UNIVERSIDAD MIGUEL HERNÁNDEZ DE ELCHE
Programa de Doctorado en Bioingeniería

MODELING OF USER BEHAVIOR USING ARTIFICIAL INTELLIGENCE TECHNIQUES FOR THE DESIGN AND ADAPTATION OF REHABILITATION THERAPIES ASSISTED BY ROBOTIC DEVICES

DAVID MARTÍNEZ PASCUAL

Director: Nicolás Manuel García Aracil

Elche, Junio 2024

This Doctoral Thesis is supported by a compendium of previously published works, two of them in an impact journal indexed according to JCR Science Edition. The body of this thesis is partially constituted by the following articles, whose complete bibliographic references are indicated below:

Martínez-Pascual, D.; Catalán, J. M.; Blanco-Ivorra, A.; Sanchís, M.; Arán-Ais, F. and García-Aracil, N. (2023), «Estimating vertical ground reaction forces during gait from lower limb kinematics and vertical acceleration using wearable inertial sensors,» *Frontiers in Bioengineering and Biotechnology* **11**, DOI: https://doi.org/10.3389/fbioe.2023.1199459

- Journal title: Frontiers in Bioengineering and Biotechnology (ISSN: 1530-437X)
- Impact Factor JCR-SCI (2022): 5,7
- Category: Multidisciplinary Sciences, Quartile Q1 (16/73).

Martínez-Pascual, D.; Catalán, J. M.; Blanco-Ivorra, A.; Sanchís, M.; Arán-Ais, F. and García-Aracil, N. (2024), «Gait Activity Classification with Convolutional Neural Network using Lower Limb Angle Measurement from Inertial Sensors,» *IEEE Sensors Journal*, pp. 1–1, DOI: 10.1109/JSEN.2024.3400296

- Journal title: IEEE Sensors Journal (ISSN: 2296-4185)
- Impact Factor JCR-SCI (2022): 4,3
- Category: Instruments and Instrumentation, Quartile Q1 (15/63).

ABSTRACT

Recent reports indicate that the prevalence of stroke has increased by 50% over the last 17 years, and at present, 1 in 4 people are at risk of getting a stroke in their lifetime. Hemiparesis is a common after-effect of stroke, leading to a decrease in limb function and significantly affecting quality of life. However, personalized therapy adapted to the needs of each patient can lead to the recovery of mobility of the affected limbs.

This doctoral thesis focuses on the study and modeling of user behavior to design and adapt robotic rehabilitation systems to the user's needs. First, an underwater sensor system has been developed to measure the kinematics of the lower limbs. In addition, a forces model acting during gait in aquatic environments has been developed with the aim of designing and subsequently controlling a lower limb exoskeleton immersed in water. Moreover, a Machine Learning model has been proposed to overcome the difficulties encountered in measuring the vertical ground reaction force in an aquatic environment. The model takes as inputs the kinematics of the lower limbs in the sagittal plane and the vertical acceleration of the user. Furthermore, a Deep Learning model has been developed to recognize a set of gait activities to generalize the previous model and to adapt the assistance of a lower limb exoskeleton to different activities of daily living. Also, in order to adapt and personalize robot-assisted upper limb therapies, two approaches are presented. First, a method has been proposed to dynamically adapt the difficulty of upper limb rehabilitation therapies. In addition, a method to provide assistance according to the needs of each patient based on demonstrations by an experienced therapist has been proposed.

Part of the main results of this thesis have been published in two journals included in the Journal Citation Reports (JCR). The publication *Estimating vertical ground reaction forces during gait from lower limb kinematics and vertical acceleration using wearable inertial sensors* focuses on the development of a Machine Learning model to estimate the vertical reaction force during gait using a motion capture system based on inertial sensors. The publication *Gait Activity Classification with Convolutional Neural Network using Lower Limb Angle Measurements from Inertial Sensors* focuses on the development of a classifier based on a Deep Learning model to recognize activities of daily living during gait, employing

techniques to understand the decision-making process and analyzing the feasibility of adapting the model to multiple situations.

RESUMEN

Informes recientes señalan que la frecuencia de los accidentes cerebrovasculares ha aumentado un 50 % durante los últimos 17 años y, en la actualidad, 1 de cada 4 personas están en riesgo de sufrir de ictus durante su vida. La hemiparesia es una secuela común, y conlleva una disminución en la función de las extremidades y afecta significativamente su calidad de vida. No obstante, una terapia personalizada y adaptada a las necesidades de cada paciente puede llevar a la recuperación de la movilidad de los miembros afectados.

Esta tesis doctoral se centra en el estudio y modelado del comportamiento del usuario para diseñar y adaptar sistemas robóticos de rehabilitación a las necesidades de cada usuario. En primer lugar, se ha desarrollado un sistema de sensores sumergible para medir la cinemática de los miembros inferiores. Además, se ha desarrollado un modelo de las fuerzas que actúan durante la marcha en entornos acuáticos con el objetivo de diseñar y posteriormente controlar un exoesqueleto de miembro inferior sumergido en agua. Asimismo, debido a las dificultades encontradas para medir la fuerza de reacción vertical con el suelo en un entorno acuático, se ha propuesto un modelo de Machine Learning que toma como entradas la cinemática de los miembros inferiores en el plano sagital y la aceleración vertical del usuario. Además, con el fin de poder generalizar el modelo anterior y poder adaptar la asistencia de un exoesqueleto de miembro inferior a diferentes actividades de la vida diaria, se ha desarrollado un modelo de Deep Learning para reconocer un conjunto de actividades de la marcha. Asimismo, con el fin de adaptar y personalizar terapias de miembro superior asistidas con robots, se han propuesto dos posibles soluciones. En primer lugar, se ha propuesto un método que permite adaptar de forma dinámica la dificultad de terapias de rehabilitación de miembro superior. Además, se ha propuesto un método para ofrecer asistencia según la necesidad de cada paciente basado en demostraciones realizadas por un terapeuta experimentado.

Parte de los resultados principales de la presente tesis se han publicado en dos revistas incluidas en el *Journal Citation Reports* (JCR). La publicación *Estimating vertical ground reaction forces during gait from lower limb kinematics and vertical acceleration using wearable inertial sensors* se centra en el desarrollo de un modelo de Machine Learning

para estimar la fuerza de reacción vertical durante la marcha empleando un sistema de captura de movimiento basado en sensores inerciales. La publicación *Gait Activity Classification with Convolutional Neural Network using Lower Limb Angle Measurements from Inertial Sensors* se centra en el desarrollo de un clasificador basado en un modelo de Deep Learning para reconocer actividades de la vida diaria durante la marcha, empleando técnicas para comprender el proceso de decisión y analizando la viabilidad de adaptar el modelo a múltiples situaciones.

AGRADECIMIENTOS

Gracias a mis padres y a mi hermana, por apoyarme cada día durante toda mi carrera académica y por procurar que nunca me faltara de nada.

Gracias a mi director de tesis, por brindarme su confianza y por enseñarme a investigar.

Gracias a todos mis compañeros por hacer que cada día en el laboratorio sea más ameno.

Gracias a Sonia, mi compañera de vida, por animarme y apoyarme cada día, haciendo que cada día contigo sea especial.

Gracias a todos vosotros porque esta tesis no es solo mía, también es vuestra.

This thesis was funded by the Spanish Ministry of Universities through the Research and Doctorate Supporting Program FPU20/05137

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

ADL	Activity of Daily Living
ABI	Acquired Brain Injury
CVA	Cerebrovascular Accident
SCI	Spinal Cord Injuries
GRF	Ground Reaction Forces
vGRF	Vertical Ground Reaction Force

HAR Human Activity Recognition
DDA Dynamic Difficulty Adjustment
AAN Assistance-As-Needed
IMU Inertial Measurement Units
LSTM Long Short-Term Memory
SVM Support Vector Machine
CNN Convolutional Neural Network
ECG Electrocardiography
EMG Electromiography
HR Heart Rate
RFQ Respiratory Frequency
EDA Electrodermal Activity
RBF Radial Basis Function
RL Reinforcement Learning
LfD Learning from Demonstration
RMSE Root Mean Squared Error
DH Denavit-Hartenberg
COM Center of Mass
BW Body Weight
LP Loading Peak
MP Mid Stance Peak
TP Terminal Stance Peak
RF Random Forest
FNN Feedforward Neural Network
NRMSE Normalized Root Mean Squared Error
1D-CNN One-dimensional Convolutional Neural Network
2D-CNN Two-dimensional Convolutional Neural Network
MLP Multilayer Perceptron
GAP Global Average Pooling
FCN Fully-Convolutional Neural Network
CE Categorical Cross-Entropy
Grad-CAM Gradient-weighted Class Activation-Map
t-SNE t-Distributed Stochastic Neighbor Embedding
SAM Self-Assessment Manikin
GSR Galvanic Skin Response
HRV Heart Rate Variability

RMSSD Root Mean Square of Successive Differences between normal heartbeats
HF High-Frequency
SCR Skin Conductance Responses
SCL Skin Conductance Level
LOWESS Locally WEighted Scatterplot Smoothing
IQR Interquartile Range
ReLu Rectified Linear Unit
PReLu Parametric Rectified Linear Unit
MSE Mean Square Error
FPR False Positive Rate
FNR False Negative Rate
ResNet Residual Network
t-LeNet Time LeNet
MCDCNN Multi-scale Convolutional Neural Network
Time-CNN Time Convolutional Neural Network

1

INTRODUCTION

1.1 THE IMPACT OF HUMAN-ROBOT INTERACTION IN REHABILITATION

Recent reports indicate that the prevalence of stroke has increased by 50% over the last 17 years, and at present, 1 in 4 people are at risk of getting a stroke in their lifetime (Kalita et al., 2023). The United Nations' World Population Prospects 2022 report indicates that the proportion of individuals aged 65 or older is projected to increase globally from 2022 to 2050 (Nations, 2017). In 2022, approximately one in ten individuals worldwide fall into this age group. By 2030, this figure is anticipated to increment to nearly 12%, and by 2050, it is expected to reach 16%. This population segment is particularly susceptible to Acquired Brain Injury (ABI) such as a Cerebrovascular Accident (CVA) or stroke. The risk of stroke doubles every decade after the age of 55, and most cases of ABI or stroke lead to profound challenges in the affected individual's daily life, as they can result in cognitive, physical, emotional, or social impairments. In most cases, a suitable rehabilitation program with appropriate intensity can aid hemiparetic patients in regaining movement in their affected limbs (Kwakkel, 2006; Winstein et al., 2016; Alessandro et al., 2020).

In this scenario, the use of robotic rehabilitation equipment has emerged as a beneficial method to aid in the post-stroke motor recovery process. Such devices can provide highly standardized and intensive therapy, and they also offer mechanisms to monitor the patient's recovery advancements. In the current landscape, there exists a diverse array of robotic systems designed to assist human limb movements in Activity of Daily Living (ADL). These systems can be broadly categorized into two types: exoskeletons and end-effector systems (Oña et al., 2018). Exoskeletons are mechanical structures with degrees of freedom that are aligned with the joints of the limb requiring assistance or rehabilitation. This alignment enables direct control over the limb's movement. Notably, exoskeletons offer the advantage of complete determination of the arm's position, as the robot is specifically adjusted to the limb being rehabilitated. In contrast, end-effector systems maintain the user's grip at a single point, which

simplifies the adjustment of the robot to the user, streamlining its usability. However, this approach does not allow for a complete determination of the limb's position that is undergoing rehabilitation.

The spectrum of proposals for the design and control of rehabilitation robotics is extensive and diverse, as evidenced by the literature (Qassim and Wan Hasan, 2020; C. Liu et al., 2022; J. Zhou et al., 2021). Nonetheless, it is imperative to prioritize the end-user to ensure the delivery of effective assistance. Initially, a comprehensive understanding of the biomechanics of the limb requiring rehabilitation is crucial, needing analysis of both its kinematics and dynamics (Hasan and Dhingra, 2022; L. Zhou et al., 2017). Moreover, there is substantiated evidence within the scientific community regarding the significant role of patient engagement in promoting neural plasticity during motor relearning scenarios (Blank et al., 2014; Losey et al., 2018). Additionally, various research studies underscore the necessity of personalizing therapy to the patient's needs to facilitate optimal recovery (Faria et al., 2020). Consequently, it will be necessary to personalize the assistance of robotic rehabilitation systems according to the needs of each patient.

The integration of Machine Learning methodologies has significantly influenced their application in the context of robotic-assisted therapies (Fong et al., 2020b; Faria et al., 2020; Ai et al., 2023). Given that rehabilitation robotic systems directly interact with patients, they have are able to collect data related to the executed movements (such as position, velocity, or tracking error) and other relevant physiological signals during therapeutic exercises (such as motion capture, cardiovascular responses, or electrodermal activity). Therefore, the acquired data can be used as valuable input for Machine Learning models, enabling the study of user behavior, learning from historical experiences, and dynamically adapting the assistance provided by robotic systems to optimize patient outcomes.

1.2 MOTIVATION

As per scientific literature, the use of robotic systems has demonstrated a favorable impact on rehabilitation therapies for both the upper and lower limbs. Nevertheless, some studies centered on the effects of robot-assisted therapy reported that individuals who suffered from severe spasticity, pain, and joint contractures were excluded from the randomized controlled trials (Veerbeek et al., 2017). Hence, this is a challenge that

is hindering the progress of robot-assisted therapies and their application to certain patient groups.

In an attempt to overcome these problems, the SPLASH research project proposes to perform robot-aided rehabilitation therapy in a supportive environment. Water therapy has great advantages, including weightlessness or joint unloading, and has been shown to be effective in the treatment of several disorders, such as arthritis (Roper et al., 2013), or Spinal Cord Injuries (SCI) (Kesiktas et al., 2004).

It is essential to acknowledge that the execution of rehabilitation exercises during immersion in the aquatic environment causes the occurrence of hydrodynamic forces. Therefore, the design and control of robotic devices must be adapted and take into account the aforementioned forces produced in the aquatic environment. A principal force exerted on the limbs during aquatic immersion is the buoyancy force, which is an upward force exerted by a fluid that opposes the weight of a partially or fully immersed object. Additionally, the drag force emerges as a significant factor, acting in opposition to the relative motion of an object traversing through the fluid.

In addition to the forces mentioned above, the Ground Reaction Forces (GRF) must be considered when designing a lower limb rehabilitation device. Among these forces, the Vertical Ground Reaction Force (vGRF) stands out due to its significant magnitude. However, there is an important challenge: measuring vGRF can be financially costly, especially when dealing with aquatic environments. Due to this limitation, one potential solution is to develop a Machine Learning model. This model would use information about the user's movements (kinematics) and estimate the vGRF as its output.

It is worth noting that understanding the actions performed by the user, as well as their state and behavior, is essential to adapt robotic systems to the needs of each patient. Related to the rehabilitation of the lower limbs, different activities for gait rehabilitation could be considered during aquatic therapies. Therefore, a Human Activity Recognition (HAR) system could help to adapt the assistance of exoskeleton-type rehabilitation devices.

Regarding upper limb rehabilitation, instead of using an upper limb exoskeleton, an end-effector robotic device can also be employed. As introduced, these systems are an easy-to-use solution to assist during upper limb rehabilitation, and they are commonly used together with serious games to enhance patient engagement and motivation (Heins et al., 2017; Catalán; Blanco-Ivorra, et al., 2023a). In this context, different approaches can be considered to adapt the therapy. On the one hand, a Dynamic

Difficulty Adjustment (DDA) system can be proposed to modulate the difficulty of the serious games according to the users' skills to provide an appropriate level of challenge and keep motivation . On the other hand, the assistance provided by the robotic system can be adapted to each patient by using an Assistance-As-Needed (AAN) intelligent system.

In this thesis, diverse objectives are pursued within the field of rehabilitation and assistive robotics context. First, a model that describes the forces during underwater gait is proposed to study the effects on the human body and to design underwater lower limb exoskeleton devices. In this context, Machine Learning models have been trained to estimate the vGRF and recognize different gait ADLs, which can be used to adapt the robotic systems accordingly. Additionally, a DDA system has been developed to adapt the difficulty of serious games during upper limb rehabilitation exercises. Finally, an AAN architecture has been proposed to perform assistance decision-making during robot-aided rehabilitation therapies.

1.3 MAIN CONTRIBUTIONS OF THE THESIS

The present work contains diverse original contributions in the domain of rehabilitation and assistive robotics. Specifically, it focuses on the development of models to enhance the design and adaptability of robotic systems. The main contributions are highlighted below:

- The hydrodynamic forces acting on the human body during gait have been modeled in order to design an exoskeleton-type device for assistance during aquatic therapies
- A Machine Learning model to estimate the vGRF during ground-level walking from the lower limb joint angles has been trained. Two different models have been compared and the impact of introducing vertical acceleration to estimate this force has been evaluated
- An adaptation of the model for estimating vGRF during diverse gait ADLs has been proposed. In addition, the impact of knowing the performed ADLs has been evaluated
- A Deep Learning model has been trained to recognize different ADLs during gait from the joint trajectories of the lower limbs. Diverse methods have been used to

understand the decision-making of the model, and the impact of reducing the number of inputs to adapt the model to different situations has been studied

- A DDA system has been developed to adjust the difficulty of serious games and has been evaluated with able-bodied users during five consecutive sessions. The proposed system is based on a genetic algorithm that varies different parameters of a game according to the user's movements and the game score. In addition, the psychophysiological state of the users has been evaluated using self-assessment tests and several physiological signals.
- An AAN architecture has been proposed to decide when to assist during upper limb robot-assisted therapies. The architecture is based on demonstrations from a therapist, which have been used to train a Deep Learning model. In addition, it has been proposed to fine-tune the assistance model to each patient, evaluating the impact on the model performance

1.4 THESIS STRUCTURE

This thesis is framed in the field of rehabilitation and assistance robotics and has multiple objectives. The first objective is to present a model that describes the forces that arise during walking in water, to study their effects on the human body, and to design a robotic device that provides assistance during gait. With this purpose, several Machine Learning models have been trained to estimate the vGRF and detect different ADLs, which can be then used to adapt the robotic systems accordingly. Additionally, it is intended to develop architectures that can adjust the difficulty of serious games and provide assistance when necessary during robot-assisted upper limb rehabilitation therapies.

Part of the results presented in this thesis have been published in a journal paper included in the *Journal Citation Reports* (JCR).

This document consists of eight Chapters, and is structured as described below:

- **CHAPTER** 1. Description of the impact of human-robot interaction in rehabilitation scenarios, as well as the importance of Machine Learning models for decision-making. The motivation of this doctoral thesis and a summary of the main contributions are also presented.

- **CHAPTER** 2. A State of the Art review of methods used to perform biomechanical analysis in aquatic environments, Machine Learning-based methods to estimate the vGRF, methods used to perform HAR, systems proposed to adapt serious games dynamically, and AAN methods with robots in rehabilitation scenarios are reviewed.
- **CHAPTER** 3. Description of the proposal and motivation of the SPLASH project, detailing the sensor systems developed to analyze the biomechanics of the lower limbs. An underwater motion capture system is also detailed, and the theoretical calculations to estimate the forces involved during walking in an aquatic environment are described.
- **CHAPTER** 4. A model is presented for estimating the vGRF during ground-level walking from the measurement of the flexion/extension angles of the lower limbs. A feasibility study is also included to estimate the vGRF during different gait ADLs.
- **CHAPTER** 5. A model for recognizing different gait ADLs based on the joint trajectories of the lower limbs in the sagittal plane is presented. In addition, diverse techniques to understand the decision-making process of the model are used.
- **CHAPTER** 6. The development of a system for dynamically adapting the difficulty of serious games with rehabilitation robots is described. The proposed method has been evaluated during several consecutive sessions, and an analysis of the psychophysiological state of the users is performed.
- **CHAPTER** 7. An AAN method based on the demonstrations of a therapist is presented to decide when to assist during robot-aided rehabilitation therapies.
- **CHAPTER** 8. Conclusions obtained during the development of the doctoral thesis and possible future work.

1.5 PUBLICATIONS

The work carried out during the doctoral thesis has resulted in the following publications:

- Martínez-Pascual, D.; Catalán, J. M.; Blanco-Ivorra, A.; Sanchís, M.; Arán-Ais, F. and García-Aracil, N. (2023), «Estimating vertical ground reaction forces during

gait from lower limb kinematics and vertical acceleration using wearable inertial sensors,» *Frontiers in Bioengineering and Biotechnology* **11**, DOI: https://doi.org/10.3389/fbioe.2023.1199459

- Martínez-Pascual, D.; Catalán, J. M.; Blanco-Ivorra, A.; Sanchís, M.; Arán-Ais, F. and García-Aracil, N. (2024), «Gait Activity Classification with Convolutional Neural Network using Lower Limb Angle Measurement from Inertial Sensors,» *IEEE Sensors Journal*, pp. 1–1, DOI: 10.1109/JSEN.2024.3400296

2

STATE OF THE ART

In this Chapter, a comprehensive State of the Art review has been conducted to contextualize the main contributions of this thesis concerning existing research. Firstly, an examination of studies focused on biomechanical analysis and modeling of hydrodynamic forces during aquatic rehabilitation therapies are described. Secondly, current research that estimates vGRF using Machine Learning is presented. Thirdly, recent studies related to HAR during gait for the identification of different ADLs are described. Fourthly, a review of studies presenting DDA systems to adapt serious games during rehabilitation therapies is included. Finally, a section with a review of AAN systems for robot-assisted upper limb rehabilitation therapies is presented.

2.1 BIOMECHANICAL ANALYSIS AND HYDRODYNAMIC FORCES MODELING DURING GAIT

As commented in Chapter 1 section 1.2, performing aquatic rehabilitation therapies has great advantages such as joint unloading and a decrease of the apparent weight. Due to these advantages, different authors have proposed diverse methods to analyze both the kinematics and kinetics of the lower limbs during gait.

Concerning the kinematic analysis, D. Volpe et al. (Volpe et al., 2017) employed a vision-based system fixed over the surface water to analyze the effects of hydrotherapy in people with Parkinson's disease. In that study, the authors performed a comparison between underwater and on-land gait, and they showed a reduction in stride length, cadence, and speed. Despite vision-based motion capture systems are considered the gold standard, the equipment can be financially costly, data acquisition can be time-consuming, and markers occlusions can be a source of error (Jakob et al., 2021). As an alternative, Inertial Measurement Units (IMU) can be used to acquire kinematic data during gait. C. Monoli et al. (Monoli et al., 2021) designed an underwater sensor for gait analysis, validating their system with different vision-based systems (optoelectronic and single camera) while measuring the knee angle during gait.

Apart from analyzing the impact of immersion on lower limb kinematics, certain studies delve deeper and suggest approaches for analyzing the effects of joint loading by modeling the hydrodynamic forces. J. Lauer simulated the hydrodynamic forces acting on the human body while performing rehabilitation exercises (flexion, abduction, cycling) (Lauer, 2023). The author proposed a four-stage analysis workflow, which are represented in Figure 2.1. First, a video-based 2D pose estimation was performed, and estimated poses were then converted to 3D coordinates for body mesh estimation. Then, this information was introduced to a fluid flow simulation software to estimate the forces exerted on the user. Finally, the kinematics, hydrodynamic forces, and scaled models are input into OpenSim, a musculoskeletal simulation software (Delp et al., 2007), which can be used to estimate the lower limb joint forces and torques, as well as the muscular activation.

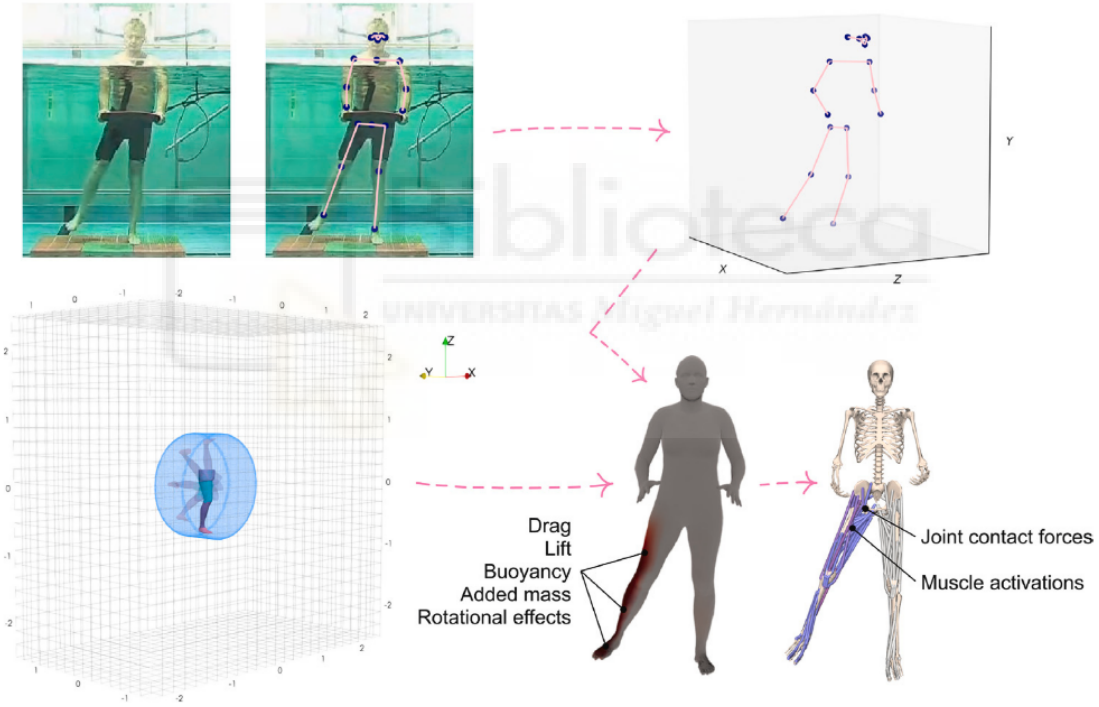


Figure 2.1: Biomechanical analysis modeling process proposed by J. Lauer. Initially, a video-based 2D pose estimation was conducted, and the estimated poses were subsequently transformed into 3D coordinates for body mesh estimation. Subsequently, this data was incorporated into a fluid flow simulation software to calculate the forces exerted on the user. Finally, the kinematics, hydrodynamic forces, and scaled models are input into OpenSim for further analysis. Source: (Lauer, 2023)

Although diverse studies related to the analysis of underwater biomechanics have been conducted, to the best of my knowledge only M. I. V. Orselli and M. Duarte (Orselli and Duarte, 2011) estimated the internal net joint forces and torques on lower limbs when walking in water. The authors took into account the buoyancy and the drag forces generated by the movement of their bodies in the water. In addition, a force plate was used to measure the vGRF, and a video camera was employed to perform a 2D gait analysis at the sagittal plane while walking at comfortable speeds on land and in water at a chest-high level.

In conclusion, IMUs and vision-based methods have been used to analyze the kinematics of the lower limbs during underwater gait. Nevertheless, an underwater motion capture system composed of diverse IMUs can be more appropriate to perform gait analysis whether a method that allows and easy placement and calibration procedures is proposed. In addition, instead of relying on complex physical equations to estimate joint loads, as suggested by M. I. V. Orselli and M. Duarte, an alternative approach involves simplifying the modeling process. Specifically, one can provide a detailed description of the forces exerted on the user during walking. By adopting a methodology similar to the one proposed by J. Lauer, the estimation of the buoyancy force, the drag force, and the vGRF can be introduced into a simulation software such as OpenSim. The main advantage of this method relies on its flexibility: the musculoskeletal models can be adapted based on the user's physical characteristics, enabling assessments of joint loads and muscular activity.

2.2 MACHINE LEARNING FOR VERTICAL GROUND REACTION FORCE ESTIMATION

The advancement of Machine Learning algorithms has prompted researchers to use these methods for the analysis of gait kinematics and the estimation of vGRF, especially during ground-level walking. In (Oh et al., 2013; Choi et al., 2013) human body kinematics together with artificial neural networks were used to predict vGRF. Nonetheless, a significant limitation of this approach is the reliance on multiple infrared cameras to collect motion capture data, which may constrain the space of analysis.

The exploration of IMUs for the estimation of vGRF has been a subject of research. E. Sahabpoor and A. Pavic demonstrated the feasibility of estimating vGRF using a single IMU to measure the vertical acceleration of the subject's body mass (Shahabpoor and Pavic, 2018). Their findings suggest that the placement of an IMU over the C7 vertebra

results in the most accurate vGRF estimations. However, this method cannot determine the force produced on the right and left lower limbs.

In contrast, other works propose to position IMUs directly on the lower limbs to facilitate movement analysis. For instance, X. Jiang et al. used a Random Forest (RF) algorithm to predict vGRF during gait, employing a single IMU to record acceleration and gyroscope data (Jiang et al., 2020). F. J. Wouda et al. also used inertial sensors over the lower limbs, placed on the pelvis and the shank (Wouda et al., 2018) (Figure 2.2). Firstly, the authors trained an artificial neural network to estimate the knee angle. Then, an additional neural network was used to estimate the vGRF during running, introducing the estimated knee angle and the vertical accelerations from the placed IMUs as inputs of the model.

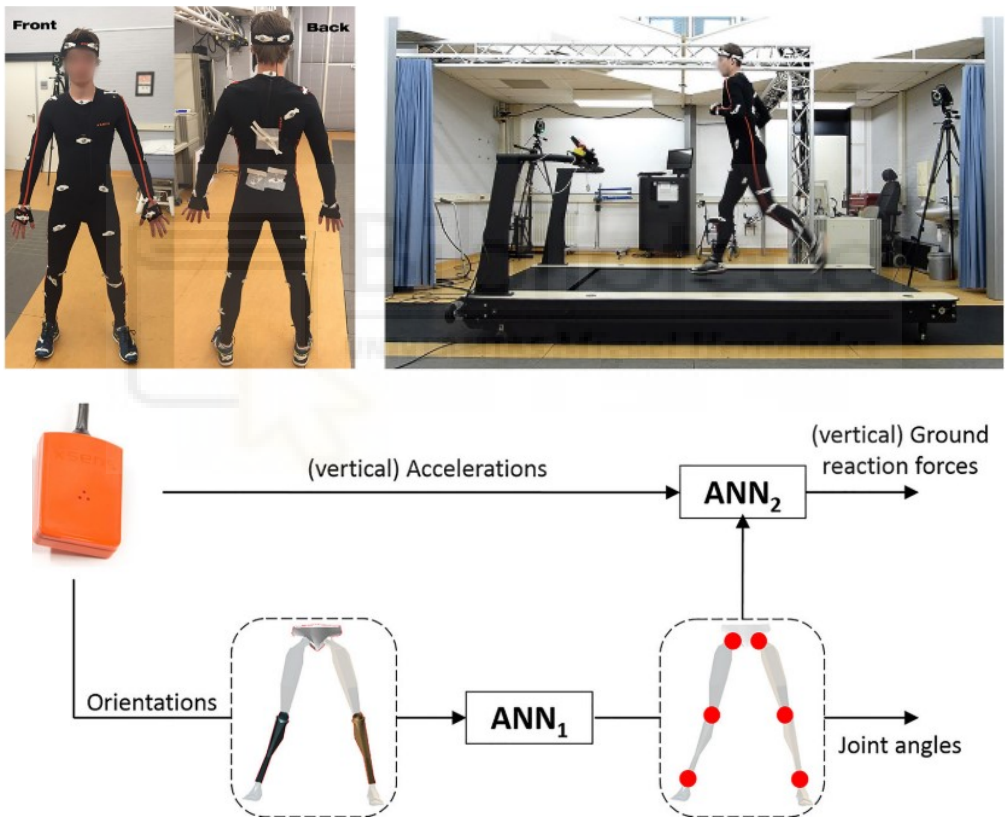


Figure 2.2: Experimental setup and IMU-based approach to estimate the vGRF during running proposed by F. J. Wouda. First, the knee angle is estimated and later input together with vertical accelerations to estimate the vGRF. Source: (Wouda et al., 2018)

It is also noteworthy that the work presented by D. Liu et al. (D. Liu et al., 2022) is the first study to estimate accurately the GRF during stair walking. The authors acquired motion capture data with a set of infrared cameras to calculate the whole-body kinematics and two force plates were embedded into the stair steps (Figure 2.3). With this data, the authors trained a bidirectional Long Short-Term Memory (LSTM) model to estimate the GRF with high accuracy.

Regarding the studies presented, the use of Machine Learning algorithms has been shown to be feasible when estimating the vGRF by different approaches. When performing a biomechanical analysis, the objective often extends beyond vGRF measurements, since this force is typically measured together with the kinematics of the lower limbs. Consequently, a Machine Learning model can be trained with the kinematics of lower limb joints for the estimation of vGRF. It is also noteworthy that the studies presented in this section extract different features directly from the raw IMU signals. Hence, variations in device placement would affect negatively the vGRF estimation accuracy. Additionally, as previously indicated, vertical acceleration has been employed to estimate the vGRF due to the established correlation with lower limb kinematics. Nevertheless, the combined impact of lower limb kinematics and vertical acceleration on the prediction of vGRF remains unexamined.

It is also worth mentioning that there is a lack of studies that estimate the vGRF during different scenarios (ground-level, ramp, and stairs). In this context, only the study proposed by D. Liu et al. presented a method to estimate the GRF during stair

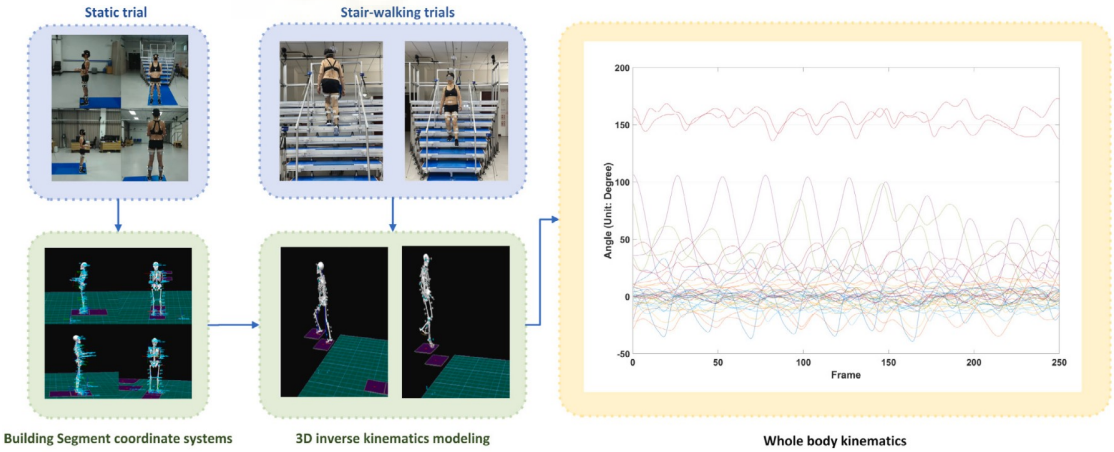


Figure 2.3: Experimental setup used by D. Liu et al. for whole-body kinematics and GRF measurement. Source: (D. Liu et al., 2022)

walking. However, as introduced, the main limitation of vision-based approaches could limit the space of analysis. Therefore, IMU-based models to estimate the vGRF during different tasks can be presented as an alternative approach.

2.3 HUMAN ACTIVITY RECOGNITION DURING GAIT

HAR is a complex problem that has many practical applications such as human-robot interaction, rehabilitation, or health monitoring (Yadav et al., 2021). With recent advances in wearable technology and Machine Learning, analyzing human biomechanics and recognizing different ADLs has become more affordable, accurate, and feasible.

Diverse vision-based methods have been investigated for activity recognition and adaptation of prosthetic limbs. In this context, H. A. Varol and Y. Massalin (Varol and Massalin, 2016) investigated the feasibility of using a leg-mounted depth camera to adapt lower limb prostheses. The authors used a Support Vector Machine (SVM) to detect five different situations (standing, walking, level walking, stair climbing, and stair descending). By contrast, B. Laschowski et al. (Laschowski et al., 2022) proposed using a chest-mounted camera, which provides less relative body motion than lower limb mounting. In that work, the authors used a Deep Learning model to analyze the images from the camera to detect ground-level walking and ascending and descending stairs (Figure 2.4).

The use of inertial sensors for HAR has been widely investigated in addition to vision-based methods. C. Pham et al. (Pham et al., 2017) used accelerometers embedded in shoe insoles to detect different activities, using the raw signals as inputs to a Convolutional Neural Network (CNN) to detect six activities (running, walking, cycling, kicking, standing, and jumping). The study presented by A. A. Badawi (Badawi et al., 2018) used the public dataset HuGaDB (Chereshnev and Kertész-Farkas, 2018) to classify different gait ADLs. A RF model was trained using as inputs different features extracted from the accelerometer and gyroscope signals placed over the thigh, shin, and foot. A Deep Learning approach was presented by I. H. Lopez-Nava et al. in (Lopez-Nava et al., 2020) to avoid extracting features directly from the IMU signals. The authors used a Two-dimensional Convolutional Neural Network (2D-CNN) to detect ground-level walking, ascending/descending a ramp, and ascending/descending stairs by transforming accelerometer and gyroscope signals from the OU-ISIR public dataset

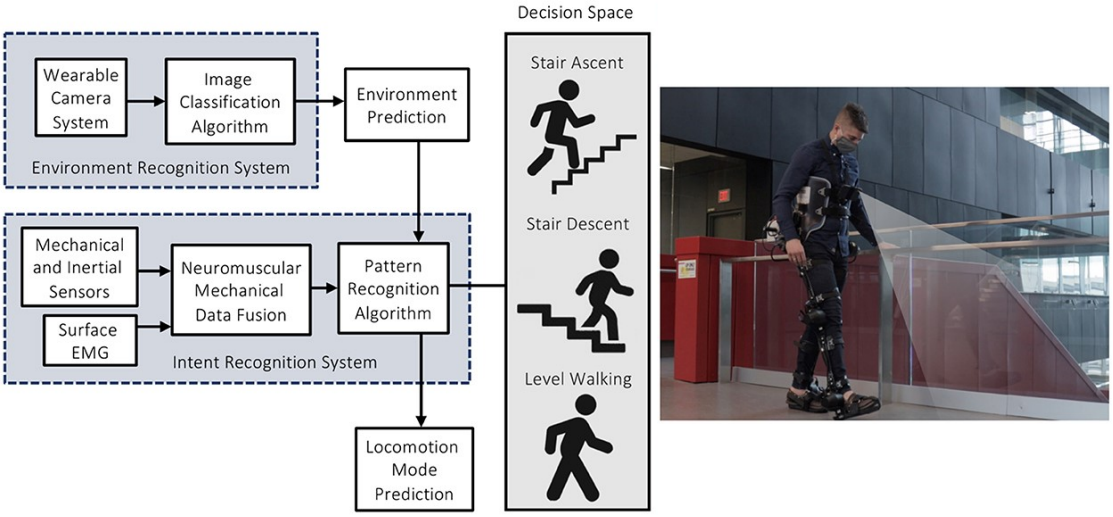


Figure 2.4: Automated locomotion mode recognition system for robotic leg prostheses and exoskeletons proposed by B. Laschowski et al. A vision-based HAR approach is used to adapt the lower limb assistive systems. Source: (Laschowski et al., 2022)

(Makihara et al., 2012). An alternative approach was presented by V. B. Semwal et al. (Semwal et al., 2021), where a single IMU placed at the Center of Mass (COM) of each participant was used to detect different gait activities (jogging, ground-level walking, standing, and stair climbing). In that work, the authors proposed to use a hybrid model based on ensemble learning. The main limitation of this method would be its significant computational cost due to the ensemble learning architecture making it impractical for real-time inference applications.

Although various methods have been proposed for HAR using wearable sensors, certain limitations can be identified. Due to the need for image analysis to detect user activity, vision-based techniques can be computationally costly. In contrast, using IMUs appears to be cost-effective, and computationally affordable to allow HAR using wearable devices. Nevertheless, in the proposed IMU-based methods, the accurate placement of the devices is essential, since any error in the positioning of the IMU can cause variations in the signals and reduce classification accuracy.

2.4 DYNAMIC DIFFICULTY ADJUSTMENT SYSTEMS FOR SERIOUS GAMES

DDA is a method of adjusting a game's features in real-time based on the player's skill level (Dziedzic, 2016; Dziedzic and Włodarczyk, 2018). The goal of DDA is to

enhance the player's satisfaction by creating a state of "flow," a concept from flow theory (Czikszentmihalyi, 1990). Flow refers to a state where the player is fully engaged and enjoying the activity. DDA aims to provide the appropriate level of challenge for the player to achieve this state. If the difficulty level exceeds the player's ability, they may feel anxious, while if their ability surpasses the difficulty, they may feel bored (Csikszentmihalyi, 2000).

In the field of rehabilitation, one of the primary methods employed to adjust game difficulty involves considering the game score. In this context, different approaches have been proposed to adapt the game difficulty level based on the patient's skills. M. Goršič et al. (Goršič et al., 2017) presented a competitive rehabilitation game that was dynamically adjusted based on individual patient scores. Specifically, the difficulty level increased when users achieved a certain score rate and decreased vice versa. Additionally, M. Pirovano et al. (Pirovano et al., 2014) proposed a system architecture where therapists identify a game parameter directly related to the game's difficulty. Using therapist-defined score criteria, Bayesian optimization techniques were implemented to adapt the difficulty level.

In addition to the assessment of exercise performance, an alternative approach for adapting serious games involves the estimation of user preferences through the capture of players' emotions using biofeedback devices. For the implementation of these DDA systems, physiological responses are typically correlated with user preferences as determined by questionnaires (Byrne and Parasuraman, 1996). In this context, Verhulst et al. acquired Electrodermal Activity (EDA) and Electrocardiography (ECG) data from children diagnosed with cerebral palsy to adapt game difficulty according to their emotional state (Verhulst et al., 2015). Building on this concept, Darzi and Novak proposed a DDA system that adjusted a competitive rehabilitation game according to diverse physiological signals, including ECG, Respiratory Frequency (RFQ), skin conductance, and facial Electromiography (EMG) (Darzi and Novak, 2021). Within the field of rehabilitation robotics, F. J. Badesa et al. proposed the use of Heart Rate (HR), RFQ, and EDA to recognize three stress levels during robot-aided rehabilitation therapies, as represented in Figure 2.5 (Badesa; Morales; N. Garcia-Aracil, et al., 2014; Badesa; Morales; N. M. Garcia-Aracil, et al., 2016). Furthermore, Xu et al. proposed to use features derived from ECG, facial EMG, RFQ, EDA, and game performance metrics to identify anxiety levels in patients while using a robotic system for therapy across three levels of exercise intensity (Xu et al., 2018). Although these studies show that adapting exercises to the user's emotions can help to maintain motivation, the

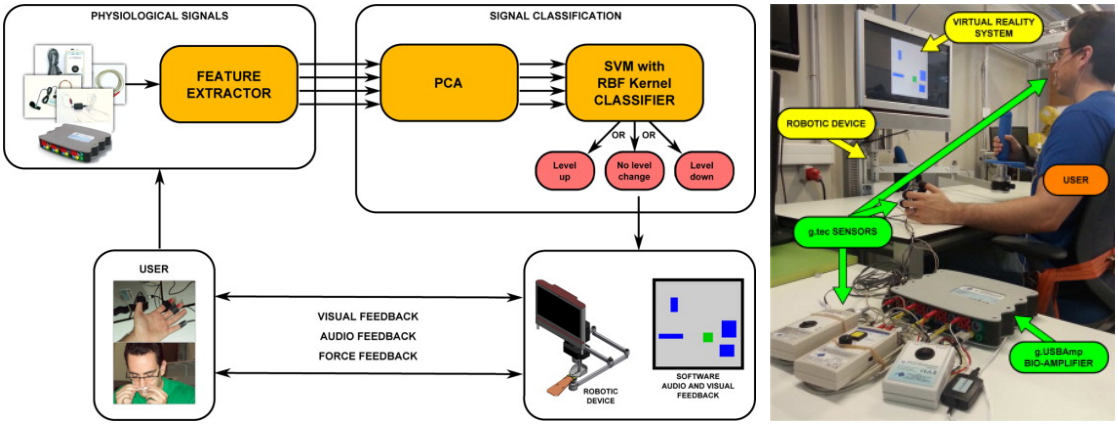


Figure 2.5: DDA system proposed by J. Badesa et al. to adapt serious games during robot-aided rehabilitation therapies. Diverse physiological signals were acquired to assess the psychophysiological state of the user and change the game difficulty. Source: (Badesa; Morales; N. Garcia-Aracil, et al., 2014)

main drawback of these methods is that they only guarantee adaptation to the user's emotions. Hence, they may not guarantee intensive therapy.

An alternative strategy for adapting rehabilitation games involves the modeling of the user's abilities. N. Hocine et al. (Hocine et al., 2015; Tresser et al., 2021) suggested carrying out an initial assessment of the patient to establish a model of the reachable space (Figure 2.6). Based on this evaluation, the DDA system positioned targets within the patient's ability range, changing the number of targets in each zone to preserve a determined score level. By contrast, Andrade et al. proposed to dynamically model the user's behavior and skills (K. d. O. Andrade et al., 2016; K. O. Andrade et al., 2018). The authors used a genetic algorithm to determine game parameters that optimize performance in robotic rehabilitation therapies. Nevertheless, it is important to note that the studies were validated with a virtual player, which does not demonstrate the practicality of applying evolutionary algorithms with real users.

Despite various DDA approaches have been proposed in the literature, the State of the Art review reveals that only the DDA system proposed by Xu et al. takes into account the game score, the therapy performance, and the psychophysiological state of the users. Nevertheless, a limitation of the method proposed is worth mentioning. The system presented by Xu et al. is limited to selecting from only three difficulty levels. Consequently, while the DDA system ensures that anxiety remains at an appropriate level, there is no guarantee that the chosen difficulty level will consistently push users to achieve their maximum performance.

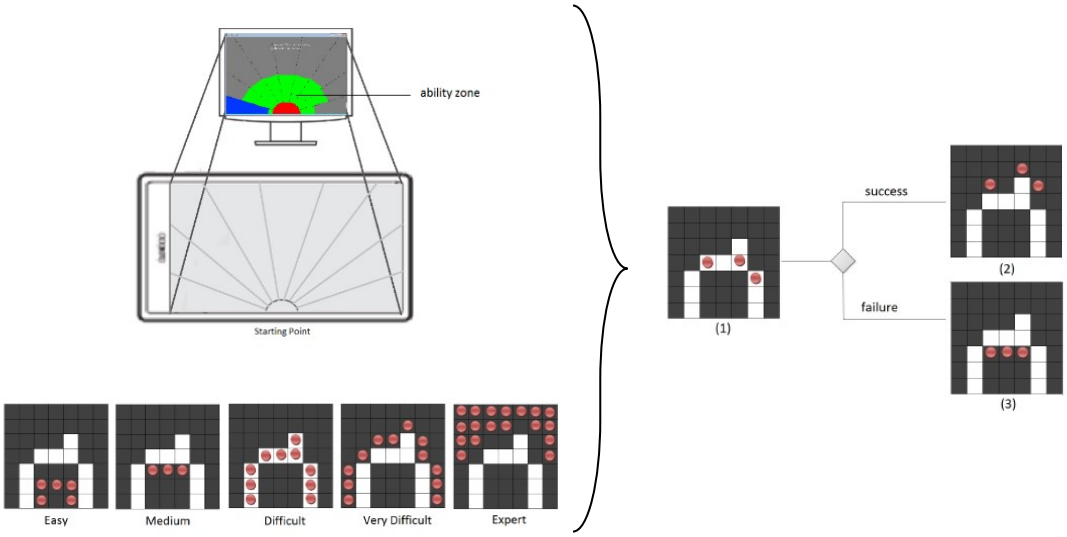


Figure 2.6: DDA system proposed by N. Hocine et al. to adapt rehabilitation games. The system positioned targets within the patient's ability range, changing the number of targets in each zone to preserve a determined score level. Source: (Hocine et al., 2015; Tresser et al., 2021)

It is also worth noting that DDA studies typically evaluate the behavior of the method in a single session. However, the skills and abilities are expected to change over different sessions in rehabilitation scenarios. Therefore, there is a need to develop a method to adapt the game difficulty concerning the user's abilities without performing a recalibration for each therapy session. Furthermore, according to (Maier et al., 2019), increasing the difficulty of rehabilitation exercises has shown beneficial effects on motor recovery. Therefore, a decrease in the target score level between sessions can be used to increase the difficulty of the game.

2.5 ASSISTANCE-AS-NEEDED ARCHITECTURES FOR ROBOT-AIDED REHABILITATION THERAPIES

Recent applications of Machine Learning in rehabilitation robotics attempt to build intelligent devices that can adjust assessment tasks or provide interventions based on the features of a patient to build AAN systems (Ai et al., 2023). L. Luo et al. presented a scheme called GAAN controller for upper extremity rehabilitation training (Luo et al., 2019). The GAAN control paradigm includes a baseline controller and a Gaussian Radial Basis Function (RBF) network. The baseline controller models the

subject's functional ability and provides appropriate task challenges. The weight vectors of the RBF networks are updated based on a greedy strategy that allows the networks to progressively learn the maximum forces provided by the subjects, to encourage active engagement. An alternative AAN method was proposed by S. Pareek et al. (Pareek; Nisar, et al., 2023), where Reinforcement Learning (RL) was used to provide adaptive assistance during a handwriting rehabilitation task. The AAN method provides real-time modulation of robotic support based on subject-tracking errors to minimize robotic support. For the training of the RL agent, the authors used a virtual patient for the generalization of the agent's behavior in different situations. However, the principal limitation of the study is that the proposed method was validated with virtual patients and able-bodied humans.

Despite the great potential of the aforementioned AAN methods, the vast majority of research does not focus on learning the actions and interventions demonstrated by an experimented therapist. According to Fong et al., robots can learn from therapists' movements or assistance criteria, combining the advantages of robotic systems with the expertise and experience of human therapists (Fong et al., 2020a). With this in mind, Learning from Demonstration (LfD) techniques could be applied to robot-aided rehabilitation sessions. Specifically, LfD is a family of Machine Learning methods in which a robot observes a task demonstration by a human operator and learns rules to describe the desired task-oriented actions (Ravichandar et al., 2020).

