Engelen, H. T. Hendricks, G. J. van der Wilt, and R. A. Oostendorp, "Exercise Therapy and Other Types of Physical Therapy for Patients With Neuromuscular Diseases: A Systematic Review," *Archives of Physical Medicine and Rehabilitation*, vol. 88, pp. 1452–1464, nov 2007.

[13] I. M. Lee, H. D. Sesso, Y. Oguma, and R. S. Paffenbarger, "Relative intensity of physical activity and risk of coronary heart disease," *Circulation*, vol. 107, pp. 1110–1116, mar 2003.

[14] L. Oldervoll, S. Kaasa, M. Hjermstad, J. Lund, and J. Loge, "Physical exercise results in the improved subjective well-being of a few or is effective rehabilitation for all cancer patients?," *European Journal of Cancer*, vol. 40, no. 7, pp. 951–962, 2004.

[15] L. Fleig, S. Lippke, S. Pomp, and R. Schwarzer, "Exercise maintenance after rehabilitation: How experience can make a difference," *Psychology of Sport and Exercise*, vol. 12, no. 3, pp. 293–299, 2011.

[16] W. Göhner, H. Seelig, and R. Fuchs, "Intervention effects on cognitive antecedents of physical exercise: A 1-year follow-up study," *Applied Psychology: Health and Well-Being*, vol. 1, no. 2, pp. 233–256, 2009.

[17] S. Balducci, M. Sacchetti, J. Haxhi, G. Orlando, V. D'Errico, S. Fallucca, S. Menini, and G. Pugliese, "Physical exercise as therapy for type 2 diabetes mellitus," *Diabetes/metabolism research and reviews*, vol. 30, no. S1, pp. 13–23, 2014.

[18] H. M. Abd-Elfattah, F. H. Abdelazeim, and S. Elshennawy, "Physical and cognitive consequences of fatigue: A review," *Journal of Advanced Research*, vol. 6, pp. 351–358, may 2015.

[19] L. Baussard, M. Carayol, B. Porro, F. Baguet, and F. Cousson-gelie, "European Journal of Oncology Nursing Fatigue in cancer patients : Development and validation of a short form of the Multidimensional Fatigue Inventory (MFI-10)," *European Journal of Oncology Nursing*, vol. 36, no. July, pp. 62–67, 2018.

[20] A. F. Alghannam, K. Tsintzas, D. Thompson, J. Bilzon, and J. A. Betts, "Exploring mechanisms of fatigue during repeated exercise and the dose dependent effects of carbohydrate and protein ingestion: study protocol for a randomised controlled trial," *Trials*, vol. 15, no. 1, p. 95, 2014.

[21] O. Ozalp, D. Inal-Ince, E. Calik, N. Vardar-Yagli, M. Saglam, S. Savci, H. Arikan, M. Bosnak-Guclu, and L. Coplu, "Extrapulmonary features of bronchiectasis: muscle function, exercise capacity, fatigue, and health status," *Multidisciplinary respiratory medicine*, vol. 7, no. 1, p. 3, 2012.

[22] Z. S. Maman, Y.-J. Chen, A. Baghdadi, S. Lombardo, L. A. Cavuoto, and F. M. Megahed, "A data analytic framework for physical fatigue management using wearable sensors," *Expert Systems with Applications*, vol. 155, p. 113405, 2020.

[23] A. Aguirre, M. J. Pinto, C. A. Cifuentes, O. Perdomo, C. A. Díaz, and M. Múnera, "Machine learning approach for fatigue estimation in sit-to-stand exercise," *Sensors*, vol. 21, no. 15, p. 5006, 2021.

[24] C. E. Bouchard, R. J. Shephard, and T. E. Stephens, "Physical activity, fitness, and health: international proceedings and consensus statement.," *Human Kinetics Publishers*, 1994.

[25] P.-O. Åstrand and G. Grimby, *Physical activity in health and disease.* Coronet Books, 1986.

[26] t. World Health Organization *et al.*, *Global recommendations on physical activity for health.* World Health Organization, 2010.

[27] T. M. Burkhalter and C. H. Hillman, "A narrative review of physical activity, nutrition, and obesity to cognition and scholastic performance across the human lifespan," *Advances in Nutrition*, vol. 2, no. 2, pp. 201S–206S, 2011.

[28] R. F. DeBusk, C. B. Taylor, and W. S. Agras, "Comparison of treadmill exercise testing and psychologic stress testing soon after myocardial infarction," *The American journal of cardiology*, vol. 43, no. 5, pp. 907–912, 1979.

[29] W. Nieuwland, M. A. Berkhuysen, D. J. van Veldhuisen, J. Brügemann, M. L. Landsman, E. van Sonderen, K. Lie, H. J. Crijns, and P. Rispens, "Differential effects of high-frequency versus low-frequency exercise training in rehabilitation of patients with coronary artery disease," *Journal of the American College of Cardiology*, vol. 36, no. 1, pp. 202–207, 2000.

[30] R. Belardinelli, F. Lacalaprice, G. Cianci, R. Piva, and A. Purcaro, "Exercise training intervention after coronary angioplasty (the etica trial): Ten-year follow-up," *Am Heart Assoc*, 2007.

[31] M.-A. Dalzell, N. Smirnow, W. Sateren, A. Sintharaphone, M. Ibrahim, L. Mastroianni, L. Zambrano, and S. O'Brien, "Rehabilitation and exercise oncology program: translating research into a model of care," *Current Oncology*, vol. 24, no. 3, pp. 191–198, 2017.

[32] R. R. Spence, K. C. Heesch, and W. J. Brown, "Exercise and cancer rehabilitation: a systematic review," *Cancer treatment reviews*, vol. 36, no. 2, pp. 185–194, 2010.

[33] L. Ignarro, M. Balestrieri, and C. Napoli, "Nutrition, physical activity, and cardiovascular disease: An update," *Cardiovascular Research*, vol. 73, pp. 326–340, jan 2007.

[34] G. O. Dibben, H. M. Dalal, R. S. Taylor, P. Doherty, L. H. Tang, and M. Hillsdon, "Cardiac rehabilitation and physical activity: systematic review and meta-analysis," *Heart*, vol. 104, no. 17, pp. 1394–1402, 2018.

[35] M. A. Spruit, F. Pitta, E. McAuley, R. L. ZuWallack, and L. Nici, "Pulmonary rehabilitation and physical activity in patients with chronic obstructive pulmonary disease," *American journal of respiratory and critical care medicine*, vol. 192, no. 8, pp. 924–933, 2015.

[36] Y.-s. Lee and S.-w. Ahn, "The effects of kinesio taping and neuromuscular rehabilitation exercise for patients with acute whiplash-associated disorder," *The Journal of Korean Academy of Orthopedic Manual Physical Therapy*, vol. 22, no. 2, pp. 41–49, 2016.

[37] Y. Escalante, J. M. Saavedra, A. Garcia-Hermoso, A. J. Silva, and T. M. Barbosa, "Physical exercise and reduction of pain in adults with lower limb osteoarthritis: a systematic review," *Journal of back and musculoskeletal rehabilitation*, vol. 23, no. 4, pp. 175–186, 2010.

[38] W. R. Frontera and H. G. Knuttgen, "Exercise and musculoskeletal rehabilitation: restoring optimal form and function," *The physician and sportsmedicine*, vol. 31, no. 12, pp. 39–45, 2003.

[39] Y. Oguma and T. Shinoda-Tagawa, "Physical activity decreases cardiovascular disease risk in women: review and meta-analysis," *American journal of preventive medicine*, vol. 26, no. 5, pp. 407–418, 2004.

[40] I. Vuori, "Physical inactivity is a cause and physical activity is a remedy for major public health problems," *Kinesiology*, vol. 36, no. 2, pp. 123–153, 2004.

[41] A. C. of Sports Medicine *et al.*, "The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults," *Med Sci Sports Exerc*, vol. 30, pp. 975–911, 1998.

[42] S. R. Colberg, R. J. Sigal, B. Fernhall, J. G. Regensteiner, B. J. Blissmer, R. R. Rubin, L. Chasan-Taber, A. L. Albright, and B. Braun, "Exercise and type 2 diabetes: the american college of sports medicine and the american diabetes association: joint position statement," *Diabetes care*, vol. 33, no. 12, pp. e147–e167, 2010.

[43] L. B. Andersen, P. Schnohr, M. Schroll, and H. O. Hein, "All-cause mortality associated with physical activity during leisure time, work, sports, and cycling to work," Tech. Rep. 11, 2000.

[44] A. F. Manley, *Physical activity and health: a report of the surgeon general.* Diane Publishing, 1996.

[45] A. C. of Sports Medicine, *ACSM's Health-Related Physical Fitness Assessment Manual*. Lippincott Williams & Wilkins, 2013.

[46] Y. Dun, J. R. Smith, S. Liu, and T. P. Olson, "High-intensity interval training in cardiac rehabilitation," *Clinics in geriatric medicine*, vol. 35, no. 4, pp. 469–487, 2019.

[47] I.-M. Lee, H. D. Sesso, Y. Oguma, and R. S. Paffenbarger Jr, "Relative intensity of physical activity and risk of coronary heart disease," *Circulation*, vol. 107, no. 8, pp. 1110–1116, 2003.

[48] M. Tanasescu, M. F. Leitzmann, E. B. Rimm, W. C. Willett, M. J. Stampfer, and F. B. Hu, "Exercise type and intensity in relation to coronary heart disease in men," *Jama*, vol. 288, no. 16, pp. 1994–2000, 2002.

[49] A. C. of Sports Medicine *et al.*, *ACSM's health-related physical fitness assessment manual*. Lippincott Williams & Wilkins, 2013.

[50] D. E. Warburton, N. Gledhill, and A. Quinney, "Musculoskeletal fitness and health," *Canadian journal of applied physiology*, vol. 26, no. 2, pp. 217–237, 2001.

[51] L. M. Ross, R. R. Porter, and J. L. Durstine, "High-intensity interval training (hiit) for patients with chronic diseases," *Journal of sport and health science*, vol. 5, no. 2, pp. 139–144, 2016.

[52] G. J. Balady, M. A. Williams, P. A. Ades, V. Bittner, P. Comoss, J. A. M. Foody, B. Franklin, B. Sanderson, and D. Southard, "Core components of cardiac rehabilitation/secondary prevention programs: 2007 update - A sci. statement from the Am. Heart Assoc. exercise, cardiac rehabilitation, and prevention comm., the council on clinical cardiology; the councils on cardiovascular nu," *Circulation*, vol. 115, pp. 2675–2682, may 2007.

[53] M. Tanasescu, M. F. Leitzmann, E. B. Rimm, W. C. Willett, M. J. Stampfer, and F. B. Hu, "Exercise type and intensity in relation to coronary heart disease in men," *Journal of the American Medical Association*, vol. 288, pp. 1994–2000, oct 2002.