Within the LfD field, M. Najafi et al.(Najafi et al., 2020) developed a force field-based LfD method using a planar robot to assist patients (Figure 2.7). First, the therapist's movements and interaction forces are recorded while performing an exercise. Then, a non-parametric potential field function is modeled using a convex optimization algorithm. Finally, using the potential field function and a velocity field controller, the robot mimics the therapist's level of interaction. C. Lauretti et al. (Lauretti et al., 2017) proposed a motion planning system for robotic devices to be used in either assistive or rehabilitation scenarios. The system was developed based on dynamic motor primitives for learning the therapist-demonstrated paths for activities of daily living. C. Manuel Martínez et al. (Martinez et al., 2019) proposed an approach to deliver rehabilitation using a Gaussian Mixture Model and Gaussian Mixture Regression-based LfD framework. The rehabilitation exercises were performed collaboratively, with the therapist performing motion commands to a master robot, which then delivered the intervention, while the slave robot and the patients directly interacted with the physical tasks according to their upper limb position and speed. S. Hu et al. (Hu et al., 2021)

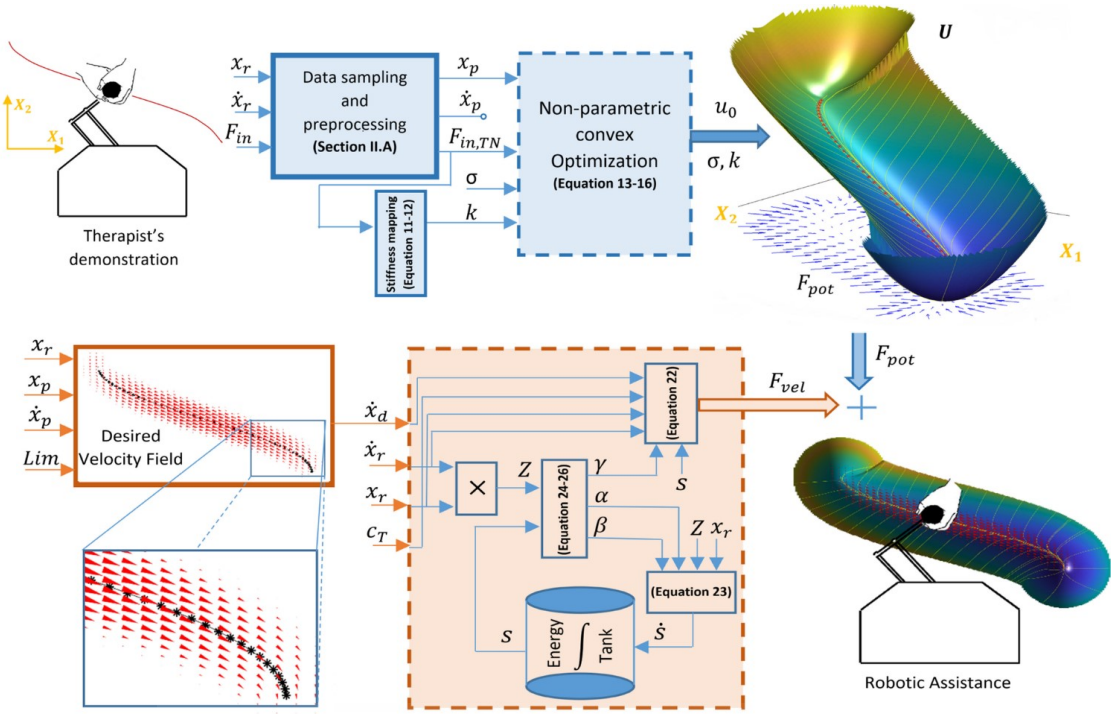


Figure 2.7: LfD architecture proposed by M. Najafi et al. The process involves recording the movements and forces of the therapist. A non-parametric potential field function is then created using a convex optimization algorithm and the robot uses the potential field function and a velocity field controller to replicate the therapist's interaction level. Source: (Najafi et al., 2020)

proposed an LfD approach that allows a humanoid robot to work with people during physical exercises. In this approach, the therapist first demonstrates the exercise by teleoperating the robot to perform the task with the patient. Then, from the state of the robot's end effector, which includes its position and velocity, a model is fitted to infer the robot's effort (force output). This approach allows the robot to learn the exercises and replicate the force interaction with the user.

However, it is worth mentioning that it is of particular interest to allow the patients to perform the exercise by themselves and learn from the therapist the situations in which assistance is needed, rather than using LfD techniques to replicate predetermined assistance movements instructed by the therapist. With this methodology, the robotic system would assist when necessary, encouraging the patient to perform the exercises. Upon thorough review, it appears that the sole study presented by S. Pareek and T. Kesavadas (Pareek and Kesavadas, 2020) proposes to employ this methodology (Figure 2.8). In that study, an AAN architecture based on a LSTM model was presented. A

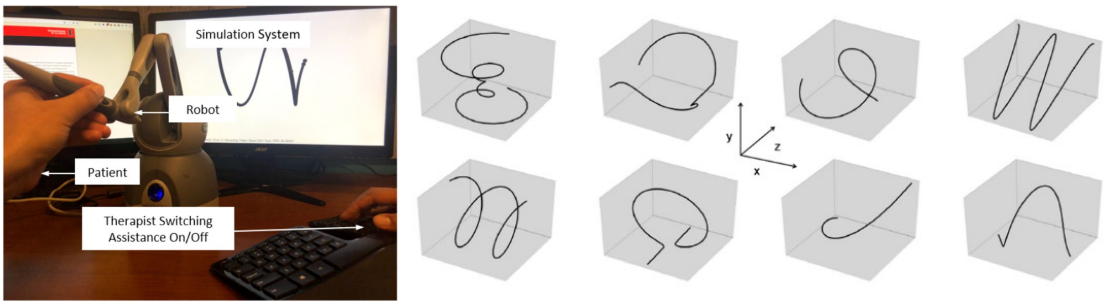


Figure 2.8: LfD architecture proposed by S. Pareek and T. Kesavadas, where a haptic device was used to interact with the game and assist the user. Source: (Pareek and Kesavadas, 2020)

haptic device was used to perform exercises based on following different trajectories. The force-based assistance could be enabled or disabled concerning the observed movements. This information was then used as the demonstration phase to learn when to assist. However, one of the primary limitations of the study proposed by S. Pareek and T. Kesavadas is that the AAN proposed was validated with able-bodied participants, and one of the authors acted as the therapist.

3

AUTO-ADAPTIVE AND MULTIMODAL REHABILITATION ROBOTIC SYSTEM IN CONTROLLED SUPPORTIVE ENVIRONMENTS

3.1 THE SPLASH PROJECT PROPOSAL

As introduced in Chapter 1 section 1.2, some studies that investigated the effects of robot-assisted therapy reported that individuals who suffered from severe spasticity, pain, and joint contractures were excluded from the randomized controlled trials (Veerbeek et al., 2017). In the stroke patients case, spasticity is a relevant topic, arising in about 30% of patients (Thibaut et al., 2013). Additionally, it has been shown that the prevalence of spasticity does not differ between the upper and the lower limbs. In SCI patients, spasticity affects about 70%, causing a considerable disability (Rekand et al., 2012). Moreover, in individuals with SCI, the development of joint contractures and stiffness is one of the most relevant complications, being reported at least one joint contracture in about 66% of patients (Nas et al., 2015). A systematic review summarized the effects of aquatic therapy on mobility in individuals with neurological diseases and showed "fair" evidence that aquatic therapy increases dynamic balance in participants with some neurological disorders (Marinho-Buzelli et al., 2015). One randomized controlled trial and two before-and-after tests demonstrated "fair" evidence for improvement of gait speed after aquatic therapy. In addition, water therapy has been shown to be effective in the treatment of several disorders, such as arthritis (Roper et al., 2013), or SCI (Kesiktas et al., 2004).

Within the SPLASH project, the use of robot-assisted therapies in a controlled water environment has been proposed to address the challenge of clinical trials being inaccessible to certain individuals. However, using rehabilitation robotic devices in water presents challenges, and there have only been preliminary attempts to demonstrate their use in aiding patients' gait in such environments. A proof-of-concept study with nine healthy subjects showed that a robotic orthosis immersed in water could be viably and usefully used for this purpose (Miyoshi; Hiramatsu, et al., 2008). Another

study by the same authors presented an aquatic exercise device that applies assistive or resistive forces to the limb, similar to aquatic elliptical bikes, for those with gait disorders (Miyoshi; Komatsu, et al., 2015). However, the SPLASH concept goes beyond the results reported on those studies. Specifically, it is proposed to understand, model and objectively monitor how a subject performs rehabilitation therapies immersed in water together with the development of a new rehabilitation robotic paradigm.

The envisioned architecture of the SPLASH project is represented in Figure 3.1, and the following objectives need to be addressed for delivering aquatic robot-aided therapies:

- Understanding and modeling the human body effects of water immersion during aquatic rehabilitation therapies.
- Development of aquatic robotic devices based on a radically new type of actuators that are able to exploit the properties of the aquatic medium to get the systems more efficient and lighter than the application of the current actuator technologies
- Development of an intelligent control system capable of understanding the user's motor abilities, intentions, and general physiological state to adapt the assistance provided by the robotic devices according to the the patient's needs

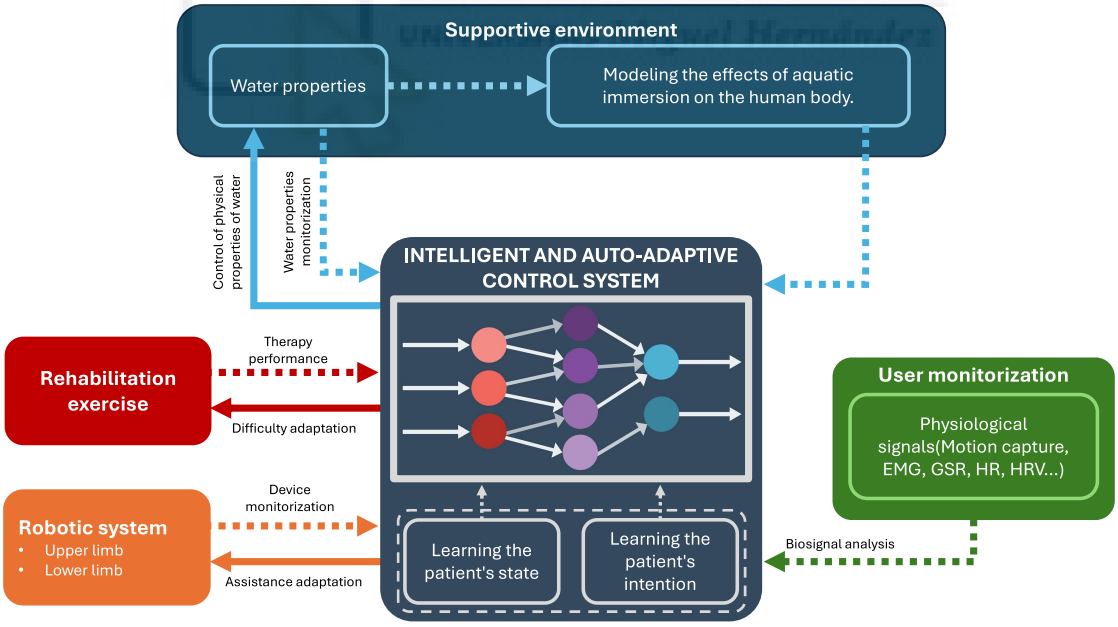


Figure 3.1: Envisioned architecture of the SPLASH project for the development of an aquatic robot-aided rehabilitation environment.

To fulfill the specified objectives, within the framework of the SPLASH project, a multimodal underwater sensor system for user monitorization has been developed (Figure 3.2). On the one hand, a set of EMG sensors are included to estimate muscle activity. It is worth noting that the electronics responsible for data acquisition have been encapsulated to ensure functionality during water immersion. Furthermore, to safeguard against water interference, waterproof adhesive sheets have been strategically positioned over each surface electrode to prevent water contact. On the other hand, a motion capture system based on several XSens Dot IMUs was implemented to analyze the kinematics during gait (see section 3.2). These wireless devices are IP68-protected and allow synchronized data acquisition and storage. These sensors will be used to study the effects of water on the human body and to develop and validate a model capable of describing the forces exerted by the presence of water. A potential approach is to use the methodology proposed by J. Lauer (Lauer, 2023), where the kinematics of the human body together with the calculation of the forces exerted by the aquatic environment were introduced into the OpenSim simulation software to estimate the loads on the user's joints.

The information provided by human kinetics and kinematics study will play a crucial role in the subsequent design and control of an underwater lower limb exoskeleton,

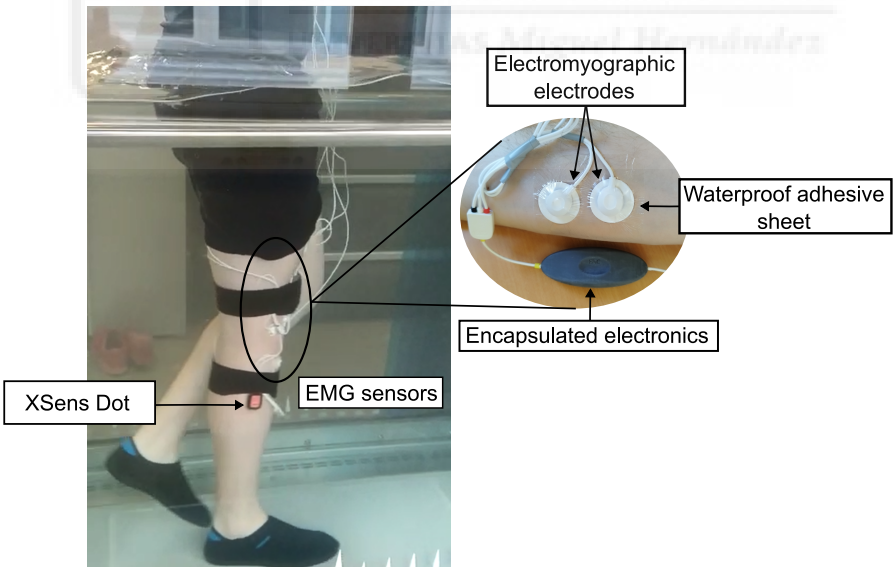


Figure 3.2: Sensor system developed within the SPLASH project. An IMU-based motion capture data was implemented using diverse XSens Dot devices. In addition, commercial EMG sensors have been encapsulated to make them waterproof.

which relies on an innovative hybrid actuator design. The proposed hybrid actuation system will employ two actuators: a small electric motor, and a novel actuation system based on a set of micro water jets in opposite directions (Figure 3.3). The plan is to use a water-jet-based actuation system to provide a large torque contribution, while a small electric actuator will provide a small torque contribution. This approach allows us to use a low transmission, which keeps the actuator's mechanical impedance low. This hybrid actuator solution is a missing element in robotics, and it can innovate the way robots for underwater applications are designed and built. It is noteworthy that this actuator is not only designed for the development of active exoskeletons since it is also conceived to be part of other robotic solutions. As shown in Figure 3.4, the proposed hybrid actuator can be used to develop a lower limb exoskeleton or could also be part of an end-effector upper limb rehabilitation robotic system.

The SPLASH project aims to propose robotic aquatic therapies for both upper and lower limbs, as well as different types of robots. Therefore, the adaptive intelligent architecture can be addressed from different approaches. Regarding a lower limb exoskeleton, a possible adaptation of the device may be due to the gait activity performed. In the SPLASH project, instead of only performing therapy on a submerged

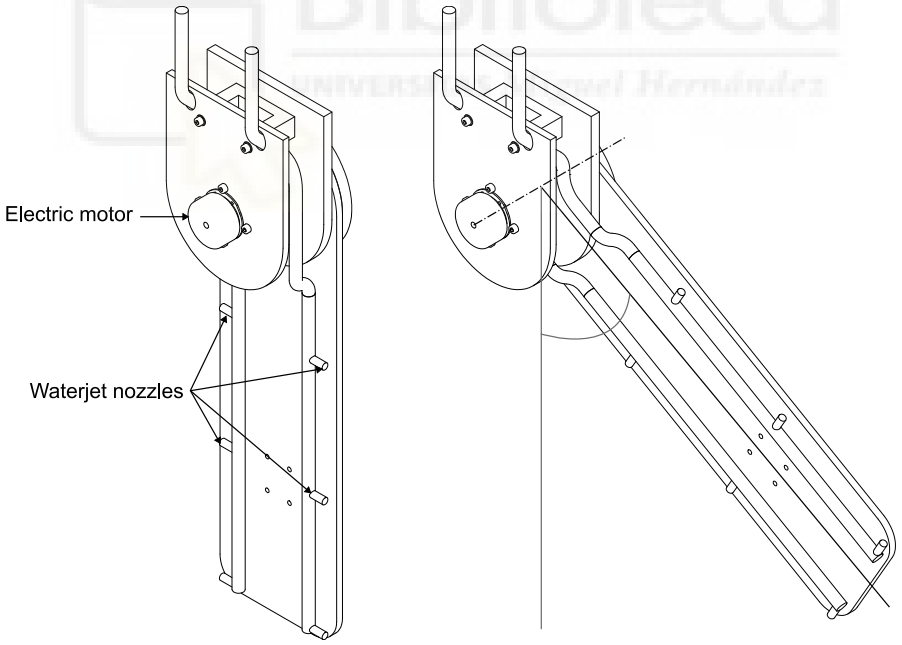


Figure 3.3: The novel hybrid actuator proposed in the SPLASH project. The actuator is composed of an electric motor and diverse waterjet nozzles to produce a rotational movement.

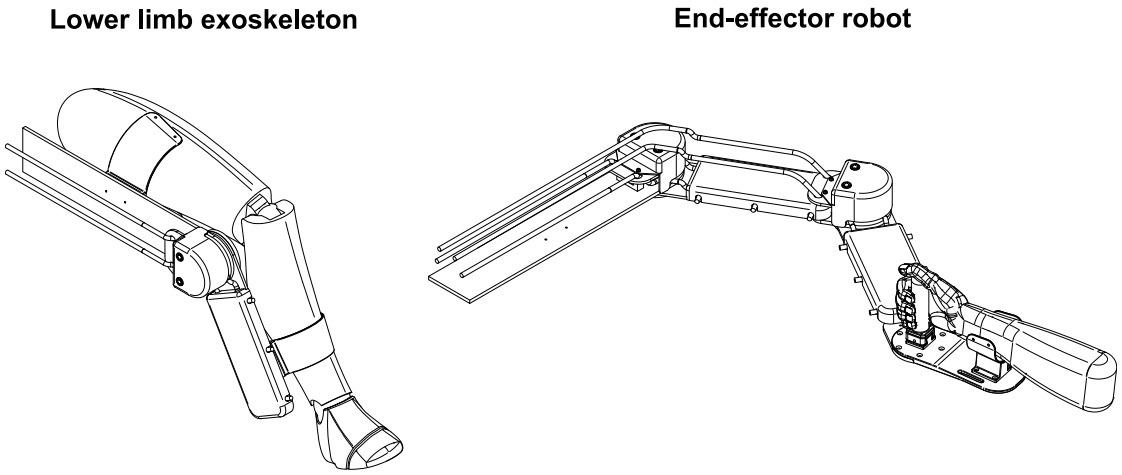


Figure 3.4: Concepts of use of the hybrid actuator proposed in the SPLASH project. The actuator can be used to develop an exoskeleton to assist during walking. In addition, it is conceived for the development of an aquatic end-effector robot.

treadmill, it is proposed to generalize aquatic therapies to multiple activities. Figure 3.5 shows a concept of therapy with several ADLs, where the lower limb exoskeleton assists the patient during ground-level, ramp, and stair walking. For this purpose, a classifier based on a Deep Learning model has been developed in Chapter 5 to differentiate between different activities.

In the context of upper limb rehabilitation, end-effector robots can be presented as an alternative to exoskeletons. With this type of robot, robotic systems are commonly employed together with serious games. In this thesis, two adaptation methods are proposed to personalize therapy activities. The first method involves adjusting the

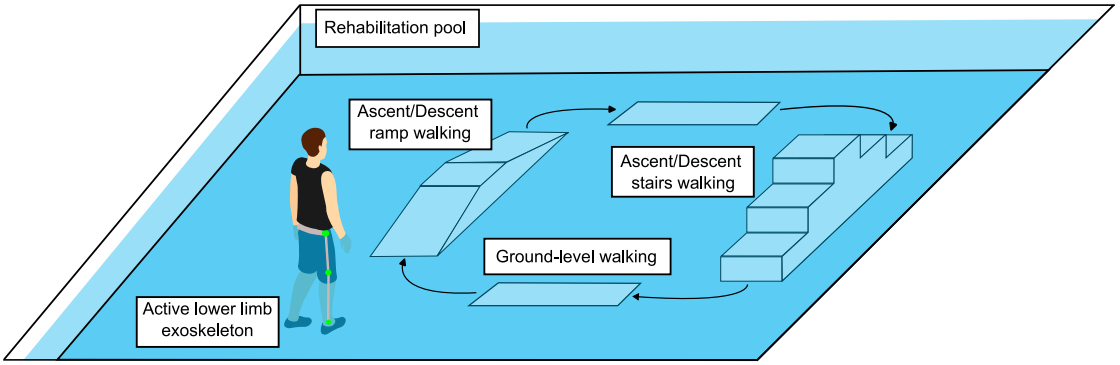


Figure 3.5: Concept of aquatic rehabilitation therapy with several ADLs, where the lower limb exoskeleton assists the patient during ground-level, ramp, and stair walking.

difficulty level of the games according to the user's abilities. Specifically, a genetic algorithm is proposed and detailed in Chapter 6. The second method relies on an AAN system to decide when to assist during a point-to-point serious game. This method is described in Chapter 7, and it is based on the therapist's interventions. The system incorporates a Deep Learning model that learns from the therapist's interactions to determine when the robot should provide assistance.

3.2 LOWER LIMBS ANGLE MEASUREMENT

As introduced in section 3.1, a set of XSens Dot IMUs has been used to develop an underwater motion capture system. The system uses the method proposed by T. Seel et al. to estimate the joint angles of the lower limbs during gait (Seel; Raisch, et al., 2014). This method relies on the analysis of joint movements and can accurately obtain joint angles, even at low acquisition rates (from 40 Hz) (Seel; Schauer, et al., 2012). In this section, a description of the implemented method is detailed, including the calibration and angle measurement processes.

3.2.1 *Identification of the joint axis and position coordinates*

The method used to develop a motion capture system avoids the need for an algorithm that makes assumptions about the position and orientation of different IMUs related to the user's lower limb. Additionally, to estimate joint angles using this algorithm, we rely on gyroscope and accelerometer data from the IMUs, which eliminates the dependence on a uniform magnetic field. In this method, two IMUs attached to the upper and lower segments of each joint are assumed, which measure the accelerations $a_1(t), a_2(t) \in \mathbb{R}^3$ and the angular rates $g_1(t), g_2(t) \in \mathbb{R}^3$ at a sample period Δt.

The data from the gyroscope is employed to determine the unit direction vectors $j_1, j_2 \in \mathbb{R}^3$ corresponding to the flexion/extension axes of the hip, knee, and ankle. It is worth mentioning that vectors j_1 and j_2 remain constant and are only influenced by the sensors' orientation relative to the joint. Geometrically, $g_1(t)$ and $g_2(t)$ are only differentiated by the angular velocity of the joint and a rotational matrix. As a result, the length of their projections onto the plane of the joint remains consistent at every moment, which can be described as:

$$||g_1(t) \times j_1||_2 - ||g_2(t) \times j_2||_2 = 0 \; \forall t \tag{3.1}$$

where $||\cdot||_2$ is the Euclidean norm.

In addition to determining direction vectors, it is essential to calculate the joint center position in the local coordinates of the IMUs. These vectors, denoted as $o_1, o_2 \in \mathbb{R}^3$ remain constant, and their values depend on the mounting position and orientation relative to the leg segments. To compute these vectors, T. Seel et al.(Seel; Raisch, et al., 2014) propose considering the acceleration of each sensor as a combination of two components: the center of joint acceleration, and the acceleration due to the rotation of the sensor around the joint center. Furthermore, the acceleration of the joint center should be consistent in both IMU local frames. This consistency can be expressed mathematically as:

$$||a_1(t) - \Gamma_{g_1(t)}(o_1)||_2 - ||a_2(t) - \Gamma_{g_2(t)}(o_2)||_2 = 0 \; \forall t \tag{3.2}$$

$$\Gamma_{g_i(t)}(o_i) := g_i(t) \times (g_i(t) \times o_i) + \dot{g}_i(t) \times o_i \; , i = 1, 2$$

where $\Gamma_{g_i(t)}(o_i)$ is the radial and tangential acceleration due to the IMU rotation around the joint center.

All previously established constraints must be satisfied by any movement of the joint. Therefore, j_1, j_2, o_1, and o_2 can be identified by minimizing the left side of the Equations (3.1) and (3.2). To determine j_1 and j_2, the problem can be reduced to a four-dimensional space by limiting the axis to a unit length. This allows us to represent j_1 and j_2 using spherical coordinates.

$$j_1 = \begin{bmatrix} cos(\phi_1) \cdot cos(\theta_1) \\ cos(\phi_1) \cdot sin(\theta_1) \\ sin(\phi_1) \end{bmatrix}, j_2 = \begin{bmatrix} cos(\phi_2) \cdot cos(\theta_2) \\ cos(\phi_2) \cdot sin(\theta_2) \\ sin(\phi_2) \end{bmatrix} \tag{3.3}$$

where ϕ_i and θ_i are inclination and azimuth. In addition, let the sum of squared errors for N samples be defined as:

$$\Psi(\phi_1, \theta_1, \phi_2, \theta_2) := \sum_{i=1}^{N} (||g_1(t_i) \times j_1||_2 - ||g_2(t_i) \times j_2||_2)^2 \tag{3.4}$$

In the same manner, o_1 and o_2 can also be estimated by:

$$\Psi(o_1, o_2) := \sum_{i=1}^{N} (||a_1(t_i) - \Gamma_{g_1(t_i)}(o_1)||_2 - ||a_2(t_i) - \Gamma_{g_2(t_i)}(o_2)||_2)^2 \tag{3.5}$$

In order to minimize $\Psi(\phi_1, \theta_1, \phi_2, \theta_2)$ and $\Psi(o_1, o_2)$ the Trust Region Reflective algorithm have been applied (Byrd et al., 1988).

Once the IMUs are placed, the participants should execute diverse movements as a calibration process to obtain o_1, o_2,j_1, j_2, as represented in Figure 3.6. First, to determine the coordinates for the hip, knee, and ankle joints o_1, o_2 participants can be instructed to execute random circling movements. Additionally, to estimate the coordinates of the joint axes j_1, j_2, data from the participants' walking patterns can be collected to identify the primary axis of motion. In other words, these gait data are used to determine flexion/extension axes of the hip, knee, and ankle joints.

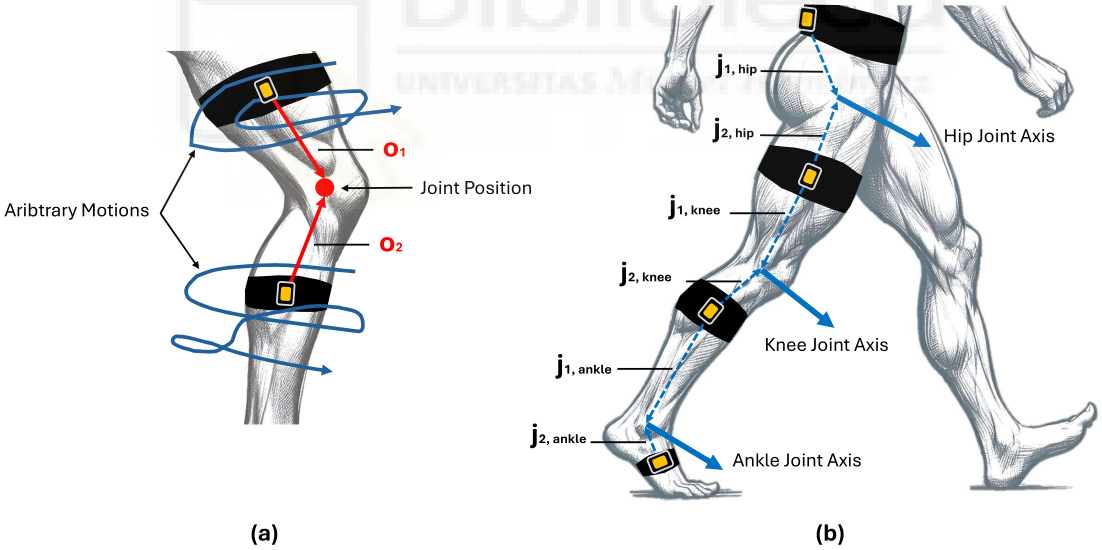


Figure 3.6: The placement of inertial sensors on the human body and calibration movements used to measure the angle of the lower limb joints. **(a)** Arbitrary movements are performed to estimate the coordinates for the hip, knee, and ankle joint centers position o_1, o_2 in the local coordinates of the IMUs. **(b)** Data collected during gait are used to determine the unit direction vectors j_1, j_2 corresponding to the flexion/extension axes of the hip, knee, and ankle.

3.2.2 *Flexion/extension angle measurement*

After successfully determining the coordinates of the joint axes (j_1, j_2) and the joint position coordinates (o_1, o_2), the angle of the lower limb joints can be measured from accelerations and angular rate data from the placed IMUs. By employing the data from the gyroscopes the flexion/extension angles can be calculated by integrating the difference of the angular rates around the joint axis:

$$\alpha_{gyr}(t) = \int_0^t (g_1(\tau) \cdot j_1 - g_2(\tau) \cdot j_2) d\tau \tag{3.6}$$

Furthermore, the knowledge of the joint axis coordinates allows us to employ many of the restrictive methods from the literature which require the sensor axes to coincide with joint axes or segment axes, as detailed by T. Seel et al. (Seel; Raisch, et al., 2014). In particular, an extension of the approach proposed in (K. Liu et al., 2009) has been extended to three-dimensional space, where the measured accelerations are shifted onto the joint axis by applying:

$$\tilde{a}_1 = a_1(t) - \Gamma_{g_1(t)}(o_1), \;\; \tilde{a}_2 = a_2(t) - \Gamma_{g_2(t)}(o_2) \tag{3.7}$$

The flexion/extension angles can be estimated through the angle between $\tilde{a_1}$ and $\tilde{a_2}$ projections into the joint plane. Subsequently, for each local frame a pair of joint plane axes $x_1, y_1, x_2, y_2 \in \mathbb{R}^3$ are defined:

$$x_1 = j_1 \times c, \;\; y_1 = j_1 \times x_1, \;\; x_2 = j_2 \times c, \;\; y_2 = j_2 \times x_2 \tag{3.8}$$

where $c \in \mathbb{R}^3$ represents any vector that ensures none of the products to be zero. Finally, the joint angle can be estimated through the accelerometer data:

$$\alpha_{acc}(t) = \angle \left(\begin{bmatrix} \tilde{a_1}(t) \cdot x_1 \\ \tilde{a_1}(t) \cdot y_1 \end{bmatrix}, \begin{bmatrix} \tilde{a_2}(t) \cdot x_2 \\ \tilde{a_2}(t) \cdot y_2 \end{bmatrix} \right) \tag{3.9}$$

The resulting angle α_{acc} is not affected by drift as α_{gyr} since any integration has been used to compute the angle. However, it is worth noting that α_{acc} is affected by the accelerometer noise. Furthermore, α_{gyr} is accurate in short time scales despite the drift. Therefore, it could be appropriate to combine both obtained angles using a standard

tool of sensor fusion. In the proposed system, a complementary filter (Young, 2009) has been used to fuse α_{acc} and α_{gyr}, and it is defined as:

$$\alpha(t) = \lambda \cdot \alpha_{acc}(t) + (1 - \lambda) \cdot (\alpha(t - \Delta t) + \alpha_{gyr}(t) - \alpha_{gyr}(t - \Delta t)) \tag{3.10}$$

By using the filter defined above, the flexion/extension angles are estimated with a sample period of $\Delta t = 0.016$ (60 Hz) and $\lambda = 0.02$.

Finally, it is worth mentioning that the conducted study by T. Seel et al. and studies presented by other authors (Kumar et al., 2018) have shown that this method presents a Root Mean Squared Error (RMSE) between 3 to 5 degrees for the hip, knee, and ankle flexion/extension movements. Consequently, it may be assumed that the accuracy of this method is sufficient to conduct a satisfactory gait analysis.

3.3 FORCES MODELING DURING UNDERWATER GAIT

As explained in section 3.1, understanding and modeling the forces exerted on the human body during gait to design an active lower limb exoskeleton has been investigated in the SPLASH project context. With this purpose, the motion capture system implemented in section 3.1 can be used to model the hydrodynamic forces.

To facilitate the analysis and ensure that the suggested method is applicable for designing and controlling an underwater exoskeleton, different reference systems have been established in the lower extremities. This approach allows us to model the forces applied during gait within the local reference systems of each segment, which can then be transformed into a global reference frame.

In Figure 3.7 each of the reference systems has been represented. First, the global reference system G has been defined at the hip joint, with a positive forward direction in the X-axis. In addition, three reference systems have been placed at the hip H, knee K, and ankle A joints following the Denavit-Hartenberg (DH) rules (Corke, 2007). Table 3.1 includes the DH parameters that solve the forward kinematics of the chain, where $L_i, i \in \{1, 2, 3\}$ are the longitudes of the user-measured thighs, shanks, and feet.

The human lower limbs have been considered rigid bodies in the sagittal plane, composed of three segments (the thigh, the shank, and the foot), and three net forces have been taken into account (buoyancy force, drag force, and vGRF). In addition, it should be noted that aquatic therapy is usually performed at chest level (Morer et al.,

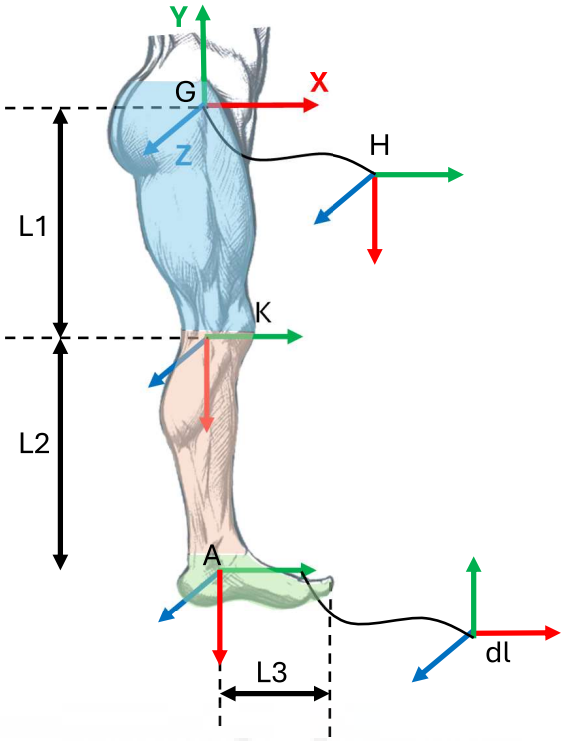


Figure 3.7: System references placed on the lower limb legs following the Denavit-Hartenberg convention.

Table 3.1: Denavit-Hartenberg parameters to solve the forward kinematics of the leg kinematics chain in the sagittal plane.

	θ	**d**	**a**	α
G-H	-90°	0	0	0
H-K	θ_{hip}	0	L1	0
K-A	θ_{knee}	0	L2	0
A-dl	θ_{ankle} + 90°	0	L3	0

2020; Veldema and Jansen, 2021). Hence, apart from the lower limb segments, the abdominopelvic volume has also been taken into account to estimate the buoyancy and vGRF when walking underwater.

3.3.1 *Buoyancy force*

The buoyancy force is an upward force exerted by a fluid that opposes the weight of an immersed body, which can be defined in the system reference G as:

$$\vec{F}_b = \rho_w \cdot g \cdot V_i \cdot \hat{y} \tag{3.11}$$

where ρ_w is the fluid density, g is the gravitational constant, V_i is the body segment volume, and $\hat{y}$ is the unitary vector in the Y-axis direction.

The estimation of $\vec{F}_b$ needs to account for the dimensions of each body segment and the corresponding volume. There exists a proportional relationship between the size of the human body segments, where anthropometric measurements can be regarded as an essential parameter appropriate to determine the volume of each segment (Drillis et al., 1964; Clauser et al., 1969; Xia et al., 2014). Concerning the study conducted by J. Sendroy and H. A. Collison the human body volume can be estimated according to height (H), weight (W), and body surface area (BSA) for male and female subjects as:

$$V_{male} = BSA \cdot 60.20 \cdot (W/H)^{0.562} \tag{3.12}$$

$$V_{female} = BSA \cdot 62.90 \cdot (W/H)^{0.578} \tag{3.13}$$

where V is the total body volume in l, H is the height in cm, W is the weight in kg, and BSA is the body surface area in m^2. The DuBois formula can be used to estimate the body surface area (DuBois, 1916) according to the height and weight:

$$BSA = 0.007184 \cdot W^{0.425} \cdot H^{0.725} \tag{3.14}$$

Once the total body volume is estimated, the volume for the thigh, the shank, and the foot can be estimated with the relative volumes defined by C. Clauser et al. (Clauser et al., 1969). In addition, as the authors obtained the relative volume of the total trunk, the abdominopelvic and the thorax volumes have been extracted from the study conducted by W. Dempster et al. (Dempster and Gaughran, 1967). Apart from the segment body volumes, COM positions must be estimated to determine the F_b application point. In Table 3.2 the relative volumes of each segment to the total body volumes and the relative COM positions to the total segment lengths are collected.

Table 3.2: Volume ratios and relative COM distance to the proximal joint of each segment.

	Volume Ratio	COM Distance Ratio
Abdominopelvic (Dempster and Gaughran, 1967)	26.3%	-
Thigh (Clauser et al., 1969)	10.3%	37.2%
Shank (Clauser et al., 1969)	4.3%	37.1%
Foot (Clauser et al., 1969)	1.5%	44.9%

After the COM pose of each segment is estimated in i-th the local reference system (${}^{L,i}T_{COM,i}$), a transformation can be applied to determine the pose of each COM in the global reference system:

$$ {}^{G}T_{COM,i} = {}^{G}T_{L,i} \cdot {}^{L,i}T_{COM,i} \tag{3.15} $$

3.3.2 *Drag forces*

Drag force F_D refers to the resistance offered by a fluid to the motion of an object moving through it and can be described as:

$$ \vec{F_D} = -\frac{C_D \cdot \rho_w \cdot A_\perp \cdot v^2}{2} \hat{v} \tag{3.16} $$

where C_D is the drag coefficient, ρ_w is the density of the water, $A_\perp$ is the projection of the area perpendicular to the velocity of the body, v is the body velocity, and $\hat{v}$ is a unit vector that indicates the direction velocity.

Each segment can be divided into an infinitesimal length dl for the analysis of drag forces. Thus, applied to our problem, the drag force can also be defined as:

$$ \vec{F_D} = -\int_0^l \frac{C_D \cdot \rho_w}{2} \cdot v^2 \cdot \varnothing \cdot dl_\perp \vec{v_{dl}} \tag{3.17} $$

where $\varnothing$ is the segment diameter and $dl_\perp \vec{v_{dl}}$ is the differential length segment perpendicular to its velocity.

Each of the lower limb dl will have a certain $\vec{v_{dl}}$, which is assumed to be due to two causes:

- **User's gait velocity ($\vec{v_g}$).** The subject imparts a certain velocity $\vec{v_g}$ to each lower limb segment during underwater gait. This velocity will be assumed equal for the entire human body.
- **Angular velocity of the joints (ω_i).** During gait, each lower limb joint has a certain position θ_i and angular velocity ω_i. This rotation will cause the different dl to have a given tangential velocity $\vec{v}_t$, depending on the distance to the rotation axis of the joint.

Therefore, the estimation of both $\vec{v}_t$ and the forward velocity of the segment $\vec{v_g}$ are required to estimate the velocity $\vec{v}_{dl}$ at each dl.

Using the IMU-based system developed to measure the joint angles of the lower limbs, $\vec{v_g}$ can be estimated in the sagittal plane from the device placed over the pelvis such that:

$$\vec{v}_x = K_x + \int_0^t \vec{a}_x \tag{3.18}$$

$$\vec{v}_y = K_y + \int_0^t \vec{a}_y \tag{3.19}$$

$$\vec{v}_g = v_x \cdot \hat{x} + v_y \cdot \hat{y} \tag{3.20}$$

Generally, $\vec{v}_{dl}$ can be expressed for the thigh, shank, and foot segments as:

$$\vec{v}_{dl,thigh} = \vec{v_g} + \vec{v}_{t,hip} \tag{3.21}$$

$$\vec{v}_{dl,shank} = \vec{v_g} + \vec{v}_{t,hip} + \vec{v}_{t,knee} \tag{3.22}$$

$$\vec{v}_{dl,foot} = \vec{v_g} + \vec{v}_{t,hip} + \vec{v}_{t,knee} + \vec{v}_{t,ankle} \tag{3.23}$$

where $\vec{v}_{t,hip}$, $\vec{v}_{t,knee}$, and $\vec{v}_{t,ankle}$ are the tangential velocities due to the hip, knee, and ankle rotations.

Note that the velocities must all be expressed in the same reference frame. Therefore, the following transformations must be applied to calculate each $\vec{v}_t$ in G:

$$^{G}\vec{v}_{t,hip} = {}^{G}T_{H} \cdot ({}^{H}\vec{\omega}_{hip} \times {}^{H}\vec{r}_{hip}) \tag{3.24}$$

$$^{G}\vec{v}_{t,knee} = {}^{G}T_{K} \cdot ({}^{K}\vec{\omega}_{knee} \times {}^{K}\vec{r}_{knee}) \tag{3.25}$$

$$^{G}\vec{v}_{t,ankle} = {}^{G}T_{A} \cdot ({}^{A}\vec{\omega}_{ankle} \times {}^{A}\vec{r}_{ankle}) \tag{3.26}$$

Once the velocity $\vec{v}_{dl}$ is estimated, the unit vector $\hat{v}_{dl}$ can be found by:

$$\hat{v}_{dl} = \frac{\vec{v}_{dl}}{||\vec{v}_{dl}||} \tag{3.27}$$

The projection of the area perpendicular to the velocity vector $dA_{\perp}$ is required to calculate the applied drag force at each dl. This projection can be deduced through trigonometric methods, which are illustrated in Figure 3.8. First, the direction of $\vec{dl}$ needs to be known. This direction will be given in the form of a unit vector in the local coordinate system, and will subsequently be transformed into G. Regarding the segment, and according the DH reference frames, $\vec{dl}$ can be expressed in the global reference system as:

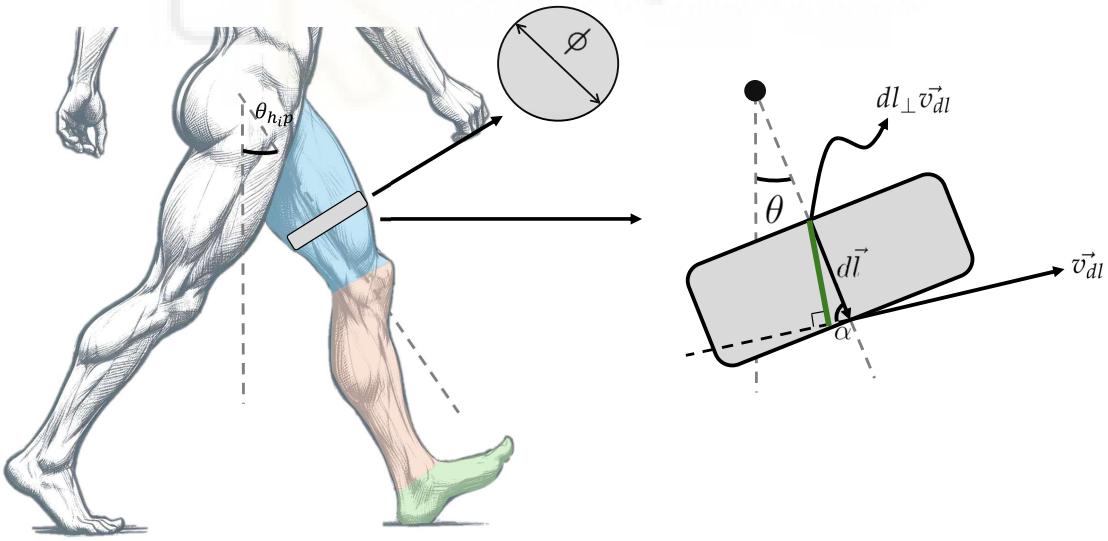


Figure 3.8: Representation of trigonometric methods used to calculate $dl_{\perp}\hat{v_{dl}}$.

$$\hat{dl}_{thigh} = {}^{G}T_{dl} \cdot \hat{x} \tag{3.28}$$

$$\hat{dl}_{shank} = {}^{G}T_{dl} \cdot \hat{x} \tag{3.29}$$

$$\hat{dl}_{foot} = {}^{G}T_{dl} \cdot \hat{y} \tag{3.30}$$

After obtaining $\hat{dl}$, the angle α between vectors $\hat{dl}$ and $\vec{v}_{dl}$ can be calculated as:

$$\alpha = cos^{-1}\left(\frac{\hat{dl} \cdot \hat{v}_{dl}}{||\hat{dl}|| \cdot ||\hat{v}_{dl}||}\right) \tag{3.31}$$

Finally, $dA_{\perp}$ can be calculated by $\varnothing$ and $\vec{dl}$:

$$dA_{\perp} = \varnothing \cdot ||\vec{dl}|| \cdot sin(\alpha) \tag{3.32}$$

3.3.3 *Vertical ground reaction force*

As introduced, the vGRF is the greatest magnitude of the GRF during gait. Hence, a force plate to measure this force or a method to estimate the vGRF is needed to perform an adequate underwater gait analysis. Due to the difficulties that can be caused by using an underwater force plate, in this thesis a Machine Learning model is proposed to estimate the vGRF.

The Machine Learning algorithm will use as inputs the kinematics of the lower limb joints and the vertical acceleration measured at the *C7* vertebra, which training and evaluation can be found in Chapter 4. Data from a force plate embedded in a treadmill and pressure insoles have been used to train the model, which results in:

$$\widehat{vGRF} = f(x_{kin}) \tag{3.33}$$

where $\widehat{vGRF}$ is the output of the model (i.e. the estimated vGRF) and x_{kin} is a vector that contains data from the lower limb kinematics and the user's vertical acceleration.

It should be noted that the model has been trained without immersion. Hence, the output $\widehat{vGRF}$ is normalized to the user's Body Weight (BW) to be later adapted to the aquatic environment such that:

$$\widehat{vGRF_w} = \widehat{vGRF} \cdot (F_{BW} - F_B) \cdot \hat{y} \tag{3.34}$$

where $\widehat{vGRF_w}$ is the estimated vGRF during immersion expressed in G coordinates, F_{BW} is the user's BW, and F_B is the total buoyancy force exerted on the user.

3.4 CONCLUSION

In this Chapter, the SPLASH project proposal has been presented, which intends to address the limitations of robot-aided therapies caused by spasticity, pain, or joint stiffness. As described in section 3.1, the SPLASH project aims to develop an aquatic rehabilitation environment by understanding and modeling the effects of water immersion, and developing aquatic rehabilitation devices together with intelligent control systems to adapt the assistance provided by the robotic devices to the user's needs.

In this thesis, to study the effects of water immersion in human gait, an IMU-based motion capture system has been developed to measure the lower limb flexion/extension angles (section 3.2). Based on the measurement of the lower limb kinematics, a model describing the buoyancy and drag forces during gait is described in section 3.3. Moreover, a Machine Learning model has been developed to overcome the difficulties encountered in measuring the vGRF in an aquatic environment (Chapter 4). In order to generalize the aquatic therapies to diverse gait ADLs, a classifier has been developed to detect a set of activities based on the trajectories of the lower limb angles (Chapter 5). Furthermore, two different approaches are presented to develop an intelligent auto-adaptive control system for upper limb rehabilitation. First, a DDA system has been developed to adapt the difficulty of serious games to the patient's abilities (Chapter 6). Finally, a Deep Learning model has been developed to identify when assistance is needed by learning from a therapist's interventions (Chapter 7).

4

VERTICAL GROUND REACTION FORCE ESTIMATION DURING GAIT FROM LOWER LIMB KINEMATICS AND VERTICAL ACCELERATION USING WEARABLE INERTIAL SENSORS

4.1 INTRODUCTION

As explained, there is a need to estimate the vGRF to carry out an appropriate gait analysis due to the difficulties presented by employing an underwater force plate, since it represents an important part of the pattern and magnitude of the mechanical loading (Jacobs et al., 1972; Zadpoor and Nikooyan, 2012). It is noteworthy that three main peaks can be distinguished in the vGRF during gait (Marasović et al., 2009). First, the Loading Peak (LP) is produced, a local maximum corresponding to the loading response. Second, a local minimum appears, corresponding to the Mid Stance Peak (MP). Finally, a local maximum is produced during the terminal stance, named as Terminal Stance Peak (TP). As the magnitude and timing of these peaks influence the muscle and joint loads, the vGRF estimated signal must retain the highest accuracy at these peaks.

With this purpose, different Machine Learning algorithms have been proposed and analyzed to estimate the vGRF during gait. The IMU-based method presented in Chapter 3, section 3.2 has been used to measure the kinematics of the lower limbs, which data are used together with the vertical acceleration as inputs of the model. In this context, a study has been conducted first to estimate the vGRF during ground-level walking, with special emphasis on the importance of each of the inputs and the estimation of the peaks. In addition, as it is intended to study the biomechanical effects during different activities (Chapter 3, section 3.1), the estimation of the vGRF must be evaluated and analyzed whether an adaptation of the algorithms is required.

4.2 VERTICAL GROUND REACTION FORCE ESTIMATION DURING GROUND-LEVEL WALKING

4.2.1 *Materials and Methods*

4.2.1.1 *Experimental sessions*

In this section, the participants involved, the experimental setup, and the protocol followed are detailed.

Twelve volunteer non-disabled users participated in the experimental sessions, ten male and two female. The ages were between 23 and 52 years old (29.8 $\pm$ 7.4), with heights ranging from 1.65m to 1.87m (176.2 $\pm$ 7.4cm) and weights from 56.1kg to 90.2kg (76 $\pm$ 12.5kg). Written informed consent was obtained from the individuals before participating in the sessions, and the study was approved by the INESCOP Footwear Technology Center.

The devices involved in the experimental sessions are shown in Figure 4.1. The IMU-based motion capture system described in Chapter 3, section 3.2 has been used to acquire lower limb kinematics data at 60 Hz. An additional IMU was placed at the C7 vertebra to measure the vertical acceleration of the users, as proposed in (Shahabpoor

Figure 4.1: In the left image, the placement of the IMUs is represented, which are utilized to measure the vertical acceleration (C7 IMU) and the joint angles of the lower limb (hip, thigh, leg, and foot). The right image shows the treadmill h/p/cosmos 150/50, which is employed for capturing the vGRF during gait.

and Pavic, 2018). In addition, a treadmill model h/p/cosmos 150/50 with an embedded force plate was used to measure the vGRF at 100 Hz.

Before starting the experimental session, IMUs were attached to the user's back hip, thigh, leg, and foot, and their height and weight were measured. The calibration data was then collected to determine the joint angles of the lower limb. Initially, the user was asked to perform arbitrary circling movements with the hip, knee, and ankle for one minute to obtain the center position coordinates (o_1, o_2) of each joint. After that, the user walked on a treadmill at a comfortable speed for another minute to obtain the flexion/extension joint axis coordinates (j_1, j_2).

Once the calibration process was completed, the participants walked on the treadmill at four different speeds, ranging from 1.5 km/h to 4.5 km/h, to cover a wide range of normal gait speeds. For each speed, kinematics and vGRF data were recorded for 5 minutes. There was a 2-minute break between each activity for rest.

4.2.1.2 *Acquired data and processing*

In this section, the data acquisition process from each device and the subsequent signal processing are described. The data processing pipeline is represented in Figure 4.2.

The vGRF was captured using the plate-instrumented treadmill, which has been normalized relative to the BW of each participant. Data from IMUs located at the hip, thigh, shank, and foot were used to acquire accelerometer and gyroscope measurements to estimate the lower limb joint angles in the sagittal plane. Following the computation of the hip, knee, and ankle angles, a forward-backward low-pass filter was applied (Gustafsson, 1996) with a 6 Hz cutoff frequency and the angular velocities for each joint were calculated.