[54] S. L. Halson, "Monitoring training load to understand fatigue in athletes," *Sports medicine*, vol. 44, no. 2, pp. 139–147, 2014.

[55] M. J. Pinto-Bernal, A. Aguirre, C. A. Cifuentes, and M. Munera, "Wearable sensors for monitoring exercise and fatigue estimation in rehabilitation," in *Internet of Medical Things*, pp. 83–110, CRC Press, 2021.

[56] 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 us manufacturing workers," *Applied ergonomics*, vol. 65, pp. 139–151, 2017.

[57] A. R. Zamunér, M. A. Moreno, T. M. Camargo, J. P. Graetz, A. C. Rebelo, N. Y. Tamburús, and E. da Silva, "Assessment of subjective perceived exertion at the anaerobic threshold with the Borg CR-10 scale," *Journal of Sports Science and Medicine*, vol. 10, pp. 130–136, mar 2011.

[58] G. A. Curt, W. Breitbart, D. Cella, J. E. Groopman, S. J. Horning, L. M. Itri, D. H. Johnson, C. Miaskowski, S. L. Scherr, R. K. Portenoy, *et al.*, "Impact of cancer-related fatigue on the lives of patients: new findings from the fatigue coalition," *The oncologist*, vol. 5, no. 5, pp. 353–360, 2000.

[59] J. Annett, "Subjective rating scales: Science or art?," *Ergonomics*, vol. 45, pp. 966–987, nov 2002.

[60] N. Williams, "The borg rating of perceived exertion (rpe) scale," *Occupational Medicine*, vol. 67, no. 5, pp. 404–405, 2017.

[61] S. Pettersson, I. E. Lundberg, M. H. Liang, J. Pouchot, and E. Welin Henriksson, "Determination of the minimal clinically important difference for seven measures of fatigue in swedish patients with systemic lupus erythematosus," *Scandinavian journal of rheumatology*, vol. 44, no. 3, pp. 206–210, 2015.

[62] L. Whitehead, "The measurement of fatigue in chronic illness: a systematic review of unidimensional and multidimensional fatigue measures," *Journal of pain and symptom management*, vol. 37, no. 1, pp. 107–128, 2009.

[63] H. J. Michielsen, J. De Vries, G. L. Van Heck, F. J. Van de Vijver, and K. Sijtsma, "Examination of the dimensionality of fatigue," *European Journal of Psychological Assessment*, vol. 20, no. 1, pp. 39–48, 2004.

[64] L. B. Krupp, N. G. LaRocca, J. Muir-Nash, and A. D. Steinberg, "The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus," *Archives of neurology*, vol. 46, no. 10, pp. 1121–1123, 1989.

[65] N. Williams, "The Borg Rating of Perceived Exertion (RPE) scale," in *Occupational Medicine*, vol. 67, pp. 404–405, jul 2017.

[66] H. J. Michielsen, J. De Vries, and G. L. Van Heck, "Psychometric qualities of a brief self-rated fatigue measure: The fatigue assessment scale," *Journal of psychosomatic research*, vol. 54, no. 4, pp. 345–352, 2003.

[67] T. R. Mendoza, X. S. Wang, C. S. Cleeland, M. Morrissey, B. A. Johnson, J. K. Wendt, and S. L. Huber, "The rapid assessment of fatigue severity in cancer patients: use of the brief fatigue inventory," *Cancer*, vol. 85, no. 5, pp. 1186–1196, 1999.

[68] T. Chalder, G. Berelowitz, T. Pawlikowska, L. Watts, S. Wessely, D. Wright, and E. Wallace, "Development of a fatigue scale," *Journal of psychosomatic research*, vol. 37, no. 2, pp. 147–153, 1993.

[69] J. D. Fisk, A. Pontefract, P. G. Ritvo, C. J. Archibald, and T. Murray, "The impact of fatigue on patients with multiple sclerosis," *Canadian Journal of Neurological Sciences*, vol. 21, no. 1, pp. 9–14, 1994.

[70] G. Borg, "Borg's range model and scales.," *International Journal of Sport Psychology*, 2001.

[71] A. Sehle, M. Vieten, S. Sailer, A. Mündermann, and C. Dettmers, "Objective assessment of motor fatigue in multiple sclerosis: the Fatigue index Kliniken Schmieder (FKS)," *Journal of Neurology*, vol. 261, pp. 1752–1762, sep 2014.

[72] J. Qi, P. Yang, A. Waraich, Z. Deng, Y. Zhao, and Y. Yang, "Examining sensor-based physical activity recognition and monitoring for healthcare using internet of things: A systematic review," *Journal of biomedical informatics*, vol. 87, pp. 138–153, 2018.

[73] A. I. Zeni, M. D. Hoffman, and P. S. Clifford, "Relationships among heart rate, lactate concentration, and perceived effort for different types of rhythmic exercise in women," *Archives of Physical Medicine and Rehabilitation*, vol. 77, pp. 237–241, mar 1996.

[74] D. C. Poole, M. Burnley, A. Vanhatalo, H. B. Rossiter, and A. M. Jones, "Critical power: An important fatigue threshold in exercise physiology," *Medicine and Science in Sports and Exercise*, vol. 48, pp. 2320–2334, nov 2016.