In addition, the acceleration and orientation of the IMU placed over the C7 vertebra were acquired to estimate the vertical acceleration of the user (Esser et al., 2009). First, the acceleration vector from the IMU system was transformed into the global system by using the orientation. Finally, the gravity from the measured vertical acceleration was removed by subtracting the Z-axis acceleration.

It is worth mentioning that identifying the onset and the end of the gait cycle is commonly used to transform the data from the temporal domain to the gait cycle domain (0-100%) when performing a gait analysis. Hence, all signals have been

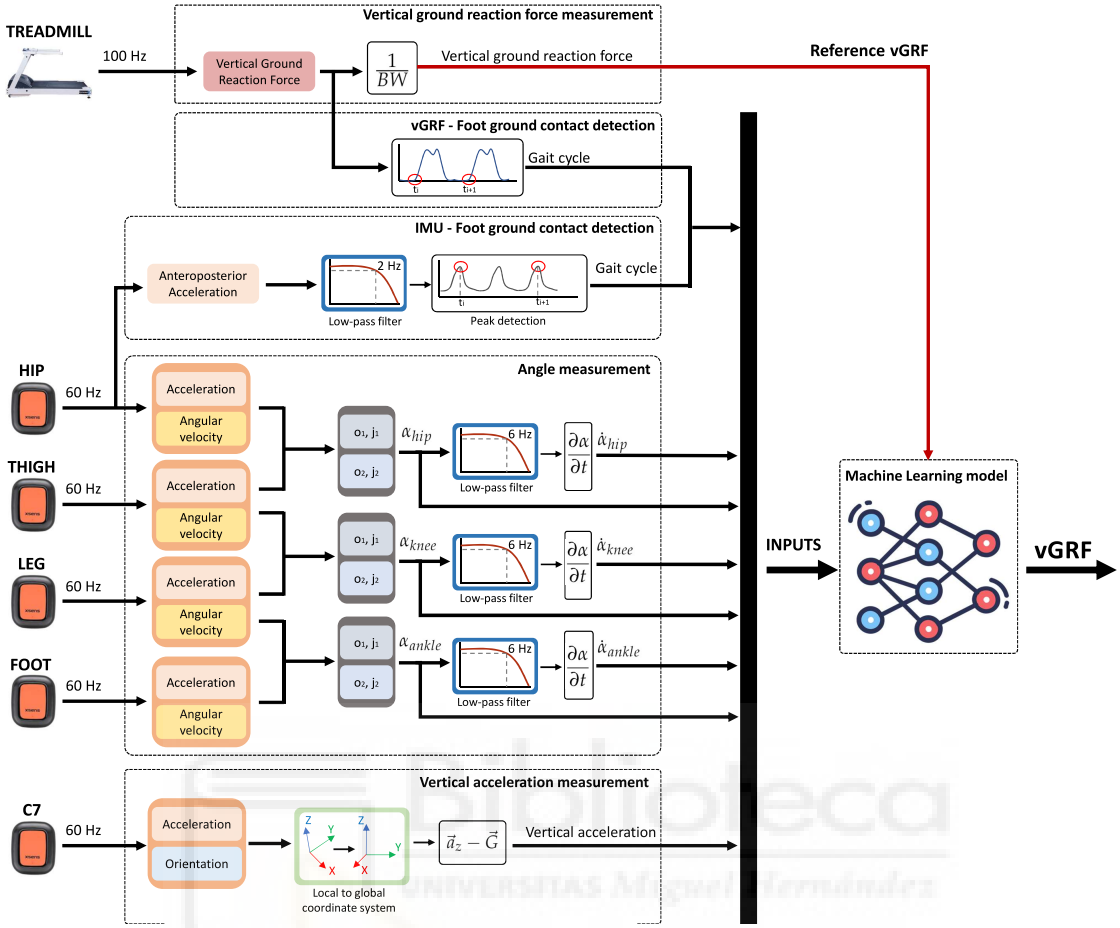


Figure 4.2: A treadmill and diverse IMUs were used to collect data to estimate the vGRF. The IMUs were placed on the lower limbs to measure acceleration and angular velocity, which allowed us to estimate the flexion/extension angles using the motion capture system described in Chapter 3, section 3.2. In addition, an IMU placed over the C7 vertebra was employed to measure the vertical acceleration of the users. The foot-ground contact was detected using the anteroposterior acceleration measured at the hip to compute the gait cycle. The vGRF was measured during gait and normalized by the body weight of each user. To detect the ground-foot contact in the force signal, the end of the stance phase (null force) was detected. The measured vGRF and lower limb kinematics data were synced by using the gait cycle. The flexion/extension joint angles and velocities, the vertical acceleration, and the gait cycle have been used as inputs of the proposed models, and the measured vGRF have been used as the ground truth values for model learning.

transformed into the gait cycle domain. With this purpose, the foot-ground contact was detected by employing two different methods. First, the foot-ground contact was identified in the IMU-based motion capture system by using the hip acceleration

(Zijlstra and Hof, 2003). A forward-backward low-pass filter with a 2 Hz cutoff was applied, and local maxima were used for ground-contact detection, which correspond to the left and right ground foot contacts. The right foot contacts were used to calculate the gait cycle as a matter of convention. Additionally, foot-ground contact was detected in the vGRF signal when the force value started increasing.

Signal synchronization starts by detecting the initial ground-foot contact in the vGRF and the IMU signals. Additionally, by using the foot-ground contact detection in both systems, the beginning and the end of each step can be determined. For each step, the vGRF and the lower limb kinematics data are labeled according to the gait cycle percentage in which it occurs. Finally, to train the Machine Learning models, the vGRF and leg kinematics data are synchronized from the gait cycle percentage in which they occur, ensuring that each lower limb kinematics data is assigned a vGRF value.

Figure 4.3 represents the flexion/extension angles, the vertical acceleration, and the vGRF acquired for each gait speed.

4.2.1.3 *Selected Machine Learning algorithms*

The vGRF can be estimated by using a regression model. In this study, a RF and a Feedforward Neural Network (FNN) have been trained since their use has been previously evaluated by other authors, but their performance has not been compared (see Chapter 2, section 2.2).

RF models are composed of an ensemble of decision trees (Zhang and Ma, 2012). Each tree is independently trained on a different subset of the data, and the final output of the RF model is based on the most effective outcome produced by the trees. The strength of RF models relies on their ability to make accurate predictions and to achieve good generalization across different datasets. This is largely due to the bagging approach they employ, which reduces variance and enhances the robustness of the model. The SciKit Learn Python Library (Pedregosa et al., 2011) was used to train the RF models. The training process involved the creation of 170 decision trees, each with a maximum depth of 35. The trees were configured to require at least one sample at the leaf nodes and a minimum of two samples to split an internal node. For determining the best split at each node, the models considered a number of inputs equal to the base-2 logarithm of the total number of inputs.

FNN models refer to a type of artificial neural network consisting of neurons that take in an input signal, which is processed by an activation function to generate an

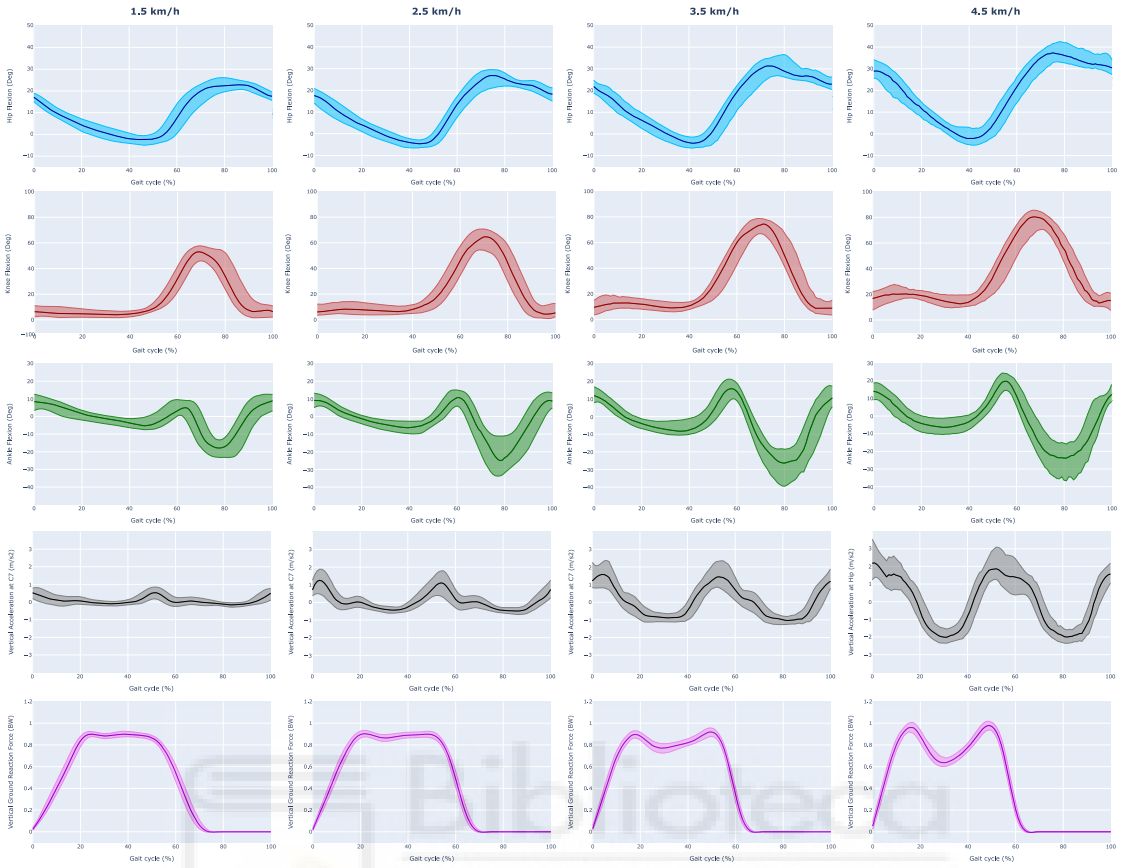


Figure 4.3: The flexion/extension angles of the hip, knee, and ankle are represented in the first, second, and third rows, respectively. The measured vertical acceleration at the C7 vertebra is shown in the fourth row, and the measured vGRF is presented in the fifth row. These data are represented across different gait speeds ranging from 1.5 to 4.5 km/h and they are standardized based on the gait cycle.

output signal (Svozil et al., 1997; Sharma et al., 2017). These neurons are organized into layers, with the first layer called the input layer, the last layer known as the output layer, and the layers between them known as hidden layers. Each neuron in a specific layer is connected to all the neurons in the next layer through weight coefficients. The optimization of the neural network's weights is typically performed via back-propagation, which utilizes a non-linear optimization technique like the gradient descent algorithm (Andrychowicz et al., 2016). The Keras Python Library (Gulli and Pal, 2017a) has been used to train the FNN models. The architecture of the FNN models trained is composed of an input layer with a number of neurons equal to the number of inputs, five hidden layers with ten neurons, and an output layer with

one neuron. The Rectified Linear Unit (ReLu) has been used as the activation function, the Adam algorithm has been used to optimize the network weights (Kingma and Ba, 2014), and the RMSE serves as the loss function. To mitigate the risk of overfitting, the dropout regularization technique is applied (Srivastava et al., 2014). Additionally, in line with best practices for using dropout, a weight constraint is enforced on each hidden layer, ensuring that the weight norms do not surpass a threshold of four.

It must be noted that the RF and FNN hyperparameters have been tuned to achieve the best performance based on the RMSE with the training data.

4.2.1.4 *Preprocessing and training*

To train the RF and FNN models the flexion/extension angles and the angular velocities of the hip, knee, and ankle, the vertical acceleration measured at the C7 vertebra, and the instant of the gait cycle were introduced as inputs. Additionally, the measured vGRF with the treadmill was used as the ground truth value.

Data collected from the participants was divided into three different datasets to evaluate the performance of the RF and FNN models. Data from ten users have been randomly divided to evaluate the accuracy of the models by cross-validation: 80% of the data was used for training the models and the remaining 20% was used for validation (intra-participant data). Additionally, the performance of the models has been evaluated on users who were not part of the learning process (inter-participant data) using data from two additional participants.

Before training the models, the training dataset was used to build a scaler, which normalized the input values within the range of 0 to 1. This scaler has also been employed to normalize the validation and inter-participants dataset to evaluate the accuracy of the regression models.

As previously introduced, one of the primary objectives of this study is to investigate the impact of vertical acceleration and lower limb kinematics when estimating the vGRF. To analyze these contributions, two different training approaches have been conducted for each model, using the following inputs:

- **Training 1 (FNN-Kinematics, RF-Kinematics).** Joint angles, angular velocities, and gait cycle.
- **Training 2 (FNN-C7, RF-C7).** Joint angles, angular velocities, vertical acceleration, and gait cycle.

4.2.1.5 *Performance evaluation of the trained models*

The performance of the RF and FNN models has been evaluated as follows:

- vGRF RMSE. A lower RMSE in the estimation of the vGRF signal has been considered as a better performance of the models.
- vGRF Normalized Root Mean Squared Error (NRMSE). The NRMSE is defined as (Shahabpoor and Pavic, 2018):

$$NRMSE = \frac{RMSE}{maxvGRF_{measured} - minvGRF_{measured}} \tag{4.1}$$

  A lower NRMSE in the estimation of the vGRF signal has been considered as a better performance of the model.
- Correlation. A higher correlation (ρ) between the ground truth and the estimated vGRF signal has been considered as a better performance of the models. The Person's correlation coefficient has been used to calculate the correlation between estimated and ground truth data.
- LP, MP, and TP RMSE. A lower RMSE between ground truth and estimated vGRF peaks has been considered as a better performance of the models. To identify the LP and TP in gait analysis, the local maxima during the first and second periods of foot support were detected. For determining the MP, it was specifically looked for the local minimum value situated between the LP and TP peaks.
- LP, MP, and TP delay. When the delay or advance between ground truth and estimated vGRF peaks is close to zero, it has been considered as a better model performance. The delay is calculated as the difference between the estimated and measured peaks. Specifically, a positive value means that the estimated peak lags behind the measured peak, and a negative value indicates that the estimated peak occurs earlier than the measured peak.

4.2.1.6 *Statistical data analysis*

Statistical data analysis was conducted to evaluate the accuracy of the models in estimating vGRF with intra and inter-participant datasets. Initially, a comparative of the four trained models (FNN-Kinematics, FNN-C7, RF-Kinematics, and RF-C7) was

performed in terms of RMSE. Subsequently, the FNN and RF models with a lower RMSE were selected, and a comparison of vGRF estimations and peak reproduction accuracy was carried out using both the intra and inter-participant datasets.

In the statistical analysis, a normality test was performed using the Kolmogorov-Smirnov test (Berger and Y. Zhou, 2014). The results of the test showed that data distribution was not normal ($p > 0.05$).

The Friedman test was used to study the differences between the accuracy of the models (Sheldon et al., 1996). In the post-hoc analysis, pairwise comparisons have been studied by the Wilcoxon signed-rank test with the zero-method proposed by Pratt (Pratt, 1959).

4.2.2 *Results*

4.2.2.1 *Feature importance*

The impact of the angles, the angular velocities, the vertical acceleration, and the gait cycle on the training of the RF and FNN models have been studied by applying the method known as feature permutation (Altmann et al., 2010). Feature permutation is a technique that quantifies the increase in the prediction error when an input feature is randomly shuffled. Hence, the dependency of the vGRF prediction on each input can be quantified. The feature permutation results are represented in Figure 4.4, which shows the increase of the RMSE in the vGRF estimation when a feature is shuffled with the FNN-C7 model and the RF-C7 model.

The results of the feature permutation analysis suggest that the gait cycle has the most substantial impact on the performance of the model, as shown by an increment of 0.2590 BW in the RMSE for the FNN-C7 model and 0.1927 BW in the RF-C7 model. Concerning the kinematics of the lower limb joints, the knee and ankle angles can be identified as the most critical inputs influencing the FNN-C7 model, followed by the angular velocities of the hip and knee. By contrast, the training of the RF-C7 model demonstrates a higher importance of the hip, knee, and ankle angles, surpassing their respective angular velocities. In both models, the feature permutation results suggest that the ankle angular velocity has a low significance compared to the kinematics of the hip and knee.

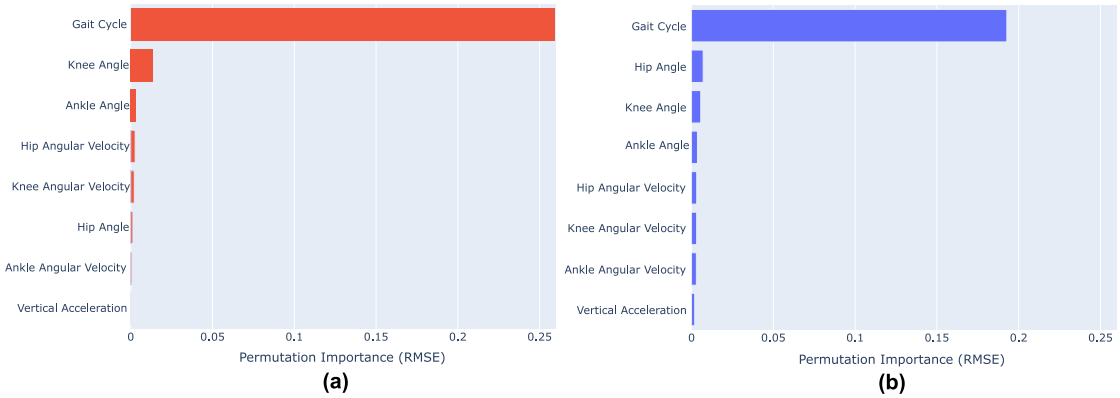


Figure 4.4: Feature importance calculated with feature permutation for the **(a)** FNN-C7 model and **(b)** RF-C7 model

It is worth mentioning that feature permutation results suggest that vertical acceleration has a low impact on the training of the models compared to the lower limb kinematics. In the FNN-C7 model, the vertical acceleration has the second lowest importance, and in the RF-C7 model, it has the lowest relevance.

4.2.2.2 *Validation and test results*

The performance of the RF and FNN models has been evaluated with the intra and inter-participant data by calculating the RMSE for each of the recorded steps. The mean and standard deviations of the RMSE results are represented in Figure 4.5, where Figure 4.5.a show the RMSE with the intra-participants, and 4.5.b the results with the inter-participants. Furthermore, Table 4.1 and Table 4.2 collect the intra and inter-participant RMSE results.

The statistical analysis revealed significant differences in the performance of the models when evaluated with validation data across different gait speeds (Friedman test $p < 0.0001$ for all speeds). In pairwise comparisons, no significant difference was observed in model accuracy when users' vertical acceleration was included as an input for each gait speed ($p > 0.05$). However, the inclusion of vertical acceleration led to a significant reduction in global RMSE when using the FNN models ($p < 0.05$). Comparing the FNN and RF performance, a significant reduction in RMSE is observed when employing the RF models with intra-participant data for global RMSE across all gait speeds. Although there was no statistical significance, a reduction of the RMSE errors in the RF training when the vertical acceleration is included can be observed,

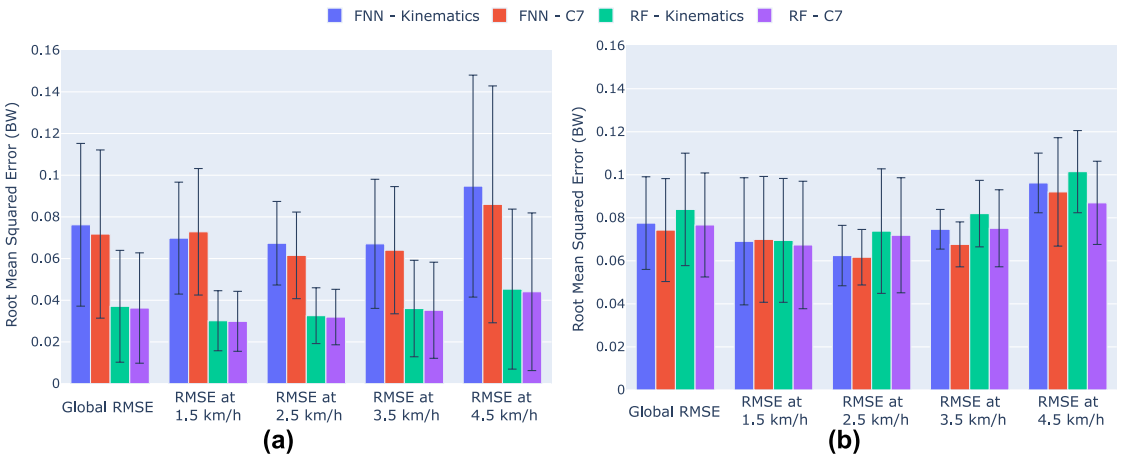


Figure 4.5: Mean RMSE and standard deviation for each of the FNN and RF models trained. Different gait speeds and the global RMSE are represented when the models are evaluated with the **(a)** intra-participants and **(b)** inter-participants.

particularly at gait speeds of 3.5 and 4.5 km/h. This can also be noticed when using the FNN models. However, when employing a FNN model, the inclusion of the vertical acceleration negatively impacted performance at 1.5 km/h.

The analysis of RMSE obtained with inter-participant data shows that there is no significant difference in the accuracy of the models. The Friedman test for global RMSE yields a $p = 0.24$, and at speeds of 1.5 km/h, 2.5 km/h, 3.5 km/h, and 4.5 km/h the p-values are $p = 0.80$, $p = 0.80$, $p = 0.17$, and $p = 0.24$, respectively. However, including vertical acceleration leads to a reduction in errors in both models. Additionally, the FNN and RF models produce similar RMSE results.

The results suggest that both the FNN-C7 and RF-C7 models have higher accuracy in terms of RMSE. Consequently, both models are deeply analyzed in the next section.

4.2.2.3 *Validation and test vGRF signal reconstruction*

Figure 4.6 represents the measured vGRF, together with the vGRF predictions made by the FNN-C7 and RF-C7 models. The represented graphs include the median vGRF throughout the gait cycle, with the shaded regions indicating the range of vGRF from the first to the third quartile.

Table 4.3 collects the correlation (ρ) and NRMSE between the ground truth and predicted vGRF using the selected models with the intra and inter-participant data. In addition, Figure 4.7 presents heatmaps representing the average magnitude errors

Table 4.1: For each training of the FNN and RF models, the table shows the mean (standard deviation) RMSE (BW) and the global RMSE for each model with the intra-participant data. Additionally, the table lists the RMSE errors obtained for each gait speed.

	FNN - Kinematics (a)	**FNN - C7 (b)**	**RF - Kinematics (c)**	**RF - C7 (d)**
Global	0.076 (0.039) ^{***c,d}	0.072 (0.040) ^{***c,d}	0.037 (0.027) ^{***a,b,d}	0.036 (0.026) ^{***a,b,c}
1.5 km/h	0.070 (0.027) ^{***c,d}	0.073 (0.030) ^{***c,d}	0.030 (0.014) $^{***a,b\ *d,}$	0.030 (0.015) $^{***a,b\ *c,}$
2.5 km/h	0.067 (0.020) ^{***c,d}	0.062 (0.020) ^{***c,d}	0.033 (0.013) ^{***a,b,d}	0.032 (0.013) ^{***a,b,c}
3.5 km/h	0.067 (0.030) ^{***c,d}	0.064 (0.030) ^{***c,d}	0.036 (0.023) ^{***a,b,d}	0.035 (0.023) ^{***a,b,c}
4.5 km/h	0.095 (0.050) ^{***c,d}	0.086 (0.056) ^{***c,d}	0.045 (0.038) ^{***a,b,d}	0.044 (0.037) ^{***a,b,c}

Statistical differences are represented by * ($p <= 0.05$), ** ($p <= 0.01$), *** ($p <= 0.001$), and **** ($p <= 0.0001$).

Table 4.2: For each training of the FNN and RF models, the table shows the mean (standard deviation) RMSE (BW) and the global RMSE for each model with the inter-participant data. Additionally, the table lists the RMSE errors obtained for each gait speed.

	FNN - Kinematics	**FNN - C7**	**RF - Kinematics**	**RF - C7**
Global	0.077 (0.022)	0.074 (0.028)	0.083 (0.026)	0.077 (0.024)
1.5 km/h	0.069 (0.030)	0.070 (0.029)	0.070 (0.029)	0.067 (0.030)
2.5 km/h	0.062 (0.014)	0.062 (0.013)	0.074 (0.029)	0.072 (0.027)
3.5 km/h	0.075 (0.010)	0.068 (0.010)	0.082 (0.012)	0.075 (0.018)
4.5 km/h	0.096 (0.014)	0.092 (0.025)	0.102 (0.019)	0.087 (0.019)

No statistical differences were found.

and delays associated with the characteristic peaks for both intra-participant and inter-participants. Furthermore, the Appendix A provides detailed information on the magnitude error and delay of these peaks, together with pairwise comparison results between the chosen models (FNN-C7 and RF-C7).

The analysis of the intra-participant data shows that ρ values for the vGRF reconstruction ranged from 0.9676 to 0.9886 when using the FNN-C7 model, and exhibited coefficients from 0.9966 to 0.9988 with the RF-C7 model. Moreover, the NRMSE results range from 0.0399 BW to 0.0651 BW for the FNN-C7 model, and from 0.0127 BW to 0.0210 BW for the RF-C7 model. However, when predicting vGRF with the inter-participants, the correlation is reduced, with ρ values ranging between 0.9705 and 0.9876 for the FFN-C7 model, and between 0.9746 and 0.9849 for the RF-C7 model.

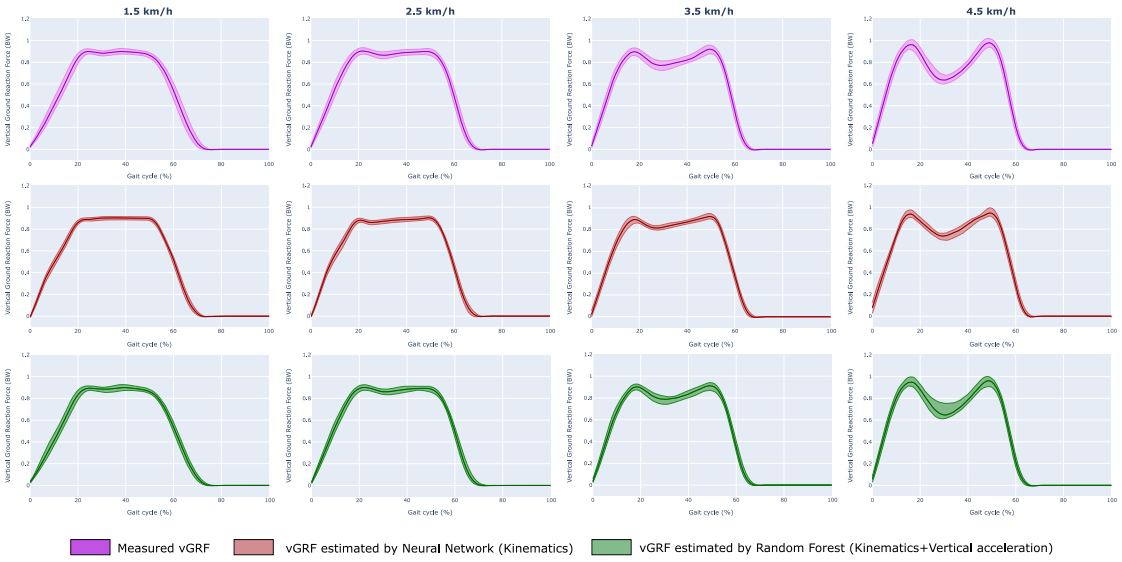


Figure 4.6: The ground truth and estimated vGRF (intra-participants and inter-participants) have been represented according to the gait cycle (%) for different gait speeds. The median values have been represented, and the values between the first and third quartiles have been colored. The first row shows the measured vGRF, the second row shows the estimated vGRF using the selected FNN model (Training 2, FNN-C7), and the third row shows the estimated vGRF using the selected RF model (Training 2, RF-C7). These vGRF representations include all participant users.

In terms of NRMSE, the values ranged from 0.0560 BW to 0.0853 BW with the FNN-C7 model, and from 0.0691 BW to 0.0792 BW with the RF-C7 model.

The analysis of the vGRF peaks shows statistically significant differences across all the characteristic peaks ($p < 0.0001$). As represented in Figure 4.7.a, the RF-C7 model consistently produced smaller peak magnitude RMSE compared to the FNN-C7 model. Specifically, the RF-C7 achieves magnitude errors lower than 0.0210 BW for the LP, 0.0161 BW for the MP, and 0.0171 BW for the TP. In contrast, the FNN-C7 model errors reached up to 0.0900 BW in the LP, 0.0793 BW for the MP, and 0.0491 BW for the TP. By contrast, in the evaluation of the models with the inter-participant data, the statistical analysis indicates differences in the prediction of the peak magnitude errors between models except for the MP at 3.5 km/h ($p = 0.1825$) and TP at 4.5 km/h ($p = 0.0834$). For the FNN-C7 model, the errors range from 0.0206 BW to 0.0619 BW in the LP, 0.0263 BW to 0.0810 BW for the MP, and 0.0249 BW to 0.0922 BW for the TP. By contrast, the RF-C7 model errors range from 0.0225 BW to 0.0611 BW for the LP, 0.0399 BW to 0.0983 BW for the MP, and 0.0316 BW to 0.1256 BW for the TP.

Table 4.3: Correlation (ρ) and NRMSE results obtained between the ground truth and estimated vGRF with the inter-participant data using the FNN-C7 and RF-C7 models.

	Model	Gait Speed (km/h)	ρ	NRMSE
Intra-participants	**FNN-C7**	1.5	0.9817	0.0483
		2.5	0.9886	0.0399
		3.5	0.9847	0.0449
		4.5	0.9676	0.0651
	RF-C7	1.5	0.9988	0.0118
		2.5	0.9987	0.0127
		3.5	0.9980	0.0155
		4.5	0.9966	0.0210
Inter-participants	**FNN-C7**	1.5	0.9777	0.0718
		2.5	0.9876	0.0560
		3.5	0.9859	0.0604
		4.5	0.9705	0.0853
	RF-C7	1.5	0.9776	0.0716
		2.5	0.9797	0.0708
		3.5	0.9797	0.0691
		4.5	0.9746	0.0792

The timing delay of the predicted vGRF peaks with the intra-participants shows that the lags achieved by the RF-C7 model are nearly 0%, with mean lags lower than 0.18% (advance in LP at 1.5 km/h). By contrast, the timing errors with the FNN-C7 range from -0.21% to 1.6% in the LP, from -0.3% to 3.4% in the MP, and -0.92% to -3% in the TP. The timing errors increased for both models with inter-participant data. The timing errors obtained with the FNN-C7 model range from -0.84% to -2.2% in the LP, from 1.3% to 4.1% in the MP, and -2.3% to 0.73% in the TP, while the RF-C7 model errors range from -2.3% to -4.9% in the LP, from -2.6% to 3% in the MP, and -2.4% to 5.4 in the TP estimation.

4.2.3 *Discussion*

The implemented motion capture system described in section 3.2 was used to estimate the joint angles. A significant advantage of this method relies on its independence from

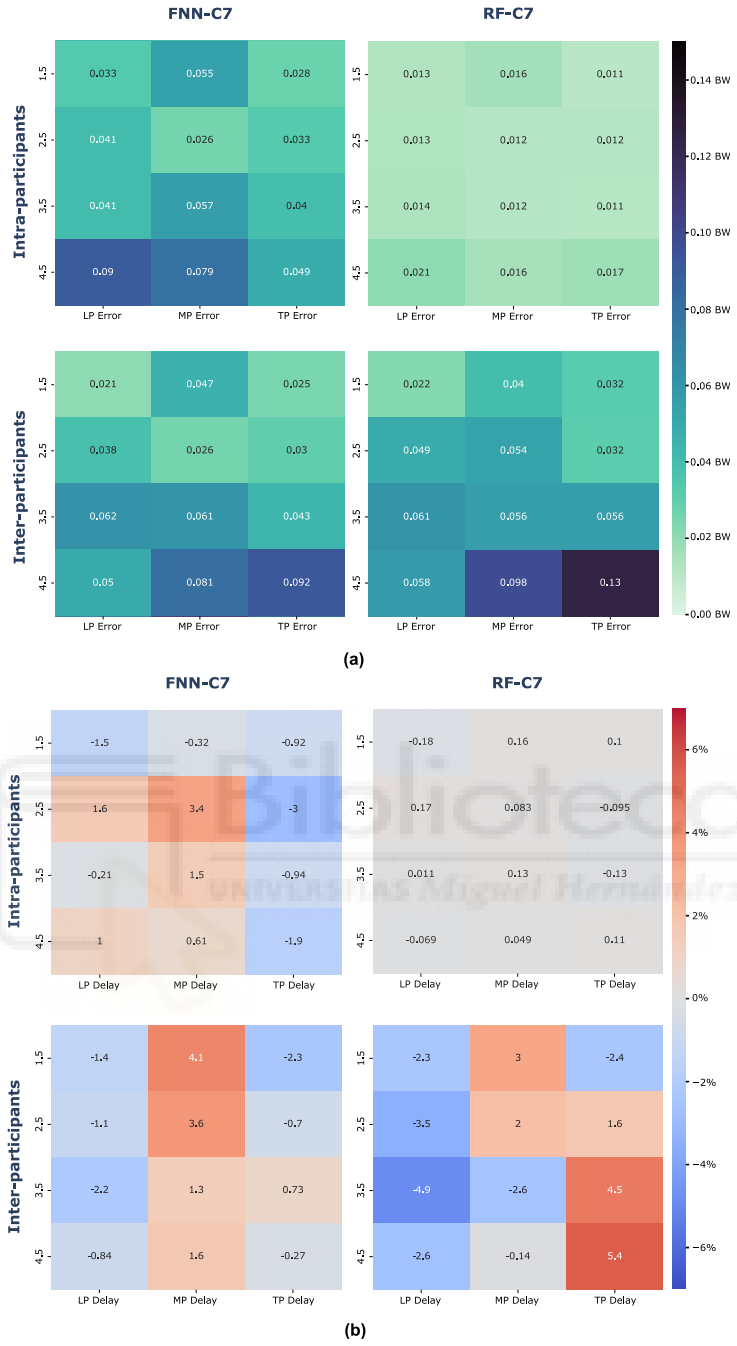


Figure 4.7: The vGRF peak errors for four gait speeds ranging from 1.5 km/h to 4.5 km/h have been represented as heatmaps. **(a)** Mean magnitude errors of the vGRF **(b)** Mean delays of the vGRF peaks

the position and orientation of the IMUs. Consequently, variations in sensor placement can be assumed to be unlikely to impact the accuracy of the estimated vGRF.

Two Machine Learning models have been trained and evaluated for estimating vGRF during gait. Initially, feature importance was determined for both the FNN-C7 and RF-C7 models to study the impact of each input (Figure 4.4). The results suggest that the gait cycle has the highest impact in model training, underscoring a strong temporal dependency in the vGRF estimation with the method proposed. This behavior can be explained by considering that angle measurements, angular velocities, and vertical accelerations may take the same values at different moments within the gait cycle, whereas vGRF values differ. Additionally, the results suggest that the kinematics of the hip and knee joints have a more relevant impact on the models than ankle kinematics.

While the strong relationship between vertical acceleration and vGRF was initially emphasized , the subsequent analysis reveals an interesting pattern. When calculating feature importance, it can be observed that vertical acceleration has a limited influence on model training compared to the kinematics of the lower limbs. Despite the vertical acceleration seems to have low relevance in the models, the results suggest that this input could still impact the performance of the FNN and RF models (Figure 4.5, Table 4.1, and Table 4.2). Although no statistically significant differences were observed, the results with both intra and inter-participant data suggest that including vertical acceleration improves the performance of both models in terms of RMSE. Interestingly, when vertical acceleration is introduced as input of the FNN models, the RMSE appears to increase. In summary, while vertical acceleration is not crucial for estimating vGRF using the proposed method (given accurate measurement of lower limb kinematics), it may still enhance the performance of specific models.

Furthermore, when comparing the evaluation metrics of the four models trained (FNN-Kinematics, FNN-C7, RF-Kinematics, and RF-C7) certain differences can be observed concerning the RMSE. When using the validation data derived from the intra-participants, the RMSE values for the RF models range from 0.030 BW to 0.045 BW, which is approximately half the RMSE for the FNN models (0.067-0.095 BW). However, when the models are evaluated with the inter-participant data, these differences are not statistically significant, yielding comparable RMSE values across the four models. Although an increase in the error with inter-participant data can be acceptable, this increasing tendency of the RF models might indicate that they tend to overfit more than the FNN models, which is a known problem in algorithms based on decision trees (Dietterich, 1995). Despite the overfitting, the results suggest that the RF models

achieve similar accuracy, with a global RMSE of 0.077 BW for the RF-C7 model, and 0.074 BW for the FNN-C7 model, which are the best FNN and RF models in terms of RMSE.

A deeper analysis of the FNN-C7 and the RF-C7 (Table 4.3) shows that the estimated vGRF is strongly correlated with the ground truth values for both intra and inter-participants. In terms of RMSE, the RF-C7 model outperforms the FNN-C7 model in achieving higher ρ with intra-participant data. However, when considering inter-participant data, the FNN-C7 correlations are higher than those achieved by the RF-C7, specifically for gait speeds between 1.5 km/h and 4.5 km/h. The NRMSE results follow a similar pattern. While the FNN-C7 model exhibits higher errors overall, it performs better than the RF-C7 model for inter-participant data at gait speeds of 2.5 km/h and 3.5 km/h. In contrast, the RF-C7 model achieves lower errors for gait speeds of 1.5 km/h and 4.5 km/h.

The estimated vGRF signals shown in Figure 4.6 indicate that both FNN-C7 and RF-C7 are capable of accurately replicating the measured vGRF signals. It is worth noting that these models can effectively distinguish between gait phases, as evidenced by their ability to estimate the characteristic peaks and swing phase (vGRF = 0 BW) from the graphs.

It is worth noting that the magnitude and timing of the vGRF peaks play a crucial role in gait analysis. Consequently, a deeper analysis of these peaks is conducted (Figure 4.7, Appendix A). Comparing the two selected models, the RF-C7 model demonstrates significantly higher accuracy than the FNN-C7 model in terms of magnitude errors with intra-participant data across all gait speeds. However, when evaluating the models with inter-participant data, the peak magnitude errors remain similar between both RF-C7 and FNN-C7. Notably, the errors obtained with the FNN-C7 in the LP at 1.5, 2.5, and 4.5 km/h, and the TP at 1.5 and 2.5 km/h are lower with the inter-participant data than the intra-participants. However, the RF-C7 model shows higher errors for inter-participants across all peaks and gait speeds. These findings would be in agreement with the previous observations related to RMSE and may be attributed to the potential overfitting of the RF models.

According to the timing errors of the vGRF peaks, the RF-C7 model demonstrates high accuracy in predicting the peak timings across all gait speeds using intra-participant data, with a maximum lag of 0.18% of the gait cycle. In contrast, the FNN-C7 model exhibits larger errors in peak timing, particularly in the MP with mean delays up to 3.4%. Concerning the results obtained for the inter-participants,

both models show delays and advances in peak estimation compared to the measured vGRF. It must be noted that all the mean lags obtained with the RF-C7 are greater than the FNN-C7 except for the MP peak at 1.5 and 2.5 km/h.

Concerning the results obtained, it can be highlighted that both the FNN-C7 and RF-C7 models exhibit robustness and good accuracy in estimating vGRF when using the method proposed in this thesis. This can be concluded by the RMSE and ρ obtained with the proposed models, which are in a similar range as reported by other authors when estimating the vGRF during gait. Specifically, in (Oh et al., 2013) the authors reported an RMSE of 0.066 BW and $\rho = 0.991$ with intra-participants; in (Choi et al., 2013) the authors obtained an RMSE of around 0.074 BW and $\rho = 0.99$ with intra-participants; in (Jiang et al., 2020) the proposed models achieved an RMSE of 0.02 BW and $\rho = 1.00$ with intra-participants, and an RMSE of 0.10 BW and $\rho = 0.97$ with inter-participants. Although the RF model trained by Jiang et al. achieves higher accuracy when evaluated with intra-participant data, the models presented in this Chapter achieve a lower RMSE and higher ρ with inter-participant data. Hence, these findings suggest that the FNN-C7 and RF-C7 models exhibit greater generalization capabilities compared to those proposed by Jiang et al.

It is important to note that while both the FNN-C7 and RF-C7 models achieve high accuracy compared to previous studies, the results show certain performance differences between them. The RF-C7 model yields a lower mean RMSE with intra-participant data, whereas the FNN-C7 model achieves a lower global RMSE compared to the RF models when they are evaluated with inter-participants. In addition, while the ρ and NRMSE show that a higher correlation and a lower error are obtained at 1.5 and 4.5 km/h with the RF-C7, better results are achieved at walking speeds of 2.5 and 3.5 km/h with the FNN-C7 when evaluated with inter-participant data. For inter-participants, the mean magnitude errors for characteristic peaks are comparable between the two models. Moreover, the FNN-C7 model appears to surpass the RF-C7 in accurately estimating the timing of the peaks. Given these variances, it can be inferred that the FNN-C7 model surpasses the trained RF-C7 model in terms of accuracy.

4.2.4 *Conclusions*

In this Chapter, a method to estimate the vGRF during ground-level walking has been presented. The IMU-based motion capture system presented in Chapter 3, section

3.2 has been employed to estimate the kinematics of the lower limbs and the vertical acceleration of the users. Two different Machine Learning models were evaluated for estimating the vGRF: FNN and RF models.

The findings indicate that the trained models are significantly impacted by the temporal variable (gait cycle). While the vertical acceleration has only a minor influence on the training of the FNN and RF models compared to other inputs, the results suggest that its inclusion can improve the performance of the models. Notably, in the case of the FNN and RF models, incorporating this input seems to enhance performance.

Additionally, the accuracy achieved with the method proposed for estimating the vGRF is accurate for performing a biomechanical analysis. This can be assumed since the vGRF estimates, including peak RMSE values, are aligned with the accuracy range reported in comparable studies, and even the results suggest that a lower RMSE is obtained in the estimation of vGRF. Furthermore, the RMSE, ρ, and the accuracy in the characteristic peaks point to the FNN-C7 model as the most accurate among the Machine Learning models that were trained.

4.3 FEASIBILITY STUDY FOR ESTIMATING VERTICAL GROUND REACTION FORCE DURING DIFFERENT GAIT ACTIVITIES OF DAILY LIVING

In this section, a feasibility study for estimating the vGRF during different gait activities has been conducted. Based on the methodology described in the preceding section, the IMU-based motion capture system has been used to acquire data from the lower limb kinematics and vertical acceleration to be used as inputs of a Machine Learning model. As the FNN-C7 has been shown to achieve the highest performance for ground-level walking vGRF estimation, an adaptation of this model is proposed to estimate this force during different gait ADLs. In addition, it is discussed whether the performance of the model proposed improves when the activity performed is known.

4.3.1 *Materials and Methods*

4.3.1.1 *Experimental sessions*

Twelve users (nine male, three female) aged from 23 to 52 years old (29.8 ± 7.4), with heights ranging from 165cm to 187cm (176.2 ± 7.4cm) and weights from 56.1kg to 90.2kg (76 ± 12.5kg) were involved in the study. Written informed consent was obtained from the individuals before participating in the sessions, and the study was approved by the INESCOP Footwear Technology Center.

The performed gait ADLs together with the systems used to simulate the activities are shown in Figure 4.8. A treadmill model h/p/cosmos 150/50 was used to simulate three different scenarios: ground-level walking, ramp ascent walking, and ramp descent walking. In addition, a stair treadmill was employed to simulate stair ascent climbing.

Lower limb kinematics were collected using four XSens Dot IMUs, and an additional device was placed over the C7 vertebra to measure the vertical acceleration of the users (see Chapter 3, section 3.2 and section 4.2). Moreover, the vGRF was measured by using the force plate embedded in the h/p/cosmos 150/50 for ground-level, ramp ascent,

Figure 4.8: The performed gait activities and the experimental setup used to estimate the vGRF are represented in the images.

and ramp descent walking. During stair climbing, the pressure insoles model *novel.de pedar* based on capacitive sensors were used to measure the vGRF.

After the placement of the aforementioned devices, and after calibration of the implemented IMU-based motion capture and pressure insoles, the vGRF and the flexion/extension angles of the hip, knee, and ankle were measured during different gait ADLs according to the following protocol:

1. Ground-level walking at 4.5 km/h for 5 minutes.
2. 12% positive slope walking at 2.5 km/h for 5 minutes.
3. 12% negative slope walking at 2.5 km/h for 5 minutes.
4. Climbing stairs at 50 stairs/minute for 3 minutes.

The INESCOP Footwear Technology Center protocols for footwear certification were followed to set the walking speeds and slope of the treadmill. The stair-climbing time was shortened to 3 min to prevent excessive fatigue.

4.3.1.2 *Acquired data and processing*

The data acquisition process to estimate the vGRF during different gait activities is based on the methodology described in section 4.2.1.2. Concerning the angle and angular velocities of the lower limb joints in the sagittal plane, data from the IMUs placed on the hip, the thigh, the shank, and the foot were employed. In addition, the IMU placed at the C7 vertebra was used to measure the vertical acceleration of the users during gait.

Regarding the vGRF, the force plate embedded in the h/p/cosmos treadmill was used to measure the exerted force when the participants performed the ground-level and ramp exercises. In the case of the stair walking, the pressure insoles were used to estimate the vGRF. Despite the treadmill-embedded force plate provides the vGRF magnitude during gait when employing the pressure insoles further processing is required. According to the foot size, the pressure insoles are composed of a pressure sensor matrix (Figure 4.9), where each sensor covers a certain area. Therefore, the exerted vGRF can be estimated as:

$$vGRF(t) = \sum_{i=1}^{n} P_i(t) \cdot A_i \tag{4.2}$$

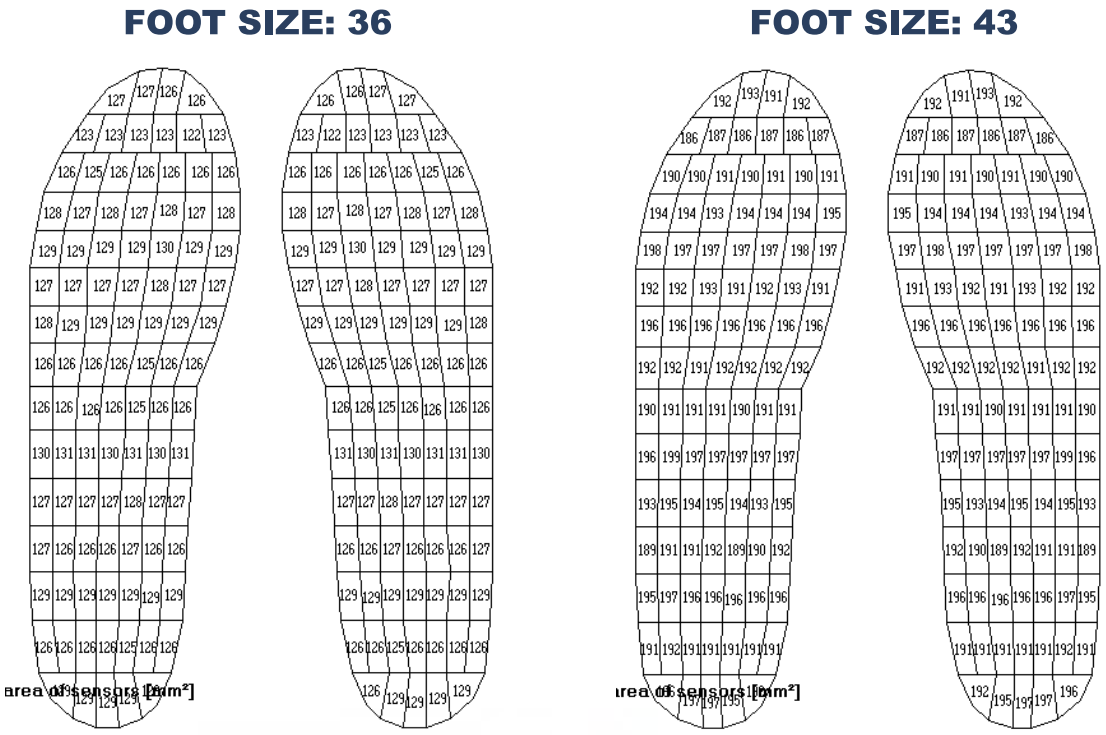


Figure 4.9: Representation of two pressure sensor matrices used to measure the vGRF during ascent stairs. Each of the numeric values indicates the area covered by each sensor in mm^2.

where P_i is the pressure measured by each sensor and A_i is the sensor measurement area.

In the same manner as described in section 4.2.1.2, the foot-foot ground contacts have been detected in the motion capture system and the acquired vGRF in order to calculate the gait cycle and subsequently synchronize all the signals. In Figure 4.10, the flexion/extension angles, the vertical acceleration, and the vGRF signals have been represented for each gait ADL.

4.3.1.3 *Vertical Ground Reaction Force estimation architecture*

Based on the analysis performed in section 4.2, the feasibility of estimating the vGRF during different gait tasks using the FNN-C7 model architecture is studied. Therefore, the flexion/extension angles of the hip, knee, and ankle, their angular velocities, the

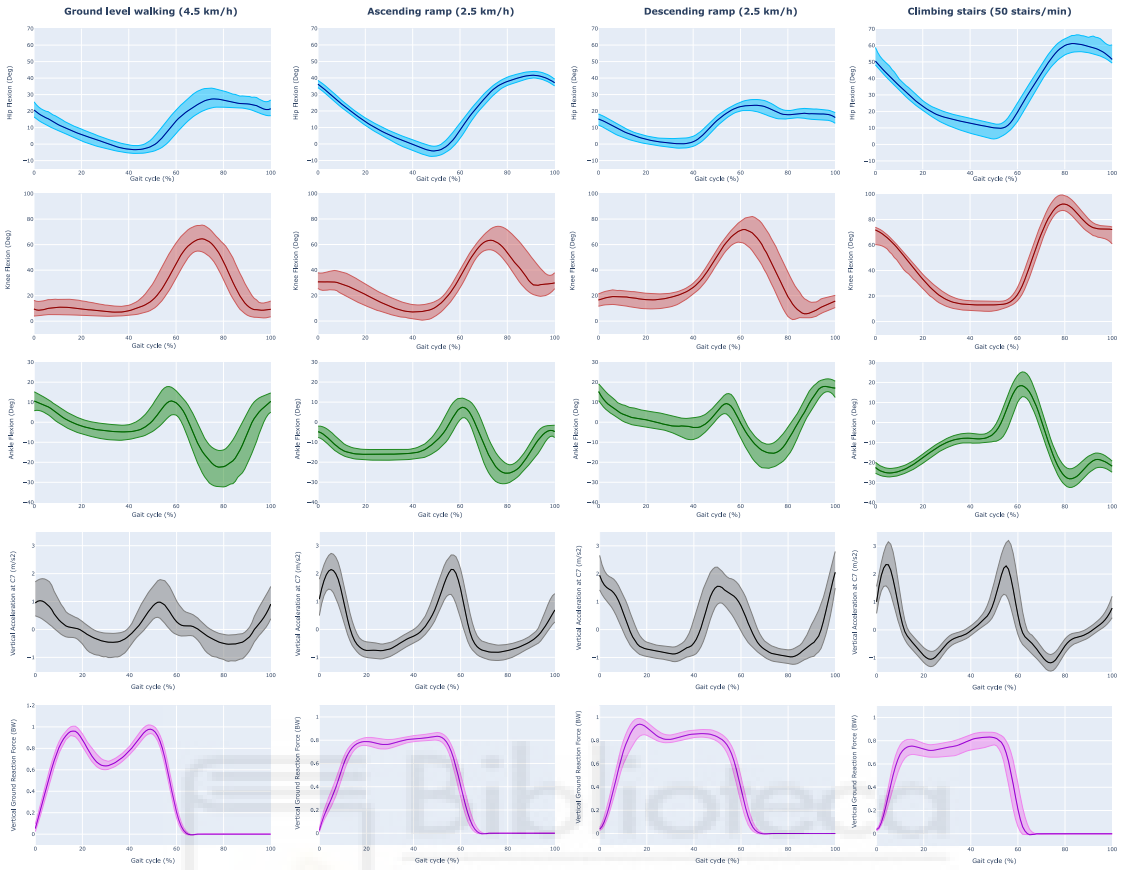


Figure 4.10: The flexion/extension angles of the hip, knee, and ankle are represented in the first, second, and third rows, respectively. The measured vertical acceleration at the C7 vertebra is shown in the fourth row, and the measured vGRF is presented in the fifth row. These data are represented across the different gait activities performed during the experimental sessions.

vertical acceleration measured at the C7 vertebra, and the gait cycle are introduced as inputs.

As introduced in Chapter 2, section 2.2, there is a lack of studies proposing the estimation of the vGRF during different gait ADLs. Hence, apart from introducing kinematic measurements, it is intended to analyze whether the knowledge of the performed ADL has a positive impact on the vGRF prediction. With this purpose, three different approaches are proposed to estimate the vGRF, which are represented in Figure 4.11:

- **General model.** Joint angles, angular velocities, vertical acceleration, and gait cycle data from the four ADLs performed are used as inputs to train a general FNN model.
- **Activity-labeled model.** Joint angles, angular velocities, vertical acceleration, gait cycle, and the current activity performed are used as inputs to train an FNN model.
- **Ensembled models.** Joint angles, angular velocities, vertical acceleration, and gait cycle data from the four ADLs performed are used as inputs to four different FNN models (i.e. one per ADL).

Before model training, data from ten users were randomly divided to evaluate the accuracy of the models by cross-validation: 80% of the data was used for training and the remaining 20% was used for validation (intra-participant data). Additionally, the performance of the models has been evaluated on users who were not involved in the learning process (inter-participant data) using data from the two remaining participants. In addition, it is worth mentioning that the training dataset was used to build a scaler, which normalized the input values of the Machine Learning models within the range of 0 to 1. This scaler was also used to normalize the validation and inter-participants dataset to evaluate the accuracy of the vGRF estimation models. Moreover, to train the activity-labeled model, a OneHot encoder has been used to create four additional inputs to indicate the ADL performed.

The Keras Python Library (Gulli and Pal, 2017a) has been used to train the FNN models. The FNN models trained are composed of an input layer with a number of neurons equal to the number of inputs, five hidden layers with ten neurons, and an output layer with one neuron. The ReLu has been used as the activation function, the Adam algorithm has been used to optimize the network weights, and the RMSE serves as the loss function. Furthermore, the dropout regularization technique is applied to avoid overfitting.

The performance of the proposed approaches to estimate the vGRF during different gait ADLs has been evaluated as follows:

- vGRF RMSE. A lower RMSE in the estimation of the vGRF signal has been considered as a better performance of the models.
- vGRF NRMSE. A lower NRMSE in the estimation of the vGRF signal has been considered as a better performance of the model.

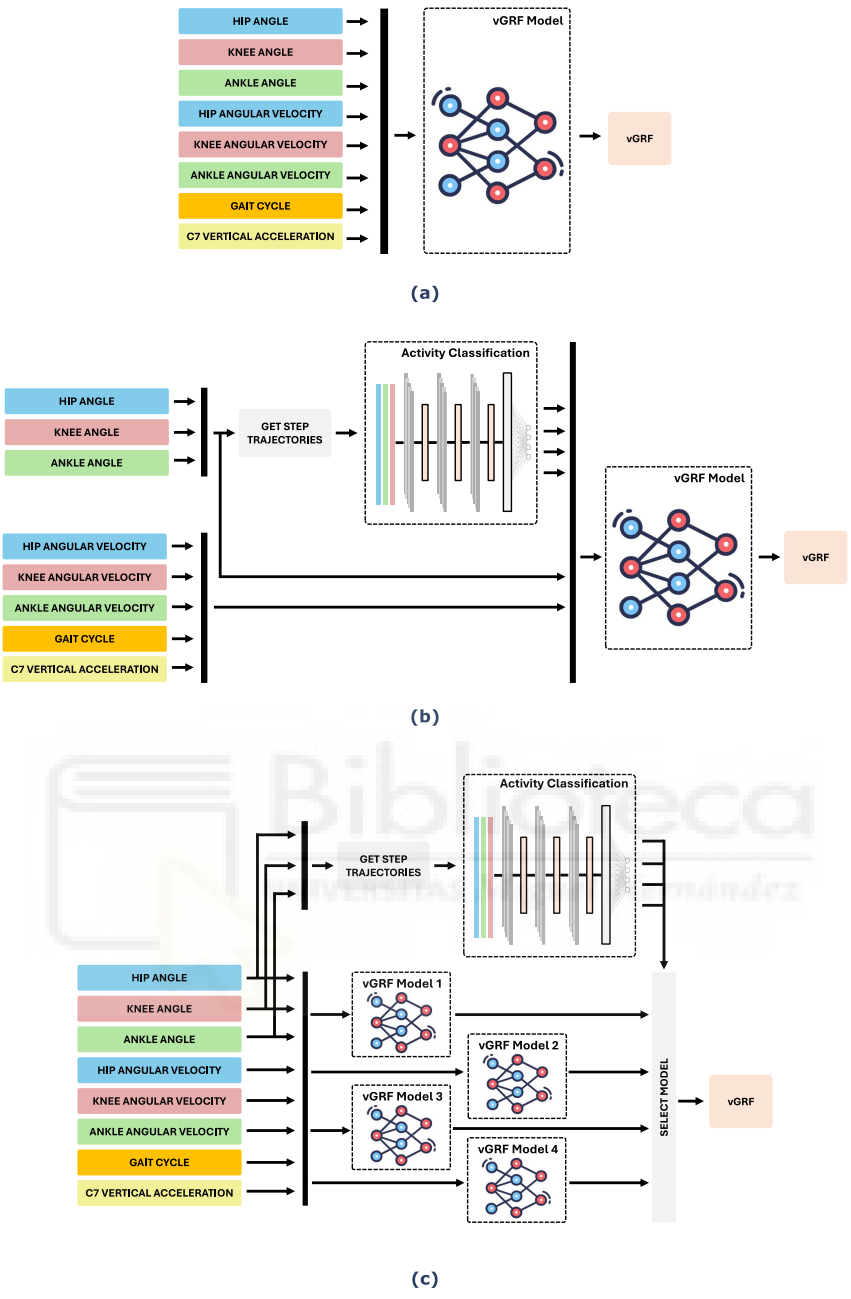


Figure 4.11: Approaches proposed to estimate the vGRF during different gait activities. **(a)** A general model is trained to estimate the vGRF **(b)** A classifier can be proposed to introduce the performed activity to the model **(c)** A model per activity can be trained, and a classifier selects the output of the model according to the activity performed

- Correlation. A higher correlation (ρ) between the ground truth and the estimated vGRF signal has been considered as a better performance of the models. The Person's correlation coefficient has been used to calculate the correlation between estimated and ground truth data.

4.3.2 *Results and Discussion*

Table 4.4 collects the mean and standard deviation of the RMSE for each of the performed activities when estimating the vGRF with the three proposed approaches using the intra-participant and inter-participant data. In addition, these results are represented in Figure 4.12.

Concerning the RMSE results, it can be observed that the knowledge of the performed task reduces the error when estimating the vGRF for different gait ADLs. Generally, using an FNN model without knowledge of the activity performed leads to a higher RMSE in the vGRF estimation. Only when ascending a ramp, the general model achieves a lower error than the ensembled approach for both intra and inter-participants. In contrast, when the activity-labeled model and the ensembled architecture are compared, similar RMSE results obtained are obtained. First, it can be observed that the global RMSE obtained are lower with the activity-labeled using the intra-participant

Table 4.4: The table shows the mean (standard deviation) RMSE (BW) and the global RMSE for each vGRF estimation approach and activity with the intra-participant and inter-participant data.

	Activity	General	Activity-labeled	Ensembled
Intra-participants	**Global**	0.0960 (0.0154)	0.0819 (0.0178)	0.0839 (0.0165)
	Ground-level	0.0945 (0.0171)	0.0794 (0.0186)	0.0791 (0.0206)
	Ascent ramp	0.0729 (0.0147)	0.0653 (0.0160)	0.0784 (0.0153)
	Descent ramp	0.1218 (0.0289)	0.1062 (0.0297)	0.1020 (0.0309)
	Ascent stairs	0.1084 (0.0203)	0.1029 (0.0210)	0.1078 (0.0183)
Inter-participants	**Global**	0.1031 (0.0021)	0.0883 (0.0103)	0.0827 (0.0055)
	Ground-level	0.1027 (0.0055)	0.0832 (0.0103)	0.0717 (0.0089)
	Ascent ramp	0.0892 (0.0072)	0.0802 (0.0113)	0.0936 (0.0064)
	Descent ramp	0.1132 (0.0167)	0.1073 (0.0183)	0.0964 (0.0193)
	Ascent stairs	0.1194 (0.0012)	0.1070 (0.0070)	0.1148 (0.0262)

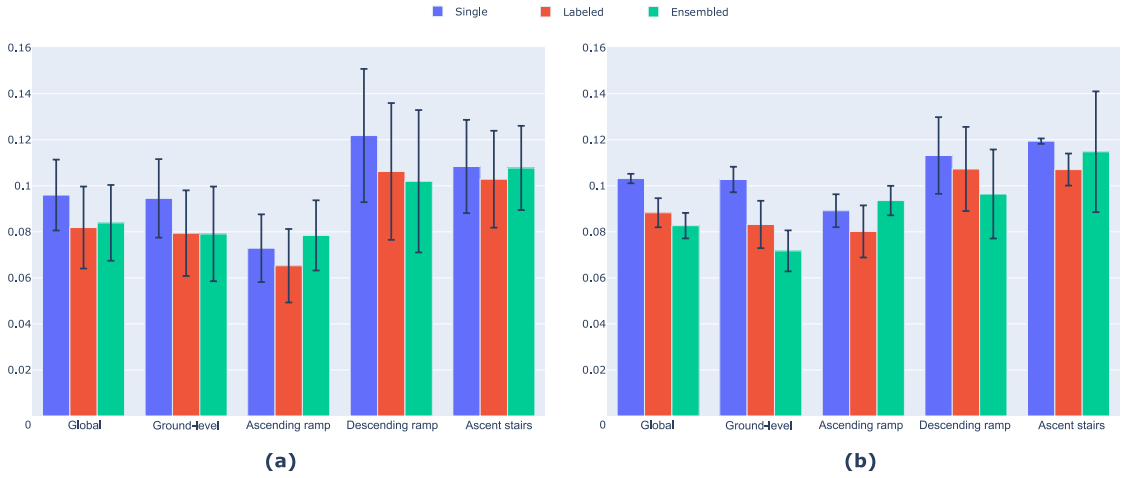


Figure 4.12: Mean RMSE and standard deviation for each of the approaches proposed to estimate the vGRF during different activities. The global RMSE and the errors obtained for each of the activities are represented when the approaches are evaluated with the **(a)** intra-participants and **(b)** inter-participants.

data, while the ensembled models achieve a lower RMSE with the inter-participant data. Furthermore, on the one hand, the RMSE errors obtained with the activity-labeled model are lower than the ensembled approach when ascending a ramp and ascending stairs. On the other hand, the ensembled models achieve a lower RMSE during ground-level walking and descending a ramp.

In addition, Table 4.5 collects the NRMSE and ρ values for each activity and estimation approach. From these results can also be deduced that recognizing the activity performed by the users would improve the vGRF estimation performance. Furthermore, as for the RMSE, the ρ and NRMSE values are similar for the activity-labeled and ensembled.

In addition, the ground truth and the estimated vGRF signals during gait are represented in Figure 4.13 for each of the proposed approaches. The representation of the vGRF signal also suggests that the knowledge of the activity performed can improve the performance of Machine Learning models for estimating vGRF. However, the graphs also reveal the main limitation of the trained FNN models. Although ground-level walking allows for a faithful estimation of the characteristic peaks, there is a noticeable decrease in accuracy for other tasks. This is especially evident when descending a ramp, where the proposed models fail to estimate the LP accurately.

Table 4.5: Correlation (ρ) and NRMSE results between the ground truth and estimated vGRF with the intra-participant and inter-participant data for each of the prediction architectures proposed.

	Architecture	Activity	ρ	NRMSE
Intra-participants	**General**	Ground-level	0.9603	0.0768
		Ascent ramp	0.9837	0.0446
		Descent ramp	0.9625	0.0773
		Ascent stairs	0.9554	0.0691
	Activity-labeled	Ground-level	0.9639	0.0682
		Ascent ramp	0.9850	0.0398
		Descent ramp	0.9658	0.0676
		Ascent stairs	0.9549	0.0663
	Ensembled	Ground-level	0.9626	0.06937
		Ascent ramp	0.9848	0.0480
		Descent ramp	0.9700	0.0657
		Ascent stairs	0.9529	0.0689
Inter-participants	**General**	Ground-level	0.9511	0.1139
		Ascent ramp	0.9782	0.0781
		Descent ramp	0.9649	0.0995
		Ascent stairs	0.9521	0.1045
	Activity-labeled	Ground-level	0.9603	0.9797
		Ascent ramp	0.9792	0.0704
		Descent ramp	0.9633	0.0941
		Ascent stairs	0.9585	0.0943
	Ensembled	Ground-level	0.9740	0.0794
		Ascent ramp	0.9759	0.0820
		Descent ramp	0.9711	0.0851
		Ascent stairs	0.9504	0.1033

While the results suggest that the errors committed are in a similar range, further analysis is required to affirm that the estimation of the vGRF is accurate in the rest of the gait ADLs using a FNN architecture. Nevertheless, the results of this study suggest that it is feasible to develop a model that allows estimating the vGRF during different tasks using IMUs. Furthermore, as introduced, there seems to be an absence of research offering a methodology that facilitates the estimation of the vGRF across different terrains. Therefore, it is also intended that the results of this study serve as a

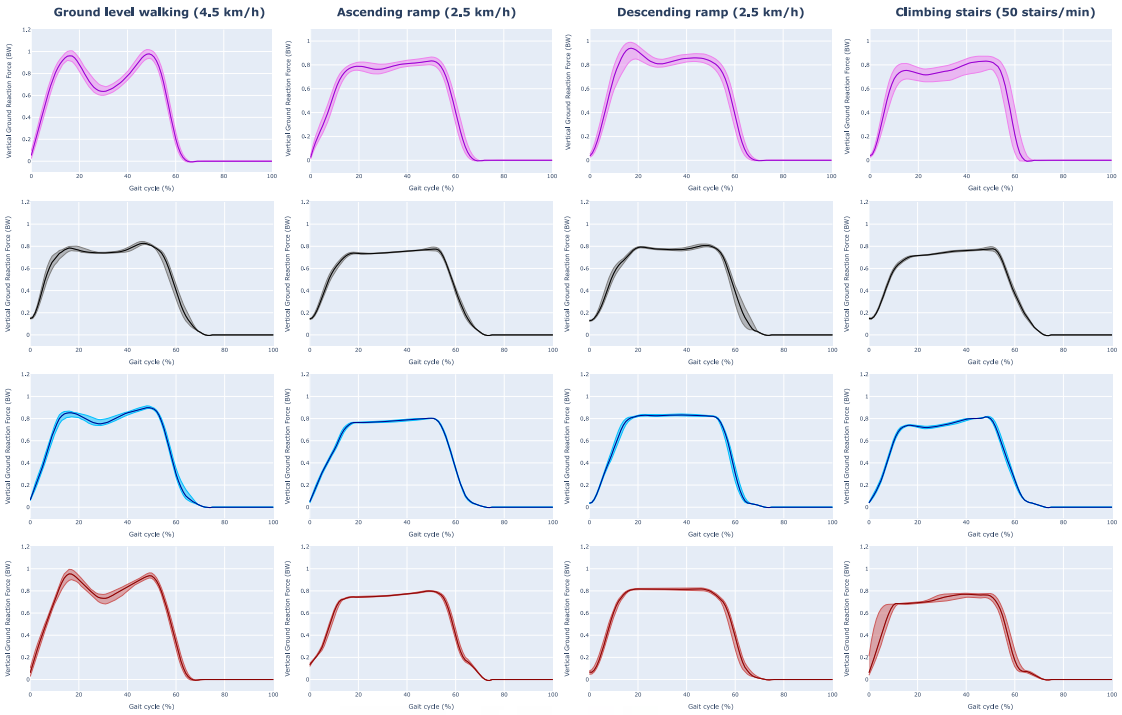


Figure 4.13: The ground truth and estimated vGRF have been represented according to the gait cycle (%) for different gait activities. The median values have been represented, and the values between the first and third quartiles have been colored. The first row shows the measured vGRF, the second row shows the estimated vGRF using the general model, the third row shows the estimated vGRF using the activity-labeled model, and the fourth row represents the estimated vGRF using the ensembled models approach. These graphs include all the participants.

baseline for the development of future Machine Learning models that allow estimating the vGRF during different ADLs.

4.3.3 *Conclusions*

In this Chapter, based on the motion capture system (Chapter 3, section 3.2) and the FNN models developed for estimating the vGRF during ground-level walking, a feasibility study has been conducted to estimate this force during different gait ADLs. Instead of only introducing kinematic data from the users, it has been evaluated whether the knowledge of the performed activity has a positive impact on the vGRF

estimation. With this purpose, three different Machine Learning approaches have been evaluated.

The results suggest that a higher accuracy is achieved in the vGRF estimation when the activity performed is known. Therefore, a classification system would help to recognize the gait ADL performed by the user, and subsequently adapt the estimation of the vGRF autonomously. In addition, although an acceptable accuracy seems to be achieved with the FNN architecture, the accuracy when estimating the characteristic peaks of the vGRF signal must be improved. Hence, as future work, it is intended to evaluate different Machine Learning model architectures.

5

GAIT ACTIVITY CLASSIFICATION WITH A CONVOLUTIONAL NEURAL NETWORK USING LOWER LIMB ANGLE MEASUREMENT FROM INERTIAL SENSORS

5.1 INTRODUCTION

As introduced in Chapter 3, section 3.1, the SPLASH project proposes to develop an underwater lower limb exoskeleton. In addition, it is intended to not only assist the patient's gait during ground-level scenarios but also it is proposed to generalize the assistance to different gait ADLs. Therefore, there is a need to develop a system that can recognize the performed activity, and adapt the assistance accordingly.

Furthermore, to study and understand the human lower limb kinetics during underwater gait, either the measurement or the estimation of the vGRF is required to understand the loads in the joints of the lower limbs. As described in Chapter 4, diverse approaches have been evaluated to estimate the vGRF during gait in an attempt to provide a solution to estimate the vGRF in those cases where using an underwater force plate is a limitation. Hence, as concluded in Chapter 4, section 4.3, the knowledge of the activity performed is required to maximize the performance in the vGRF estimation.

With this purpose, a Deep Learning model is presented in this Chapter to autonomously classify different gait activities. The data collected during the experimental sessions described in Chapter 4, section 4.3 have been used to develop the model. The flexion/extension angle trajectories during each step have been introduced as inputs of the models. Hence, as the inputs of the model are one-dimensional windows, a One-dimensional Convolutional Neural Network (1D-CNN) is proposed to recognize different gait ADLs. In addition, diverse techniques have been used to understand the decision-making process of the model. Finally, possible variations of the model inputs have been studied in order to adapt the model to multiple lower limb exoskeletons or gait analysis scenarios.

5.2 MATERIALS AND METHODS

As introduced, the data acquired during the experimental sessions described in Chapter 4, section 4.3.1.1 have been used to propose a Deep Learning-based classifier approach. In this section, the Deep Learning model architecture, the methods use to understand the model decision-making, the data processing, and the model training are described.

5.2.1 *Proposed Convolutional Neural Network architecture*

As introduced, it is intended to develop a 1D-CNN to recognize four gait activities: ground-level walking, ascending a ramp, descending a ramp, and ascending stairs. Typically, the CNN models consist of two parts (Zhao et al., 2017a): the convolutional and pooling layers, and the fully-connected layers or Multilayer Perceptron (MLP).

The CNN models have an input layer of Nxk, where N is the length of the univariate time series and k is the number of input series. Subsequent to this layer, the model employs convolution and pooling operations to extract deep features from the input data. Within the convolutional layers, the model applies convolutional operations to the temporal series data from the precedent layer using convolutional filters, also known as kernels, creating an output known as a feature map. After the convolution operations, a non-linear activation function is applied. In the pooling layers, the feature maps are divided, and each segment is represented by its average or maximum value. Finally, after the convolution and pooling layers, the original time series is represented by a series of feature maps.

In the second part of the model, the feature maps are typically connected to a MLP. These MLP layers perform the classification based on the features and filters learned in the preceding layers. Alternatively, a Global Average Pooling (GAP) layer can replace this step. The primary goal of the GAP layer is to generate one feature map per class (Lin et al., 2013). This layer computes the average value for each feature map, resulting in a vector that directly feeds into the softmax layer. Notably, the GAP layer aligns well with the convolutional structure, establishing meaningful associations between feature maps and distinct categories. Furthermore, it mitigates the risk of overfitting since it lacks parameters for optimization

The proposed architecture of the 1D-CNN model developed to recognize four gait ADLs is a Fully-Convolutional Neural Network (FCN), which has been represented in Figure 5.1. The model inputs are the normalized flexion/extension angles of the hip, knee, and ankle, scaled according to their maximum and minimum values. The architecture includes three convolutional layers, with each layer applying 64 filters of a kernel size of 3. Padding is also utilized during the convolution processes to ensure the resulting feature maps maintain a length N that matches the input length. Following each convolutional layer, the ReLu function is employed as the activation function, complemented by Batch Normalization. After the convolutional layers, a GAP layer is connected to a dense layer consisting of four neurons, each corresponding to a gait ADL, and uses the softmax function for activation. The hyperparameters of the model, including the number of convolutional layers, filters, kernel size, and others, have been tuned to optimize performance.

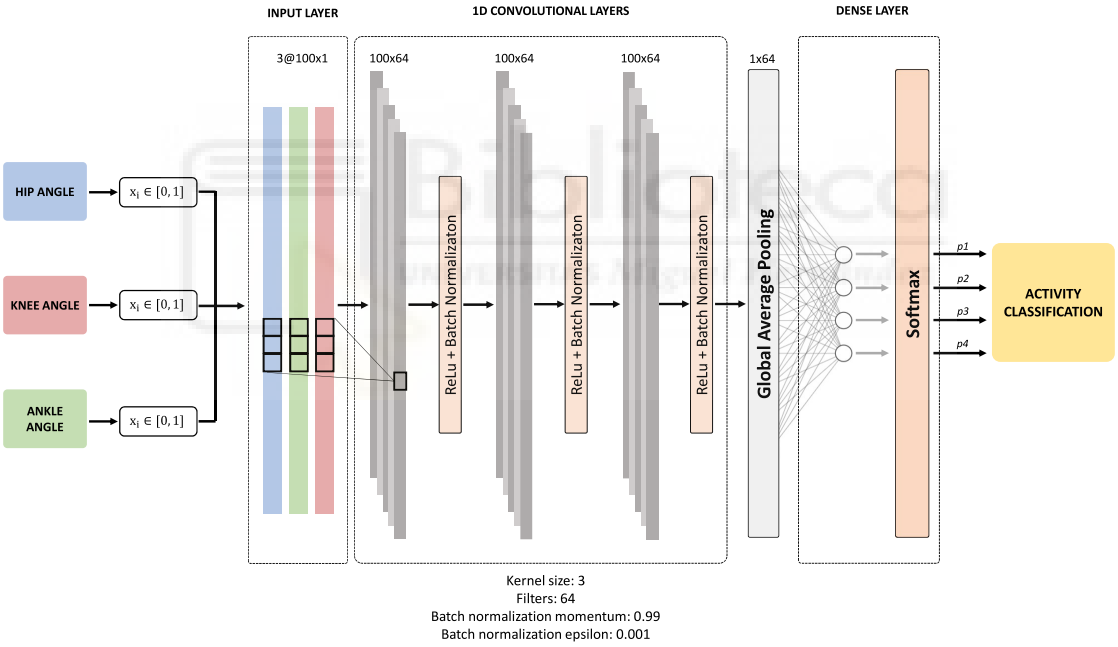


Figure 5.1: A FCN model is proposed to classify four gait activities based on data from the hip, knee, and ankle flexion/extension joints. The input data are normalized between their minimum and maximum values before being introduced to the model. The model consists of three one-dimensional convolutional layers with a kernel size of 3 and 64 filters. Padding is applied to ensure that the feature maps have the same length as the inputs. The ReLU activation function is used for the convolutional layers, and batch normalization is applied. A GAP layer is used to reduce the generated feature maps connected to the output layer.

5.2.2 *Data processing and model training*

The data acquisition and processing pipeline is represented in Figure 5.2. The accelerometer and gyroscope data from IMUs placed on the hip, thigh, shank, and foot were used to estimate the flexion/extension angles of the hip, knee, and ankle using the method described in Chapter 3, section 3.2. In addition, the acceleration of the hip was used to detect the ground-foot contacts. These contacts are used to identify the onset and end of each step and subsequently transform the data from the temporal to the gait cycle domain (0-100%).

Furthermore, it is worth mentioning that the CNN models require a predetermined window size for input. Given that the objective is to classify the activity performed by the user, selecting a window size that encompasses 100 joint measurements derived from the gait cycle appears to be a logical choice to classify each step within one of the ADLs. Consequently, three different windows have been used as inputs of the model, each containing 100 data points, which represent the flexion/extension angles of the hip, knee, and ankle throughout a single step.

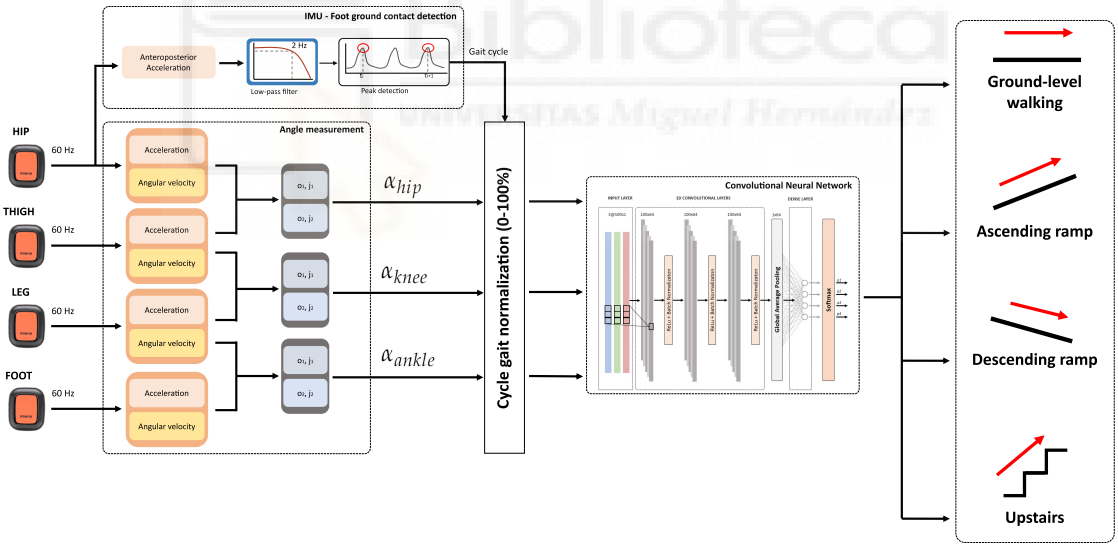


Figure 5.2: From the IMUs placed on the lower limbs, the acceleration and angular velocity were used to measure the flexion/extension of the hip, knee, and ankle. Additionally, foot-ground contact is detected using hip acceleration to calculate the gait cycle and splitting the joint angles step by step. The gait cycle is used to represent each step in a 100-size window (cycle-gait normalization). Finally, the normalized trajectories of the hip, knee, and ankle are used as inputs for the model.

The data collected from the twelve participants were divided into three different datasets. The time series data from nine participants were randomly segmented by individual steps for training and validating the accuracy of the model (intra-participant evaluation): 80% for training and 20% for validation. The data from the remaining three participants were utilized to evaluate the performance of the model with individuals who were not part of the training process (inter-participant evaluation). Furthermore, the training dataset was normalized by scaling the hip, knee, and ankle angle values to a range from 0 to 1. This same normalization scaler was then applied to both intra-participant and inter-participant datasets during the model evaluation phase.

The proposed 1D-CNN model has been developed and trained with the Keras Python Library. The model was trained for 50 epochs and a batch size of 8. The Adam algorithm was used as the optimizer to adjust the weights of the network, and the Categorical Cross-Entropy (CE) was used as the loss function, which is defined as:

$$CE = -\sum_{i}^{C} y_i \cdot log(\hat{y}_i) \tag{5.1}$$

where y_i is the actual class, $\hat{y}_i$ is the CNN score for each class, and C is the number of classes.

5.2.3 *Model explainability*

Machine Learning models are often perceived as 'black boxes' since their decision-making processes are mostly not transparent to humans. However, understanding this process can be critical in fields like healthcare. For this study, it would be particularly valuable to comprehend the mechanisms by which our trained CNN model differentiates among the four ADLs simulated in the experimental trials.

To address the challenge of model interpretability, various techniques have emerged in an attempt to understand the decision-making process of black-box Machine Learning models. One such method we employed is the Gradient-weighted Class Activation-Map (Grad-CAM) algorithm, which was originally proposed for image analysis. Grad-CAM provides visual explanations by highlighting important regions in an image that contribute to a specific prediction. By extending Grad-CAM to time series analysis using a 1D-CNN, the segments of joint trajectories that play a crucial role in distinguishing different gait activities can be highlighted.

In addition to the use of Grad-CAM, the representation of the feature map extracted by the proposed 1D-CNN may provide information about the decision-making of the model. With this purpose, the t-Distributed Stochastic Neighbor Embedding (t-SNE) has been used to represent the 64-dimensional feature map, since this technique visualizes high-dimensional data by giving each datapoint a location in a two or three-dimensional map (Van der Maaten and Hinton, 2008).

5.3 RESULTS

5.3.1 *Model evaluation*

The performance of the proposed 1D-CNN model has been evaluated with data from both intra-participants and inter-participants. The evaluation and inference process was conducted on a PC with an Intel(R) Core(TM) i7-1165G7 @ 2.80GHz with 16 GB of RAM, where each inference was completed in 1 ± 0.37 milliseconds. Furthermore, the performance is compared with four State of the Art models proposed in recent works: a Residual Network (ResNet) model (Wang et al., 2017a), a Time LeNet (t-LeNet) model (Le Guennec et al., 2016), a Multi-scale Convolutional Neural Network (MCDCNN) model (Zheng et al., 2016), and a Time Convolutional Neural Network (Time-CNN) model (Zhao et al., 2017b). The accuracy and CE metrics for both intra-participant and inter-participant data are collected in Table 5.1. The results indicate that the model proposed achieves an accuracy of 99.65% and a CE of 0.0054 with intra-participant data, and an accuracy of 99.56% and a CE of 0.0025 with inter-participant data.

Confusion matrices are shown in Figure 5.3 for both intra-participant and inter-participant data to better understand the classification results. The findings indicate that the CNN could sometimes misclassify ground-level walking, ascending a ramp, and descending a ramp. Specifically, for intra-participant data, 99% of the descent ramp steps are classified accurately, with 0.32% of the steps being confused with ground-level walking and 0.97% confused with ascending a ramp. The inter-participant results demonstrate similar trends, with 1.5% of the descent ramp steps being incorrectly classified as ground-level walking.

Table 5.1: Accuracy and Categorical Cross-Entropy (CE) obtained with the intra-participants and inter-participants.

Model	Dataset	Accuracy(%)	CE
Proposed 1D-CNN	Intra-participants	99.65	0.0054
	Inter-participants	99.56	0.0025
ResNet (Wang et al., 2017a)	Intra-participants	100.00	0.0004
	Inter-participants	99.46	0.0082
t-LeNet (Le Guennec et al., 2016)	Intra-participants	99.98	0.0132
	Inter-participants	99.21	0.0287
MCDCNN (Zheng et al., 2016)	Intra-participants	99.98	0.0036
	Inter-participants	99.46	0.0039
Time-CNN (Zhao et al., 2017b)	Intra-participants	99.98	0.0003
	Inter-participants	98.09	0.0102

5.3.2 *Model interpretability*

The t-SNE has been applied to project the 64-dimensional feature map obtained by the convolutional layers into a two-dimensional graph. In Figure 5.4 the representation of the generated feature maps for two intra-participants and two inter-participants are represented using perplexity and early exaggeration values of 50. Different groups of points (clusters) can be observed, where each point represents the feature map obtained by the CNN for each step, and clusters indicate groups of similar steps.

The mean values of the hip, knee, and ankle trajectories have been represented together with the class activation maps obtained with Grad-CAM for the last convolutional layer in Figure 5.5. The activation map is visualized as a heatmap, emphasizing the significant regions of the input data. Areas shaded in blue indicate regions with less attention while those in red denote areas with lower attention. For ground-level walking, the activation map reveals high activation during the terminal stance phase (45%-65%), the swing phase, and at initial contact. The map for ascending ramps shows pronounced activation during the loading response and mid-stance phase (approximately 10%-50%), as well as during the terminal swing phase. In the case of descending ramps, there is a notable activation during the terminal swing phase

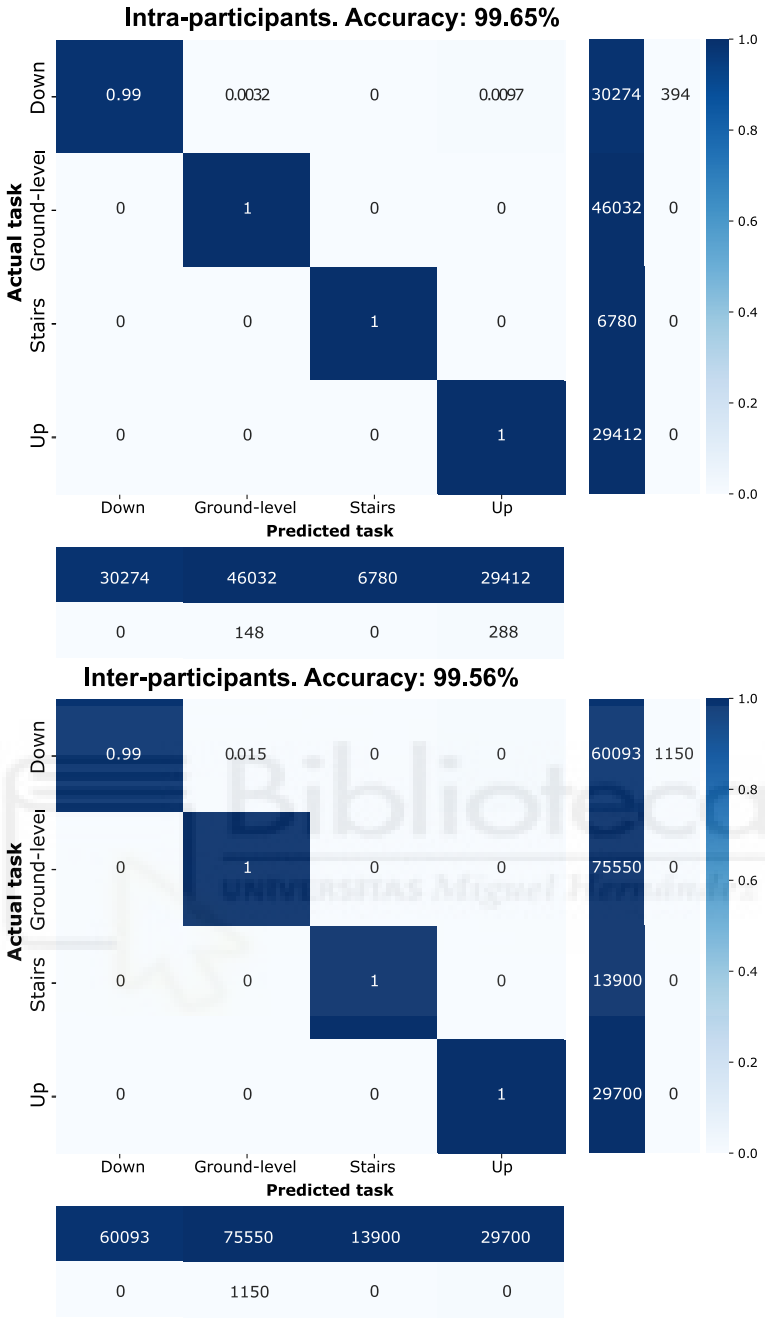


Figure 5.3: The confusion matrices obtained with the proposed CNN classifier show the results for intra-participant and inter-participant scenarios. The rows of the matrix indicate the actual ADLs performed, while the columns show the predicted activity. These matrices include the four gait ADLs conducted in the experiment: ground-level walking, ascending a ramp (Up), descending a ramp (Down), and climbing stairs.

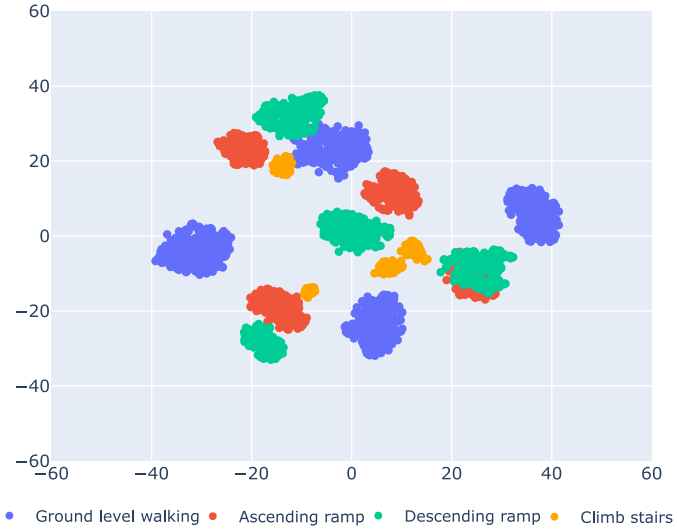


Figure 5.4: The two-dimensional representation of the feature map generated by the CNN model proposed for two intra-participants and two inter-participants. The map was created using the t-SNE embedding, and each activity was assigned a specific color: blue for ground-level walking, red for ascending a ramp, green for descending a ramp, and yellow for climbing stairs.

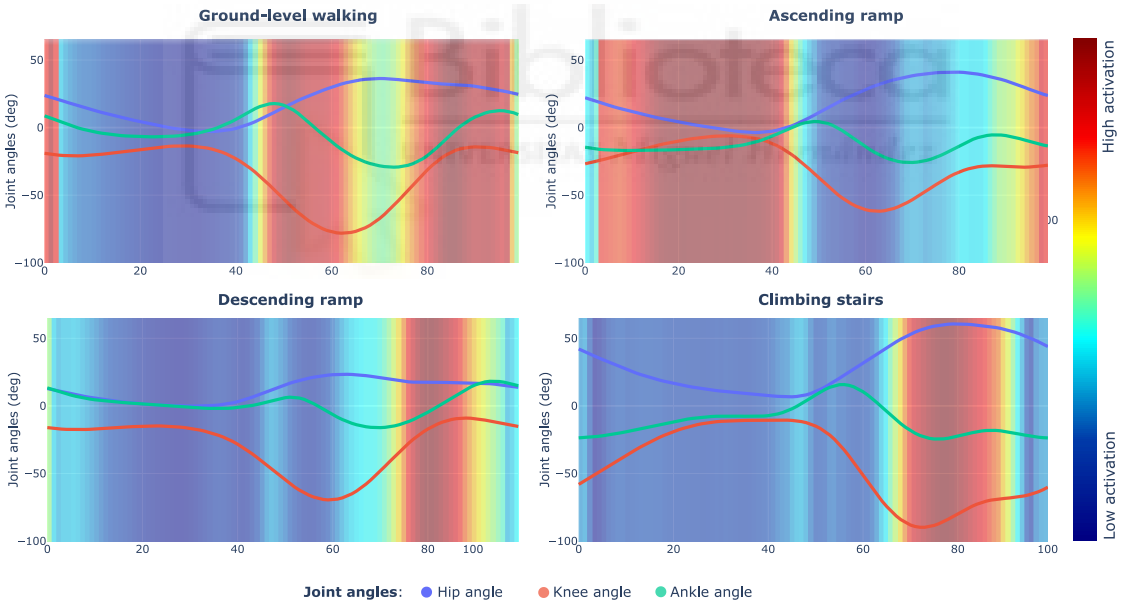


Figure 5.5: Mean trajectories of lower limb joints during four different gait activities are presented alongside mean class activation maps. These maps were generated using the Grad-CAM method, highlighting areas of low activation in blue and areas of high activation in red.

(75%-90%). Finally, for the ascending stairs activity, the map shows a strong activation during the swing phase (70%-90%).

Moreover, the Grad-CAM technique can also offer insights into the misclassification of certain steps. Figure 5.6 presents a pair of steps that were wrongly identified as descending a ramp. The graph shows the joint angles and the class activation maps, with the softmax probabilities for the true category and the predicted category indicated.

5.3.3 *Model adaption*

To assess the adaptability of the CNN model, diverse models with the same proposed architecture have been trained while varying the number of inputs. Specifically, seven 1D-CNN models have been trained for a different number of epochs, all following the proposed architecture. The accuracy of these models is summarized in Table 5.2 and Table 5.3, categorized by the number of inputs and training epochs.

Using intra-participant data, the performance of the proposed 1D-CNN model varied based on the number of joint angle inputs and the number of training epochs. For the hip angle, accuracy ranged from 26.91% after 25 epochs to 97.82% following 200 epochs. Training with the knee angle ranged accuracies from 38.59% at 200 epochs to 93.21% after 250 epochs. When the ankle angle was used, accuracy ranged from 43.64% with just 10 epochs to 99.91% by 250 epochs. Combining hip and knee angles, the accuracy ranged from 38.41% at 10 epochs to 98.87% at 250 epochs. With

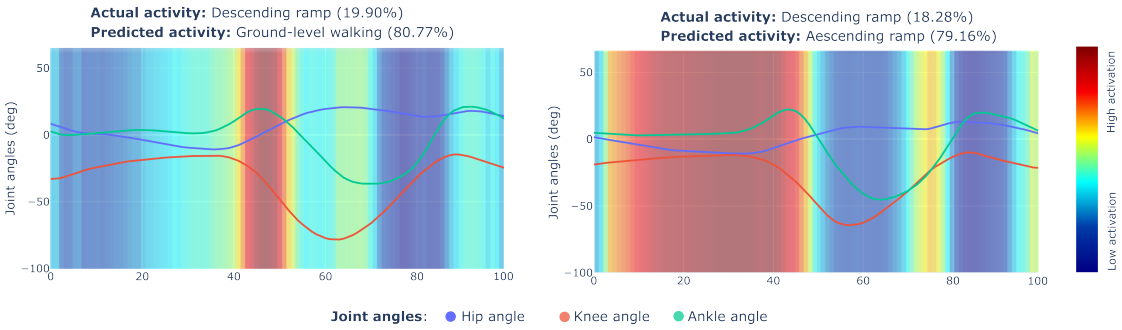


Figure 5.6: Examples of incorrect classifications with the model proposed. The representation includes the class activation maps obtained with Grad-CAM and the mean joint trajectories. Moreover, the softmax values for both the predicted and actual classes are included.

Table 5.2: Intra-participant accuracy (%) according to training epochs

	Epochs							
	10	**25**	**50**	**75**	**100**	**150**	**200**	**250**
Hip, knee, and ankle	100.00	100.00	100.00	99.91	100.00	100.00	99.65	100.00
Hip	37.54	26.91	70.73	76.05	46.34	93.73	97.82	96.52
Knee	47.30	65.85	70.12	51.74	45.12	71.52	38.59	93.21
Ankle	43.64	57.49	81.10	86.24	60.28	96.43	44.16	99.91
Hip and knee	38.41	45.56	91.46	67.07	72.30	99.39	72.83	98.87
Hip and ankle	92.42	99.65	99.91	82.32	100.00	100.00	100.00	100.00
Knee and ankle	41.20	30.49	69.60	47.91	69.33	31.97	99.91	84.84

Table 5.3: Inter-participant accuracy (%) according to training epochs

	Epochs							
	10	**25**	**50**	**75**	**100**	**150**	**200**	**250**
Hip, knee, and ankle	100.00	99.80	99.90	100.00	100.00	100.00	99.90	100.00
Hip	37.63	25.71	70.17	55.30	39.99	68.45	98.28	85.97
Knee	37.44	60.84	81.31	39.60	44.46	67.22	35.23	85.77
Ankle	37.24	46.37	89.11	85.33	59.67	98.18	34.44	97.55
Hip and knee	37.63	43.67	79.00	57.85	62.71	92.89	75.47	89.94
Hip and ankle	86.16	96.42	99.46	63.54	99.71	99.95	99.26	99.31
Knee and ankle	44.50	31.99	69.92	43.92	60.94	29.83	85.57	90.72

the hip and ankle angles, the accuracy ranged from 82.32% for 75 epochs to 100% for epochs 100, 150, 200, and 250. Lastly, using both knee and ankle angles, the accuracy of the model ranged from 30.49% at 25 epochs to 99.91% at 200 epochs.

With the inter-participant data, when the proposed model was trained with the hip angle, the accuracy ranged from 25.71% for 25 epochs to 98.28% for 200 epochs. When using the knee angle, the accuracy ranged from 35.23% for 200 epochs to 85.97% for 250 epochs. With the ankle angle, the accuracy ranged from 37.24% for 10 epochs to 98.18% for 150 epochs. By using the hip and knee angles, the accuracy ranged from 37.63% for

10 epochs to 89.94% for 250 epochs. When using the hip and ankle angles, the accuracy ranged from 63.54% for 75 epochs to 99.95% for 150 epochs. Finally, with the knee and ankle angles, the accuracy ranged from 29.83% for 150 epochs to 90.72% for 200 epochs.

5.4 DISCUSSION

A 1D-CNN model has been trained to recognize four different gait ADLs: ground-level walking, ascending a ramp, descending a ramp, and ascending stairs. The proposed 1D-CNN demonstrates promising results, achieving 99.65% accuracy for intra-participants and 99.56% accuracy for inter-participants. Additionally, the results summarized in Table 5.1 reveal that the proposed model outperforms other State of the Art architectures when handling inter-participant data. It is worth mentioning that accuracy and CE remain within a similar range across all trained models. This suggests that introducing lower limb joint trajectories can be a robust approach for gait ADLs recognition, surpassing previously proposed methods (see Chapter 2, section 2.3). Furthermore, the proposed 1D-CNN model demonstrates efficient computational performance, with inference times of approximately 1 ms on a laptop. This suggests that this model can be effectively employed for real-time HAR. Despite the high accuracy achieved by the proposed CNN model for both intra-participants and inter-participants, the confusion matrices reveal occasional misclassifications, particularly for ground-level activities and walking on ramps. To address this, exploring t-SNE and Grad-CAM methods could provide insights into the model's behavior.

The projection of the 64-dimensional feature maps extracted has been represented with data from four users in Figure 5.4. This graph represents diverse clusters, where each data point symbolizes the feature map produced by the model at every step, and the clusters correspond to sets of similar steps. Specifically, the feature maps for stair climbing obtained by the proposed model are organized into four different clusters. The remaining clusters are likely composed of one or two ADLs, which can be attributed to the similarity in joint movement patterns observed during ground-level walking and walking on a ramp (see Chapter 4, Figure 4.10). Moreover, the graph reveals that, while some clusters contain two ADLs, they are grouped into four separate clusters, which is aligned with the number of participants represented in the projection. This result suggests that each participant exhibits a distinct walking pattern, leading to unique feature maps from the CNN for each user. However, if t-SNE does not show a clear separation of the extracted features, one might wonder why such a good classification

result is achieved. This could be due to the ability of the model to discern nonlinear boundaries, allowing ADLs to be separated in the 64-dimensional space.

The t-SNE embedding method provides insights into how the features extracted by the CNN are grouped. However, this method cannot be used to determine which segments of the joint trajectories play a crucial role in identifying the performed activity. The Grad-CAM method can be used to highlight which parts of the joint trajectories are critical to the proposed model to classify the different gait ADLs. Figure 5.5 represents the mean class activation maps together with the mean hip, knee, and ankle trajectories for each activity. For instance, during stair climbing, the highest activation occurs between 70% and 90% of the gait cycle. By comparing the hip, knee, and ankle trajectories during these periods with those of other gait ADLs, significant differences can be observed. However, despite this analysis, the t-SNE graph and confusion matrices reveal that incorrect classifications are more frequent during ground-level walking, ascending a ramp, and descending a ramp. Specifically, certain steps while descending a ramp are misclassified as ground-level walking or ascending a ramp.

The Grad-CAM method can offer valuable insights into understanding the incorrect predictions of the CNN model proposed. By using joint trajectories and activation maps, two incorrect classifications can be analyzed (Figure 5.6). The graph on the left shows a step produced while descending a ramp, although it is classified as ground-level walking. The class activation map reveals significant activation between 40% to 50%. This activation aligns with the peak of the ankle angle, indicating that the gait pattern is similar to ground-level walking. The reason for this similarity relies on the fact that the ankle angle exceeds the hip angle at this moment. This occurrence does not happen when descending a ramp. When analyzing the step classified as ascending a ramp, high activation can be observed during the loading and stance phase (10% - 50%). This activation is aligned with the activation map associated with ascending ramp steps and may be attributed to the similarity in knee trajectory during this specific section of the gait cycle.

In addition, the feasibility of reducing the number of inputs to the 1D-CNN was also studied to adapt the model to multiple situations. The results from Table 5.2 and Table 5.3 demonstrate that high accuracy can still be achieved even with fewer inputs. Although the accuracy is slightly lower compared to when considering hip, knee, and ankle flexion/extension, these findings suggest that the model can be adapted to various scenarios while maintaining an accuracy of over 90%. Therefore, the 1D-CNN classifier proposed could be used to recognize different activities while analyzing the

biomechanics of a single lower limb joint and detect gait activities for different types of lower limb exoskeletons.

5.5 CONCLUSION

This Chapter presents a Deep Learning model to perform HAR during gait. The classifier relies on a 1D-CNN and employs the flexion/extension angles of the hip, knee, and ankle, which are measured using the IMU-based motion capture system described in Chapter 3, section 3.2. Notably, the method employed in this study to measure lower limb angles is independent of sensor position or location, ensuring a highly reliable classification approach.