[75] S. Pettersson, I. Lundberg, M. Liang, J. Pouchot, and E. Welin Henriksson, "Determination of the minimal clinically important difference for seven measures of fatigue in Swedish patients with systemic lupus erythematosus," *Scandinavian Journal of Rheumatology*, vol. 44, pp. 206–210, may 2015.

[76] F. Yu, A. Bilberg, E. Stenager, C. Rabotti, B. Zhang, and M. Mischi, "A wireless body measurement system to study fatigue in multiple sclerosis," *Physiological Measurement*, vol. 33, pp. 2033–2048, dec 2012.

[77] F. Möhler, S. Ringhof, D. Debertin, and T. Stein, "Influence of fatigue on running coordination: A ucm analysis with a geometric 2d model and a subject-specific anthropometric 3d model," *Human movement science*, vol. 66, pp. 133–141, 2019.

[78] S. R. Kang, J.-Y. Min, C. Yu, and T.-K. Kwon, "Effect of whole body vibration on lactate level recovery and heart rate recovery in rest after intense exercise," *Technology and Health Care*, vol. 25, pp. 115–123, jul 2017.

[79] A. J. Glynn and H. Fiddler, *The Physiotherapist's Pocket Guide to Exercise E-Book: Assessment, Prescription and Training*. Elsevier Health Sciences, 2009.

[80] T. Reybrouck, L. Mertens, S. Brusselle, M. Weymans, B. Eyskens, J. Defoor, and M. Gewillig, "Oxygen uptake versus exercise intensity: A new concept in assessing cardiovascular exercise function in patients with congenital heart disease," *Heart*, vol. 84, pp. 46–52, jul 2000.

[81] M. Jetté, K. Sidney, and G. Blümchen, "Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity," *Clinical Cardiology*, vol. 13, no. 8, pp. 555–565, 1990.

[82] F. A. da Cunha, P. d. T. V. Farinatti, and A. W. Midgley, "Methodological and practical application issues in exercise prescription using the heart rate reserve and oxygen uptake reserve methods," *Journal of Science and Medicine in Sport*, vol. 14, no. 1, pp. 46–57, 2011.

[83] J. B. Kostis, A. E. Moreyra, M. T. Amendo, J. Di Pietro, N. Cosgrove, and P. T. Kuo, "The effect of age on heart rate in subjects free of heart disease. Studies by ambulatory electrocardiography and maximal exercise stress test," *Circulation*, vol. 65, no. 1 I, pp. 141–145, 1982.

[84] L. E. Cruz-Martínez, J. T. Rojas-Valencia, J. F. Correa-Mesa, and J. C. Correa-Morales, "Maximum heart rate during exercise: Reliability of the 220-age and Tanaka formulas in healthy young people at a moderate elevation," *Revista Facultad de Medicina*, vol. 62, no. 4, pp. 579–585, 2014.

[85] M. L. Goodwin, J. E. Harris, A. Hernández, and L. B. Gladden, "Blood lactate measurements and analysis during exercise: a guide for clinicians," *Journal of diabetes science and technology*, vol. 1, no. 4, pp. 558–569, 2007.

[86] T. C. Jansen, J. van Bommel, and J. Bakker, "Blood lactate monitoring in critically ill patients: a systematic health technology assessment," *Critical care medicine*, vol. 37, no. 10, pp. 2827–2839, 2009.

[87] D. Saey, A. Michaud, A. Couillard, C. H. Côté, M. J. Mador, P. LeBlanc, J. Jobin, and F. Maltais, "Contractile fatigue, muscle morphometry, and blood lactate in chronic obstructive pulmonary disease," *American Journal of Respiratory and Critical Care Medicine*, vol. 171, pp. 1109–1115, may 2005.

[88] H. Ishii and Y. Nishida, "Effect of Lactate Accumulation during Exercise-induced Muscle Fatigue on the Sensorimotor Cortex," *Journal of Physical Therapy Science*, vol. 25, no. 12, pp. 1637–1642, 2013.

[89] J. L. Helbostad, D. L. Sturnieks, J. Menant, K. Delbaere, S. R. Lord, and M. Pijnappels, "Consequences of lower extremity and trunk muscle fatigue on balance and functional tasks in older people: A systematic literature review," *BMC Geriatrics*, vol. 10, p. 56, dec 2010.

[90] J.-j. Wan, Z. Qin, P.-y. Wang, Y. Sun, and X. Liu, "Muscle fatigue: general understanding and treatment," *Experimental & molecular medicine*, vol. 49, no. 10, pp. e384–e384, 2017.

[91] P. A. Karthick, D. M. Ghosh, and S. Ramakrishnan, "Surface electromyography based muscle fatigue detection using high-resolution time-frequency methods and machine learning algorithms," *Computer Methods and Programs in Biomedicine*, vol. 154, pp. 45–56, feb 2018.

[92] A. Subasi and M. K. Kiymik, "Muscle fatigue detection in EMG using time-frequency methods, ICA and neural networks," *Journal of Medical Systems*, vol. 34, pp. 777–785, aug 2010.

[93] A. Muro-de-la Herran, B. Garcia-Zapirain, and A. Mendez-Zorrilla, "Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications," *Sensors*, vol. 14, pp. 3362–3394, feb 2014.

[94] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of emg signal analysis: detection, processing, classification and applications," *Biological procedures online*, vol. 8, no. 1, pp. 11–35, 2006.

[95] E. M. Wojtys, B. B. Wylie, and L. J. Huston, "The effects of muscle fatigue on neuromuscular function and anterior tibial translation in healthy knees," *The American journal of sports medicine*, vol. 24, no. 5, pp. 615–621, 1996.

[96] M. Cifrek, V. Medved, S. Tonković, and S. Ostojić, "Surface emg based muscle fatigue evaluation in biomechanics," *Clinical biomechanics*, vol. 24, no. 4, pp. 327–340, 2009.

[97] M. Melnyk and A. Gollhofer, "Submaximal fatigue of the hamstrings impairs specific reflex components and knee stability," *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 15, no. 5, pp. 525–532, 2007.

[98] M. R. Al-Mulla, F. Sepulveda, and M. Colley, "A Review of Non-Invasive Techniques to Detect and Predict Localised Muscle Fatigue," *Sensors*, vol. 11, pp. 3545–3594, mar 2011.

[99] N. Céspedes, M. Múnera, C. Gómez, and C. A. Cifuentes, "Social human-robot interaction for gait rehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 6, pp. 1299–1307, 2020.

[100] V. Camomilla, E. Bergamini, S. Fantozzi, and G. Vannozzi, "Trends Supporting the In-Field Use of Wearable Inertial Sensors for Sport Performance Evaluation: A Systematic Review," *Sensors*, vol. 18, p. 873, mar 2018.

[101] A. Ejupi, Y. J. Gschwind, T. Valenzuela, S. R. Lord, and K. Delbaere, "A Kinect and Inertial Sensor-Based System for the Self-Assessment of Fall Risk: A Home-Based Study in Older People," *Human-Computer Interaction*, vol. 31, pp. 261–293, jul 2016.

[102] A. Aguirre, J. Casas, N. Céspedes, M. Múnera, M. Rincon-Roncancio, A. Cuesta-Vargas, and C. A. Cifuentes, "Feasibility study: Towards estimation of fatigue level in robot-assisted exercise for cardiac rehabilitation," in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, pp. 911–916, IEEE, 2019.

[103] M. D. Sánchez Manchola, M. J. P. Bernal, M. Munera, and C. A. Cifuentes, "Gait phase detection for lower-limb exoskeletons using foot motion data from a single inertial measurement unit in hemiparetic individuals," *Sensors*, vol. 19, p. 2988, 2019.

[104] D. B. Segala, D. Chelidze, A. Adams, J. M. Schiffman, and L. Hasselquist, "Tracking physiological fatigue in prolonged load carriage walking using phase space warping and smooth orthogonal decomposition," in *ASME International Mechanical Engineering Congress and Exposition*, vol. 48630, pp. 323–331, 2008.

[105] M. Mugnosso, F. Marini, M. Holmes, P. Morasso, and J. Zenzeri, "Muscle fatigue assessment during robot-mediated movements," *Journal of neuroengineering and rehabilitation*, vol. 15, no. 1, pp. 1–14, 2018.

[106] V. C. Chan, S. M. Beaudette, K. B. Smale, K. H. Beange, and R. B. Graham, "A subject-specific approach to detect fatigue-related changes in spine motion using wearable sensors," *Sensors*, vol. 20, no. 9, p. 2646, 2020.

[107] B. Hanley and A. K. Mohan, "Changes in Gait During Constant Pace Treadmill Running," *Journal of Strength and Conditioning Research*, vol. 28, pp. 1219–1225, may 2014.

[108] S. C. Winter, J. B. Lee, R. I. Leadbetter, and S. J. Gordon, "Validation of a single inertial sensor for measuring running kinematics overground during a prolonged run," *J. Fit. Res*, vol. 5, pp. 14–23, 2016.

[109] J. García-López, J. C. Morante, A. Ogueta-Alday, and J. A. Rodríguez-Marroyo, "The Type Of Mat (Contact vs. Photocell) Affects Vertical Jump Height Estimated From Flight Time," *Journal of Strength and Conditioning Research*, vol. 27, pp. 1162–1167, apr 2013.

[110] H. S. Longpré, J. R. Potvin, and M. R. Maly, "Biomechanical changes at the knee after lower limb fatigue in healthy young women," *Clinical Biomechanics*, vol. 28, pp. 441–447, apr 2013.

[111] X. Qu and J. C. Yeo, "Effects of load carriage and fatigue on gait characteristics," *Journal of Biomechanics*, vol. 44, pp. 1259–1263, apr 2011.

[112] M. Yung, "Fatigue at the workplace: Measurement and temporal development," *UWSpace*, 2016.

[113] M. Yung, P. L. Bigelow, D. M. Hastings, and R. P. Wells, "Detecting within-and between-day manifestations of neuromuscular fatigue at work: an exploratory study," *Ergonomics*, vol. 57, no. 10, pp. 1562–1573, 2014.

[114] W. E. Craighead and C. B. Nemeroff, *The concise Corsini encyclopedia of psychology and behavioral science*. John Wiley & Sons, 2004.

[115] S. Chernbumroong, S. Cang, A. Atkins, and H. Yu, "Elderly activities recognition and classification for applications in assisted living," *Expert Systems with Applications*, vol. 40, no. 5, pp. 1662–1674, 2013.

[116] H. Banaee, M. U. Ahmed, and A. Loutfi, "Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges," *Sensors*, vol. 13, no. 12, pp. 17472–17500, 2013.

[117] I. H. Lopez-Nava and A. Munoz-Melendez, "Wearable inertial sensors for human motion analysis: A review," *IEEE Sensors Journal*, vol. 16, no. 22, pp. 7821–7834, 2016.