The proposed CNN classifier demonstrates exceptional accuracy for both intra-participant (99.65%) and inter-participant (99.56%) data, indicating a high generalization. To understand the decision-making process of the model, the t-SNE and Grad-CAM algorithms have been used. The analysis led to the conclusion that the model can find nonlinear boundaries so that each ADL can be separated in an n-dimensional space. In addition, it can be concluded that misclassifications are attributable to either unusual step patterns or a significant similarity in joint movement trajectories across different activities.

Furthermore, diverse pieces of training for the proposed models were conducted while varying the number of inputs to study the feasibility of adapting the model to multiple scenarios. The results suggest that the proposed 1D-CNN architecture can be adapted for analyzing the biomechanics of an individual lower limb joint and detecting gait activities across various types of lower limb exoskeletons.

6

A GENETIC ALGORITHM-BASED METHOD TO MODULATE THE DIFFICULTY OF SERIOUS GAMES ALONG CONSECUTIVE ROBOT-ASSISTED THERAPY SESSIONS

6.1 INTRODUCTION

As described in Chapter 3, the SPLASH project not only proposes to design an underwater lower limb exoskeleton but also aims to develop end-effector robotic solutions based on the novel hybrid actuators proposed within the context of the project (see Chapter 3, Figure 3.4).

As introduced, the end-effector rehabilitation robots are commonly used together with serious games to enhance patient engagement and motivation (Heins et al., 2017; Catalán; Blanco-Ivorra, et al., 2023a). However, it must be taken into account that rehabilitation therapies should be adapted to each patient (Faria et al., 2020). Therefore, the difficulty of the serious games employed during robot-aided rehabilitation therapies must be adapted.

The DDA systems described in Chapter 2, section 2.4 were designed to adapt serious games according to the user's abilities, the psychophysiological state, or the game score. However, it is noteworthy that the conducted studies were only evaluated in a single session, which does not demonstrate the feasibility of employing the proposed systems for consecutive rehabilitation sessions. Hence, as the ability of the patient is expected to change across the sessions, there is a need to propose a DDA system for adapting the difficulty of the serious games while avoiding a recalibration before starting the robot-aided rehabilitation sessions.

In this Chapter, a genetic algorithm-based DDA system has been proposed to enhance the user to perform intensive rehabilitation therapy while maintaining a determined game score to keep him/her engaged. Based on these objectives, the system proposed searches for which parameters of the serious game are optimal for the user. In addition, the DDA system designed is intended to be used during consecutive sessions. To avoid recalibration, the parameters that best suited the user in the previous session

have been introduced as the starting point for the next session. Furthermore, a recent study demonstrated that an increase in the difficulty of the rehabilitation exercises has a beneficial effect on motor recovery (Maier et al., 2019). Hence, in the study conducted to evaluate the proposed DDA system, the desired game score has been reduced between sessions in order to increase the difficulty of the games. Additionally, the Self-Assessment Manikin (SAM) test together with different physiological signals have been used to evaluate the psychophysiological state during the sessions.

6.2 MATERIALS AND METHODS

6.2.1 *Experimental setup*

The experimental setup shown in Figure 6.1 has been employed to conduct the experiments to develop and evaluate the DDA system proposed. The Rubidium robot was employed to interact with the serious games, developed by the research group and commercialized by the spin-off iDRhA (iDRhA, n.d.). Furthermore, two additional devices were utilized to capture diverse physiological signals.

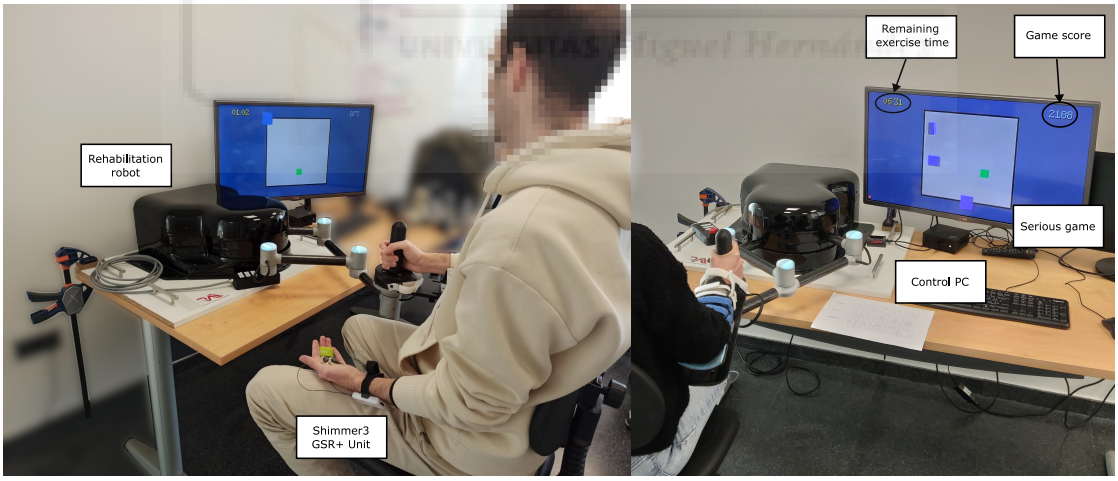


Figure 6.1: Setup employed during the experimental sessions to evaluate the DDA system proposed.

6.2.1.1 *Robotic rehabilitation system*

The Rubidium robot, previously used in research studies of the group (Vales; Catalán, et al., 2023; Catalán; García-Pérez; Blanco; Martínez, et al., 2021; Catalán; Blanco-Ivorra, et al., 2023b; Vales; Catal, et al., 2022), was used to play serious games during the experimental sessions. This portable robotic platform focuses on upper limb rehabilitation and is composed of two actuated degrees of freedom. Its design includes an articulated parallelogram mechanism driven by two electric motors, enabling movements within a horizontal plane. In addition to the robot, a control PC was employed to execute the rehabilitation games and run the DDA system.

The serious games utilized during the experimental sessions are represented in Figure 6.2, referred to as Game A and Game B. In these games, participants can control a green square through the use of the Rubidium rehabilitation robot. It is worth mentioning that the interface provided real-time feedback to the participants by displaying the remaining time for the exercise and the current game score.

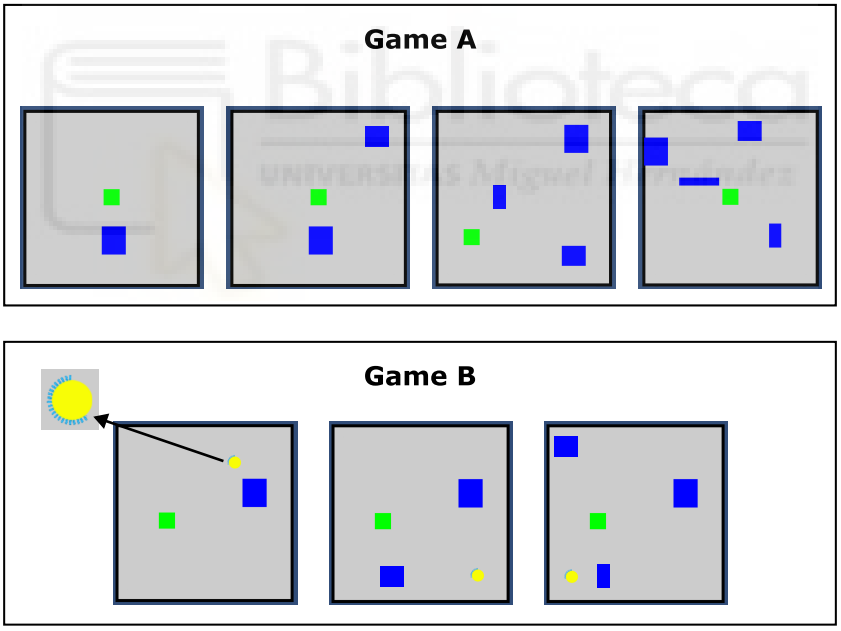


Figure 6.2: The serious games used during the experimental sessions. The green square can be controlled using the Rubidium robot. In Game A, the objective is to evade the enemies (represented by blue rectangles) while remaining within the limits of the safe area (gray square). In Game B, the user must reach the target (yellow circle) while avoiding the enemies.

In Game A, the player needs to avoid the adversaries, represented as blue rectangles, while remaining within the designated safe zone, a grey square. While within this zone, the score increases by 10 points/second. If the green square remains outside the safe zone, the score stops increasing. The contact with an enemy not only prevents score gains but also results in a deduction of 5 points/second.

The objective in Game B is to reach the yellow targets while the user avoids the blue rectangles. Each target appears at a random location within the designated safe area and remains visible for 5 seconds. The maximum score awarded by each target is 100 points, which is indicated by a blue circle surrounding the target. The score awarded for each target decreases at a rate of 20 points/second. Therefore, reaching targets quickly contributes to a higher overall score. If the green square remains outside the safe zone, the score decreases by 10 points/second. The contact with an enemy results in a deduction of 50 points.

The proposed serious games allow for dynamic adjustments of two parameters to adapt the difficulty of the game to each user. These parameters are the speed of the enemies v_e and the number of enemies n_e. In Game A, it was let $n_e \in \{1, 2, 3, 4\}$. However, in Game B, due to the high difficulty of achieving objectives with four enemies, it was decided to let $n_e \in \{1, 2, 3\}$. Finally, for both games, the speed of the enemies was let $v_e \in [0.1, 2.0]$.

6.2.1.2 *Physiological measurements*

In addition to the Rubidium robot, a Shimmer3 GSR+ Unit has been employed to acquire Galvanic Skin Response (GSR) measurements. Furthermore, a Zephyr BioHarness was used to acquire cardiorespiratory responses. The data collected from both devices serve as conventional metrics within psychophysiological paradigms during rehabilitation (Palumbo et al., 2021).

The Shimmer3 GSR+ Unit incorporates a built-in signal-processing module that handles measurement acquisition, signal cleaning, and subsequent transmission via Bluetooth. To capture the GSR signal, two circular adhesive electrodes with pre-gelled Ag/AgCl material, each having a diameter of 26 mm, were strategically positioned on the proximal phalanges of the index and middle fingers of the non-dominant hand.

The Zephyr BioHarness incorporates an internal unit that handles measurement acquisition, signal processing, and transmitting via Bluetooth the ECG and the respiratory signal of the user measured with the elastic belt.

6.2.2 *Therapy performance definition*

The objective of the DDA system during therapy is to maximize the therapy performance η. This performance metric is influenced by both the movements made during the game and the score achieved. To achieve this goal, η is evaluated within a time window $t_w \in [0, k]$, where k represents the number of samples in the window. Let η be defined as:

$$\eta = \sum_{i=0}^{n} \omega_i \cdot \alpha_i \tag{6.1}$$

where α_i are the evaluations of the user movements or score, and ω_i are constants used as weights of the proposed metrics.

Given the unique dynamics of each game, it is proposed to employ different definitions of η for each game, which are detailed in the following subsections.

6.2.2.1 *Game A therapy performance formulation*

Regarding the characteristics of Game A, the objective is to evade the enemies while staying within the designated safe area. However, the scoring mechanism can allow the user to increase the score for maintaining a position in the safe zone whether no enemy contacts the green square. Therefore, the definition of η will incorporate terms to assess the amount of game space covered by the user and the duration of these movements to avoid the user remaining still. In addition, to ensure that the game has an appropriate difficulty level, a term related to the desired score is also integrated into the definition of η. Thus, η_A is defined as:

$$\eta_A = w_s \cdot \alpha_s + w_g \cdot \alpha_g + w_t \cdot \alpha_t \tag{6.2}$$

where α_w is the rate of covered game space, α_g is the weighted score relative to the desired score, and α_t is the move rate during the evaluation window. It was let w_s, w_g, and w_t to be equal with a value of $1/3$.

The term α_s can be defined by segmenting the safe zone of the game into a grid composed of i rows and j columns, whose value is determined by calculating the ratio of the number of squares traversed to the total number of squares in the grid n_s, and the value of each square being $s(i, j)$:

$$\alpha_s = \sum_{i=0}^{n} \sum_{j=0}^{n} \frac{s_{i,j}}{n_s} \tag{6.3}$$

Concerning α_t, the aim is to evaluate whether the user is performing a continuous movement or he/she is remaining still. This can be determined by setting a velocity threshold v_r, below which the user is considered to be at rest. Given that the game operates within a 2D workspace, the overall velocity of the end-effector of the robot can be computed as:

$$v = ||\vec{v}|| = \sqrt{v_x^2 + v_y^2} \tag{6.4}$$

With v, the user's movement rate during the evaluation window can be estimated as:

$$\alpha_t = \sum_{i=0}^{k} \frac{v_i \cdot I(v_i)}{k} \tag{6.5}$$

where $I(v_i)$ is a piecewise function defined as:

$$I(v_i) = \begin{cases} 0 & \text{if } v_i \leq v_r \\ 1 & \text{if } v_i > v_r \end{cases} \tag{6.6}$$

As introduced, during rehabilitation therapies is essential to maintain a balance in therapy difficulty to challenge and motivate the patient. Hence, the difficulty of the game should allow for errors while maintaining the state of flow. Consequently, a high score may not reflect an optimal performance within the DDA system. To address this, a metric that evaluates the proximity of the user's score to a predefined target score g_o can be integrated into the definition of η.

The game score rate g_r can be expressed as the relation between the score obtained and the maximum possible score in the time window:

$$g_r = \sum_{i=0}^{k} \frac{g_{s,i} - g_{s,i-1}}{r \cdot (t_{w,i} - t_{w,i-1})} \tag{6.7}$$

where r is the ratio of punctuation obtained per second (100 points/s) and g_s is the score achieved at each instant of t_w.

By using g_r a Gaussian function can be utilized to weight the score rate obtained by the user concerning g_o. With this aim, a Gaussian function was defined with mean $\mu = g_o$ and standard deviation $\sigma = 0.25$:

$$\alpha_g = e^{-\frac{(g_r - g_o)^2}{2 \cdot \sigma^2}} \tag{6.8}$$

6.2.2.2 *Game B therapy performance formulation*

In Game B, the goal is to reach targets within the safe zone to increment the score while evading the enemies. This game mechanic differs from Game A in that it requires the player to move to increase the score. It is also worth mentioning that targets will spawn at random locations within the safe zone, meaning the player's coverage of the game space will be largely influenced by the position and the number of targets the player successfully reaches. As a result, the terms of η that describe the player's movement need to be redefined.

In Game B, it can be assumed that a specific set of game parameters forces the user to execute a longer path to reach the targets. Moreover, with the intention of sustaining a particular score level throughout the exercise, the therapy performance for Game B can be defined as:

$$\eta_B = w_p \cdot \alpha_p + w_g \cdot \alpha_g \tag{6.9}$$

where α_p is the path overlength to reach the targets, and α_g is the weighted score relative to the desired score. The w_p and w_g values are established equal to 0.5.

First, the minimum distance needed to reach the targets d_t can be estimated as:

$$d_t = \sum_{i=0}^{n_t - 1} \sqrt{(x_{t,i+1} - x_{t,i})^2 + (y_{t,i+1} - y_{t,i})^2} \tag{6.10}$$

where x_t, y_t represents the Cartesian coordinates of the targets and n_t the number of targets reached.

In addition, the path length taken to reach the targets can be calculated as:

$$d_p = \sum_{i=0}^{k-1} \sqrt{(x_{p,i+1} - x_{p,i})^2 + (y_{p,i+1} - y_{p,i})^2} \tag{6.11}$$

where x_p, y_p represents the position of the green square for each instant k in Cartesian coordinates.

Once d_t and d_p are estimated, the path overlength caused by the presence of the enemies can be calculated as:

$$\alpha_p = \frac{d_p - d_t}{d_t} \tag{6.12}$$

As for the Game A, the game score for each analysis window g_r can be calculated as the ratio of the obtained score to the maximum possible score within that time window. In Game B, g_r is defined as:

$$g_r = \sum_{i=0}^{n_t} \frac{g_{t,i}}{m_t} \tag{6.13}$$

where n_t is the number of targets reached, g_t is the score given by each target, and m_t is the maximum possible score given by each target.

Given g_r, a Gaussian function is defined to weight the score rate regarding the desired game score g_o:

$$\alpha_g = e^{-\frac{(g_r - g_o)^2}{2 \cdot \sigma^2}} \tag{6.14}$$

6.2.3 *Dynamic difficult adjustment system design*

For a personalized game difficulty level, the optimal values of v_e and n_e need to be searched to maximize η. In this section, genetic algorithms are introduced to develop a DDA system. Furthermore, the usage protocol is described to employ the designed DDA system across consecutive therapy sessions.

6.2.3.1 *A genetic algorithm as a method to develop a DDA system*

Genetic algorithms are an optimization technique that imitates natural selection and evolution (Kramer and Kramer, 2017; Lambora et al., 2019; Mirjalili and Mirjalili, 2019). They can be used to find effective solutions for complex problems that are challenging to solve using other conventional methods. These algorithms operate by generating and adjusting a population of potential solutions referred to as individuals. The fitness of each individual is evaluated using a predefined function. Those individuals who are best suited to the environment have a higher chance of survival and reproduction,

passing on their genes to the next generation. This iterative process, known as elitist selection, drives the population toward improved solutions over time.

To find the optimal solution, genetic algorithms employ a key mechanism known as crossover. This process involves combining the genetic material from two or more individuals. In this context, the crossover mechanism merges the v_e and n_e genes of individuals that exhibit higher values of η.

Following the crossover, genetic algorithms implement a mutation mechanism. This process involves random alterations to a solution by mutation operators, thereby injecting new genetic material into the population. The main purpose of this mechanism is to avoid converging on a local minimum, ensuring a diverse gene pool among the individuals (Hassanat et al., 2019). In the context of the games proposed, considering that v_e is a continuous variable, a mutation operator that introduces a noise factor x_v with a uniform probability has been employed:

$$x_v \sim U(-0.25, 0.25) \tag{6.15}$$

By contrast, considering that n_e is a discrete parameter, a mutation operator that either adds or removes an enemy with a probability σ_n has been used. It is worth mentioning that introducing or eliminating an enemy can significantly impact the game's difficulty. To ensure the stability of the game difficulty and facilitate convergence toward an optimal solution, the mutation probability σ_n was dynamically adjusted based on the session time. This approach, known as deterministic control, has been used in previous studies (Bäck and Schütz, 1996; Fogarty, 1989; Karafotias et al., 2014).

The implemented genetic algorithm used to adapt dynamically the serious games during the experimental sessions has been represented in Figure 6.3.

6.2.3.2 *Usage of DDA system protocol for different sessions*

The protocol followed to dynamically adapt the serious games across the consecutive experimental sessions is represented in Figure 6.4.

For the first session, the target score rate g_o is established at 80%. Additionally, a standard initial population is created and shared across all users for each game. This population is derived by assigning $v_e \in \{0.25, 0.50, 0.75, 1.00\}$, ensuring a broad spectrum of speeds along with the possible n_e values. Each individual undergoes a 20-second evaluation period. Following this initial assessment, the two best individuals

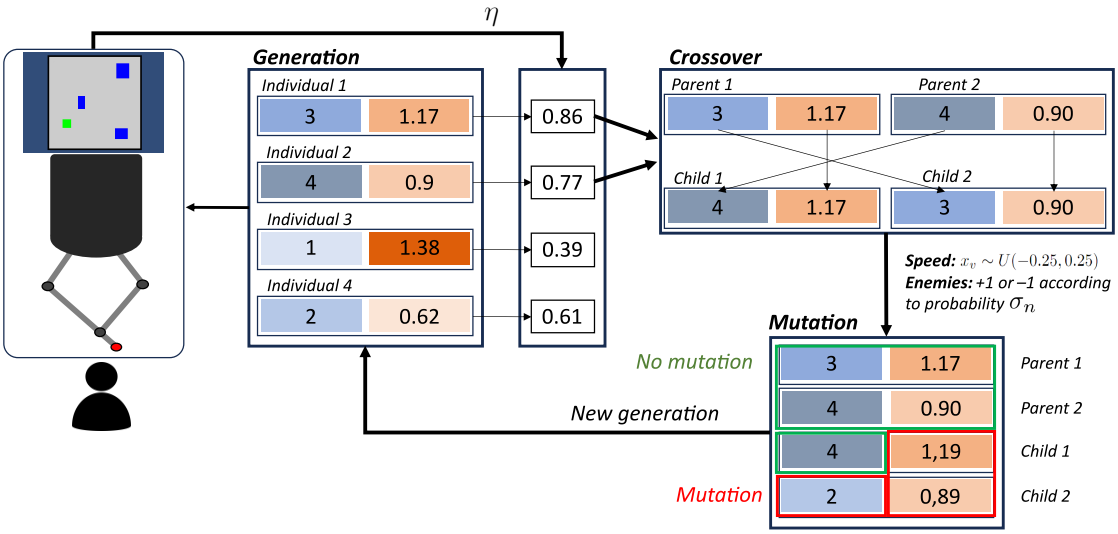


Figure 6.3: The genetic algorithm implemented is used to dynamically adapt serious games to find the game parameters that maximize η. A population of individuals is created, with each individual containing two genes: the number of enemies n_e (blue) and the speed of the enemies v_e (orange). Each individual is introduced into the game, and η is evaluated for a period t_w. After evaluating the entire population, the best individuals are selected, and the crossover and mutation operators are applied. A new generation is created, and the process is repeated with the individuals from the new population.

are selected based on η. Subsequently, the crossover operation is applied, resulting in the generation of two new individuals.

Once the new individuals are created, mutation operators are applied. In the case of v_e, a uniform noise has been introduced:

$$v'_e = v_e + x_v \tag{6.16}$$

For n_e, the mutation probability σ_n was set to 50% at the start of each session. This choice ensures that the algorithm can explore a diverse range of game parameter combinations. As new generations evolve, σ_n is gradually reduced to achieve stable difficulty levels during the session:

$$\sigma = 0.5 \cdot e^{0.05 \cdot p_i} \tag{6.17}$$

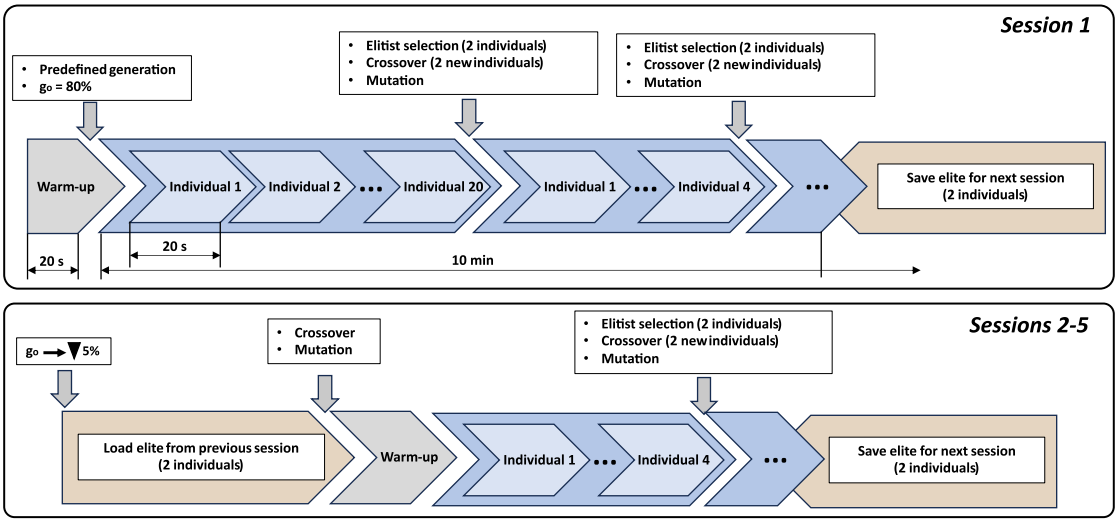


Figure 6.4: The protocol for dynamically adjusting the serious games across various sessions begins with utilizing a pre-established population to assess the user's capabilities during the initial session. Following this preliminary assessment, the parameters with the highest η are identified. The next step involves the application of crossover and mutation processes. To maintain continuity and progress, the elite individuals from each session are preserved and used as the initial population for the subsequent session.

where p_i is the i-th generation. Furthermore, a baseline mutation probability of 25% has been established. In case a mutation is applied to n_e, the number of enemies will be randomly increased or decreased by one unit.

After the application of crossover and mutation operators, the newly formed generation undergoes evaluation, and this cycle is repeated until the session concludes. The final elite is preserved for the next session.

For subsequent sessions, the target score rate g_o is decreased by 5%. An initial group of four individuals is then created based on the elite saved from the previous session. Throughout these sessions, mutation operators continue to be used to explore changes that can improve the therapy performance.

6.2.4 *Experimental sessions*

The conducted study involved 13 able-bodied users (8 males and 5 females), aged between 20 and 44 years (28.7 ± 5.7 years), with body masses ranging from 49 to 85 kg (67.3 ± 10.9 kg) and heights ranging from 159 to 187 cm (172.2 ± 8.6 cm). The

participants provided written, informed consent before starting the experimental sessions. This study was approved by the Hospital La Pedrera Ethical Committee and registered with reference number HLP-35/2021.

The experiments were carried out for five consecutive days to evaluate the proposed DDA system. The participants were consistently scheduled for the same time slot throughout the week, from Monday to Friday. At the beginning of the initial session, it was clarified to the users that they would realize two different exercises using the Rubidium rehabilitation robot, spread across five separate sessions. While explaining the activities, it was indicated that both the quantity and velocity of the enemies would vary. It is important to highlight that details regarding the mechanics and rationale behind the variability of these game parameters were not mentioned to prevent any potential influence on the users' in-game behavior.

Before each exercise, the physiological signals were recorded for three minutes at rest to establish a baseline state. The users played Game A and Game B for 10 minutes, and 20 seconds were left as a warm-up. The sequence of the games was shuffled for every session and all participants. Following each exercise, participants completed the SAM test (Bradley and Lang, 1994).

6.2.5 *Data acquired and processing*

A set of data collected with the Rubidium robot was employed to evaluate the DDA system proposed. Specifically, the end-effector position in Cartesian coordinates is collected to assess the movements of the participants during games.

In addition to evaluating the DDA system, it is also intended to analyze the psychophysiological state of the participants during sessions. With this aim, the SAM test was used to collect the self-reported emotional state of the users, which is an effective measure for self-reporting emotion recognition. In the test, the emotions are rated on a nine-point scale by three dimensions: valence, arousal, and dominance.

In addition to SAM, GSR and cardiorespiratory data were collected. The physiological signals have been cleaned and processed using the Python toolbox NeuroKit2 (Makowski et al., 2021).

The Zephyr BioHarness device acquired ECG and respiratory signals at a frequency of 250 Hz. The RFQ was computed through the local maxima and minima caused by

inspirations and expirations. Similarly, the R peaks in the ECG signal were identified to calculate the HR and Heart Rate Variability (HRV) indices, which are indicative biomarkers of both physiological and psychological states (Gilgen-Ammann et al., 2019). Typically, a lower HR is associated with relaxation or light exercise, whereas higher HR values are indicative of more intense physical activity. HRV is commonly utilized as an indicator of mental workload and stress levels associated with various activities. The time-domain metrics, particularly the Root Mean Square of Successive Differences between normal heartbeats (RMSSD), have been extensively used in HRV studies and are considered a reliable measure of the vagal tone. (Laborde et al., 2017; Rajendra Acharya et al., 2006). In addition, low parasympathetic activity, characterized by a decrease in High-Frequency (HF) power, is frequently reported as a factor associated with variation in HRV variables (Kim et al., 2018). It is generally observed that higher HRV values suggest a state of lower stress or intensity.

The Shimmer GSR+ device acquired the GSR signal at a sampling rate of 512 Hz. The GSR signal can be divided into two different components: the Skin Conductance Level (SCL), which represents the baseline level of skin conductance in the absence of any discrete environmental event, and the Skin Conductance Responses (SCR), which are transient increases in skin conductance. From these signals, the mean SCL and the SCR frequency values were calculated.

To minimize the differences across participants and sessions, the physiological measurements were adjusted to their respective mean baseline values. During the three-minute rest period, the mean basal value was obtained from the minute preceding the start of the activity. For the SCR frequency, the baseline value was subtracted from the signal to standardize the data. The rest of the physiological signals were normalized as:

$$x_{norm} = \frac{x - x_{baseline}}{x_{baseline}} \tag{6.18}$$

6.2.6 *Statistical data analysis*

A statistical data analysis was conducted to compare the variables involved in the DDA architecture and the psychophysiological states across different sessions. Significant results are considered for $p < 0.05$. The analysis included a normality test using the Shapiro-Wilk test, which indicated that the distribution of certain variables

deviated from normality. Consequently, non-parametric methods have been used to identify differences among groups of variables.

The Friedman test was used to study the differences among sessions. Subsequently, in the post-hoc analysis, pairwise comparisons were performed using the Wilcoxon signed-rank test with the zero method proposed by Pratt. In the pairwise comparisons, special interest has been taken into account between the first and the rest of the sessions. Consequently, four pairwise comparisons are analyzed.

6.3 RESULTS

6.3.1 *Exercise adaptation*

The heatmaps collected in Figure 6.5 represent the activity of one user playing Game A and Game B for predefined difficulty levels during the initial assessment. In Game A, the user's movement and coverage of the game space are influenced by the game parameters. For instance, with low values of n_e and v_e the user may remain stationary. However, with higher values more movement is required to avoid the obstacles. By contrast, in Game B, the game's parameters appear to have low impact on how the user

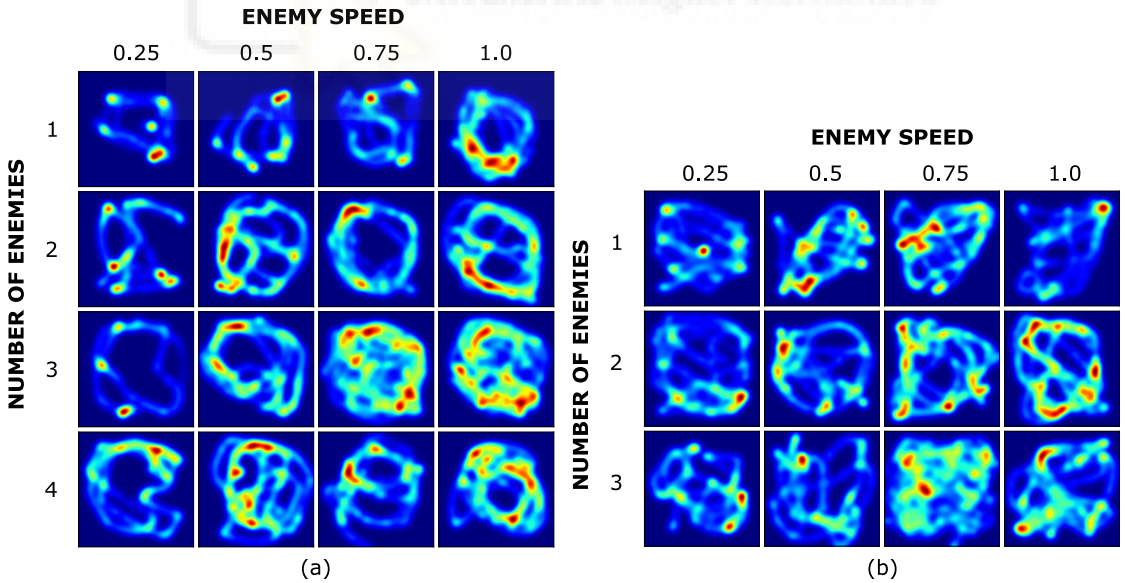


Figure 6.5: The heatmaps represent the movements of one user for different game difficulties during **(a)** Game A, and **(b)** Game B.

navigates the game space, probably due to the random emergence of targets within this mode of play.

Figure 6.6 represents two examples of the progression of η_A and η_B throughout the initial and the last session. Each of the points included in the graph represents the value of η, each reflecting an evaluation over the preceding 20-second interval. Moreover, trendlines have been calculated and included in the graphs using the Locally WEighted Scatterplot Smoothing (LOWESS) method, with a smoother span of 0.2.

The mean and standard deviation of each of the defined game metrics in section 6.2.2 has been represented in Figure 6.7 and summarized in Table B.1. According to the Friedman test, a significant increase in η ($p = 0.0019$), α_s ($p < 0.0001$), and α_t ($p < 0.0001$) is shown for Game A. Regarding the Game B, the Friedman test indicates that there is a significant increase in both η_B ($p < 0.0001$) and α_p ($p = 0.0019$). For a detailed comparison, the results of pairwise comparisons obtained using the Wilcoxon test are summarized in Table 6.1 for both games.

Figure 6.8 represents a violin plot with the average values of n_e and v_e across sessions. For a more detailed analysis of the data distribution of these parameters, refer to Table B.2 for Game A and Table B.3 for Game B.

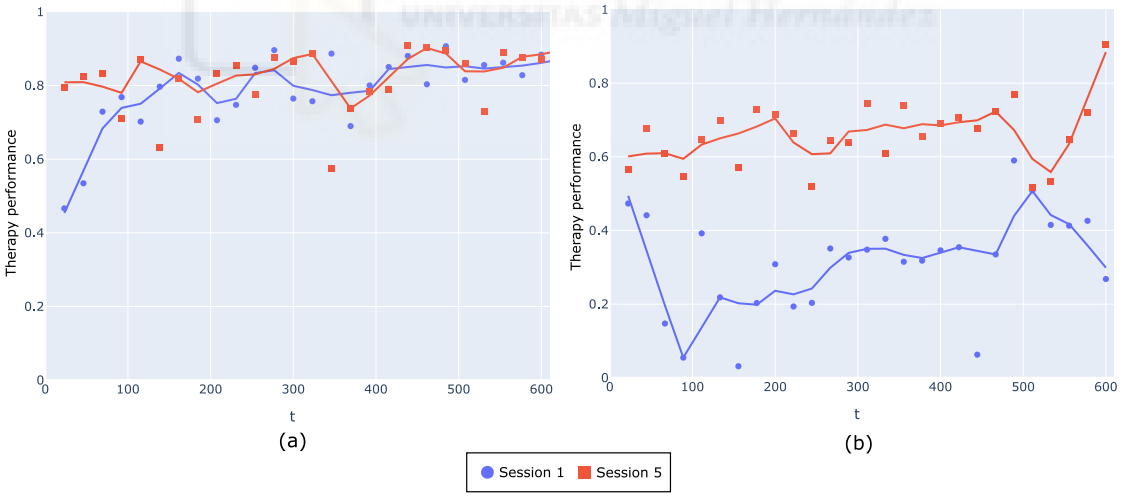


Figure 6.6: Therapy performance (η) evolution during the initial and the last session for **(a)** Game A, and **(b)** Game B. The included trendlines have been calculated using the LOWESS method, with a smoother span of 0.2

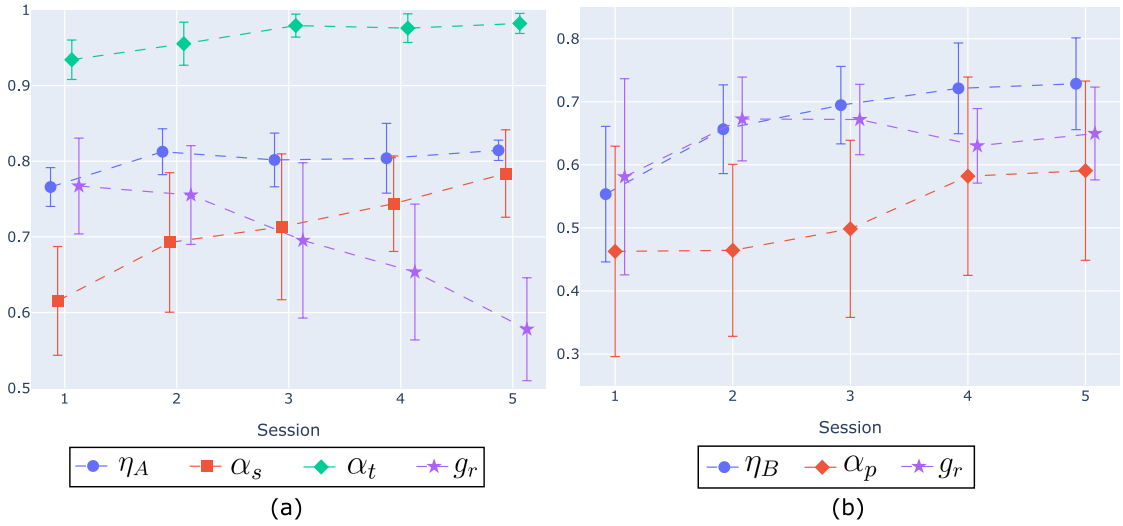


Figure 6.7: Graphical representation of the mean and standard deviation of therapy performance (η), movement metrics, and score rate during the five sessions for **(a)** Game A, and **(b)** Game B.

Table 6.1: Wilcoxon signed-rank test results for therapy performance measurements. The p-values of the pairwise comparisons between the first and the n session are summarized.

		S1-S2	S1-S3	S1-S4	S1-S5
Game A	η_A	0.0093	0.0161	0.0210	0.0005
	α_s	0.0640	0.0015	0.0005	0.0005
	α_t	0.0771	0.0024	0.0005	0.0005
Game A	η_B	0.0002	0.0002	0.0002	0.0007
	α_P	0.7869	0.3396	0.0105	0.0681

6.3.2 *Psychophysiological state*

The results of self-reported emotions during sessions, assessed using the SAM test, are summarized in Figure 6.9 and Table B.4. For Game A, the Friedman test indicates no significant differences in valence ($p = 0.1058$), arousal ($p = 0.9861$), and dominance ($p = 0.6895$). However, for Game B, there are significant differences in arousal ($p = 0.0399$) among sessions, while no significant differences are observed in valence ($p = 0.2878$) and dominance ($p = 0.2164$). In the pairwise comparison, significant differences

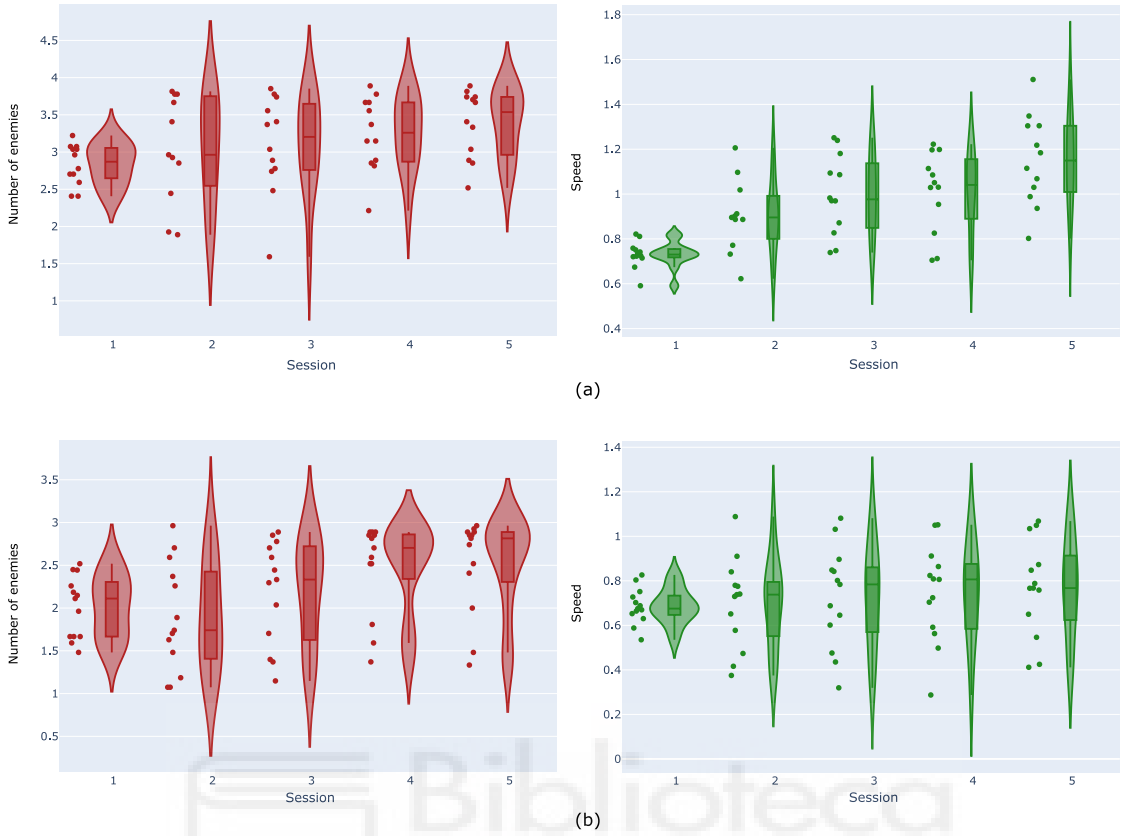


Figure 6.8: Violin plots representing the mean values of n_e and v_e across sessions for **(a)** Game A, and **(b)** Game B.

were found in arousal between the first and the third session ($p = 0.0209$), and the first and the fifth session ($p = 0.0215$).

The cardiorespiratory responses are represented in Figure 6.10. No statistical differences between sessions were found in HR ($p = 0.2548$), RF ($p = 0.3084$), RMSSD ($p = 0.6033$), or HF ($p = 0.2805$) for Game A across the experimental sessions. The same results apply to Game B, where statistical differences were not found across sessions in HR ($p = 0.3971$), in RF ($p = 0.3084$), in RMSSD ($p = 0.1291$), and HF ($p = 0.0681$).

The average values of SCL and the SCR frequency have been represented in Figure 6.11 across sessions for both serious games. In the case of Game A, no significant differences were found in SCL ($p = 0.0645$) or SCR frequency ($p = 0.1712$) among sessions. A similar pattern was noted for Game B, where no significant changes were observed in SCL ($p = 0.1669$) and SCR frequency ($p = 0.4530$) among the sessions.

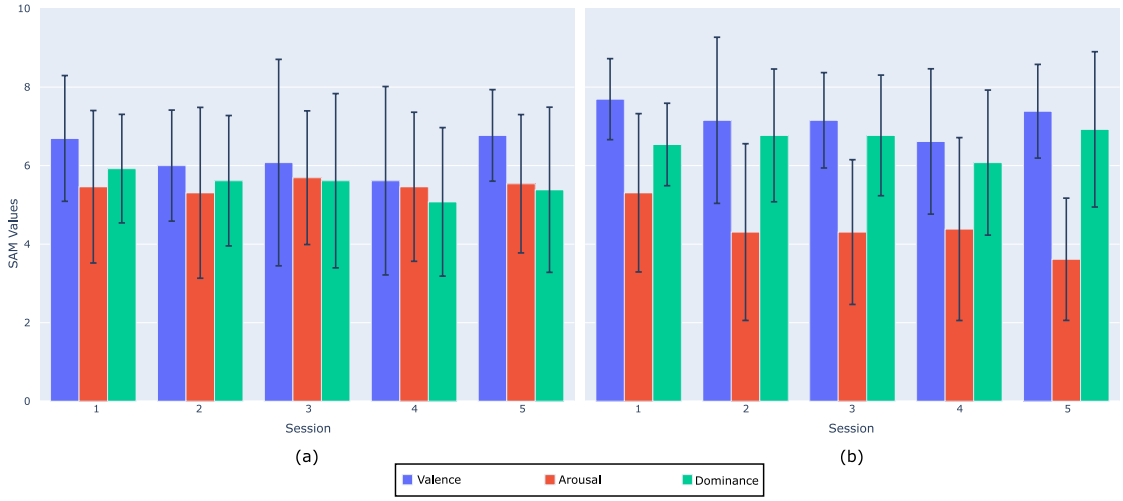


Figure 6.9: Graphical representation of the mean and standard deviation of the SAM test across sessions for **(a)** Game A and **(b)** Game B.

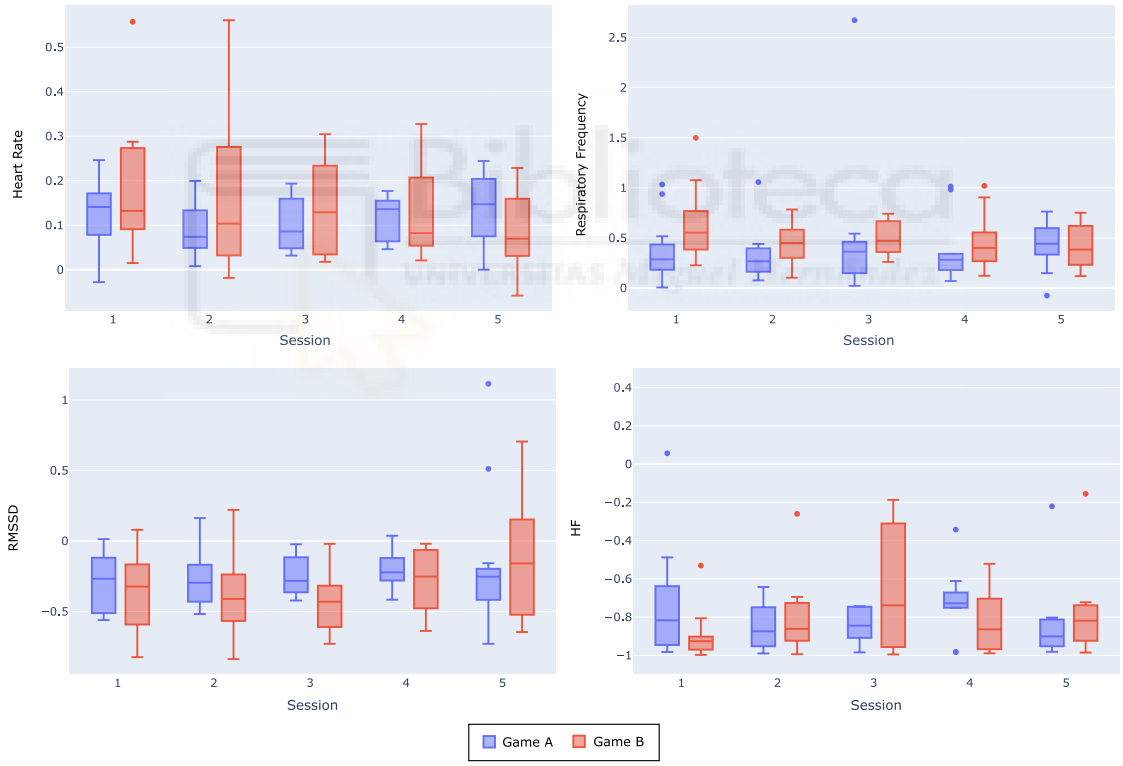


Figure 6.10: The normalized results for cardiorespiratory responses are presented through boxplots. These graphs represent HR, RFQ, RMSSD, and HF power of successive RR intervals, categorized by session for both Game A and Game B.

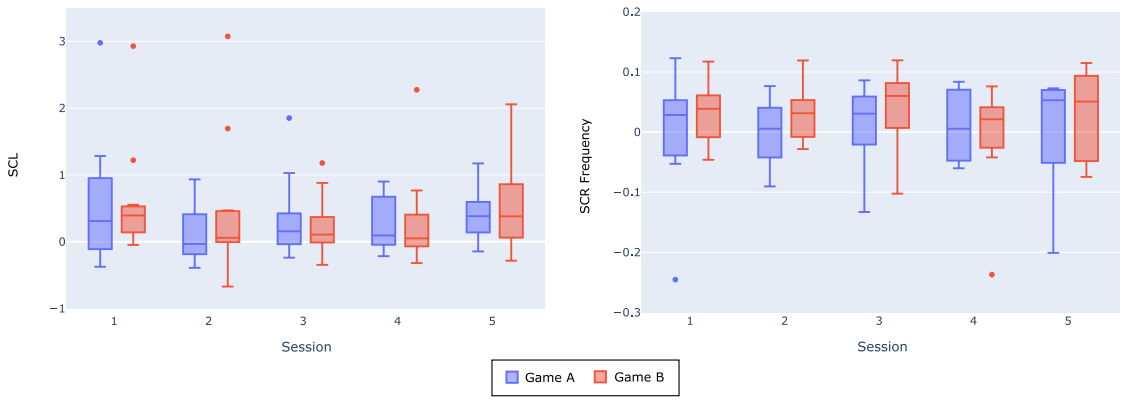


Figure 6.11: The normalized results for GSR responses are presented through boxplots. These graphs represent the SCL and the SCR frequency, categorized by session for both Game A and Game B.

6.4 DISCUSSION

In neurorehabilitation therapies, adapting the intensity and difficulty to each patient is crucial to keep the patients motivated and enhance the motor recovery of the affected limbs (Kwakkel, 2006; Winstein et al., 2016). A DDA system has been proposed to adjust robotic therapy exercises to encourage patient movement and maintain an optimal challenge level. To ensure that the difficulty level is appropriate for each user, it is proposed to guarantee a certain success percentage during the sessions. It is worth mentioning that determining an appropriate difficulty or success rate to enhance engagement can be complex, involving variables like the patient's skill level and preferences. Despite this challenge, recent research demonstrated that gradually increasing the difficulty of exercises can positively impact motor recovery (Maier et al., 2019). Hence, this study proposes to evaluate the developed DDA system and the psychophysiological responses when applying a success trajectory that begins at 80% success in the initial session and decreases to 60% by the final session.

In an attempt to personalize the difficulty of serious games, a genetic algorithm-based DDA system has been proposed. The method aims to determine optimal values for two parameters (n_e and v_e), based on η_A (Equation 6.2) and η_B (Equation 6.9). Both η_A and η_B incorporate terms that emphasize game parameters related to user movement. Specifically, they give more weight to parameters that encourage active participation. To demonstrate the importance of the terms α_s and α_t in Game A, and α_p in Game B, the movements of one user for some predefined game parameters are represented

in Figure 6.5. In Game A, it can be observed that some parameter combinations (e.g. $n_e = 1$ and $v_e = 0.25$) allow the user to remain relatively stationary and cover a small game space. In contrast, other parameter combinations (e.g. $n_e = 3$ and $v_e = 0.75$) seem to force the user to cover a larger game area. Regarding Game B, the heatmaps suggest that parameters do not significantly impact the amount of game space covered. This behavior may be attributed to the presence of randomly appearing targets that users must reach. Notably, certain parameter values (e.g. $n_e = 3$ and $v_e = 0.75$) result in more evenly distributed heat across the game area, suggesting continuous movements or higher speed.

Figure 6.6 represents two examples with the progression of η_A and η_B for one user during the initial and final experimental session. At the beginning of the first session, the initial assessment is conducted to estimate the user's skills. It can be observed that η increases when v_e and n_e are aligned with the user's needs, while it decreases when the values deviate from the optimal parameters. Following the initial evaluation, the implemented genetic algorithm searches for a combination of n_e and v_e to maximize η. During the fifth experimental session, η exhibits a more stable value compared to the first session. In addition, it is worth mentioning that some instances of η are abnormally low. This phenomenon can be attributed to the inherent characteristics of the algorithm, particularly when mutations in v_e and n_e prove to be unsuitable for the user. However, an ascending trend in η is noticeable throughout both sessions for each game.

The mean and standard deviation of the terms defining η are represented in Figure 6.7 and summarized in Table B.1. In Game A, there is a marked improvement in η_A during the first session, which is maintained until the last session. The strategy implemented significantly increased α_s, leading to a significant expansion of the game space covered by users (from 62% in the first session to 78% in the final session). In addition, α_t maintains a high level, causing users to move from 93% of the time in the first session to 98% in the last two. Furthermore, it can be observed that the DDA system can find game parameters that approach the score rate requirement. Concerning Game B, there is also a significant increase in η_B from the initial session. Despite the method proposed can identify game parameters that increase α_p (46% in the first session, 59% in the last session), the results evidence that the DDA system is not able to find game parameters to align the game difficulty with the trajectory proposed. One possible explanation for this discrepancy can be the presence of a game element beyond the control of the DDA system, i.e. the target location. Given that targets spawn randomly within the game environment and scores are time-dependent, increasing player movement seems

to make achieving the target score g_o more difficult. These findings may indicate that an effective adaptation strategy should not only focus on the method itself but also on the adaptability of the game design. Therefore, the game design should exclude elements that cannot be adapted by the DDA system. Moreover, it is essential to limit the number of adjustable game parameters to guarantee a time-affordable adaptation process.

By examining the median values of n_e and v_e during the game sessions (as represented in Figure 6.8 and detailed in Tables B.2 and B.3) the degree of personalization may be assessed. Across both games, the median values of n_e and v_e increased as the sessions progressed. These findings suggest that the proposed DDA system tries to increase the difficulty level to maintain a specific score rate. In addition, the variation in the values of the game parameters seems to increase, which suggests that each user played with a personalized difficulty level. Furthermore, except n_e in Game B, the Interquartile Range (IQR) increased compared to the initial session. This implies that while the personalization was initially minimal, starting from identical game parameters, it became more pronounced over time, aligning more closely with each user's skill level. Notably, even though the initial maximum value for v_e was set to 1.0, the DDA system determined that for some users, it was necessary to elevate this parameter beyond the initial limit to follow the specified difficulty curve.

The proposed method enhances the intensity of activities, but it's crucial to examine users' psychophysiological states during sessions. The valence, arousal, and dominance were self-evaluated with the SAM test (Figure 6.9, Table B.4). The Friedman test results revealed that for Game A, there were no significant differences in any of the SAM variables. This suggests that users experienced pleasure, excitement, and a medium-high sense of dominance throughout all sessions, despite the progressive difficulty increase. However, in Game B, significant differences can be observed in arousal. Specifically, arousal levels were significantly lower in session 3 compared to session 1, and in session 5 compared to session 1. No significant differences were found in valence and dominance. A possible explanation is that this decrease in arousal may be attributed to the consistent score rate maintained during Game B across sessions. With the progression of the sessions, the users get more experienced and their arousal may decrease. This is aligned with the high levels of dominance reported by the users in Game B.

In addition, no significant differences were found across sessions for the cardiorespiratory (Figure 6.10) and GSR signals (Figure 6.11). These results are

aligned with the self-reported emotions reflected in the SAM test, which suggests no changes in the psychophysiological responses despite a progressive increase in game difficulty.

The results of this study are highly promising, suggesting that the proposed DDA system could be used across consecutive rehabilitation sessions. However, it's essential to recognize that the experiments conducted to evaluate the adaptation method were carried out with able-bodied users. As a next step, it is intended to evaluate the adaptation architecture with patients. Furthermore, it is intended to investigate how this architecture can be combined with other methods to offer effective robot assistance to patients with limited upper limb mobility.

6.5 CONCLUSION

In this Chapter, a genetic algorithm-based approach for adapting serious games during robot-aided rehabilitation therapies has been proposed. The DDA system has demonstrated efficacy in maximizing the therapy performance (η) across sessions, suppressing the recalibration between sessions. This achievement can be attributed to the use of mutations to explore variations in user behavior. Notably, the results show that the method proposed excels in personalizing game parameters, ensuring wide player movements while maintaining specific score levels, especially for Game A. The personalization performed by the proposed DDA system is further supported by the observed dispersion in the game parameters. In addition, from the results obtained is derived that game design is crucial for performing an appropriate game adaptation. Therefore, the deliberate exclusion of random elements from the game design is recommended to enhance overall effectiveness.

In terms of psychophysiological responses, it's noteworthy that despite the escalating difficulty of the game, there is a remarkable consistency in valence, arousal, and dominance throughout the sessions. The physiological reactions of players also show uniformity across different sessions, reinforcing the conclusion that the adaptive game design, which is enabled by the DDA system proposed, does not markedly influence the players' emotional or physiological states.

7

A DEEP LEARNING MODEL FOR ASSISTIVE DECISION-MAKING IN ROBOT-AIDED REHABILITATION THERAPY

7.1 INTRODUCTION

In addition to personalizing the difficulty of the serious games, patients with a higher degree of disability may need assistance to perform the movements required by the games. Therefore, there is a need to develop an assistance system to find out when rehabilitation robots should provide assistance.

Diverse AAN methods have been presented in recent studies (Chapter 2, section 2.4) to provide assistance during robot-aided rehabilitation therapies. However, it is observed that such studies typically overlook the expertise of experienced therapists in determining the appropriate moments for intervention. A previous study of the group highlighted that therapists possess a superior ability to determine the most suitable manner and timing for providing assistance during exercises, compared to the standard fixed assistance methods found in robotic rehabilitation devices (Catalán; García-Pérez; Blanco; Ezquerro, et al., 2021). In essence, the assistance offered by therapists is better adapted to the individual requirements of patients, thereby optimizing their exertion and potentially enhancing their recovery, a factor identified as crucial by professionals for successful rehabilitation outcomes (Maclean et al., 2000; Friedrich et al., 1998).

Only S. Pareek and T. Kesavadas (Pareek and Kesavadas, 2020) propose to develop an AAN architecture based on a therapist's expertise to determine optimal moments for providing assistance during rehabilitation exercises. However, the validation of the AAN was conducted with able-bodied participants, and one of the authors assumed the role of the therapist.

An AAN approach for determining in which scenarios the robot should assist is presented in this Chapter. With this aim, data from multiple patients were collected during real robot-aided rehabilitation sessions, where an experienced therapist demonstrated when assistance should be provided. Furthermore, a series of transformations have been applied to ensure the applicability of the method across

different point-to-point activities. Unlike S. Pareek and T. Kesavadas, who advocate training individual LSTM models for each user, the presented approach involves training a Deep Learning model using data from diverse patients to create a more generalized method. Finally, the trained model is fine-tuned to adapt the therapy to each patient's specific needs.

7.2 MATERIALS AND METHODS

7.2.1 *Subjects*

The data acquisition used to develop the AAN model has been performed in La Pedrera Hospital (Denia, Spain) with patients suffering from neurological conditions (Table 7.1). The rehabilitation therapists were responsible for defining the inclusion criteria, which included:

- Adults with hemiparesis/hemiplegia
- Orientation in the three spheres (social, temporal, and spatial)
- Capacity for collaboration and comprehension of task instructions and study-related information

Table 7.1: Information about the patients who participated in the study.

Patients	Sex	Age (years)	Diagnostic	Laterality	STREAM
1	Male	72	Abdominal Dystension	Right	10/16
2	Male	51	Basal Ganglia hematoma	Right	11/16
3	Male	70	Bilateral SARS-CoV2 pneumonia	Left	N/A
4	Male	78	Ischemic stroke atherothrombotic	Right	11/16
5	Female	84	Ischemic stroke	Right	7/16
6	Male	69	Protuberance ischemic stroke	Left	7/16
7	Male	69	Thalamic capsule-lacunar stroke	Right	0/16
8	Male	62	Chronic ischemic heart disease	Right	8/16

Note. **N/A**, not available value; **ACM**, Artery Cerebral Middle; **STREAM**, Stroke Rehabilitation Assessment of Movement, the value indicate the score obtained for the upper limb movement.

To evaluate patients, the STREAM test was used (Ahmed et al., 2003). This assessment evaluates coordination, functional voluntary movement, and range of motion. For our purposes, only the score obtained from the upper limb was considered.

7.2.2 *Experimental setup and data collection*

The data acquisition was conducted in a previous study of the group (Catalán; García-Pérez; Blanco; Ezquerro, et al., 2021). The experimental setup has been represented in Figure 7.1. The Rubidium robot was employed to interact with a serious game based on a point-to-point modality. During the game, the patients needed to reach a peripheral point of the roulette starting from its center, as shown in Figure 7.1.a.

In the conducted study, a master-slave assistance approach was employed, where the Rubidium robot itself does not assist the patient. Instead, the therapist provides direct assistance to the patient. The therapist exerted force on the end effector of the Rubidium, which was equipped with an onRobot HEX force sensor to record the assistance provided (Figure 7.1.b). This setup allowed for the measurement of both the force applied by the therapist and the timing of the assistance to determine when the assistance was provided.

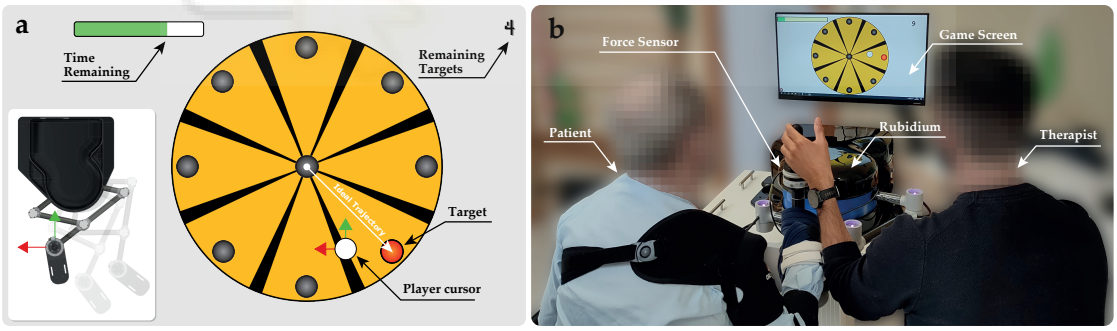


Figure 7.1: Set of devices used during the experimental sessions (Catalán; García-Pérez; Blanco; Ezquerro, et al., 2021). **(a)** The point-to-point game employed during the rehabilitation sessions together with the Rubidium robot, an end-effector upper-limb rehabilitation robot. **(b)** Image of the experimental setup used in La Pedrera Hospital.

7.2.3 *Data processing*

The proposed method in this Chapter is designed to be used in diverse serious games or exercises based on the dynamics of goal achievement. Consequently, a sequence of transformations must be applied to the data obtained from the Rubidium robot before incorporating them as input to the Deep Learning model. The applied data processing is represented in Figure 7.2.

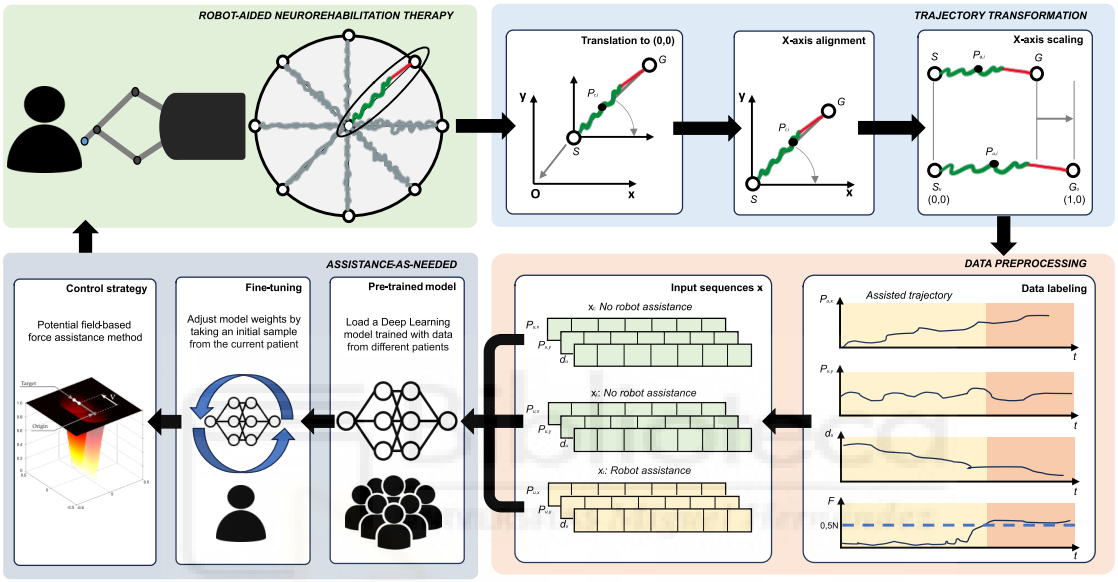


Figure 7.2: AAN proposed strategy during robot-aided therapy. As the patient executes a movement, the Cartesian coordinates of the end-effector are acquired. The trajectory, considering the starting point S_r and the goal position G_r, is aligned with the X-axis and scaled. The trajectories performed by the patients are labeled according to the measurements of the force sensor to train a model that learns from the therapist's interventions. These trajectories are then segmented into sequences to serve as input for the Deep Learning model. The model is trained using data collected from diverse patients and then fine-tuned before its application during robot-aided therapies. In the final stage of the pipeline, the model evaluates whether assistance is required. The assistance provided by the robot can be facilitated by the force control-based method presented in (Catalán; García-Pérez; Blanco; Ezquerro, et al., 2021).

Each of the trajectory points P_r between the starting point S_r and the target point G_r has been considered, and the transformation ${}_rT^u$ has been applied. By ${}_rT^u$, we transform $S_r \rightarrow S_u = (0,0)$ and P_r is aligned with the X-axis, which can be defined as:

$$ {}_rT^u = \begin{bmatrix} cos(\theta_x) & -sin(\theta_x) & 0 & -S_{r,x} \\ sin(\theta_x) & cos(\theta_x) & 0 & -S_{r,y} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7.1} $$

where θ_x is the angle between vector $\vec{PG}$ and the X-axis. The trajectory performed by the user aligned with the X-axis P_a and the target point G_a can be obtained by applying ${}_rT^u$:

$$ P_a =_r T^u \cdot P_r \tag{7.2} $$

$$ G_a =_r T^u \cdot G_r \tag{7.3} $$

A scale factor is finally applied to P_a to transform $G_a \rightarrow G_u = (1,0)$:

$$ P_u = \begin{bmatrix} \frac{1}{G_{a,x}} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot P_a \tag{7.4} $$

Figure 7.3 represents a between the original trajectories and the post-transformation trajectories. Following the application of these transformations, the distance to the target d_u is calculated for each point:

$$ d_u = \|\overrightarrow{P_u G_u}\| \tag{7.5} $$

The exerted force F on the end-effector of the robot was measured and employed to identify the assistance provided by the therapist. Each trajectory point $P_u = (X_u, Y_u)$

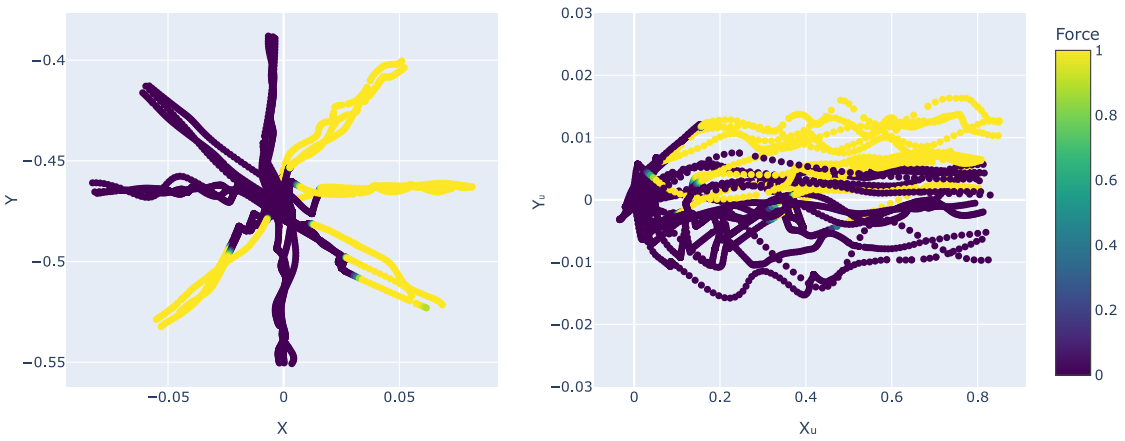


Figure 7.3: On the left, the trajectories executed by a patient during the rehabilitation activity are represented. On the right, the transformed trajectories are presented, having undergone alignment and scaling along the X-axis. Moreover, the force applied by the therapist to the end-effector is visually represented by color coding along the trajectories.

is labeled as assisted whether $F > 0.5$ N. Otherwise, the trajectory performed by the patient is labeled as unassisted. Therefore:

$$Y = \begin{cases} 0 & \text{if } F \leq 0.5 \\ 1 & \text{if } F > 0.5 \end{cases} \tag{7.6}$$

To determine whether a patient requires assistance, the dataset $\mathcal{D} : (x, y)$ has been used for training the model, where x represents the input from the set $x \in X = \{X_u, Y_u, d_u\}$, and y is the corresponding output. As the problem is addressed as a multivariate time-series classification, the input vector x_t includes measurements from the preceding two seconds to predict y_t. The data collected from the Rubidium robot have been acquired at 30 Hz, meaning that x_t consists of the last 60 samples $(x_{t-59}, x_{t-58}, ..., x_{t-1}, x_t) \in X$.

It is worth mentioning that the model only receives data from the movements executed by the patients. Once the assistance was detected, any movement made by the therapist to aid the patient was excluded from the input data. Consequently, the x_t before the therapist's assistance during the trial is categorized as assisted.

Before training, the dataset $\mathcal{D}$ was divided into three different subsets. Data from six users (intra-participants) were utilized for both training and validation datasets. Specifically:

- **Training Data.** 80% of the assisted trials and 80% of the non-assisted trials were used to train the model.
- **Validation Data.** The remaining 20% of assisted and non-assisted trials from the intra-participants were reserved for model validation.
- **Test Data.** Data from the other 2 patients served as the test dataset.

Furthermore, considering that $X_u, d_u \in (0, 1)$ due to the applied transformations, the input Y_u was scaled between the minimum and maximum values of the training dataset. This ensures that all inputs are within a similar range.

In addition, as it is intended to fine-tune the model for each patient, a subset comprising 20% of the trajectories (equivalent to 5 trajectories) was split for each test user from the test dataset. These data will be used to fine-tune the model, while the rest of the data will be employed to evaluate the impact of fine-tuning.

7.2.4 *Deep Learning Model*

7.2.4.1 *Model architecture*

As introduced, it is intended to develop a classifier based on a Deep Learning model in order to learn from the therapist when to assist during robot-aided rehabilitation therapies. The problem can be defined as learning a classification function $\hat{y} = f(x_t)$, such that $\hat{y}_t \approx y_t, \forall x_t \in X$.

The proposed approach involves training a Deep Learning model based on the ENCODER architecture proposed by Serrà et al. (Serra et al., 2018), which was specially designed for time series classification. The architecture of the model proposed has been represented in Figure 7.4, and consists of one-dimensional convolutional layers, a convolutional attention mechanism to summarize the time axis, and a final GAP layer.

The first convolutional layer is composed of 128 filters of length 5; the second convolutional layer consists of 256 filters of length 11; and the third convolution is composed of 512 filters of length 21. Each convolutional layer is followed by an instance

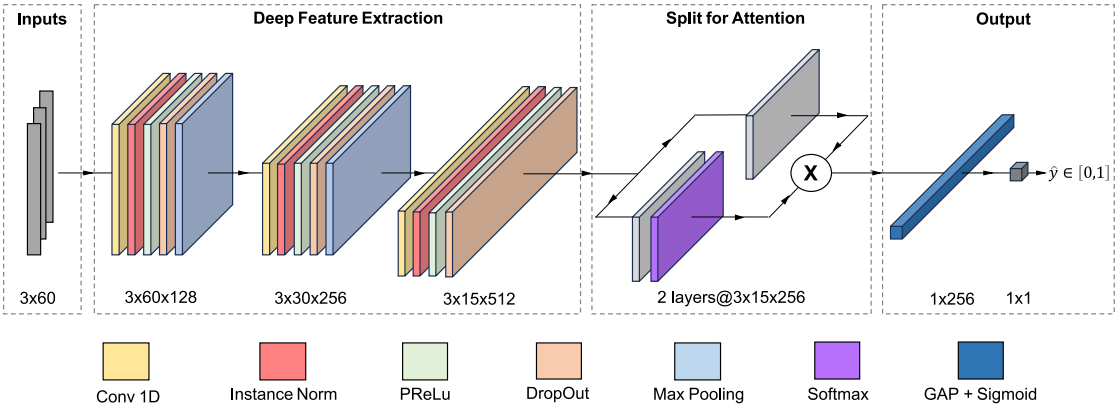


Figure 7.4: The proposed architecture of the Deep Learning model to provide AAN during robot-aided rehabilitation therapies. Feature extraction is conducted through three convolutional blocks. Subsequent to these layers, half of the extracted filters are used as inputs of a time-wise softmax activation, which acts as an attention mechanism for the remaining filters. The architecture concludes with a GAP layer and a neuron with the sigmoid activation function to infer whether a patient needs assistance according to the performed movements.

normalization operation (Ulyanov et al., 2016) and then a Parametric Rectified Linear Unit (PReLu) activation function (He et al., 2015). The output of PReLu is then subjected to a dropout operation with a rate of 0.2, followed by a final max-pooling layer of length 2. After these layers, half of the filters are input to a time-wise softmax activation, serving as an attention mechanism(Niu et al., 2021) for the remaining filters. The result of the attention mechanism is finally passed through a GAP layer. Finally, a layer with one neuron is used as the output of the model, employing the Sigmoid activation function (Pratiwi et al., 2020) for the binary classification problem. The total number of trainable parameters of the model is 3,118,723.

7.2.4.2 *Model training and fine-tuning*

The AAN model has been developed and trained using Keras and Tensorflow Python packages (Gulli and Pal, 2017b). The Mean Square Error (MSE) was chosen as the loss function, diverging from the typical Binary Cross-Entropy approach (Canbek, 2023). The Adam algorithm was used as the optimizer with an initial learning rate of 0.001. To prevent overfitting, the loss on the validation dataset was observed, and training was halted if no improvement was seen for five epochs, with the epoch limit set at 50. In the event of early stopping, the model reverted to the parameters that achieved the lowest loss. A batch size of 8 was employed throughout the training process.

In order to adjust the assistance provided by the robotic device it is intended to fine-tune the general model. The fine-tuning process is designed to transfer the knowledge gained from a Deep Learning model that has been pre-trained on a separate dataset (Beghriche et al., 2023). Although fine-tuning necessitates a degree of learning, it would be significantly quicker than building a model from scratch for each patient's assistance needs (Too et al., 2019). Moreover, fine-tuning has the advantage of incorporating broad knowledge acquired from the initial training of the model. During fine-tuning, the initial learning rate for the Adam algorithm was reduced from 0.001 to 0.0001 to prevent overfitting. The fine-tuning process was conducted during 10 epochs with a batch size of 1.

7.2.4.3 *Model evaluation*

To assess the performance of the proposed model in classifying assistance during robot-aided exercises different evaluation metrics have been used, including accuracy, recall, specificity, False Positive Rate (FPR), False Negative Rate (FNR), and the F1-Score. During the evaluation, a sequence is classified as assisted when $\hat{y} > 0.5$.

Initially, a sequence-to-sequence evaluation was conducted using the training and validation datasets. This involved comparing the predicted assistance for each data sequence against the actual assistance provided. Moreover, to benchmark the model proposed, seven additional models have been trained for a comparative analysis of sequence-to-sequence classification accuracy: a ResNet model (Wang et al., 2017a), a t-LeNet model (Le Guennec et al., 2016), a MCDCNN (Zheng et al., 2016), an ENCODER model (Serra et al., 2018), a Time-CNN model (Zhao et al., 2017b), and a FCN model (Wang et al., 2017b). Furthermore, an LSTM model has been trained to evaluate the performance of the model proposed in comparison to the approach presented by S. Pareek and T. Kesavadas (Pareek and Kesavadas, 2020).

In addition to sequence-to-sequence evaluation, a method for assistance decision-making across multiple sequences rather than a single one has been proposed. This involves employing the model to make continuous inferences throughout the point-to-point rehabilitation activity. If it is inferred that assistance should be provided for a certain period of time, the robot assistance will be enabled. Different decision times were considered to select the optimal period with the validation dataset: 0.5 seconds, 1.0 seconds, 1.5 seconds, and 2.0 seconds. After finding the optimal decision time, the test dataset was used to evaluate the classification performance in new patients.

7.3 RESULTS

This section describes the evaluation results of the AAN method proposed. Initially, the sequence-to-sequence experiments are analyzed. Subsequently, the applicability results of the AAN method within a robot-assisted rehabilitation context are examined. The model training, fine-tuning, and inference were executed on an Nvidia RTX 4070Ti GPU. The duration for model training was approximately 35 seconds, fine-tuning required about 10 seconds per user, and each inference was completed in 0.017 milliseconds.

7.3.1 *Sequence-to-sequence evaluation*

Table 7.2 and Table 7.3 summarize the accuracy, recall, specificity, FPR, FNR, and F1-Score of the trained models when evaluating sequences from the training and validation datasets. On the training dataset, the model proposed achieves the highest accuracy (97.47%) and F1-Score (94.55%). The ENCODER architecture achieves the highest specificity (99.50%) and the lowest FPR (0.50%), while the ResNet architecture obtains the highest recall (97.8%) and the lowest FNR (2.17%). When evaluating the models with the validation dataset, the model proposed achieves the highest recall (90.55%), the lowest FNR (9.45%), and the highest F1-Score (75.15%), while the ENCODER model achieves the highest accuracy (91.81%), specificity (93.64%), and the lowest FPR (6.35%).

In addition, an analysis of the impact of the sequence length on the model performance has been conducted. Table 7.4 and Table 7.5 summarize the results of the analysis with the training and validation datasets, respectively. The metric values suggest that performance tends to improve when the length of the sequences increases. For the training dataset, the accuracy increased from 93.43% (15 samples) to 97.91% (45 samples), and the F1-Score increased from 83.81% (15 samples) to 95.36% (45 samples). For the validation dataset, the accuracy increased from 89.27% (15 samples) to 91.72% (60 samples), and the F1-Score increased from 61.87% (15 samples) to 74.15% (60 samples).

Table 7.2: The evaluation of the model using the training dataset is collected. The accuracy, the recall, the specificity, the FPR, the FNR, and the F1-Score are used to evaluate the performance of the proposed model while compared with other State of the Art models.

Model	Accuracy	Recall	Specificity	FPR	FNR	F1-Score
ResNet	0.9310	**0.9783**	0.9172	0.0828	**0.0217**	0.8647
t-LeNet	0.9007	0.8523	0.9148	0.0852	0.1477	0.7947
MCDCNN	0.9570	0.8885	0.9768	0.0232	0.1115	0.9029
ENCODER	0.9656	0.8646	**0.9950**	**0.0050**	0.1354	0.9190
Time-CNN	0.9071	0.8831	0.9140	0.0860	0.1169	0.8108
FCN	0.8750	0.8365	0.8861	0.1139	0.1635	0.7510
LSTM	0.9177	0.8242	0.9449	0.0550	0.1757	0.8188
Our Model	**0.9747**	0.9754	0.9744	0.0256	0.0246	**0.9455**

Table 7.3: The evaluation of the model using the validation dataset is collected. The accuracy, the recall, the specificity, the FPR, the FNR, and the F1-Score are used to evaluate the performance of the proposed model while compared with other State of the Art models.

Model	Accuracy	Recall	Specificity	FPR	FNR	F1-Score
ResNet	0.8773	0.8046	0.8895	0.1105	0.1954	0.6536
t-LeNet	0.8763	0.8046	0.8885	0.1115	0.1954	0.6519
MCDCNN	0.8981	0.7815	0.9178	0.0822	0.2185	0.6883
ENCODER	**0.9181**	0.8088	**0.9364**	**0.0635**	0.1912	0.7397
Time-CNN	0.8748	0.8803	0.8740	0.1260	0.1197	0.6693
FCN	0.7812	0.6765	0.7988	0.2012	0.3235	0.4708
LSTM	0.8996	0.7626	0.9227	0.0773	0.2373	0.6862
Our Model	0.9139	**0.9055**	0.9153	0.0847	**0.0945**	**0.7515**

7.3.2 *Assistance during trajectory execution*

During robot-aided rehabilitation sessions, it will be considered that assistance should be provided when $\hat{y} > 0.5$ for a certain time t_{assist}. Table 7.6 collects the performance of the model when classifying complete trajectories according to t_{assist}. The results suggest that the ideal duration is approximately $t_{assist} = 1.0$ s, achieving an accuracy of 96% and an F1-Score of 84.93%. It is worth mentioning that FNR tends to rise with

Table 7.4: Results of the model evaluation with the training dataset in accordance with the sequence length. To evaluate the performance of the proposed model, the accuracy, recall, specificity, FPR, FNR, and F1-score values are included. The equivalence between samples and time is also given.

Sequence Length	Accuracy	Recall	Specificity	FPR	FNR	F1-Score
15 (0.5 s)	0.9343	0.7543	0.9867	0.0133	0.2457	0.8381
30 (1.0 s)	0.9711	0.8850	0.9962	0.0039	0.1150	0.9325
45 (1.5 s)	0.9791	0.9543	0.9862	0.0138	0.0457	0.9536
60 (2.0 s)	0.9766	0.9001	0.9986	0.0014	0.0992	0.9455

Table 7.5: Results of the model evaluation with the validation dataset in accordance with the sequence length. To evaluate the performance of the proposed model, the accuracy, recall, specificity, FPR, FNR, and F1-score values are included. The equivalence between samples and time is also given.

Sequence Length	Accuracy	Recall	Specificity	FPR	FNR	F1-Score
15 (0.5 s)	0.8927	0.6050	0.9411	0.0590	0.3950	0.6187
30 (1.0 s)	0.9102	0.7962	0.9294	0.0706	0.2038	0.7185
45 (1.5 s)	0.9030	0.8970	0.9040	0.0960	0.1029	0.7268
60 (2.0 s)	0.9172	0.8256	0.9326	0.0674	0.1744	0.7415

an increase in t_{assist}, with an FNR of 20% for $t_{assist} = 0.5$ s up to 30% for $t_{assist} = 2.0$ s. Meanwhile, the FPR remains at 0.85% for $t_{assist} \geq 1.0$.

After setting $t_{assist} = 1.0$, the performance of the model proposed has been analyzed with the test dataset. Table 7.7 collects a comparative analysis of the impact of fine-tuning on the performance. Notably, the accuracy improved from 63.04% to 76.09%, the FPR decreased from 44.23% to 26.92%, the FNR decreased from 27.50% to 20.0%, and the F1-Score improved from 63.04% to 74.42%.

In addition, Figure 7.5 represents the classification results as confusion matrices. Within the intra-participants dataset, 233 out of 235 unassisted trajectories and 31 out of 40 assisted trajectories were accurately identified. With the test dataset, 38 out of 52 unassisted trajectories and 32 out of 40 assisted trajectories were successfully classified.

The inputs of the model X_u, Y_u, and d_u have been represented together with the inference $\hat{y}$, the moment when robot assistance would be enabled, and the intervention

Table 7.6: Complete trajectory classification results by decision-making time. Accuracy, recall, specificity, FPR, FNR, and F1-score are calculated using the intra-participant data.

t_{assist} (s)	Accuracy	Recall	Specificity	FPR	FNR	F1-Score
0.5	0.9563	0.8000	0.9830	0.0170	0.2000	**0.8421**
1.0	0.9600	0.7750	0.9915	0.0085	0.2250	0.8493
1.5	0.9527	0.7250	0.9915	0.0085	0.2750	0.8169
2.0	0.9491	0.7000	0.9915	0.0085	0.3000	0.8000

Table 7.7: Complete trajectory classification results by decision-making time, with and without fine-tuning. Accuracy, recall, specificity, FPR, FNR, and F1-score are calculated using the test data.

Fine-Tuning	Accuracy	Recall	Specificity	FPR	FNR	F1-Score
No	0.6304	0.7250	0.5577	0.4423	0.2750	0.6304
Yes	0.7609	0.8000	0.7308	0.2692	0.2000	0.7442

Note: Results obtained with $t_{assist} = 1.0$ s.

of the therapist. These four examples correspond to different classification scenarios. In Figure 7.3.a, it is apparent that the model indicates that assistance would be required at two specific intervals. However, since t_{assist} is not fulfilled, a false positive was filtered and the trajectory is correctly classified as non-assisted. Figure 7.3.b plots a

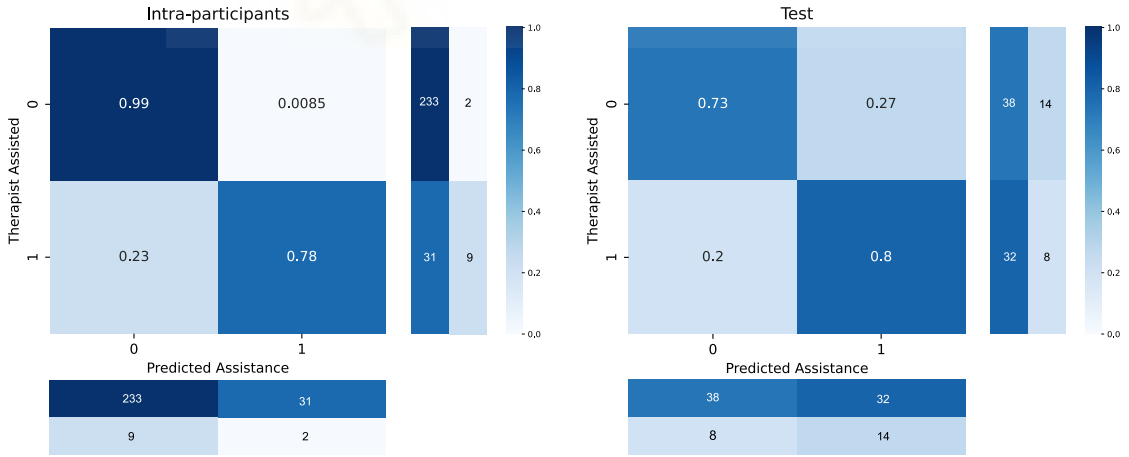


Figure 7.5: The confusion matrices have been obtained from the intra-participants and test datasets for the complete trajectories. The results collected with the test data are represented after applying fine-tuning for each patient with $t_{assist} = 1.0$ s.

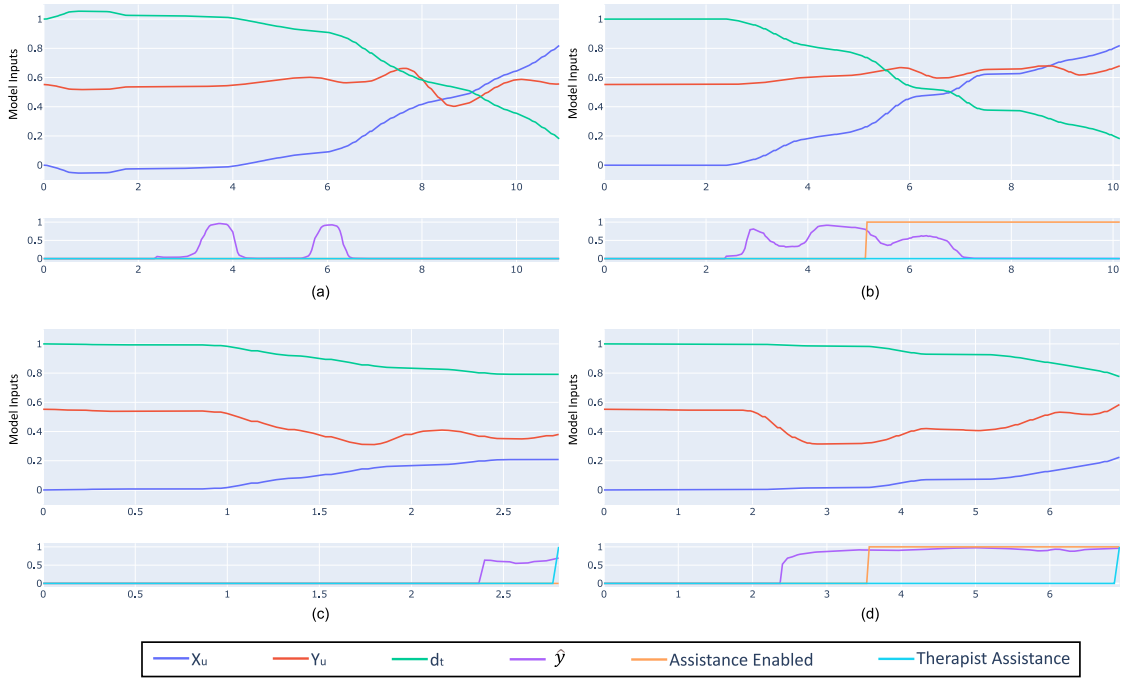


Figure 7.6: The inputs X_u, Y_u, and d_u are represented together with the inference $\hat{y}$, the moment when robot assistance would be enabled, and the intervention of the therapist from the four trajectories executed by the test subjects. The trajectories correspond to **(a)** a true negative, **(b)** a false positive, **(c)** a false negative, and **(d)** a true positive. Each trajectory is represented until the subject completes the trajectory autonomously or receives assistance from the therapist.

false positive instance, where it can be observed how $\hat{y}$ exceeds the decision time t_{assist}. Another possible situation has been represented in Figure 7.3.c, where the model infers that assistance would be required. However, the therapist assisted before t_{assist} was fulfilled, resulting in a false negative. Finally, Figure 7.3.d shows a true positive.

7.4 DISCUSSION

In this Chapter, an AAN approach has been proposed to determine whether assistance is needed during upper limb robot-aided rehabilitation scenarios. The objectives of this approach are, firstly, to encourage active patient participation and sustained engagement in their therapy, and secondly, to learn from the expertise of a therapist when assistance should be provided while maintaining a clinical perspective.

In the context of robot-assisted therapies, exercises often involve patients moving from a starting point to a designated target (Figure 7.1.a). Consequently, an AAN method has been proposed to be applied in serious games based on these mechanics. A series of transformations have been applied in an attempt to standardize the method (section 7.2.3), and the inputs of the model X_u, Y_u, and d_u are obtained after their application. It is worth mentioning that the trajectories represented in Figure 7.3 suggest that the therapist decided to assist the patients in reaching the target point when assistance was considered necessary. Hence, a classifier is proposed to enable a force-field assistive method implemented in the Rubidium robot (Catalán; Blanco-Ivorra, et al., 2023a).

The architecture of the proposed model is represented in 7.4, and its performance was evaluated using diverse methods. Initially, a sequence-to-sequence evaluation was conducted, comparing the performance of the model proposed against other State of the Art models (Table 7.2, Table 7.3). The results indicate that the proposed model outperforms other architectures, achieving the highest F1-Score on both the training and validation datasets. Notably, the ENCODER model achieves a higher accuracy (91.81%) and a lower FPR (6.35%) with the validation dataset. These findings suggest that the ENCODER architecture is robust in determining when to provide assistance based on observed trajectories. Furthermore, it is worth mentioning that our model also outperforms the LSTM architecture proposed by S. Pareek and T. Kesavadas (Pareek and Kesavadas, 2020), whose study proposes a similar AAN approach to the method proposed in this Chapter.

The previous results were obtained with a sequence length of 60 samples, corresponding to 2.0 s. The impact on the performance when reducing the sequence length was also examined using both training and validation datasets (Table 7.4, Table 7.5). The findings suggest that the model benefits from a longer sequence length.

Instead of activating the assistance when an inference $\hat{y} > 0.5$ is produced, a full-time window where the model indicates that the assistance should be activated is proposed. The impact of the decision time t_{assist} was examined using intra-participant data. The results of the classification performance (Table 7.6) indicate that the optimal value would be $t_{assist} = 1.0$ s.

Once the value of t_{assist} was set with the intra-participant data, the performance of the AAN method proposed was analyzed with the test users. Initially, the impact of fine-tuning the model has been evaluated (Table 7.7), which analysis revealed an increase in the accuracy by 13%, and by 11% in the F1-score by 11%. In addition, the

FPR decreased by 18%, and the FNR by 7.50%. Given that the fine-tuning process required only 10.0 s per user, the feasibility of adapting the model for providing better assistance based on the patient's needs.

After verifying the effectiveness of fine-tuning, a comparative analysis of the intra-participant and test dataset classification results was conducted. The analysis revealed that 99% of the unassisted trajectories and 78% of the assisted trajectories were correctly classified. The 23% FNR observed with the intra-participant data could be attributed to the disparity in the number of unassisted trajectories (235) compared to assisted ones (40). On the other hand, the classification results for the test users after the application of fine-tuning showed a more equitable distribution of results. Specifically, 73% of the unassisted trajectories and 80% of the assisted trajectories were correctly classified. This balance suggests that fine-tuning effectively mitigates the data imbalance present in the intra-participant data, ensuring that the inferences remain unaffected when applied to new patients. Consequently, this method demonstrates promising generalization capabilities in determining the appropriate times to provide assistance.

For a better comprehension of the model decision-making, four examples of performed trajectories by the test users are shown in Figure 7.6. Figure 7.6.a represents how false positives due to t_{assist} may be effectively filtered out, where the fine-tuned model infers that assistance is needed when the patient is not progressing toward the goal. This can also be observed in Figure 7.6.b and Figure 7.6.d. It is noteworthy the case represented in Figure 7.6.c. In certain cases, the therapist may have assisted the patient before reaching the decision time t_{assist}. Consequently, the trajectory is labeled as a false negative during model evaluation. However, in a real rehabilitation scenario, this situation would not occur since the model correctly infers when assistance is needed.

According to the results obtained, the feasibility of using the AAN method proposed would be demonstrated to provide assistance during robot-aided rehabilitation therapies. However, this study has some limitations that are intended to be addressed in future studies. First, during the study, only one therapist participated in the data collection and assistance during the robot-aided sessions. Ideally, the Deep Learning model described in this Chapter should learn from diverse therapists in order to incorporate a broader range of professional criteria. Additionally, it is intended to expand the study to include a larger number of patients and to examine the use of the proposed AAN method in a rehabilitation scenario.

7.5 CONCLUSION

In this Chapter, a Deep Learning-based AAN method was presented to learn from therapists' criteria when a patient needs assistance during robot-aided rehabilitation therapy. The proposed model has been trained and evaluated using data from diverse patients playing a point-to-point game. Before the introduction of the movements performed with the end-effector robot, a series of transformations were applied in an attempt to universalize the method. Moreover, a reduced set of trajectories executed by the test patients was used to fine-tune the model trained with intra-participant data.

First, a sequence-to-sequence evaluation was conducted, which revealed an accuracy of 91.39% and an F1-Score of 75.15% with the validation dataset. Furthermore, the approach presented proposes to utilize a decision-making time t_{assist} that must be fulfilled before the rehabilitation robot provides assistance, rather than activating it immediately when $\hat{y} > 0.5$. With this method, in conjunction with fine-tuning, the AAN architecture proposed demonstrated an accuracy of 76.09% and an F1-Score of 74.42% when classifying trajectories performed by the test patients.

The obtained results would demonstrate the feasibility of applying the proposed method to determine when to assist during robot-assisted point-to-point rehabilitation exercises. Furthermore, the intention is to further develop and analyze the performance of the method proposed by involving a larger number of therapists and patients.

8

CONCLUSIONS

8.1 CONCLUSION

Human modeling has been shown to be of vital importance in the field of rehabilitation robotics. In particular, understanding the user is essential to design the robotic devices taking into account the biomechanics of the upper and lower limbs, but also to adapt and personalize robot-assisted therapy to each patient needs.

Rehabilitation robotics is an emerging and dynamic area of research. However, there are no solutions to help patients with spasticity, pain, or joint stiffness use robots during rehabilitation. The SPLASH project aims to address this issue by developing robotic solutions for therapy in aquatic environments. The SPLASH project does not only focus on developing underwater robotic devices but also on studying the effects of water on user biomechanics. Specifically, in order to study the effects during gait and subsequently develop a lower limb exoskeleton, a motion capture system based on submersible IMUs has been developed to measure lower limb joint angles in the sagittal plane. The employed method does not assume a specific position or orientation of the sensors, and the calibration only requires random and gait movements for a short time. Based on the measurements of the leg kinematics, a model describing the buoyancy and drag forces has been proposed to study the loads on lower limb joints, with the goal of designing a submersible lower limb exoskeleton.

In addition to the forces arising from the presence of water, it is worth noting that knowing vGRF during gait is of great relevance for a proper study of biomechanics. Due to the drawbacks that may arise when trying to measure vGRF with an underwater force plate, a Machine Learning model has been developed to estimate the magnitude of this force by using the developed IMU-based motion capture system. First, a model has been proposed to estimate the vGRF during ground-level walking, and the results of the study suggest that the accuracy of the proposed method is higher than to previously published studies. In addition, to generalize the model and to be able to estimate the vGRF during different situations, the feasibility of adapting the model to different gait

activities has been studied, concluding that the knowledge of the activity performed by the user improves the accuracy when estimating the vGRF.

With the aim of increasing the accuracy when estimating the vGRF, in addition to proposing an approach to adapt the assistance of a submersible lower limb exoskeleton in different activities, a model has been developed to recognize a set of ADLs. Although different models using IMUs for HAR have been proposed, the presented approaches require high accuracy in the sensor placement since errors in their positioning will impact the accuracy of the models. Using the implemented motion capture system, a Deep Learning model that takes as inputs the joint trajectories of the lower limbs has been proposed to recognize different gait ADLs. The proposed model achieves high accuracy when evaluated with test users, and the feasibility of reducing the number of model inputs to adapt the system has been demonstrated.

In addition, the SPLASH project also proposes to develop a submersible end-effector robotic device for upper limb rehabilitation. In order to personalize the therapy to each patient to maximize recovery, two approaches have been proposed in this thesis: a system to dynamically adapt the difficulty of serious games, and a system to recognize when assistance should be provided during robot-aided rehabilitation through demonstrations performed by a therapist.

Several diverse DDA systems have been proposed in recent studies. However, none of them address the need to use this type of system for several consecutive sessions. To address this drawback, a DDA system has been developed for consecutive rehabilitation sessions. The system is based on a genetic algorithm and it aims to maximize the game space covered by the patient to rehabilitate the limb while maintaining a certain level of difficulty to maintain motivation.

On the other hand, according to the impairment degree of the patients, it is also necessary to identify when to assist with rehabilitation robots. Therefore, the development of intelligent AAN control algorithms, that recognize and only assist the patient in the needed moment should be deeply investigated and developed. In this context, the goal of the AAN methods is to promote the active participation of the patient, providing the necessary assistance to complete the exercises proposed during therapy. Although several solutions have been proposed in recent studies, there is a lack of systems that learn from therapists' interventions to decide when to assist. Following this principle, a model has been proposed to learn from demonstrations performed by an experienced therapist. The study has demonstrated the feasibility of developing a

model using data from different patients. In addition, the results show that fine-tuning has a positive impact on the accuracy of the model.

8.2 FUTURE WORKS

During the development of this thesis, several lines of work were left open, which could be summarized in the following points:

- To analyze the effects of the aquatic environment on gait using the submersible sensor system developed. The loads on the lower limb joints will be analyzed, and the force modeling presented will be validated using OpenSim simulation software, comparing the simulated muscular activation with the measured EMG signals.
- To study the use of new Machine Learning models to increase the accuracy when estimating the vGRF during different gait ADLs, with special interest in the accuracy when estimating the characteristic peaks of the vGRF.
- A new study can be conducted to validate the developed DDA system with patients. In addition, new serious games can be designed to be adapted based on the developed study.
- A larger number of patients and therapists can be involved to improve the AAN system proposed
- To develop a therapy adaptation architecture that unifies the presented DDA system with the proposed AAN method. Based on both systems, it is intended to develop a method that attempts to maximize patient movement while maintaining an appropriate game score and minimizing assistance with the rehabilitation robot.

9

CONCLUSIONES

9.1 CONCLUSIÓN

Se ha demostrado que el modelado del ser humano es de vital importancia en el ámbito de la robótica de rehabilitación. En concreto, comprender el usuario es de importancia para diseñar los dispositivos teniendo en cuenta la biomecánica de los miembros superiores e inferiores, pero también para adaptar y personalizar la terapia asistida con robots a las necesidades de cada paciente.

La robótica de rehabilitación es un campo de investigación emergente y dinámico. Sin embargo, no existen soluciones para ayudar a los pacientes con espasticidad, dolor o rigidez articular a utilizar robots durante la rehabilitación. El proyecto SPLASH pretende resolver este problema desarrollando soluciones robóticas para realizar la terapia en entornos acuáticos. El proyecto SPLASH no sólo se centra en el desarrollo de dispositivos robóticos sumergibles, sino también en el estudio de los efectos del agua en la biomecánica del usuario. Para estudiar los efectos durante la marcha y desarrollar posteriormente un exoesqueleto de miembro inferior, se ha desarrollado un sistema de captura de movimiento basado en IMUs sumergibles para medir los ángulos de las articulaciones de los miembros inferiores en el plano sagital. El método empleado no asume una posición u orientación específica de los sensores, y la calibración sólo requiere movimientos aleatorios y de la marcha durante un breve periodo de tiempo. A partir de las mediciones de la cinemática de las piernas, se ha propuesto un modelo que describe las fuerzas de flotabilidad y arrastre para estudiar las cargas sobre las articulaciones de las extremidades inferiores con el objetivo de diseñar un exoesqueleto sumergible.

Además de las fuerzas que surgen por la presencia del agua, cabe destacar que conocer la vGRF durante la marcha es de gran relevancia para realizar un estudio adecuado de la biomecánica. Debido a los inconvenientes que puede surgir al tratar de medir la vGRF, se ha desarrollado un modelo de Machine Learning que permite estimar la magnitud de esta fuerza a partir de las medidas tomadas con el sistema de

captura de movimiento desarrollado. En primer lugar, se ha propuesto un modelo que permite estimar la vGRF en plano, y los resultados del estudio sugieren que la precisión del método propuesto es superior a estudios publicados previamente. Además, con el fin de generalizar el modelo y poder estimar la vGRF durante diferentes situaciones, se ha estudiado la viabilidad de adaptar el modelo, y se llegado a la conclusión de que conocer la actividad realizada por el usuario mejora la precisión al estimar la vGRF.

Con el objetivo de aumentar la precisión al estimar la vGRF, así como para adaptar la asistencia de un exoesqueleto de miembro inferior sumergible en diferentes actividades, se ha desarrollado un modelo para reconocer un conjunto de ADLs. Aunque diferentes modelos que usan IMUs para HAR han sido propuestos, las soluciones presentadas requieren de una gran precisión al colocar los sensores, debido a que errores en su posicionamiento repercutirán en la precisión de los clasificadores. Empleando el sistema de captura de movimiento implementado, se ha desarollado una CNN que toma como entradas las trayectorias articulares de los miembros inferiores para reconocer diferentes ADL. El modelo propuesto logra una gran precisión cuando se ha evaluado con usuarios de test, y se ha demostrado la viabilidad de poder reducir el número de entradas del modelo para adaptar el sistema HAR a múltiples situaciones con un impacto bajo en la precisión.

Además, en el proyecto SPLASH también se propone desarrollar un dispositivo tipo efector final sumergible para la rehabilitación del miembro superior. Con el objetivo de personalizar la terapia a cada paciente para maximizar la recuperación, en esta tesis se han propuesto dos soluciones: un sistema para adaptar la dificultad de juegos serios, y un sistema para reconocer cuándo es necesario asistir a un paciente mediante demostraciones de un terapeuta.

Aunque varios sistemas DDA han sido propuestos en diferentes estudios, no se aborda la necesidad de poder usar este tipo de sistemas durante varias sesiones consecutivas. Para tratar de solventar este inconveniente, se ha desarrollado un sistema DDA pensado para poder emplearse durante sesiones de rehabilitación consecutivas. El sistema está basado en un algoritmo genético y se busca que el paciente cubra el mayor espacio de juego posible para rehabilitar el miembro afectado a la vez que se mantiene un cierto nivel de dificultad para mantener la motivación.