[118] A. Vinciarelli, A. Esposito, E. André, F. Bonin, M. Chetouani, J. F. Cohn, M. Cristani, F. Fuhrmann, E. Gilmartin, Z. Hammal, *et al.*, "Open challenges in modelling, analysis and synthesis of human behaviour in human–human and human–machine interactions," *Cognitive Computation*, vol. 7, no. 4, pp. 397–413, 2015.

[119] Y. Meng and H.-C. Kim, "A review of accelerometer-based physical activity measurement," in *Proceedings of the International Conference on IT Convergence and Security 2011*, pp. 223–237, Springer, 2012.

[120] P. Gang, J. Hui, S. Stirenko, Y. Gordienko, T. Shemsedinov, O. Alienin, Y. Kochura, N. Gordienko, A. Rojbi, J. L. Benito, *et al.*, "User-driven intelligent interface on the basis of multimodal augmented reality and brain-computer interaction for people with functional disabilities," in *Future of Information and Communication Conference*, pp. 612–631, Springer, 2018.

[121] R. LeMoyne and T. Mastroianni, "Role of machine learning for gait and reflex response classification," in *Wearable and Wireless Systems for Healthcare I*, pp. 111–120, Springer, 2018.

[122] Y. Gordienko, S. Stirenko, Y. Kochura, O. Alienin, M. Novotarskiy, and N. Gordienko, "Deep learning for fatigue estimation on the basis of multimodal human-machine interactions," *arXiv preprint arXiv:1801.06048*, 2017.

[123] Y. Kochura, S. Stirenko, O. Alienin, M. Novotarskiy, and Y. Gordienko, "Comparative analysis of open source frameworks for machine learning with use case in single-threaded and multi-threaded modes," in *2017 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT)*, vol. 1, pp. 373–376, IEEE, 2017.

[124] K. Gonzalez, F. Sasangohar, R. K. Mehta, M. Lawley, and M. Erraguntla, "Measuring fatigue through heart rate variability and activity recognition: A scoping literature review of machine learning techniques," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 61, pp. 1748–1752, SAGE Publications Sage CA: Los Angeles, CA, 2017.

[125] B. J. Shea, J. M. Grimshaw, G. A. Wells, M. Boers, N. Andersson, C. Hamel, A. C. Porter, P. Tugwell, D. Moher, and L. M. Bouter, "Development of amstar: a measurement tool to assess the methodological quality of systematic reviews," *BMC medical research methodology*, vol. 7, no. 1, pp. 1–7, 2007.

[126] M. Ulinskas, R. Damaševičius, R. Maskeliūnas, and M. Woźniak, "Recognition of human daytime fatigue using keystroke data," *Procedia computer science*, vol. 130, pp. 947–952, 2018.

[127] Z. S. Maman, M. A. A. Yazdi, L. A. Cavuoto, and F. M. Megahed, "A data-driven approach to modeling physical fatigue in the workplace using wearable sensors," *Applied ergonomics*, vol. 65, pp. 515–529, 2017.

[128] J. Zhang, T. E. Lockhart, and R. Soangra, "Classifying lower extremity muscle fatigue during walking using machine learning and inertial sensors," *Annals of biomedical engineering*, vol. 42, no. 3, pp. 600–612, 2014.

[129] M. Karg, G. Venture, J. Hoey, and D. Kulić, "Human movement analysis as a measure for fatigue: A hidden markov-based approach," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 22, no. 3, pp. 470–481, 2014.

[130] M. Lee, M. Roan, B. Smith, and T. E. Lockhart, "Gait analysis to classify external load conditions using linear discriminant analysis," *Human movement science*, vol. 28, no. 2, pp. 226–235, 2009.

[131] M. Karg, K. Kühnlenz, M. Buss, W. Seiberl, F. Tusker, M. Schmeelk, and A. Schwirtz, "Expression and automatic recognition of exhaustion in natural walking," in *IADIS Int. Conf. Interfaces and Human Computer Interaction (IHCI)*, pp. 165–172, 2008.

[132] K. Yoshino, T. Motoshige, T. Araki, and K. Matsuoka, "Effect of prolonged free-walking fatigue on gait and physiological rhythm," *Journal of biomechanics*, vol. 37, no. 8, pp. 1271–1280, 2004.

[133] J. L. Helbostad, S. Leirfall, R. Moe-Nilssen, and O. Sletvold, "Physical fatigue affects gait characteristics in older persons," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 62, no. 9, pp. 1010–1015, 2007.

[134] J. Zhang, T. E. Lockhart, and R. Soangra, "Classifying Lower Extremity Muscle Fatigue During Walking Using Machine Learning and Inertial Sensors," *Annals of Biomedical Engineering*, vol. 42, pp. 600–612, mar 2014.

[135] D. A. Winter, "Human balance and posture control during standing and walking," *Gait & posture*, vol. 3, no. 4, pp. 193–214, 1995.

[136] L. Cavuoto and F. Megahed, "Understanding fatigue and the implications for worker safety," in *ASSE Professional Development Conference and Exposition*, OnePetro, 2016.

[137] J. Taborri, S. Rossi, E. Palermo, F. Patanè, and P. Cappa, "A Novel HMM Distributed Classifier for the Detection of Gait Phases by Means of a Wearable Inertial Sensor Network," *Sensors*, vol. 14, pp. 16212–16234, sep 2014.

[138] A. Sabatini, C. Martelloni, S. Scapellato, and F. Cavallo, "Assessment of Walking Features From Foot Inertial Sensing," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 486–494, mar 2005.