Por otra parte, según el grado de movilidad de los pacientes, también es necesario identificar cuándo se debe asistir con los robots de rehabilitación. En este contexto, el objetivo de los métodos de AAN es promover una participación activa del paciente, aportando la ayuda necesaria para tratar de completar los ejercicios planteados durante

la terapia. Aunque varias soluciones han sido propuestas en estudios recientes, existe una falta de sistemas AAN que aprendan de las intervenciones de terapeutas para decidir cuándo asistir. Siguiendo este principio, se ha propuesto un modelo para tratar de aprender en qué situaciones se debe aportar asistencia a partir de las demostraciones realizadas por un terapeuta experimentado. En el estudio realizado, se ha demostrado la viabilidad de desarrollar este tipo de sistemas empleando datos de diferentes pacientes. Además, los resultados demuestran que aplicar *fine-tuning* tiene un impacto positivo en la precisión del modelo.

9.2 TRABAJOS FUTUROS

Durante el desarrollo de esta tesis han quedado abiertas diversas líneas de investigación, las cuales pueden resumirse en los siguientes puntos:

- Analizar los efectos en la marcha producidos por la inmersión en un entorno acuático empleando el sistema de captura de movimiento desarrollado. Para ello, se pretende analizar las cargas en las articulaciones, y el modelo presentado se validará empleando el software de simulación OpenSim, comparando la activación muscular simulada con señales adquiridas de EMG.
- Estudiar el uso de nuevas técnicas de Machine Learning para aumentar la precisión al estimar la vGRF durante diferentes ADL, mostrando especial interés en la precisión al estimar los picos característicos de la señal de vGRF.
- Puede llevarse a cabo un nuevo estudio para validar el sistema DDA propuesto con pacientes. Además, puede diseñarse nuevos juegos que puedan ser adaptados en base a los resultados obtenidos en el estudio realizado.
- Puede realizarse un nuevo estudio involucrando un mayor número de pacientes y terapeutas para mejorar el sistema AAN propuesto.
- Desarrollar una arquitectura de adaptación de terapias de rehabilitación que unifique el sistema DDA junto con el sistema AAN presentado. En base a ambos sistemas, se pretende desarrollar un método que trate de maximizar el movimiento del paciente mientras se mantiene una puntuación adecuada y se minimiza la asistencia aportada por los sistemas robóticos.

A

ANALYSIS OF THE VERTICAL GROUND REACTION FORCE CHARACTERISTIC PEAKS ESTIMATION

This Appendix provides an in-depth analysis regarding the estimated characteristic peaks of the vGRF. Specifically, Table A.1 and Table A.2 present the following results:

1. Table A.1. This table collects the mean magnitude error and delay associated with the LP, MP, and TP. These results are based on the FNN-C7 and RF-C7 models using intra-participant data.
2. Table A.2. This table collects the mean magnitude error and delay associated with the LP, MP, and TP. These results are based on the FNN-C7 and RF-C7 models using inter-participant data.

Additionally, the tables include pairwise comparison values between the errors obtained from the FNN and the RF models.

Table A.1: Results obtained in the vGRF estimation with the selected FNN and RF models employing data from the intra-participants. The table presents the average errors in both peak magnitude and peak delay between the measured and estimated vGRF for each gait speed. Additionally, the p-values (Wilcoxon-Pratt test) for pairwise comparisons of the selected models are summarized.

	Gait Speed (km/h)	FNN-C7	RF-C7	p-value
LP Error (BW)	1.5	0.0334 ± 0.0256	0.0128 ± 0.0075	p<0.0001
	2.5	0.0407 ± 0.0283	0.0132 ± 0.0080	p<0.0001
	3.5	0.0409 ± 0.0467	0.0139 ± 0.0138	p<0.0001
	4.5	0.0900 ± 0.1026	0.0210 ± 0.0286	p<0.0001
MP Error (BW)	1.5	0.0562 ± 0.0561	0.0161 ± 0.0147	p<0.0001
	2.5	0.0262 ± 0.0330	0.0119 ± 0.0097	p<0.0001
	3.5	0.0566 ± 0.0376	0.0120 ± 0.0102	p<0.0001
	4.5	0.0793 ± 0.0596	0.0158 ± 0.0184	p<0.0001
TP Error (BW)	1.5	0.0279 ± 0.0195	0.0110 ± 0.0062	p<0.0001
	2.5	0.0331 ± 0.0205	0.0123 ± 0.0072	p<0.0001
	3.5	0.0396 ± 0.0304	0.0114 ± 0.0117	p<0.0001
	4.5	0.0491 ± 0.0428	0.0171 ± 0.0209	p<0.0001
LP Delay (%)	1.5	-1.51 ± 6.47	-0.19 ± 2.53	p<0.0001
	2.5	1.57 ± 5.97	0.17 ± 2.96	p<0.0001
	3.5	-0.20 ± 4.04	0.01 ± 2.34	0.0772
	4.5	0.99 ± 2.05	-0.07 ± 1.36	p<0.0001
MP Delay (%)	1.5	-0.32 ± 6.80	0.16 ± 3.12	0.4353
	2.5	3.43 ± 5.10	0.08 ± 3.00	p<0.0001
	3.5	1.47 ± 4.98	0.13 ± 2.53	p<0.0001
	4.5	0.61 ± 3.38	0.05 ± 1.94	p<0.0001
TP Delay (%)	1.5	-0.92 ± 6.24	0.10 ± 2.31	p<0.0001
	2.5	-3.00 ± 7.23	-0.09 ± 3.25	p<0.0001
	3.5	-0.93 ± 2.75	-0.13 ± 1.73	p<0.0001
	4.5	-1.89 ± 2.43	0.11 ± 1.42	p<0.0001

Table A.2: Results obtained in the vGRF estimation with the selected FNN and RF models employing data from the inter-participants. The table presents the average errors in both peak magnitude and peak delay between the measured and estimated vGRF for each gait speed. Additionally, the p-values (Wilcoxon-Pratt test) for pairwise comparisons of the selected models are summarized.

	Gait Speed (km/h)	FNN-C7	RF-C7	p value
LP Error (BW)	1.5	0.0206 ± 0.0161	0.0225 ± 0.0162	0.0142
	2.5	0.0382 ± 0.0254	0.0492 ± 0.0305	p<0.0001
	3.5	0.0619 ± 0.0362	0.0611 ± 0.0315	p<0.0001
	4.5	0.0499 ± 0.0408	0.0579 ± 0.0375	p<0.0001
MP Error (BW)	1.5	0.0467 ± 0.0583	0.0399 ± 0.0540	p<0.0001
	2.5	0.0263 ± 0.0194	0.0535 ± 0.0296	p<0.0001
	3.5	0.0690 ± 0.0306	0.0563 ± 0.0315	0.1825
	4.5	0.0810 ± 0.0410	0.0983 ± 0.0418	p<0.0001
TP Error (BW)	1.5	0.0249 ± 0.0377	0.0319 ± 0.0393	p<0.0001
	2.5	0.0302 ± 0.0177	0.0316 ± 0.0221	p<0.0001
	3.5	0.0435 ± 0.0265	0.0564 ± 0.0462	p<0.0001
	4.5	0.0922 ± 0.0386	0.1256 ± 0.0466	0.0834
LP Delay (%)	1.5	-1.37 ± 6.72	-2.34 ± 5.21	0.0001
	2.5	-1.12 ± 7.17	-3.52 ± 6.35	p<0.0001
	3.5	-2.21 ± 5.30	-4.88 ± 5.49	p<0.0001
	4.5	-0.84 ± 3.34	-2.63 ± 3.15	p<0.0001
MP Delay (%)	1.5	4.14 ± 6.95	3.01 ± 6.79	0.0008
	2.5	3.56 ± 5.48	2.00 ± 6.25	p<0.0001
	3.5	1.267 ± 4.04	-2.57 ± 4.61	p<0.0001
	4.5	1.56 ± 2.37	-0.14 ± 3.66	p<0.0001
TP Delay (%)	1.5	-2.33 ± 7.40	-2.39 ± 6.40	0.7777
	2.5	-0.70 ± 6.61	1.60 ± 7.71	p<0.0001
	3.5	0.73 ± 3.27	4.51 ± 4.69	p<0.0001
	4.5	-0.27 ± 2.41	5.35 ± 2.88	p<0.0001

B

GAME METRICS INFORMATION AND SELF-ASSESSMENT MANIKIN TEST RESULTS

B.1 GAME METRICS RESULTS

Table B.1: Mean and standard deviation values of the game metrics and therapy performance across sessions.

		S1	S2	S3	S4	S5
Game A	η_A	0.77 ± 0.03	0.81 ± 0.03	0.80 ± 0.04	0.80 ± 0.05	0.81 ± 0.01
	α_s	0.62 ± 0.07	0.69 ± 0.09	0.71 ± 0.10	0.74 ± 0.06	0.78 ± 0.06
	α_t	0.93 ± 0.03	0.96 ± 0.03	0.97 ± 0.02	0.98 ± 0.02	0.98 ± 0.01
	g_r	0.77 ± 0.06	0.76 ± 0.07	0.70 ± 0.10	0.65 ± 0.09	0.58 ± 0.07
Game B	η_B	0.55 ± 0.11	0.66 ± 0.07	0.70 ± 0.06	0.72 ± 0.07	0.73 ± 0.07
	α_P	0.46 ± 0.17	0.46 ± 0.46	0.50 ± 0.14	0.58 ± 0.16	0.59 ± 0.14
	g_r	0.58 ± 0.16	0.68 ± 0.06	0.68 ± 0.06	0.64 ± 0.05	0.65 ± 0.07

Table B.2: Distribution of the mean number of enemies (n_e) and the speed of the enemies (v_e) for Game A across sessions.

		S1	S2	S3	S4	S5
n_e	min	2.40	1.89	1.59	2.21	2.52
	q_1	2.65	2.55	2.76	2.87	2.96
	median	2.89	2.96	3.20	3.26	3.54
	q_3	3.06	3.75	3.65	3.67	3.74
	max	3.22	3.81	3.85	3.89	3.89
	IQR	0.41	1.20	0.89	0.80	0.78
v_e	min	0.59	0.62	0.74	0.70	0.80
	q_1	0.72	0.80	0.85	0.89	1.00
	median	0.73	0.90	0.98	1.05	1.15
	q_3	0.75	0.99	1.14	1.16	1.30
	max	0.82	1.20	1.25	1.22	1.51
	IQR	0.03	0.19	0.29	0.27	0.51

Table B.3: Distribution of the mean number of enemies (n_e) and the speed of the enemies (v_e) for Game B across sessions.

		S1	S2	S3	S4	S5
n_e	min	1.48	1.07	1.15	1.37	1.33
	q_1	1.17	1.41	1.63	2.44	2.31
	median	2.11	1.74	2.33	2.70	2.81
	q_3	2.31	2.43	2.72	2.86	2.89
	max	2.52	2.96	2.89	2.89	2.96
	IQR	1.14	1.55	1.09	0.42	0.58
v_e	min	0.54	0.38	0.32	0.29	0.41
	q_1	0.65	0.55	0.57	0.58	0.62
	median	0.68	0.74	0.78	0.80	0.77
	q_3	0.73	0.79	0.86	0.86	0.91
	max	0.83	1.09	1.08	1.05	1.07
	IQR	0.08	0.24	0.29	0.28	0.29

B.2 SELF-ASSESSMENT MANIKIN RESULTS

Table B.4: The mean and standard deviation of valence, dominance, and arousal reported by the SAM test are summarized across sessions.

		S1	**S2**	**S3**	**S4**	**S5**
Game A	Valence	6.69 ± 1.60	6.00 ± 1.41	6.08 ± 2.63	5.62 ± 2.40	6.77 ± 1.17
	Arousal	5.46 ± 1.94	5.30 ± 2.17	5.69 ± 1.70	5.46 ± 1.90	5.54 ± 1.76
	Dominance	5.92 ± 1.38	5.62 ± 1.66	5.62 ± 2.21	5.08 ± 1.89	5.38 ± 2.10
Game A	Valence	7.69 ± 1.03	7.15 ± 2.12	7.15 ± 1.21	6.62 ± 1.85	7.39 ± 1.19
	Arousal	5.31 ± 2.02	4.31 ± 2.25	4.31 ± 1.84	4.39 ± 2.33	3.62 ± 1.56
	Dominance	6.54 ± 1.05	6.77 ± 1.69	6.77 ± 1.54	6.08 ± 1.85	6.92 ± 1.98

C

MAIN CONTRIBUTIONS

This Doctoral Thesis is supported by a compendium of previously published works, two of them in an impact journal indexed according to JCR Science Edition. The body of this thesis is partially constituted by the following articles, whose complete bibliographic references are indicated below:

Martínez-Pascual, D.; Catalán, J. M.; Blanco-Ivorra, A.; Sanchís, M.; Arán-Ais, F. and García-Aracil, N. (2023), «Estimating vertical ground reaction forces during gait from lower limb kinematics and vertical acceleration using wearable inertial sensors,» *Frontiers in Bioengineering and Biotechnology* **11**, DOI: https://doi.org/10.3389/fbioe.2023.1199459

- Journal title: Frontiers in Bioengineering and Biotechnology (ISSN: 1530-437X)
- Impact Factor JCR-SCI (2022): 5,7
- Category: Multidisciplinary Sciences, Quartile Q1 (16/73).

Martínez-Pascual, D.; Catalán, J. M.; Blanco-Ivorra, A.; Sanchís, M.; Arán-Ais, F. and García-Aracil, N. (2024), «Gait Activity Classification with Convolutional Neural Network using Lower Limb Angle Measurement from Inertial Sensors,» *IEEE Sensors Journal*, pp. 1–1, DOI: 10.1109/JSEN.2024.3400296

- Journal title: IEEE Sensors Journal (ISSN: 2296-4185)
- Impact Factor JCR-SCI (2022): 4,3
- Category: Instruments and Instrumentation, Quartile Q1 (15/63).

The documents corresponding to these publications are attached below.

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 29 September 2023
DOI 10.3389/fbioe.2023.1199459

OPEN ACCESS

EDITED BY
Marwan El-Rich,
Khalifa University, United Arab Emirates

REVIEWED BY
Robert LeMoyne,
Northern Arizona University,
United States
Abdul Aziz Hulleck,
Khalifa University, United Arab Emirates

*CORRESPONDENCE
David Martínez-Pascual,
david.martinezp@umh.es

RECEIVED 03 April 2023
ACCEPTED 20 September 2023
PUBLISHED 29 September 2023

CITATION
Martínez-Pascual D, Catalán JM, Blanco-Ivorra A, Sanchís M, Arán-Ais F and García-Aracil N (2023), Estimating vertical ground reaction forces during gait from lower limb kinematics and vertical acceleration using wearable inertial sensors.
Front. Bioeng. Biotechnol. 11:1199459.
doi: 10.3389/fbioe.2023.1199459

Estimating vertical ground reaction forces during gait from lower limb kinematics and vertical acceleration using wearable inertial sensors

David Martínez-Pascual[1]*, José M. Catalán[1], Andrea Blanco-Ivorra[1], Mónica Sanchís[2], Francisca Arán-Ais[2] and Nicolás García-Aracil[1]

[1]Biomedical Neuroengineering Research Group, Robotics and Artificial Intelligence Unit, Bioengineering Institute, Miguel Hernandez University, Elche, Spain, [2]INESCOP Footwear Technology Center, Elda, Alicante, Spain

One of the most important forces generated during gait is the vertical ground reaction force (vGRF). This force can be measured using force plates, but these can limit the scope of gait analysis. This paper presents a method to estimate the vGRF using inertial measurement units (IMU) and machine learning techniques. Four wearable IMUs were used to obtain flexion/extension angles of the hip, knee, and ankle joints, and an IMU placed over the C7 vertebra to measure vertical acceleration. We trained and compared the performance of two machine learning algorithms: feedforward neural networks (FNN) and random forest (RF). We investigated the importance of the inputs introduced into the models and analyzed in detail the contribution of lower limb kinematics and vertical acceleration to model performance. The results suggest that the inclusion of vertical acceleration increases the root mean square error in the FNN, while the RF appears to decrease it. We also analyzed the ability of the models to construct the force signal, with particular emphasis on the magnitude and timing of the vGRF peaks. Using the proposed method, we concluded that FNN and RF models can estimate the vGRF with high accuracy.

KEYWORDS

biomechanical analysis, gait analysis, magneto-inertial devices, machine learning, vertical ground reaction force

1 Introduction

The analysis of human kinematics and kinetics has great potential in the prevention of injuries (Padua and DiStefano, 2009), and the analysis or prediction of musculoskeletal disorders (Fineberg et al., 2013). The analysis of human kinematics requires the use of motion capture data. Computer vision techniques (Graci et al., 2012), but also wearable inertial sensors (Sy et al., 2020) can be used for estimating the angle of the human limbs. On the other hand, when performing a kinetic analysis, the human joint forces and torques are estimated by the forces exerted, using inverse dynamics methods (Moya-Angeler et al., 2017).

In a lower extremity biomechanical analysis, the kinetic data must include the ground reaction forces produced (Lencioni et al., 2019), which can be divided into the mediolateral ground reaction force, the anteroposterior ground reaction force, and the vertical ground reaction force (Kram et al., 1998). Vertical ground reaction force (vGRF) is the largest component of the ground-generated forces during gait. The vGRF is generated in the sagittal plane and represents the magnitude and pattern of mechanical loading in the vertical direction at the foot (Jacobs et al., 1972). During gait, three main peaks can be distinguished in this force (Marasović et al., 2009). First, the loading peak (LP) is produced, a local maximum corresponding to the loading response. Then, there is a local minimum corresponding to the mid-stance phase (MP). Finally, there is a second local maximum during the terminal stance phase (TP). The magnitude and timing of these peaks has influence in the lower limb joints and muscle loads. For this reason, the analysis of vGRF characteristic peaks could be essential to study musculoskeletal disorders (Shafizadegan et al., 2016).

The vGRF measurement is usually performed by employing force plates. The force plates can rest over the ground, but also they can be embedded in treadmills. These plates might limit the space for the gait analysis, and they also can be financially costly. As a consequence of these limitations, some authors have studied the possibility of estimating the vGRF.

To overcome these drawbacks, the use of inertial measurement units (IMU) to estimate vGRF has been investigated. E. Sahabpoor and A. Pavic showed that it is possible to estimate the vGRF with a single IMU, using the subjects' body mass and vertical acceleration to estimate the vGRF exerted on the human body. (Shahabpoor and Pavic, 2018). The authors compared several locations and concluded that placing an IMU over the C7 vertebra was the better location to achieve the lowest error in the vGRF estimation. However, the method proposed by these authors is not able to determine the force produced on the right and left lower limbs.

The development and improvement of machine learning algorithms have led some authors to apply these techniques to analyze gait kinematics and estimate vGRF.

In (Choi et al., 2013; Oh et al., 2013) the authors used human body kinematics and artificial neural networks to perform vGRF prediction. However, the main disadvantage is that they used several infrared cameras to acquire motion capture data, so this method might also limit the analysis space. On the other hand, other authors preferred to place IMUs over the lower limbs to monitor human movement. X. Jiang et al. used a Random Forest model to estimate the vGRF during gait with a single IMU using acceleration and gyroscope data (Jiang et al., 2020). F. J. Wouda et al. also used IMU sensors over the lower limbs (pelvis and lower legs) (Wouda et al., 2018). These authors used an artificial neural network to estimate the knee angle and artificial neural networks to estimate the vGRF during running using the estimated knee angle and the vertical accelerations from the placed IMUs.

The scientific literature shows that progress in machine learning algorithms has made it possible to achieve good accuracy when estimating the vGRF by different approaches. When performing a biomechanical analysis, we usually want to study not only the kinetics but also the lower limb kinematics. Hence, we intend to develop a Machine Learning model that makes use of the lower limb joint kinematics to estimate the vGRF. Moreover, it has to be mentioned that we have employed a method to estimate the hip, knee, and ankle angles in the sagittal plane by employing IMUs, so the gait analysis could be performed in any situation.

Moreover, as mentioned before, apart from the lower limb kinematics, vertical acceleration has been used to estimate the vGRF since there is a closed relation between them. However, when including these features together, the contribution to the vGRF prediction of lower limb kinematics and vertical acceleration has not been analyzed.

This work aims to develop a system to perform biomechanical analysis of the lower limbs based on wearable inertial sensors. For this purpose, we have used a method to measure the angle of the lower limb joints, and we have trained and compared the performance of machine learning models to estimate the vGRF during gait. We have introduced the lower limb joint kinematics and vertical acceleration measured at the C7 vertebra as inputs of the machine learning models to estimate the vGRF. In addition, we have analyzed the contribution to the vGRF estimation of this set of features using different machine learning methods.

2 Materials and methods

2.1 Lower limb joints measurement

The method proposed by T. Seel et al. has been used to estimate the joint angles of the lower limbs during gait (Seel et al., 2014). The proposed method is based on the analysis of the movement of the joints and it allows obtaining the joint angles with high accuracy with IMUs even at low acquisition rates (from 40 Hz) (Seel et al., 2012). This section describes the method and details the calibration and angle measurement process.

2.1.1 Identification of the joint axis and position coordinates

This method avoids developing an algorithm that assumes a position and orientation of the different IMUs concerning the user's lower limb. Moreover, to estimate joint angles by this algorithm, we must use gyroscope and accelerometer data from the IMUs, so it does not depend on a uniform magnetic field. Hence, we assume we have two IMUs attached to the upper and lower segment for each joint, so we measure the accelerations $a_1(t), a_2(t) \in \mathbb{R}^3$ and the angular rates $g_1(t), g_2(t) \in \mathbb{R}^3$ at a sample period Δt.

The gyroscope data are used to identify the direction vectors $j_1, j_2 \in \mathbb{R}$ of the hip, knee, and ankle flexion/extension axis in the local coordinates of the IMUs. Geometrically, $g_1(t)$ and $g_2(t)$ differ only by the joint angular velocity and a rotation matrix, so their projections into the joint plane have the same lengths for each instant. This can be defined as:

$$\|g_1(t) \times j_1\|_2 - \|g_2(t) \times j_2\|_2 = 0 \ \forall t \tag{1}$$

where $\| \cdot \|_2$ is the Euclidean norm.

Moreover, apart from the direction vectors, the joint center position in the local coordinates of the IMUs must be calculated. These vectors $o_1, o_2 \in \mathbb{R}$ are constant and only depend on the mounting position and orientation over the leg segments. In order to calculate them, T. Seel et al. propose that the acceleration of each sensor can be thought as the sum of the center of the joint acceleration, and the acceleration due to the rotation of the sensor around the joint center. The acceleration of the joint center should be the same in both IMU local frames, which can be expressed as:

$$\|a_1(t) - \Gamma_{g_1(t)}(o_1)\|_2 - \|a_2(t) - \Gamma_{g_2(t)}(o_2)\|_2 = 0 \ \forall t \quad (2)$$

$$\Gamma_{g_i(t)}(o_i) \coloneqq g_i(t) \times \left(g_i(t) \times o_i\right) + \dot{g}_i(t) \times o_i \ , \quad i = 1, 2 \quad (3)$$

where $\Gamma_{g_i(t)}(o_i)$ is the radial and tangential acceleration due to the IMU rotation around the joint center.

It should be noted that the constraints defined must be fulfilled by any given motion of the joint. Hence, we can identify j_1, j_2, o_1, and o_2 by minimizing the left side of Eqs 1, 2.

To determine j_1 and j_2, if we restrict the axis to the unit length, the problem becomes four-dimensional and we can express j_1 and j_2 in spherical coordinates:

$$j_1 = \begin{bmatrix} cos(\phi_1) \cdot cos(\theta_1) \\ cos(\phi_1) \cdot sin(\theta_1) \\ sin(\phi_1) \end{bmatrix}, j_2 = \begin{bmatrix} cos(\phi_2) \cdot cos(\theta_2) \\ cos(\phi_2) \cdot sin(\theta_2) \\ sin(\phi_2) \end{bmatrix} \quad (4)$$

where ϕ_i and θ_i are inclination and azimuth. We also can define the sum of squared errors for N samples as:

$$\Psi(\phi_1, \theta_1, \phi_2, \theta_2) \coloneqq \sum_{i=1}^{N} \left(\|g_1(t_i) \times j_1\|_2 - \|g_2(t_i) \times j_2\|_2\right)^2 \quad (5)$$

and in the same manner, we can also obtain o_1 and o_2 by defining another sum of squared errors:

$$\Psi(o_1, o_2) \coloneqq \sum_{i=1}^{N} \left(\|a_1(t_i) - \Gamma_{g_1(t_i)}(o_1)\|_2 - \|a_2(t_i) - \Gamma_{g_2(t_i)}(o_2)\|_2\right)^2 \quad (6)$$

To minimize $\Psi(\phi_1, \theta_1, \phi_2, \theta_2)$ and $\Psi(o_1, o_2)$ we have employed the Trust Region Reflective algorithm (Byrd et al., 1988), but the problem could be resolved by employing any other algorithm, e.g., a Gauss-Newton algorithm.

To obtain the hip, knee, and ankle joint position coordinates (o_1, o_2), before the experimental session started the users were told to perform circling arbitrary movements. Moreover, concerning the estimation of the joint axis coordinates (j_1, j_2), we got data from the users walking to identify the principal axis of motion, i.e., the flexion/extension movements of the hip, knee, and ankle.

2.1.2 Flexion/extension angle measurement

Once we have successfully obtained the joint axis coordinates (j_1, j_2) and the joint position coordinates (o_1, o_2) for each user, we can estimate the angle of the joints with the placed IMUs. These angles are calculated from accelerations and angular rates. Using the gyroscope data, the flexion/extension angles can be estimated by integrating the difference of the angular rates around the joint axis:

$$\alpha_{gyr}(t) = \int_0^t \left(g_1(\tau) \cdot j_1 - g_2(\tau) \cdot j_2\right) d\tau \quad (7)$$

As T. Seel et al. mention, by knowing the axis coordinates we can apply methods that need the axes of the sensor coincide with the joint or the segment axes (Liu et al., 2009). First, we move the measured accelerations from the local IMU coordinates to the joint axis:

$$\tilde{a}_1 = a_1(t) - \Gamma_{g_1(t)}(o_1), \quad \tilde{a}_2 = a_2(t) - \Gamma_{g_2(t)}(o_2) \quad (8)$$

The flexion/extension can be estimated by the angle between the projections of $\tilde{a_1}$ and $\tilde{a_2}$ into the joint plane. Then, we define a pair of joint plane axes $x_1, y_1, x_2, y_2 \in \mathbb{R}^3$ for the local frames:

$$x_1 = j_1 \times c, \quad y_1 = j_1 \times x_1, \quad x_2 = j_2 \times c, \quad y_2 = j_2 \times x_2 \quad (9)$$

where $c \in \mathbb{R}^3$ could be any vector that makes none of the products zero. Finally, we can estimate the joint angle by accelerometer readings:

$$\alpha_{acc}(t) = \angle\left(\begin{bmatrix} \tilde{a_1}(t) \cdot x_1 \\ \tilde{a_1}(t) \cdot y_1 \end{bmatrix}, \begin{bmatrix} \tilde{a_2}(t) \cdot x_2 \\ \tilde{a_2}(t) \cdot y_2 \end{bmatrix}\right) \quad (10)$$

This α_{acc} is not affected by drift as α_{gyr} is, since we have not employed any integration to compute the angle, but it is affected by the accelerometer noise. Moreover, despite the drift, α_{gyr} is precise in short time scales. Hence, it could be appropriate to combine α_{acc} and α_{gyr}. We have used a complementary filter (Young, 2009) defined as:

$$\alpha(t) = \lambda \cdot \alpha_{acc}(t) + (1 - \lambda) \cdot \left(\alpha(t - \Delta t) + \alpha_{gyr}(t) - \alpha_{gyr}(t - \Delta t)\right) \quad (11)$$

To obtain flexion/extension data from the participants, we have used a sample period $\Delta t = 0.016$ (60 Hz) and $\lambda = 0.02$.

The method proposed by T. Seel et al. has shown a root mean squared error (RMSE) of between 3 and 5° for the hip, knee, and ankle flexion/extension movements (Kumar et al., 2018). Therefore, we can assume that we are accurate enough to perform an adequate gait analysis and vGRF.

2.2 Experimental setup

The equipment used during the experimental sessions is shown in Figure 1.

Five XSens Dot IMUs were used to acquire kinematic data from the participants at a sampling rate of 60 Hz. Four inertial sensors were placed on the back of the hip, the thigh, the leg, and the foot to estimate the joint angles of the lower limb. The remaining IMU was placed over the C7 vertebra to provide vertical acceleration data. The hip, thigh, and leg IMUs were placed with elastic straps, and the C7 and foot sensors were attached to the skin and the user's shoes with stickers.

A treadmill with an embedded force plate model h/p/cosmos 150/50 was used to measure the generated vGRF during gait. The vGRF data were recorded at 100 Hz.

2.3 Subjects

Twelve volunteer able-bodied subjects participants were involved in the experimental sessions, 10 male and 2 female. The ages were between 23 and 52 years old (29.8 ± 7.4), with heights ranging between 1.65 m and 1.87 m (176.2 ± 7.4 cm) and weights between 56.1 kg and 90.2 kg (76 ± 12.5 kg). Written informed consent was obtained from the individuals for the publication of any potentially identifiable images or data included in this article.

2.4 Experimental protocol

At the beginning of the experimental session, the IMUs were placed on the participants at the established locations, and the height and weight of users were measured. Once the user wore the sensors, we proceeded to obtain the calibration data to estimate the joint

FIGURE 1
The devices employed during the experimental sessions. On the left image, the location of the IMUs placed to measure vertical acceleration (C7 IMU) and estimate the joint angles of the lower limb (hip, thigh, leg and foot IMUs). On the right image, the treadmill h/p/cosmos 150/50, used to acquire the vGRF during gait is shown.

angles of the lower limb. First, the users were told to perform arbitrary circling movements with the hip, knee, and ankle to obtain the center position coordinates o_1, o_2 of each joint for 1 min. Then, the participant walked for 1 min over the treadmill at a comfortable speed to obtain the flexion/extension joint axis coordinates j_1, j_2.

After the calibration was carried out, the subjects walked over the treadmill at four different speeds selected to cover a wide range of normal gait speeds Weber (2016): 1.5 km/h, 2.5 km/h, 3.5 km/h, and 4.5 km/h. When the treadmill reached the desired speed, the kinematics and vGRF data were recorded for 5 min. Between each activity, 2 min were left for rest.

2.5 Acquired data and processing

The data acquired from each device and the signals processing to be used as the machine learning models inputs are detailed in this section and illustrated in Figure 2.

The vGRF was acquired from the plate-instrumented treadmill, which has been normalized by the body weight of the participants (BW).

From the hip, thigh, leg, and ankle IMUs, we acquired accelerometer and gyroscope data to estimate the lower limb joint angles in the sagittal plane. Once the hip, knee, and ankle angles were computed, we applied a forward-backward low-pass filter (Gustafsson, 1996) with a 6 Hz cutoff and we obtain the angular velocities of each joint.

To obtain the vertical acceleration we acquired the acceleration and orientation of the IMU placed over the C7 vertebra (Esser et al., 2009). First, by using the IMU orientation, we transform the acceleration vector from the local to the global coordinates system. Then, to obtain the vertical acceleration, we remove gravity from the measured vertical acceleration, which corresponds to the Z-axis acceleration in the global coordinates system.

In addition, it is usual to identify the onset and end of the gait cycle and transform the data from the temporal domain to the gait cycle domain (0%–100%) when performing a gait analysis. Hence, all signals have been transformed to the cycle gait domain. To detect the onset of the steps we have detected the foot-ground contact detection by two methods. To identify the foot contact in the IMU data (inputs of the model) we have made use of the hip acceleration (Zijlstra and Hof, 2003). We applied a forward-backward low-pass filter with a 2 Hz cutoff. In the filtered signal, we detect the local maxima, which correspond to the left and right ground foot contacts. As a matter of convention, we employed the right foot contact to calculate the gait cycle. Moreover, we detected the foot-ground contact when the force value starts growing in the vGRF signal.

Signal synchronization begins by detecting the first foot contact with the ground in both the vGRF and the IMU system. In addition, by detecting the ground-foot contact in the vGRF and the IMU system the onset and end of each step can be detected. For each step, each vGRF and lower limb kinematics data is assigned a label according to the gait instant in which it occurs (gait cycle percentage). Finally, to train the machine learning models, the vGRF and leg kinematics data are synchronized from the gait instant in which they occur, so that each lower limb kinematics data is assigned a vGRF value.

The flexion/extension angles, the vertical acceleration, and the vGRF acquired are shown in Figure 3 for each gait speed. When the treadmill reached the desired speed, the kinematics and vGRF data were recorded for 5 min. Between each activity, 2 min were left for rest.

2.6 Machine learning models training

2.6.1 Selected machine learning algorithms

Regression techniques can be used in order to estimate the vGRF. We decided to train Random Forest (RF) models and Feed-forward Neural Networks (FNN) since their use has been validated in previous studies by different authors.

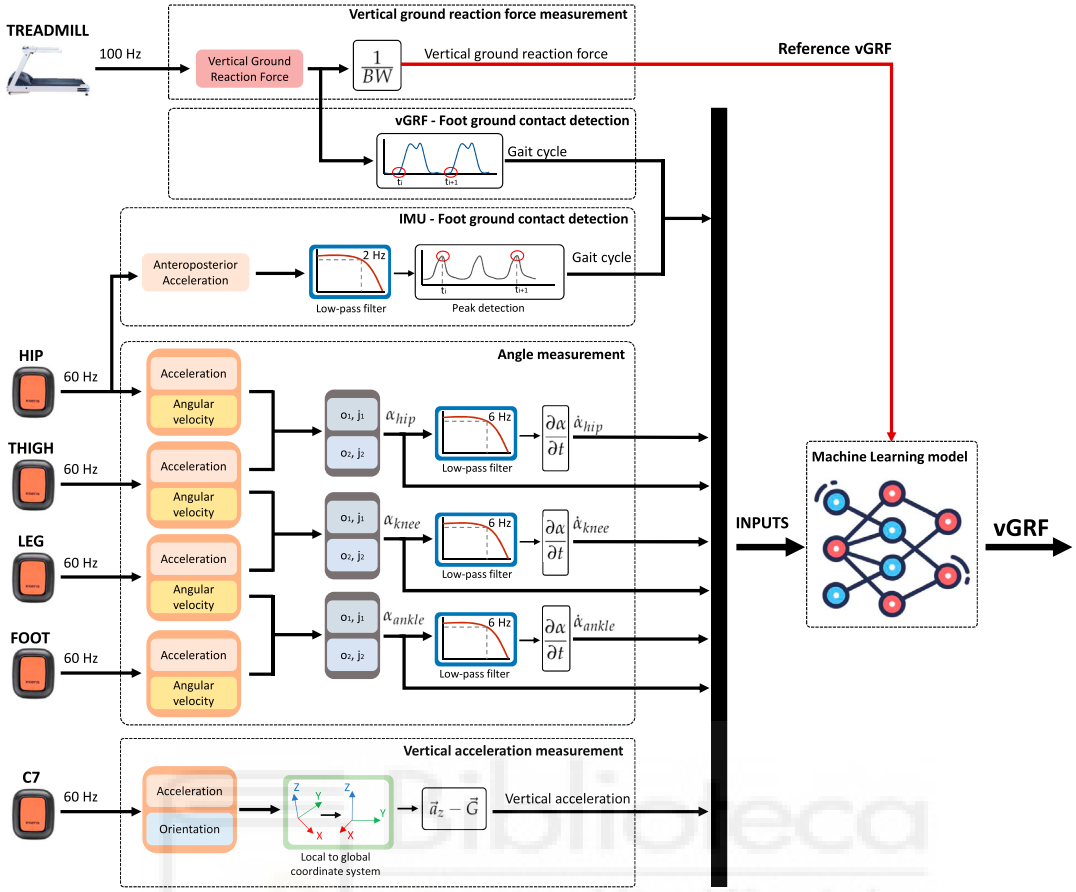


FIGURE 2
Signals obtained by the IMUs and the treadmill, and their processing before introducing them in the machine learning models. From the IMUs placed on the lower limbs, we have acquired acceleration and angular velocity to estimate the flexion/extension angles (T. Seel et al. method). With the IMU placed over the C7 vertebra, we measured the vertical acceleration. Moreover, the foot-ground contact has been detected using the anteroposterior acceleration measured at the hip to compute the gait cycle. The generated vGRF during gait has been measured and normalized by body weight. In the vGRF signal, the gait foot contact detection has been performed by finding the end of the stance phase (null force). The measured vGRF and the kinematics data have been synced by using the cycle gait. The flexion/extension joint angles and velocities, the vertical acceleration, and the cycle gait have been introduced as inputs of the models. The measured vGRF has been introduced as a reference for model learning.

The RF models consist of multiple decision trees (Zhang and Ma, 2012). These decision trees are trained independently with a data subset, and the output of the model will be determined by the best result given by the ensemble trees. This strategy allows the RF models to achieve accurate predictions as well as better generalizations, which come from the bagging scheme by decreasing variance.

The RF models have been trained with SciKit Learn Python Library (Pedregosa et al., 2011). RF training was performed with 170 decision trees, a tree maximum depth of 35, a minimum number of samples of 1, a minimum number of samples required to split an internal node of 2, and the number of inputs to consider when looking for the best split were calculated by the logarithm base 2 of the number of inputs.

The FNN models are a kind of artificial neural networks based on neurons that receive an input signal and process them by the activation function to an output signal (Svozil et al., 1997; Sharma et al., 2017). These neurons are ordered into layers, where the first layer is called the input layer, the last layer is called the output layer, and the layers between them are called the hidden layers. Each neuron of a particular layer is connected with all the neurons placed in the next layer, connected by the weights coefficients. The adjustment of the weights of the neural network is usually performed by back-propagation, employing a non-linear optimization method such as the gradient descent algorithm (Andrychowicz et al., 2016).

The FNN models have been trained with Keras Python Library (Gulli and Pal, 2017). We have employed FNN with one input layer of a number of neurons equal to the number of inputs, 5 hidden layers with 10 neurons, and an output layer with one neuron. We have used the Rectified Linear Unit (ReLu) as the activation function of the neurons, the Adam algorithm (Kingma and Ba, 2014) as the optimizer to adjust the network weights, and the RMSE as the loss function. To avoid overfitting, we have employed the dropout regularization method (Srivastava et al., 2014). Moreover, as recommended when using dropout, a constraint is imposed on the weights for each hidden layer. We have imposed that the maximum norm of the weights does not exceed a value of 4.

The hyperparameters tunned for both RF and FNN models are the best performing based on the RMSE with the training data.

10.3389/fbioe.2023.1199459

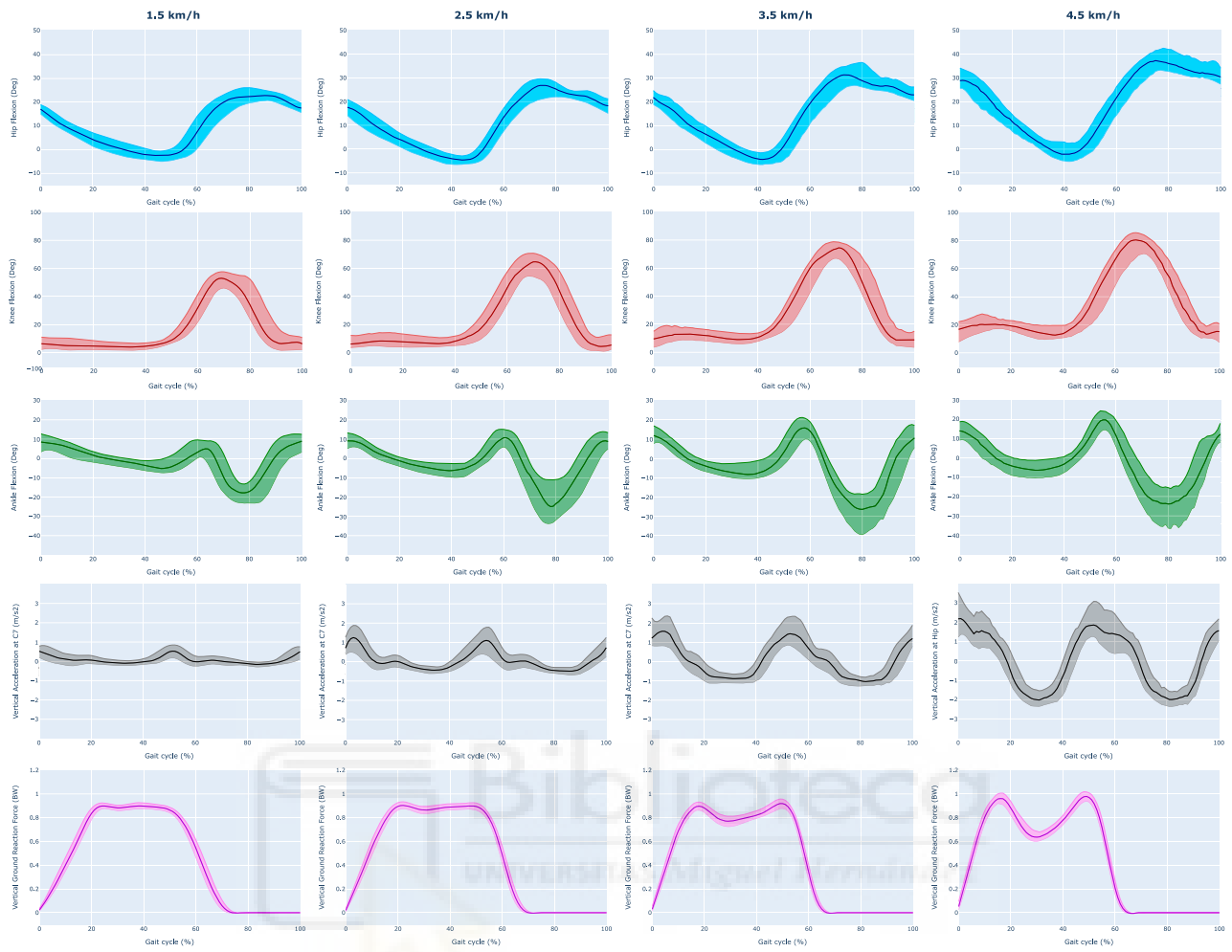


FIGURE 3
Acquired data during the experimental session. We have represented the hip, knee, and ankle flexion/extension angles (first, second, and third row), the vertical acceleration measured at the C7 vertebra (fourth row), and the measured vGRF (fifth row). The data is represented for different gait speeds (1.5–4.5 km/h) and normalized by the cycle gait.

2.6.2 Preprocessing and training

For the training of the RF and FNN models, we introduced the joint angles, angular velocities, vertical acceleration, and cycle gait as inputs (Figure 2). In addition, the vGRF measured with the instrumented treadmill was introduced as the reference of the model.

We divided the data into 3 subsets to evaluate the performance of the models. First, we used data collected from 10 participants to train the model (intra-participants). This data collection was randomly divided into two subsets to validate the accuracy of the models by cross-validation: 80% of the data for training and the remaining 20% for validation. Data from 2 participants were also used to evaluate the performance of the models on users not involved in the learning process (inter-participants).

To train the models, 8 inputs were used: the flexion/extension angles and angular velocities of the hip, knee, and ankle; the vertical acceleration measured at the C7 vertebra; and the instant of the gait cycle.

It must be noted that we have scaled each input of the training data from its minimum and maximum, so we have numerical values between 0 and 1. We have also used the training scaler to evaluate the accuracy of the classifier with the test data.

As we introduced, one of the objectives of this work is to examine the contribution of vertical acceleration and lower limb kinematics in the vGRF estimation. To analyze the contributions, we realized two different pieces of training for each model.

- **Training 1 (Kinematics).** Joint angles, angular velocities, and cycle gait have been included as inputs of the models.
- **Training 2 (C7).** Joint angles, angular velocities, vertical acceleration, and cycle gait have been included as inputs of the models.

2.6.3 Performance evaluation of the trained models

The performance achieved by the trained models has been analyzed as follows.

- **vGRF RMSE (BW).** A lower RMSE in the reproduction of the vGRF signal is understood as a better performance of the model.

- **Normalized RMSE (BW).** The Normalized RMSE (NRMSE) is defined as (Shahabpoor and Pavic, 2018):

$$NRMSE = \frac{RMSE}{max\ vGRF_{measured} - min\ vGRF_{measured}} \quad (12)$$

A lower normalized RMSE (NRMSE) in the construction of the vGRF is understood as a better performance of the model.

- **Correlation.** A higher correlation (ρ) between the measured and estimated vGRF signals is understood as a better performance of the model. To calculate the correlation, the Person's correlation coefficient has been calculated.
- **LP, MP, and TP error (BW).** A lower error between measured and estimated vGRF peaks is understood as a better performance of the model. To find the LP and TP peaks we searched the local maximum in the first and second periods of the foot support. To calculate the MP peak we searched the local minimum between LP and TP peaks.
- **LP, MP, and TP delay (%).** A delay or advance close to 0 between measured and estimated vGRF peaks is understood as a better performance of the model. The delay is calculated as the difference between the estimated and measured peaks. A positive value means a delay of the estimated peak, and a negative value means an advance of the estimated peak to the measured peak.

2.6.4 Statistical data analysis

We performed statistical data analysis to compare the accuracy of the models in estimating the vGRF with the validation and test data. First, we compared the four trained models (FNN-Kinematics, FNN-C7, RF-Kinematics, and RF-C7) in terms of RMSE. To perform a deep analysis of the vGRF signal construction, we have selected the best FNN and RF models, taking into account the evaluation detailed previously. For each model, we compared the vGRF estimation and peak reproduction with the validation and test data.

In the statistical analysis, first, a normality test was performed using the Kolmogorov-Smirnov test (Berger and Zhou, 2014). The results of the test showed that data distribution was not normal ($p < 0.05$).

The Friedman test was used to study the differences between the accuracy of the models (Sheldon et al., 1996). In the *post hoc* analysis, pairwise comparisons have been studied by the Wilcoxon signed-rank test with the zero method proposed by Pratt (1959).

3 Results

3.1 Feature importance

To understand the influence of the angles, the angular velocities, the vertical acceleration, and the cycle gait in the training of the FNN and RF models, we have calculated the feature importance by feature permutation (Altmann et al., 2010). This method is defined to be the increase in the model error when a single feature is randomly shuffled, so we could understand how much the vGRF depends on each feature. The results obtained are shown in Figure 4, which shows the increase of the RMSE in the vGRF prediction when a feature is shuffled with the FNN-C7 model (Figure 4A and the RF-C7 model (Figure 4B).

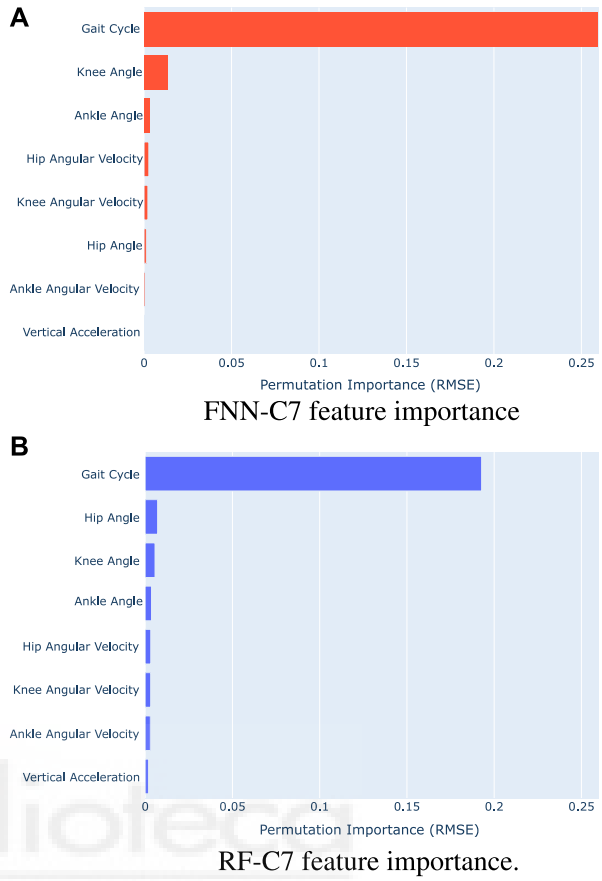


FIGURE 4
Feature importance computed with feature permutation **(A)** FNN-C7 model (Training 2) **(B)** RF-C7 model (Training 2)

The feature permutation results show that the gait cycle has the greatest influence on the performance of the models since the RMSE increases by 0.2590 BW in the FNN-C7 model and 0.1927 BW in the RF-C7 model. Concerning the lower limb kinematics, it can be observed that the knee and the ankle angles have the highest relevance on the FNN-C7, followed by the hip and knee angular velocity. By contrast, in the RF-C7 training the hip, knee, and ankle angles take on greater significance than their angular velocities. In both cases, the ankle angular velocity has lower importance than the hip and knee kinematics.

According to the results, vertical acceleration has a low impact on the training of the models compared to the lower limb kinematics. In the FNN-C7 model, the vertical acceleration has the second lowest importance, and in the RF-C7 model, it has the lowest relevance.

3.2 Validation and test results

Once we performed the two established pieces of training (Kinematics, C7), we evaluated the accuracy of the models with the validation and test data. We have calculated the RMSE for each

10.3389/fbioe.2023.1199459

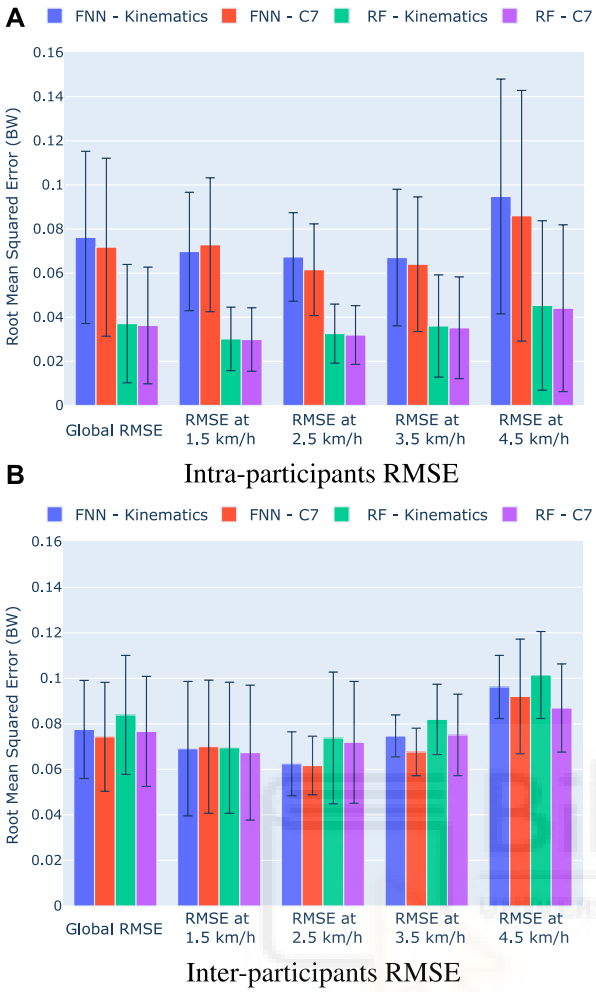


FIGURE 5
Mean RMSE (BW) and standard deviation for each training of the FNN and RF models. The graphs show the global RMSE and the RMSE per gait speed with **(A)** The intra-participant data **(B)** the inter-participant data.

of the registered steps to evaluate the performance of the models with the intra-participants and inter-participants. The mean and standard deviations of the RMSE errors are shown in Figure 5, where Figure 5A show the results with the intra-participants, and Figure 5B the results with the inter-participants. In addition, Table 1 and Table 2 show respectively the validation and test results.