[139] M. D. Sánchez Manchola, M. J. P. Bernal, M. Munera, and C. A. Cifuentes, "Gait phase detection for lower-limb exoskeletons using foot motion data from a single inertial measurement unit in hemiparetic individuals," *Sensors*, vol. 19, no. 13, 2019.

[140] D. Kotiadis, H. Hermens, and P. Veltink, "Inertial Gait Phase Detection for control of a drop foot stimulator," *Medical Engineering & Physics*, vol. 32, pp. 287–297, may 2010.

[141] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *International conference on pervasive computing*, pp. 1–17, Springer, 2004.

[142] S. Pirttikangas, K. Fujinami, and T. Nakajima, "Feature selection and activity recognition from wearable sensors," in *International symposium on ubiquitious computing systems*, pp. 516–527, Springer, 2006.

[143] Z. Sedighi Maman, M. A. Alamdar Yazdi, L. A. Cavuoto, and F. M. Megahed, "A data-driven approach to modeling physical fatigue in the workplace using wearable sensors," *Applied Ergonomics*, vol. 65, pp. 515–529, nov 2017.

[144] N. Kern, B. Schiele, and A. Schmidt, "Multi-sensor activity context detection for wearable computing," in *European Symposium on Ambient Intelligence*, pp. 220–232, Springer, 2003.

[145] W. S. Marras, S. A. Lavender, S. E. Leurgans, S. L. Rajulu, S. W. G. Allread, F. A. Fathallah, and S. A. Ferguson, "The role of dynamic three-dimensional trunk motion in occupationally-related," *Spine*, vol. 18, no. 5, pp. 617–628, 1993.

[146] T. Huynh and B. Schiele, "Analyzing features for activity recognition," in *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*, pp. 159–163, 2005.

[147] E. A. Heinz, K. S. Kunze, S. Sulistyo, H. Junker, P. Lukowicz, and G. Tröster, "Experimental evaluation of variations in primary features used for accelerometric context recognition," in *European Symposium on Ambient Intelligence*, pp. 252–263, Springer, 2003.

[148] A. Krause, D. P. Siewiorek, A. Smailagic, and J. Farringdon, "Unsupervised, dynamic identification of physiological and activity context in wearable computing.," in *ISWC*, vol. 3, p. 88, 2003.

[149] S.-W. Lee and K. Mase, "Activity and location recognition using wearable sensors," *IEEE pervasive computing*, vol. 1, no. 3, pp. 24–32, 2002.

[150] D. Lessley, J. Crandall, G. Shaw, R. Kent, and J. Funk, "A normalization technique for developing corridors from individual subject responses," tech. rep., SAE Technical paper, 2004.

[151] K. Moorhouse *et al.*, "An improved normalization methodology for developing mean human response curves," in *International Technical Conference on the Enhanced Safety of Vehicles*, 2013.

[152] N. Yoganandan, M. W. Arun, and F. A. Pintar, "Normalizing and scaling of data to derive human response corridors from impact tests," *Journal of biomechanics*, vol. 47, no. 8, pp. 1749–1756, 2014.

[153] R. Kohavi *et al.*, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Ijcai*, vol. 14, pp. 1137–1145, Montreal, Canada, 1995.

[154] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 1, pp. 4–37, 2000.

[155] H. Liu and M. Cocea, "Semi-random partitioning of data into training and test sets in granular computing context," *Granular Computing*, vol. 2, no. 4, pp. 357–386, 2017.

[156] M. W. Browne, "Cross-validation methods," *Journal of mathematical psychology*, vol. 44, no. 1, pp. 108–132, 2000.

[157] A. Dag, K. Topuz, A. Oztekin, S. Bulur, and F. M. Megahed, "A probabilistic data-driven framework for scoring the preoperative recipient-donor heart transplant survival," *Decision Support Systems*, vol. 86, pp. 1–12, 2016.

[158] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[159] Z. Y. Algamal and M. H. Lee, "Penalized logistic regression with the adaptive lasso for gene selection in high-dimensional cancer classification," *Expert Systems with Applications*, vol. 42, no. 23, pp. 9326–9332, 2015.

[160] P. Afsar, P. Cortez, and H. Santos, "Automatic visual detection of human behavior: A review from 2000 to 2014," *Expert Systems with Applications*, vol. 42, no. 20, pp. 6935–6956, 2015.

[161] P. Ghaderyan, A. Abbasi, and S. Saber, "A new algorithm for kinematic analysis of handwriting data; towards a reliable handwriting-based tool for early detection of alzheimer's disease," *Expert Systems with Applications*, vol. 114, pp. 428–440, 2018.

[162] G. Rescio, A. Leone, and P. Siciliano, "Supervised machine learning scheme for electromyography-based pre-fall detection system," *Expert Systems with Applications*, vol. 100, pp. 95–105, 2018.

[163] J. Ryu and D.-H. Kim, "Real-time gait subphase detection using an emg signal graph matching (esgm) algorithm based on emg signals," *Expert Systems with Applications*, vol. 85, pp. 357–365, 2017.

[164] H. Yigit, "A weighting approach for knn classifier," in *2013 international conference on electronics, computer and computation (ICECCO)*, pp. 228–231, IEEE, 2013.

[165] B. J. Erickson, P. Korfiatis, Z. Akkus, and T. L. Kline, "Machine learning for medical imaging," *Radiographics*, vol. 37, no. 2, pp. 505–515, 2017.

[166] G. Madzarov, D. Gjorgjevikj, and I. Chorbev, "A multi-class svm classifier utilizing binary decision tree," *Informatica*, vol. 33, no. 2, 2009.

[167] N. A. Mahmon and N. Ya'acob, "A review on classification of satellite image using artificial neural network (ann)," in *2014 IEEE 5th Control and System Graduate Research Colloquium*, pp. 153–157, IEEE, 2014.

[168] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[169] T. G. Dietterich, "Ensemble methods in machine learning," in *International workshop on multiple classifier systems*, pp. 1–15, Springer, 2000.

[170] M. Kuhn, K. Johnson, *et al.*, *Applied predictive modeling*, vol. 26. Springer, 2013.

[171] A. Fernández, S. García, M. Galar, R. C. Prati, B. Krawczyk, and F. Herrera, *Learning from imbalanced data sets*, vol. 10. Springer, 2018.

[172] B. Krawczyk, "Learning from imbalanced data: open challenges and future directions," *Progress in Artificial Intelligence*, vol. 5, no. 4, pp. 221–232, 2016.

[173] S. S. Skiena, *The data science design manual.* Springer, 2017.

[174] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," in *2014 science and information conference*, pp. 372–378, IEEE, 2014.

[175] E. M. Smets, B. Garssen, B. Bonke, and J. C. De Haes, "The multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue," *Journal of Psychosomatic Research*, vol. 39, pp. 315–325, apr 1995.

[176] P. Kakria, N. Tripathi, and P. Kitipawang, "A real-time health monitoring system for remote cardiac patients using smartphone and wearable sensors," *International journal of telemedicine and applications*, vol. 2015, 2015.

[177] A. S. Moohialdin, B. T. Suhariadi, and M. K. Siddiqui, "Practical validation measurements of a physiological status monitoring sensor in real construction activities," *Streamlining Information Transfer between Construction and Structural Engineering Edited by Shiau, J, Vimonsatit, V, Yazdani, S, and Singh, A*, 2018.

[178] D. P. Swain, C. A. Brawner, A. C. of Sports Medicine, *et al.*, *ACSM's resource manual for guidelines for exercise testing and prescription.* Wolters Kluwer Health/Lippincott Williams & Wilkins, 2014.

[179] K. Sahlin, "Muscle fatigue and lactic acid accumulation.," *Acta physiologica Scandinavica. Supplementum*, vol. 556, pp. 83–91, 1986.

[180] P. C. Douris, "The effect of isokinetic exercise on the relationship between blood lactate and muscle fatigue," *Journal of Orthopaedic & Sports Physical Therapy*, vol. 17, no. 1, pp. 31–35, 1993.

[181] M. Garnacho-Castaño, R. Dominguez, and J. Maté-Muñoz, "Understanding the meaning of lactate threshold in resistance exercises," *International journal of sports medicine*, vol. 36, no. 05, pp. 371–377, 2015.

[182] D. Abe, T. Yoshida, H. Ueoka, K. Sugiyama, and Y. Fukuoka, "Relationship between perceived exertion and blood lactate concentrations during incremental running test in young females," *BMC Sports Science, Medicine and Rehabilitation*, vol. 7, no. 1, pp. 1–9, 2015.

[183] M. A. Serrano, A. Salvador, E. González-Bono, C. Sanchís, and F. Suay, "Relationships between recall of perceived exertion and blood lactate concentration in a judo competition," *Perceptual and Motor Skills*, vol. 92, no. 3_suppl, pp. 1139–1148, 2001.

[184] J. Steed, G. A. Gaesser, and A. Weltman, "Rating of perceived exertion and blood lactate concentration during submaximal running," *Medicine and science in sports and exercise*, vol. 26, pp. 797–797, 1994.

[185] R. Hetzler, R. Seip, S. Boutcher, E. Pierce, D. Snead, and A. Weltman, "Effect of exercise modality on ratings of perceived exertion at various lactate concentrations.," *Medicine and science in sports and exercise*, vol. 23, no. 1, pp. 88–92, 1991.

[186] J. J. Demello, K. J. Cureton, R. E. Boineau, and M. M. Singh, "Ratings of perceived exertion at the lactate threshold in trained and untrained men and women," *Med Sci Sports Exerc*, vol. 19, no. 4, pp. 354–362, 1987.

[187] A. E. Aubert, B. Seps, and F. Beckers, "Heart rate variability in athletes," *Sports medicine*, vol. 33, no. 12, pp. 889–919, 2003.

[188] J. Achten and A. E. Jeukendrup, "Heart rate monitoring," *Sports medicine*, vol. 33, no. 7, pp. 517–538, 2003.

[189] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Aaai*, vol. 5, pp. 1541–1546, Pittsburgh, PA, 2005.

[190] J. Casas, B. Irfan, E. Senft, L. Gutiérrez, M. Rincon-Roncancio, M. Munera, T. Belpaeme, and C. A. Cifuentes, "Social assistive robot for cardiac rehabilitation: A pilot study with patients with angioplasty," in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '18, (New York, NY, USA), pp. 79–80, ACM, 2018.

[191] J. Casas, E. Senft, L. F. Gutierrez, M. Rincon-Rocancio, M. Munera, T. Belpaeme, and C. A. Cifuentes, "Social assistive robots: assessing the impact of a training assistant robot in cardiac rehabilitation," *International Journal of Social Robotics*, pp. 1–15, 2020.