The statistical analysis showed significant differences between the performance of the models with the validation data for each gait speed (Friedman test $p < 0.0001$ for all speeds). In the pairwise comparison, the results show that there is not a significant difference in the accuracy of the models when the users' vertical acceleration is included as an input of the model for each gait speed ($p > 0.05$). However, the inclusion of the vertical acceleration seems to reduce significantly the global error when using the FNN models ($p < 0.05$). When we compare the FNN and RF performance, we see a significant reduction of the RMSE when we employ the RF models with the validation data for the global RMSE errors and for all gait speeds. When we analyze the effect of the vertical acceleration, despite there is not an statistical significance, we observe a reduction of the RMSE errors in the RF training when the vertical acceleration is included, specially for the gait speeds of 3.5 and 4.5 km/h. This behavior can also be seen when using the FNN models. However, it should be noted that, when we employ a FNN model, the inclusion of the vertical acceleration is detrimental to the model performance at 1.5 km/h.

The analysis of the RMSE errors obtained with the inter-participants shows that there is not a significant difference between the accuracy of the models (Friedman test for global RMSE $p = 0.24$, at 1.5 km/h $p = 0.80$, at 2.5 km/h $p = 0.80$, at 3.5 km/h $p = 0.17$, and at 4.5 km/h $p = 0.24$). Nevertheless, it can be observed that there is also a reduction of errors in both models when the vertical acceleration is included. Moreover, we can observe similar results in terms of RMSE errors if we compare the FNN and RF models.

The results suggest that the FNN-C7 model and the RF-C7 model have higher accuracy in terms of RMSE errors. Hence, both models are analyzed deeply in the next section.

3.3 Validation and test vGRF signal reconstruction

In Figure 6 we have represented the measured vGRF, the vGRF estimated with the FNN-C7 model, and the vGRF estimated with the RF-C7 model. In these graphs, the median value of the vGRF is represented during the gait cycle, and the shaded areas correspond to the vGRF values between the first and third quartiles.

Table 3 and Table 4 collects the correlation (ρ) and NRMSE between the measured and estimated vGRF with the selected models for the intra-participants and the inter-participants, respectively. Figure 7 shows heatmaps with the mean of the peaks magnitude errors and delays with the intra-participants and inter-participants. Moreover, the Supplemental Data collects detailed information about the magnitude error and delay of the characteristic peaks, and the pairwise comparison results between the selected models (FNN-C7 and RF-C7).

With the intra-participants, the results show ρ values in the vGRF reconstruction between 0.9676 and 0.9886 with the FNN-C7 model, and coefficients between 0.9966 and 0.9988 with the RF-C7 model. In addition, the results show NRMSE values between 0.0399 BW and 0.0651 BW with the FNN-C7 model, and values between 0.0127 BW and 0.0210 BW with the RF-C7 model. When the vGRF is predicted with the inter-participants the correlation is reduced, with ρ values between 0.9705 and 0.9876 with the FFN-C7 model, and values between 0.9746 and 0.9849 with the RF-C7 model. According to the NRMSE results, values are between 0.0560 BW and 0.0853 BW with the FNN-C7 model, and values are between 0.0691 BW and 0.0792 BW with the RF-C7 model.

The analysis of the vGRF peaks shows statistical differences with the intra-participants for all the characteristic peaks ($p < 0.0001$). The results collected in Figure 7 show that all the peak magnitude errors obtained with the RF-C7 model are lower than errors obtained with the FNN-C7 model. The RF-C7 achieves magnitude errors lower than 0.0210 BW for the LP, 0.0161 BW for the MP, and 0.0171 BW for the TP. Otherwise, the FNN-C7 model errors are up to 0.0900 BW in the LP, 0.0793 BW for the MP, and 0.0491 BW for the TP. By contrast, when the models are

TABLE 1 Validation mean (standard deviation) RMSE (BW) for each training of the FNN and RF models. The global RMSE errors for each model are shown. In addition, the table collects the RMSE errors obtained for each gait speed.

	FNN - Kinematics (a)	FNN - C7 (b)	RF - Kinematics (c)	RF - C7 (d)
Global	0.076 (0.039) ***c,d	0.072 (0.040) ***c,d	0.037 (0.027) ***a,b,d	0.036 (0.026) ***a,b,c
1.5 km/h	0.070 (0.027) ***c,d	0.073 (0.030) ***c,d	0.030 (0.014) ***a,b*d,	0.030 (0.015) ***a,b*c,
2.5 km/h	0.067 (0.020) ***c,d	0.062 (0.020) ***c,d	0.033 (0.013) ***a,b,d	0.032 (0.013) ***a,b,c
3.5 km/h	0.067 (0.030) ***c,d	0.064 (0.030) ***c,d	0.036 (0.023) ***a,b,d	0.035 (0.023) ***a,b,c
4.5 km/h	0.095 (0.050) ***c,d	0.086 (0.056) ***c,d	0.045 (0.038) ***a,b,d	0.044 (0.037) ***a,b,c

Statistical differences are represented by * ($p < = 0.05$), ** ($p < = 0.01$), *** ($p < = 0.001$), and **** ($p < = 0.0001$).

TABLE 2 Test mean (standard deviation) RMSE (BW) for each training of the FNN and RF models. The global RMSE errors for each model are shown. In addition, the table collects the RMSE errors obtained for each gait speed.

	FNN - Kinematics	FNN - C7	RF - Kinematics	RF - C7
Global	0.077 (0.022)	0.074 (0.028)	0.083 (0.026)	0.077 (0.024)
1.5 km/h	0.069 (0.030)	0.070 (0.029)	0.070 (0.029)	0.067 (0.030)
2.5 km/h	0.062 (0.014)	0.062 (0.013)	0.074 (0.029)	0.072 (0.027)
3.5 km/h	0.075 (0.010)	0.068 (0.010)	0.082 (0.012)	0.075 (0.018)
4.5 km/h	0.096 (0.014)	0.092 (0.025)	0.102 (0.019)	0.087 (0.019)

No statistical differences were found.

TABLE 3 Correlation (ρ) and normalized RMSE (NRMSE) results obtained between the measured and estimated vGRF with the intra-participants using the FNN-C7 and RF-C7 models.

	Gait speed (km/h)	ρ	NRMSE
FNN-C7	1.5	0.9817	0.0483
	2.5	0.9886	0.0399
	3.5	0.9847	0.0449
	4.5	0.9676	0.0651
RF-C7	1.5	0.9988	0.0118
	2.5	0.9987	0.0127
	3.5	0.9980	0.0155
	4.5	0.9966	0.0210

evaluated with the inter-participant data, the statistical analysis has shown that there are differences in the prediction of the peak magnitude errors between models except for the MP at 3.5 km/h ($p = 0.1825$) and TP at 4.5 km/h ($p = 0.0834$). In the case of the FNN-C7 model, we can find errors between to 0.0206 BW and 0.0619 BW in the LP, between 0.0263 BW and 0.0810 BW in the MP, and between 0.0249 BW and 0.0922 BW in the TP. By contrast, the obtained errors with the RF-C7 model are between 0.0225 BW and 0.0611 BW in the LP, between 0.0399 BW and 0.0983 BW in the MP, and between 0.0316 BW and 0.1256 BW in the TP.

The analysis of the delay of the predicted vGRF peaks with the intra-participants shows that the lags achieved by the RF-C7 model are closely 0%, with mean lags lower than 0.18% (advance in LP at 1.5 km/h), while the timing errors with the FNN-C7 range from −0.21% to 1.6% in the LP, from −0.3% to 3.4% in the MP, and −0.92% to −3% in the TP. Furthermore, an increase in the timing errors can be appreciated for both models when they are evaluated with the inter-participant data. The timing errors obtained with the FNN-C7 model range from −0.84% to −2.2% in the LP, from 1.3% to 4.1% in the MP, and −2.3%–0.73% in the TP, while the RF-C7 model errors range from −2.3% to −4.9% in the LP, from −2.6% to 3% in the MP, and −2.4% to 5.4 in the TP estimation.

4 Discussion

This work presents a method to estimate vGRF using diverse inertial sensors placed over the lower limbs and the C7 vertebra.

The IMUs placed on the hip, thigh, shank, and foot are used to measure the flexion/extension angles of the hip, knee, and ankle joints. For this purpose, the method developed by T. Seel et al. was used to estimate the joint angles by means of the accelerometer and gyroscope data provided by the sensors (Seel et al., 2014). One of the main advantages of this method is that the position and orientation of the IMUs do not influence the estimation of joint angles. Therefore, we can assume that variations in sensor placement will not affect the error made in the estimation of vGRF. Furthermore, the calibration method of this algorithm is simple and does not require precise movements, as the calibration is based on walking and performing random movements with the legs for a short period.

We have trained and evaluated the performance of two machine learning models in order to estimate the vGRF during gait. First, we have calculated the feature importance for the FNN-C7 and RF-C7 models to understand the relevance of each input (Figure 4). For both models, the results show that the gait cycle has the biggest influence on the training of the model, which suggests that the proposed method for estimating vGRF has a strong dependence on

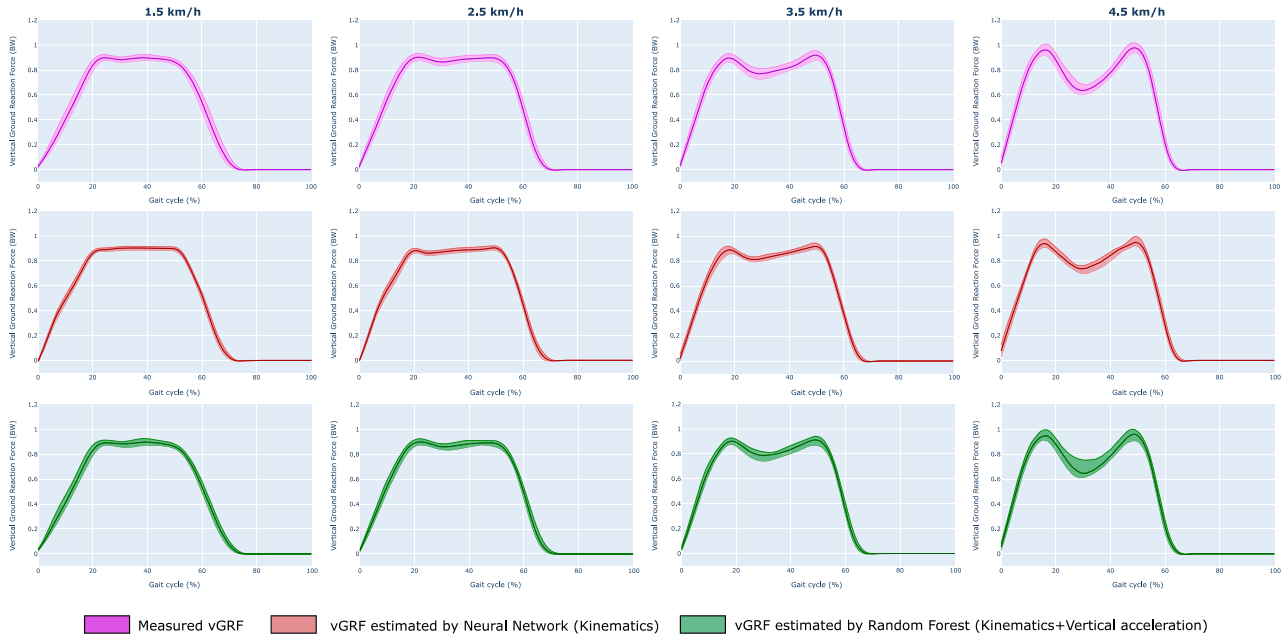


FIGURE 6
Epoch graphs with the measured and estimated vGRF (intra-participants and inter-participants) according to the cycle gait (%) for different gait speeds. The median data value has been represented, and the values between the first and third quartiles have been colored. The first row collects the measured vGRF, the second row the estimated vGRF with the selected FNN model (Training 2, C7), and the third row the estimated vGRF with the selected RF model (Training 2, C7). The vGRF representations include all participant users (intra-participants and inter-participants).

TABLE 4 Correlation (ρ) and normalized RMSE (NRMSE) results obtained between the measured and estimated vGRF with the inter-participants using the FNN-C7 and RF-C7 models.

	Gait speed (km/h)	ρ	NRMSE
FNN-C7	1.5	0.9777	0.0718
	2.5	0.9876	0.0560
	3.5	0.9859	0.0604
	4.5	0.9705	0.0853
RF-C7	1.5	0.9776	0.0716
	2.5	0.9797	0.0708
	3.5	0.9797	0.0691
	4.5	0.9746	0.0792

time. This makes sense since the values of angles, angular velocities, and vertical acceleration can take the same value for different instants of the gait cycle, but the vGRF takes different values. Moreover, from the results, it also can be extracted that hip and knee kinematics have a bigger importance than ankle kinematics.

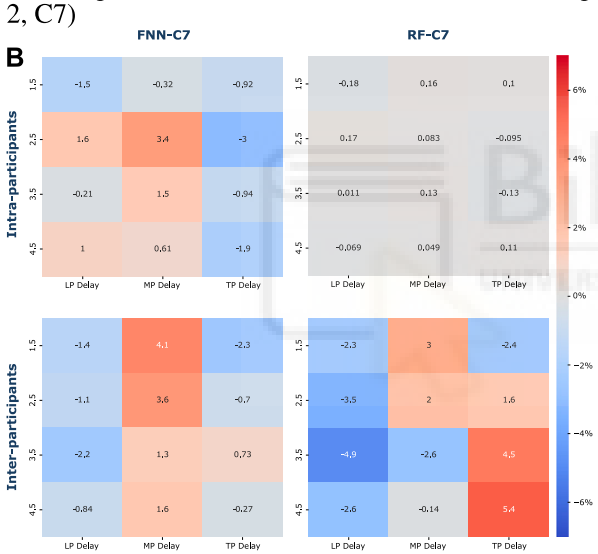
We introduced that vertical acceleration is strongly related to the vGRF. However, when the feature importance is calculated, we can observe that vertical acceleration has a small influence on the model training compared to the kinematics of the lower limbs. Despite the vertical acceleration seems to have low relevance in the models, the results suggest that this input could modify the performance of the FNN and RF models (Figure 5; Table 1; Table 2). Although no statistical differences were found, according to the validation and test results represented in Figure 5, we can observe that the vertical acceleration appears to improve the performance in the FNN and RF models in terms of RMSE. By contrast, the RMSE obtained with the FNN models when the vertical acceleration is included seems to increase. Hence, it could be concluded that vertical acceleration is not crucial to estimate the vGRF with the method proposed if we are able to measure the lower limb kinematics, but it might help to improve the performance of certain models.

In addition, when we compare the performance of the four models trained (FNN-Kinematics, FNN-C7, RF-Kinematics, and RF-C7) some differences can be observed in terms of RMSE. When we employ the validation data (i.e., data from the intra-participants), the RMSE obtained with the RF models (0.030–0.045 BW) is almost half the RMSE of obtained with the FNN models (0.067–0.095 BW). However, when we evaluate the models with the inter-participants, these differences are not statistically significant, and we obtain similar RMSE with the four models. It is usual to obtain an increase in the error with inter-participant data, but the behavior of the RF models might indicate that they tend to overfit more than the FNN models, which is a known problem on algorithms based on decision trees (Dietterich, 1995). Despite the overfitting, the results suggest that the RF models achieve similar accuracy, with a global RMSE of 0.077 BW for the RF-C7 model, and 0.074 BW for the FNN-C7 model, which are the best FNN and RF models in terms of RMSE.

If we analyze further the FNN-C7 and the RF-C7, Table 4 shows that the measured and estimated vGRF are strongly correlated for both intra-participants and inter-participants. In the same manner, as can be seen in the results obtained for RMSE, RF-C7 achieves higher correlation coefficients than the one achieved by the FNN-C7 model with the intra-participant data, but with the inter-participants, the

10.3389/fbioe.2023.1199459

Peak magnitude errors obtained with the intra-participants for the selected FNN model (Training 2, C7) and the selected RF model (Training 2, C7)



Peak delays obtained with the inter-participants for the selected FNN model (Training 2, C7) and the selected RF model (Training 2, C7)

FIGURE 7
Heatmaps with the vGRF peak errors for gait speeds of 1.5 km/h, 2.5 km/h, 3.5 km/h, and 4.5 km/h. **(A)** Mean magnitude errors of the vGRF peaks (BW) **(B)** Mean delays of the vGRF peaks (%).

FNN-C7 correlations are higher than the achieved by the RF-C7 for gait speeds between 1.5 and 4.5 km/h. The NRMSE results show the same behavior, where the FNN-C7 achieves higher errors than the RF-C7 model. However, with the inter-participant data, the NRMSE for 2.5 and 3.5 km/h are lower with the FNN-C7, while the RF-C7 obtains lower errors for 1.5 and 4.5 km/h.

The estimated vGRF signals shown in Figure 6 indicate that both selected models can replicate the measured vGRF. It can be observed that the FNN-C7 and the RF-C7 models can differentiate the cycle gait phases, as the graphs show that they might estimate the characteristic peaks and the swing phase (vGRF = 0 BW).

It must be noted that the magnitude and timing of the vGRF peaks are important parameters to perform gait analysis. For this reason, these peaks are deeply analyzed (Figure 7, Supplemental Data). The RF-C7 model has shown a significantly higher accuracy than the FNN-C7 model according to the magnitude errors in intra-participant users for all the peaks at each gait speed. Nevertheless, the errors of the peaks obtained with the inter-participants are similar between both models. We must remark that the errors obtained with the FNN-C7 in the LP at 1.5, 2.5, and 4.5 km/h, and the TP at 1.5 and 2.5 km/h are lower with the inter-participant data than the intra-participants. Although, it can be observed that the errors obtained with RF-C7 for the inter-participants are higher than the intra-participants for all the peaks and gait speeds. These results would be in agreement with what was seen previously with the RMSE and could be due to the overfitting of the RF models.

Concerning the timing errors of the vGRF peaks, the RF-C7 model presents a high accuracy in the timing prediction of the peaks at all gait speeds with the intra-participant data, with a maximum lag of 0.18% of the cycle gait. On the other hand, the FNN-C7 model appears to make a larger error at the peak timing, especially in the MP with mean delays up to 3.4%. The results obtained for the inter-participants show larger delays and advances in the estimation of the peaks with respect to the measured vGRF, specially with the RF-C7 model. We must mention that all the mean lags obtained with the RF-C7 are greater than the FNN-C7 except for the MP peak at 1.5 and 2.5 km/h.

It must be highlighted that we can assume that both FNN-C7 and RF-C7 models are robust and they have good accuracy to estimate the vGRF by using the kinematics of the lower limbs and the vertical acceleration. This conclusion can be extracted since the RMSE and ρ of our models are in a similar range to studies made by other authors when the vGRF was estimated during gait. In (Oh et al., 2013) the authors obtained an RMSE of around 0.066 BW and ρ of 0.991 with intra-participants; in (Choi et al., 2013) the authors obtained an RMSE of around 0.074 BW and ρ of 0.99 with intra-participants; in (Jiang et al., 2020) the authors obtained an RMSE of 0.02 BW and ρ of 1.00 with intra-participants, and an RMSE of 0.10 BW and ρ of 0.97 with inter-participants. Despite the RF model trained by Jiang et al. seems to achieve a higher accuracy than our models with intra-participants, the results with inter-participants suggest that our RF and FNN models achieve a lower RMSE and higher ρ, so they would have a higher generalization than those proposed by Jiang et al.

It should be noted that, although the FNN-C7 model and the RF-C7 model achieve good accuracy compared to previous works, the results show differences between them. First, the mean of the RMSE obtained with the intra-participants is lower with the RF-C7 model, but the global RMSE achieved by the FNN-C7 is lower than the RF models. However, the calculated ρ and NRMSE show that coefficients and errors obtained are lower at 1.5 and 4.5 km/h with the RF-C7, but lower at walking speeds of 2.5 and 3.5 km/h with the FNN-C7. Finally, the mean magnitude errors of the characteristic peaks show similar results for the inter-participants. Furthermore, the FNN-C7 seems to achieve higher performance than the RF-C7 in estimating the timing of the peaks. Therefore, based on these

differences, we could assume that the FNN-C7 model overcomes the accuracy of the trained RF-C7 model.

5 Conclusion

This work aims to present a method to estimate the vGRF using wearable IMUs. We used inertial sensors to estimate the kinematics of the lower limbs and the vertical acceleration of the users, which were employed to train two types of machine learning models: Feedforward Neural Networks (FNN) and Random Forest (RF).

The results show that the trained models have a big influence on the temporal variable (cycle gait), followed by the hip, knee, and ankle kinematics. The inclusion of vertical acceleration has a small influence on the training of the FNN and RF models compared to the rest of the inputs. However, the results suggest that the inclusion of vertical acceleration can modify the performance of the model. In the case of the FNN and RF models, the inclusion of this feature appears to increase the performance.

Moreover, we can assume that the method proposed to estimate the vGRF has good accuracy for biomechanical analysis, as the estimate of vGRF, including its peaks, is in a similar range of accuracy to that reported in other studies, and even the results suggest that a smaller error is obtained in the estimation of VGRF. Furthermore, the RMSE, ρ, and the errors in the characteristic peaks suggest that the FNN-C7 model achieves the highest accuracy of the trained machine learning models.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individuals for the publication of any potentially identifiable images or data included in this article.

Author contributions

DM-P, JC, AB-I, and NG-A designed the study; DM-P, JC, AB-I, and NG-A designed the methodology; DM-P developed the software; DM-P, MS, and FA-A. collected the data; MS and FA-A provided the laboratory equipment; DM-P cured the data; DM-P and JC trained the models; DM-P, JC, AB-I, MS, FA-A, and NG-A wrote and reviewed the manuscript; and NG-A performed guiding and coordination tasks.

Funding

This research was funded by the Spanish Ministry of Universities through the Research and Doctorate Supporting Program FPU20/05137; by the Ministry of Universities and European Union, "financed by European Union—Next Generation EU" through Margarita Salas grant for the training of young doctors; by the Ministry of Science and Innovation through the project PID 2019-108310RB-100; and by the Valencian Innovation Agency through the project GVRTE/2021/361542.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2023.1199459/full#supplementary-material

References

Altmann, A., Toloşi, L., Sander, O., and Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics* 26, 1340–1347. doi:10.1093/bioinformatics/btq134

Andrychowicz, M., Denil, M., Gomez, S., Hoffman, M. W., Pfau, D., Schaul, T., et al. (2016). "Learning to learn by gradient descent by gradient descent," in Proceedings of the 30th International Conference on Neural Information Processing Systems, Barcelona, Spain, December 5 - 10, 2016.

Berger, V. W., and Zhou, Y. (2014). "Kolmogorov–smirnov test: overview," in *Wiley statsref: Statistics reference online* (Hoboken: John Wiley & Sons).

Byrd, R. H., Schnabel, R. B., and Shultz, G. A. (1988). Approximate solution of the trust region problem by minimization over two-dimensional subspaces. *Math. Program.* 40, 247–263. doi:10.1007/bf01580735

Choi, A., Lee, J.-M., and Mun, J. H. (2013). Ground reaction forces predicted by using artificial neural network during asymmetric movements. *Int. J. Precis. Eng. Manuf.* 14, 475–483. doi:10.1007/s12541-013-0064-4

Dietterich, T. (1995). Overfitting and undercomputing in machine learning. *ACM Comput. Surv. (CSUR)* 27, 326–327. doi:10.1145/212094.212114

Esser, P., Dawes, H., Collett, J., and Howells, K. (2009). Imu: inertial sensing of vertical com movement. *J. Biomechanics* 42, 1578–1581. doi:10.1016/j.jbiomech.2009.03.049

Fineberg, D. B., Asselin, P., Harel, N. Y., Agranova-Breyter, I., Kornfeld, S. D., Bauman, W. A., et al. (2013). Vertical ground reaction force-based analysis of powered exoskeleton-assisted walking in persons with motor-complete paraplegia. *J. spinal cord Med.* 36, 313–321. doi:10.1179/2045772313y.0000000126

Graci, V., Van Dillen, L. R., and Salsich, G. B. (2012). Gender differences in trunk, pelvis and lower limb kinematics during a single leg squat. *Gait Posture* 36, 461–466. doi:10.1016/j.gaitpost.2012.04.006

Gulli, A., and Pal, S. (2017). *Deep learning with Keras*. Birmingham: Packt Publishing Ltd.

Gustafsson, F. (1996). Determining the initial states in forward-backward filtering. *IEEE Trans. Signal Process.* 44, 988–992. doi:10.1109/78.492552

Jacobs, N., Skorecki, J., and Charnley, J. (1972). Analysis of the vertical component of force in normal and pathological gait. *J. Biomechanics* 5, 11–34. doi:10.1016/0021-9290(72)90016-4

Jiang, X., Napier, C., Hannigan, B., Eng, J. J., and Menon, C. (2020). Estimating vertical ground reaction force during walking using a single inertial sensor. *Sensors* 20, 4345. doi:10.3390/s20154345

Kingma, D. P., and Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kram, R., Griffin, T. M., Donelan, J. M., and Chang, Y. H. (1998). Force treadmill for measuring vertical and horizontal ground reaction forces. *J. Appl. physiology* 85, 764–769. doi:10.1152/jappl.1998.85.2.764

Kumar, S., Gopinath, K., Rocchi, L., Sukumar, P. T., Kulkarni, S., and Sampath, J. (2018). "Towards a portable human gait analysis & monitoring system," in 2018 International Conference on Signals and Systems (ICSigSys), Bali, Indonesia, 01-03 May 2018, 174–180. doi:10.1109/ICSIGSYS.2018.8372660

Lencioni, T., Carpinella, I., Rabuffetti, M., Marzegan, A., and Ferrarin, M. (2019). Human kinematic, kinetic and emg data during different walking and stair ascending and descending tasks. *Sci. data* 6, 309–310. doi:10.1038/s41597-019-0323-z

Liu, K., Liu, T., Shibata, K., and Inoue, Y. (2009). "Ambulatory measurement and analysis of the lower limb 3d posture using wearable sensor system," in 2009 International Conference on Mechatronics and Automation, Changchun, 09-12 August 2009, 3065–3069. doi:10.1109/ICMA.2009.5245982

Marasović, T., Cecić, M., and Zanchi, V. (2009). Analysis and interpretation of ground reaction forces in normal gait. *WSEAS Trans. Syst.* 8, 1105–1114.

Moya-Angeler, J., Vaquero, J., and Forriol, F. (2017). Evaluation of lower limb kinetics during gait, sprint and hop tests before and after anterior cruciate ligament reconstruction. *J. Orthop. Traumatology* 18, 177–184. doi:10.1007/s10195-017-0456-9

Oh, S. E., Choi, A., and Mun, J. H. (2013). Prediction of ground reaction forces during gait based on kinematics and a neural network model. *J. Biomechanics* 46, 2372–2380. doi:10.1016/j.jbiomech.2013.07.036

Padua, D. A., and DiStefano, L. J. (2009). Sagittal plane knee biomechanics and vertical ground reaction forces are modified following acl injury prevention programs: A systematic review. *Sports Health* 1, 165–173. doi:10.1177/1941738108330971

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.

Pratt, J. W. (1959). Remarks on zeros and ties in the wilcoxon signed rank procedures. *J. Am. Stat. Assoc.* 54, 655–667. doi:10.1080/01621459.1959.10501526

Seel, T., Raisch, J., and Schauer, T. (2014). Imu-based joint angle measurement for gait analysis. *Sensors* 14, 6891–6909. doi:10.3390/s140406891

Seel, T., Schauer, T., and Raisch, J. (2012). "Joint axis and position estimation from inertial measurement data by exploiting kinematic constraints," in 2012 IEEE International Conference on Control Applications, Dubrovnik, Croatia, 03-05 October 2012, 45–49. doi:10.1109/CCA.2012.6402423

Shafizadegan, Z., Karimi, M. T., Shafizadegan, F., and Rezaeian, Z. S. (2016). Evaluation of ground reaction forces in patients with various severities of knee osteoarthritis. *J. Mech. Med. Biol.* 16, 1650003. doi:10.1142/s0219519416500032

Shahabpoor, E., and Pavic, A. (2018). Estimation of vertical walking ground reaction force in real-life environments using single imu sensor. *J. Biomechanics* 79, 181–190. doi:10.1016/j.jbiomech.2018.08.015

Sharma, S., Sharma, S., and Athaiya, A. (2017). Activation functions in neural networks. *Towards data Sci.* 6, 310–316. doi:10.33564/ijeast.2020.v04i12.054

Sheldon, M. R., Fillyaw, M. J., and Thompson, W. D. (1996). The use and interpretation of the friedman test in the analysis of ordinal-scale data in repeated measures designs. *Physiother. Res. Int.* 1, 221–228. doi:10.1002/pri.66

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958.

Svozil, D., Kvasnicka, V., and Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. *Chemom. Intelligent Laboratory Syst.* 39, 43–62. doi:10.1016/S0169-7439(97)00061-0

Sy, L. W. F., Lovell, N. H., and Redmond, S. J. (2020). Estimating lower limb kinematics using a lie group constrained extended kalman filter with a reduced wearable imu count and distance measurements. *Sensors* 20, 6829. doi:10.3390/s20236829

Weber, D. (2016). Differences in physical aging measured by walking speed: evidence from the english longitudinal study of ageing. *BMC Geriatr.* 16, 31–39. doi:10.1186/s12877-016-0201-x

Wouda, F. J., Giuberti, M., Bellusci, G., Maartens, E., Reenalda, J., van Beijnum, B.-J. F., et al. (2018). Estimation of vertical ground reaction forces and sagittal knee kinematics during running using three inertial sensors. *Front. Physiology* 9, 218. doi:10.3389/fphys.2018.00218

Young, A. (2009). "Comparison of orientation filter algorithms for realtime wireless inertial posture tracking," in 2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks, Berkeley, CA, USA, 03-05 June 2009, 59–64. doi:10.1109/BSN.2009.25

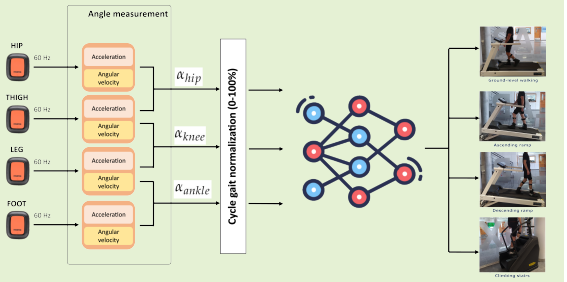
Zhang, C., and Ma, Y. (2012). *Ensemble machine learning: methods and applications*. Berlin: Springer.

Zijlstra, W., and Hof, A. L. (2003). Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture* 18, 1–10. doi:10.1016/S0966-6362(02)00190-X

Gait Activity Classification with Convolutional Neural Network using Lower Limb Angle Measurement from Inertial Sensors

David Martínez-Pascual, José M. Catalán, Andrea Blanco-Ivorra, Mónica Sanchís, Francisca Arán-Ais and Nicolás García-Aracil, *Member, IEEE*

***Abstract*— Human gait activity recognition can be crucial to adapt the assistance provided by lower limb exoskeletons, as well as for biomechanical analysis. With this purpose, Deep Learning techniques can be applied to develop a classifier based on the acquisition of the lower limb kinematics. In this paper, we present a one-dimensional Convolutional Neural Network (CNN) to classify different activities from the hip, knee, and ankle flexion/extension angles, measured with wearable inertial sensors. The proposed CNN classifier achieves 99.56% accuracy with users not involved in the learning process. In addition, the Gradient-weighted Class Activation-Map (Grad-CAM) and the t-Distributed Stochastic Neighbor Embedding (t-SNE) were used to understand the CNN model decision-making. Finally, how the accuracy of the CNN model is impacted by input reduction was analyzed to adapt the CNN model to multiple situations, and it can be concluded that the CNN maintains high accuracy with a single joint angle as input.**



***Index Terms*— Gait analysis, human activity recognition, magneto-inertial devices, Deep Learning**

I. Introduction

HUMAN locomotion is a complex and dynamic process that involves the coordination of multiple body segments and muscles. The analysis of human gait can provide valuable information about the health and functional status of an individual [1]. In this context, the analysis of human gait can be used to analyze the range of motion (ROM) of the lower limb joints, which may be reduced in patients suffering from neuromuscular pathologies such as cerebral palsy, Parkinson's disease, or hemiplegia [2], [3]. In addition, lower limb kinematics analysis could help make detailed diagnoses, plan optimal treatment, or evaluate the results of rehabilitation therapies [4]. However, traditional gait analysis methods require expensive and specialized equipment, and they are also limited by the laboratory setting, which may not reflect the natural and dynamic environment of daily living [5]. Therefore, to perform a correct biomechanical analysis in a natural environment, the recognition of the activity that the user has performed is needed.

Gait analysis and activity classification techniques can be used to personalize lower limb exoskeletons, which are wearable devices designed to assist or enhance leg movements [6]. These exoskeletons have diverse applications such as rehabilitation, assistive technology, and human augmentation [7], [8]. The exoskeleton must be able to recognize the user's intentions and adjust its parameters accordingly to ensure a natural and comfortable interaction between the user and the device, [9]. Joint sensors and inertial measurement units (IMUs) are generally used for movement recognition [10]–[12]. In addition, the activity performed is the most important factor in determining the operation and high-level behavior of assistive walking devices. Different methods have been proposed to detect and adapt the control of lower limb exoskeletons, including cameras [13], [14], infrared distance sensors [15], and a combination of infrared sensors with IMUs.

Machine learning algorithms can be employed to classify various gait activities of daily living (ADLs) based on the information provided by inertial sensors [16]. This classification is vital for biomechanical analysis and adaptation of lower limb exoskeleton assistance in a natural environment. Extracting temporal and frequency domain features from the IMUs can be used to train machine learning models for gait activity classification [17]. By contrast, Deep Learning methods can

This research was funded by the Spanish Ministry of Universities through the grant FPU20/05137; by the Ministry of Science and Innovation through the project PID2019-108310RB-100, and through the project PID2022-139957OB-I00; by the Spanish Agency of Innovation through the project PLEC2022-009424; and by the Valencian Innovation Agency through the project GVRTE/2021/361542.

D. Martínez-Pascual, J.M. Catalán, A. Blanco, and N. García-Aracil are with the Biomedical Neuroengineering Research Group, Robotics and Artificial Intelligence Unit, Bioengineering Institute, Miguel Hernandez University, 03202 Elche, Spain (email: david.martinezp@umh.es, jcatalan@umh.es, ablanco@umh.es, nicolas.garcia@umh.es)

M. Sanchís and F. Arán-Ais are with INESCOP, Elda, Alicante, Spain (email: msanchis@inescop.es, aran@inescop.es)

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3400296

autonomously extract features from the signals collected by the IMUs [18]. Convolutional Neural Network (CNN) is one of the most commonly used models for Deep Learning in image processing [19]. Some authors have transformed accelerometer and gyroscope signals from IMUs into images to apply two-dimensional (2D) CNNs for gait activity classification [20]. However, generating 2D images for the analysis of time series can be computationally expensive. Therefore, a fixed-size time series can be used instead of images to train and recognize ADLs with a one-dimensional (1D) CNN [21].

Although the accelerometer and gyroscope data provided by the IMUs have been used to develop gait ADL classifiers, the use of the angles of the lower limb joints has not been explored. This method can be easily embedded in the control system of the lower limb exoskeletons since joint sensors are generally used. In addition, this method could also be used to perform biomechanical analysis with wearable measurement devices. In this paper, we present a deep learning model that uses the angles of the lower limbs to classify different gait activities. For this purpose, we have developed a 1D-CNN architecture that can effectively extract features and classify several gait activities from the trajectories of the lower limbs measured with wearable IMUs. In addition, we have used the Gradient-weighted Class Activation-Map (Grad-CAM) and the t-Distributed Stochastic Neighbor Embedding (t-SNE) to understand the decision-making of the developed 1D-CNN model. Finally, we have analyzed possible variations of the model so that it can be adapted to multiple lower limb exoskeletons, as well as to perform a biomechanical analysis from one to three joints.

II. Related work

Human activity recognition (HAR) is a complex problem that has many practical applications such as human-robot interaction, rehabilitation, or health monitoring [22]. With recent advances in wearable technology, analyzing human biomechanics and recognizing different activities of daily living (ADLs) has become more affordable, accurate, and feasible.

Vision-based approaches have been explored to perform activity recognition and adapt lower limb exoskeletons or prostheses. In [23], the authors study the feasibility of using a depth camera placed on the leg to adapt lower limb prostheses. The authors detected five different situations (standing, running, ground-level walking, ascending stairs, descending stairs) using a Support Vector Machine (SVM) with a 99% accuracy. On the other hand, in [24], the authors proposed to use a chest-mounting camera, which provides less relative body movement than head or lower-limb mounting. In that work, the authors used a Deep Learning model to analyze the images from the camera to detect ground-level walking and ascending/descending stairs with a 94.85% accuracy.

In addition to vision-based methods, the use of inertial sensors for HAR has also been widely studied. In [25], the authors used accelerometers embedded inside the insoles of the shoes to recognize different activities. They used accelerometer signals as inputs of a CNN to recognize six activities with a 93% accuracy (running, walking, cycling, kicking, standing, and jumping).

Other authors propose to use the IMU of smartphones to recognize different activities. In [26], the authors used the gyroscope and accelerometer signals as inputs of a 1D-CNN, where the proposed model achieved a 95.57% accuracy using the Fast Fourier Transform of the input signals. Another example is the model proposed in [27], where the authors used the Euclidean norm of the accelerometer signals to detect different activities, with a 92.71%. Moreover, in [28] the authors propose to classify different activities by extracting statistical features from the accelerometer of smartphones, achieving a 95.89% accuracy with a CNN model.

More related to our work is the approach proposed in [29], where the authors used the public dataset HuGaDB [30] to classify different gait ADLs. The authors trained a Random Forest model using as inputs different extracted features from the accelerometer and gyroscope signals placed over the thigh, shank, and foot, achieving an accuracy of 98.6%. Instead of extracting features to train the models, in [20] a Deep Learning approach was used. The authors transformed the accelerometer and gyroscope signals from the OU-ISIR public dataset [31] to use a 2D-CNN. They show that their trained model achieved a 93% accuracy when detecting ground-level walking, ascending/descending a ramp, and going down/up stairs.In addition, in [32] the authors used a single IMU placed at the center of mass position of each user to recognize different gait activities (jogging, ground-level walking, standing, climbing stairs). The authors employed a hybrid model based on ensemble learning and they achieved an accuracy of 99.34%. However, the main limitation of this approach relays on the computational cost, since the ensemble learning approach is composed by four Deep Learning models, which would be not suitable for real-time inference. In addition, the authors do not include in the work an evaluation with users who were not participated in the learning process of the models. Hence, it is not clear how accurate this method would be if used with different users.

In addition to only employing inertial sensors, some studies propose to perform HAR by fusing the IMU data with other sensors. In [33], the authors propose to employ an IMU placed over the waist together with laser measurements to recognize the terrain for adapting powered lower limb prostheses. They used a Decision Tree to perform a terrain classification with a 98% accuracy. Another sensor fusion approach is presented in [34], where pressure insoles and IMUs placed on the thigh and shank were used. An intelligent fuzzy algorithm was used to recognize different gait ADLs with an accuracy of 95.52%. Another approach is presented in [35], where several IMUs and strain gauges placed on the lower limbs were used to classify different gait activities. The authors used an autoencoder to perform dimensionality reduction, followed by a Deep Learning model that combined convolutional and recursive layers, achieving an accuracy of 98.9%.

Despite different methods for HAR using wearable sensors have been proposed, they all come with certain limitations. Vision-based techniques are computationally expensive due to the need for image analysis to detect user activity. In contrast, the use of IMUs seems to be portable, cost-effective, and the

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3400296

computational cost is affordable enough to perform HAR with wearable devices. Nonetheless, in the IMU-based methods proposed, accurate placement of IMUs is essential, as any error in the IMU positioning may cause disturbances in the signals, thus reducing the accuracy of the activity classification.

To the best of our knowledge, no study has explored the use of the angles of the lower limb joints to perform HAR. In order to propose an affordable approach, while solving the possible errors due to IMU positions, we have implemented a method to measure the angle of the hip, knee, and ankle in the sagittal plane which does not depend on the position and location of the IMUs. Therefore, variations in IMU placement on the lower limbs would not impact the classification accuracy. While previous studies have proven effective in HAR, the use of explainability methods to understand model operation is not common in this field. Thus, we have used the t-SNE method and Grad-CAM to determine how our model extracts features and which parts of the input are most critical.

III. Materials and Methods

A. Lower limb joints measurement

An IMU-based algorithm was used to measure the flexion/extension angles of the hip, knee, and ankle joints. The joint motion analysis method proposed by T. Seel et al. was applied to estimate the lower limb joint angles during gait [36]. This method does not require the knowledge of the IMUs' position and orientation on the user's lower limb, and it only assumes that each joint has two inertial sensors on its upper and lower segments, and uses the measured accelerations and angular rates.

To align the IMU reference systems with the joint axis, some methods in the literature require specific and accurate calibration movements [37]. On the other hand, the method by T. Seel et al. determines the location of the sensors on the joint they measure by using arbitrary calibration movements and joint movements during gait, taking advantage of kinematic constraints. The direction vectors $j1$ and $j2$ of the hip, knee, and ankle joints' flexion/extension axes are found using gyroscope data in the local coordinates of the IMUs during the calibration. The acceleration data are also used to estimate the joint position coordinates $o1$ and $o2$.

The method by T. Seel et al. has demonstrated a root mean squared error between 3 and 5 degrees in flexion/extension movements [38]. This suggests that the method is accurate enough for gait analysis and gait ADL detection. In addition, T. Seel et al. demonstrated that this method conserves the accuracy even at low acquisition rates. Therefore, this method is robust and suitable to perform real-time measurements.

B. Experimental session

Twelve able-bodied participants (9 male, 3 female) aged 23-52 years old (29.8 ± 7.4), with heights ranging between 165cm and 187cm (176.2 ± 7.4cm) and weights between 56.1kg and 90.2kg (76 ± 12.5kg) participated in the experiment. They provided written, informed consent before participating in the experimental session.

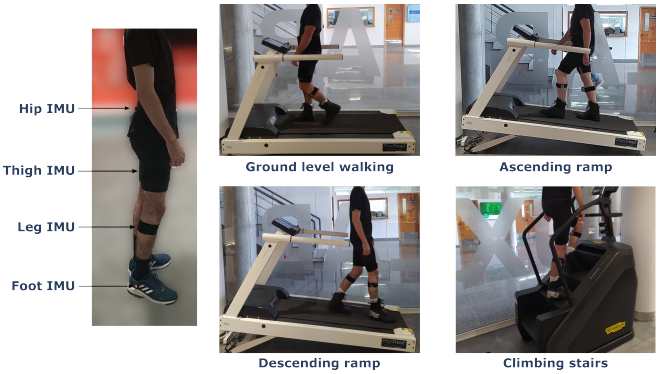


Fig. 1. Setup and performed gait activities during the experimental session. The left image shows the positions of the attached IMUs used to measure joint angles. The participants walked on a treadmill to simulate ground-level walking, as well as ascending and descending a ramp of 12% slope. In addition, stair climbing was simulated using a stair treadmill.

Four gait activities were simulated to collect lower limb movement data using four XSens Dot IMUs at 60Hz. In order to simulate and collect data during various ADLs, we used an h/p/cosmos 150/50 treadmill to simulate ground-level walking, ramp ascent, and ramp descent. A stair treadmill was used to simulate stair climbing. The experimental setup and the ADLs performed are shown in Fig. 1.

The IMUs were attached to the pelvis, thigh, shank, and foot using elastic triggers and stickers. The participants calibrated the IMUs by moving their legs for 30 s and walking on the treadmill for 1 minute at their preferred speed.

After the calibration concluded, the angles of the hip, knee, and ankle were measured during four different activities:

1) Ground-level walking at 4.5 km/h for 5 minutes.
2) 12% positive slope walking at 2.5 km/h for 5 minutes.
3) 12% negative slope walking at 2.5 km/h for 5 minutes.
4) Climbing stairs at 50 stairs/minute for 3 minutes.

We followed the INESCOP Footwear Technology Center protocols for footwear certification to set the walking speeds and slope. The stair-climbing time was shortened to 3 min to prevent excessive fatigue.

C. Acquired data

This section describes the IMU data acquisition and processing to be used as inputs of the classifier model. This procedure is shown in Fig. 2, which is based on previous studies [39], [40].

Acceleration and angular rate data were used to measure the flexion/extension angles of the hip, knee, and ankle. To perform a gait analysis, kinematics data are usually transformed from the time domain to the gait cycle domain (0-100%), which requires identifying the start and end of each gait cycle. Therefore, we applied this transformation to all signals. We used the hip acceleration [41] to detect the ground-foot contact, and applied a forward-backward low-pass filter [42] with a 2 Hz cutoff. The local maxima in the filtered signal corresponds to the left and right foot contacts. We followed the convention of using the right foot contact to calculate the gait cycle.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3400296

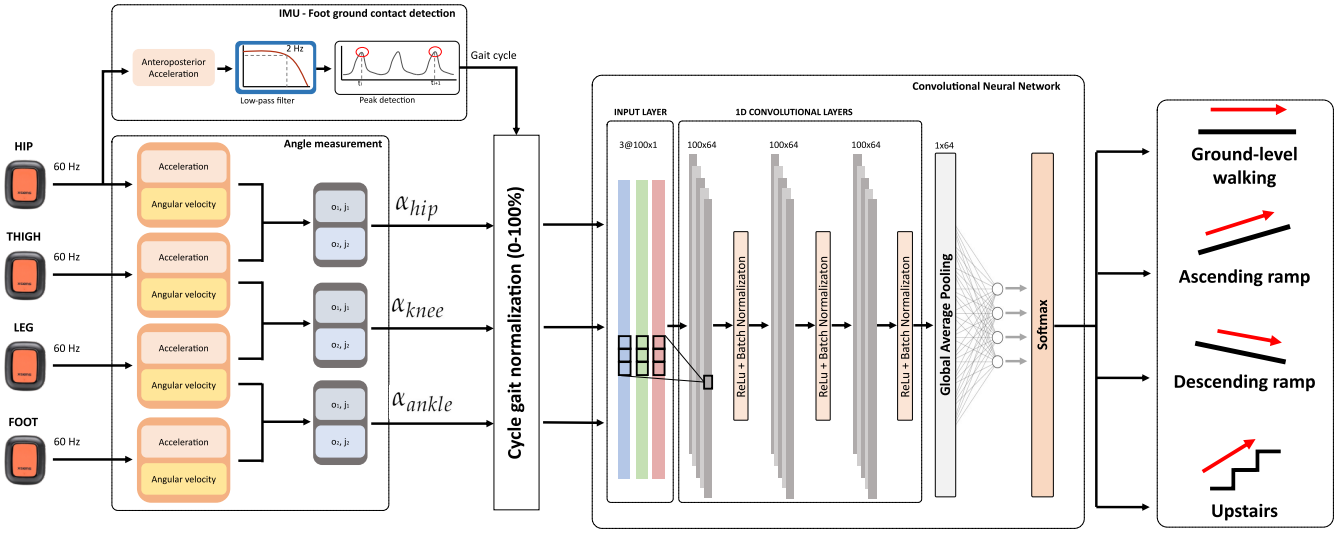


Fig. 2. The IMUs placed on the lower limbs were used to measure acceleration and angular velocity for the estimation of the flexion/extension angles. We also detected the foot-ground contact from the hip acceleration, which is used to segment the joint angles into steps and gait cycles. We normalized each gait cycle (0-100%) and fed them to the CNN as inputs. The hip, knee, and ankle flexion/extension joints are normalized between their minimum and maximum to be introduced as inputs of the model. The proposed model is composed of three 1D convolutional layers with a kernel size of 3 and 64 filters. Also, a padding operation is applied to have feature maps with lengths equal to the inputs. The ReLu is used as the activation function of the convolutional layers, and batch normalization is applied. A global average pooling layer is used to reduce the generated feature maps connected to the output layer.

It is important to note that the CNNs require a fixed window size as input. In our case, as we want to classify each step within one of the ADLs, it seems reasonable to use a window size of 100 joint values calculated from the gait cycle. Therefore, we will use as CNN inputs three windows of size 100, with the values of hip, knee, and ankle flexion/extension angles during a step.

D. Gait activity classifier

1) Proposed Convolutional Neural Network architecture: As introduced, we plan to train a CNN model to detect four gait activities. Typically, 1D-CNN classifiers consist of two parts [43]: the convolutional and pooling layers, and the fully-connected layers or Multilayer Perceptron (MLP).

In the first part of the model, the CNNs have an input layer of Nxk, where N is the length of the univariate time series and k is the number of input series. After the input layer, convolution and pooling operations are used to generate deep features of the inputs. In the convolutional layers, convolution operations are performed on the time series of the previous layer with convolutional filters (kernel), whose final output is known as a feature map. In addition, a nonlinear function, such as the Rectified Linear Unit (ReLu) function, is used after each convolution operation. In the pooling layers, the feature maps are divided, and each segment is represented by its average or maximum value. Finally, after the convolution and pooling layers, the original time series is represented by a series of feature maps.

The feature maps are usually connected to an MLP in the second part of the model. The classification task is performed by these layers based on the features and filters of the previous layers. A global average pooling layer can replace this part, which aims to produce one feature map per class [44]. The average of each feature map is taken by this layer, and the resulting vector goes straight to the softmax layer. The global average pooling layer also fits better with the convolution structure by creating correspondences between feature maps and categories, and it prevents overfitting at this layer since it has no parameter to optimize.

The architecture of the CNN model proposed to classify four gait activities is a Fully Convolutional Neural Network (Fig. 2). The hip, knee, and ankle flexion/extension angles are introduced as the CNN inputs, which have been normalized by their maximum and minimum values. We have used 3 convolutional layers and 64 filters are applied with a kernel size of 3. Moreover, we applied padding on the convolution operations to have feature maps with a length N equal to the inputs. After each convolutional layer, we operate with the ReLu as the activation function, and Batch Normalization is applied. After the convolution operations, we used a global average pooling layer connected to a dense layer with four neurons (one per gait ADL) with the softmax as the activation function. The number of convolutional layers, the number of filters, the size of the kernels, and the rest of the hyperparameters of the model have been tuned to achieve the highest performance.

2) Preprocessing and model training: We divided the data from 12 participants into 3 subsets to evaluate the performance of the model. Time series from 9 users were randomly divided by steps to train and to validate the accuracy of the model with (intra-participants): 80% was used for training and 20% for validation. Data from the remaining 3 users were used to evaluate model accuracy with participants not involved in the learning process (inter-participants). In addition, the training data was scaled between the minimum and maximum angle values of the hip, knee, and ankle so that the values of the inputs are between 0 and 1. The training scaler was also

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3400296

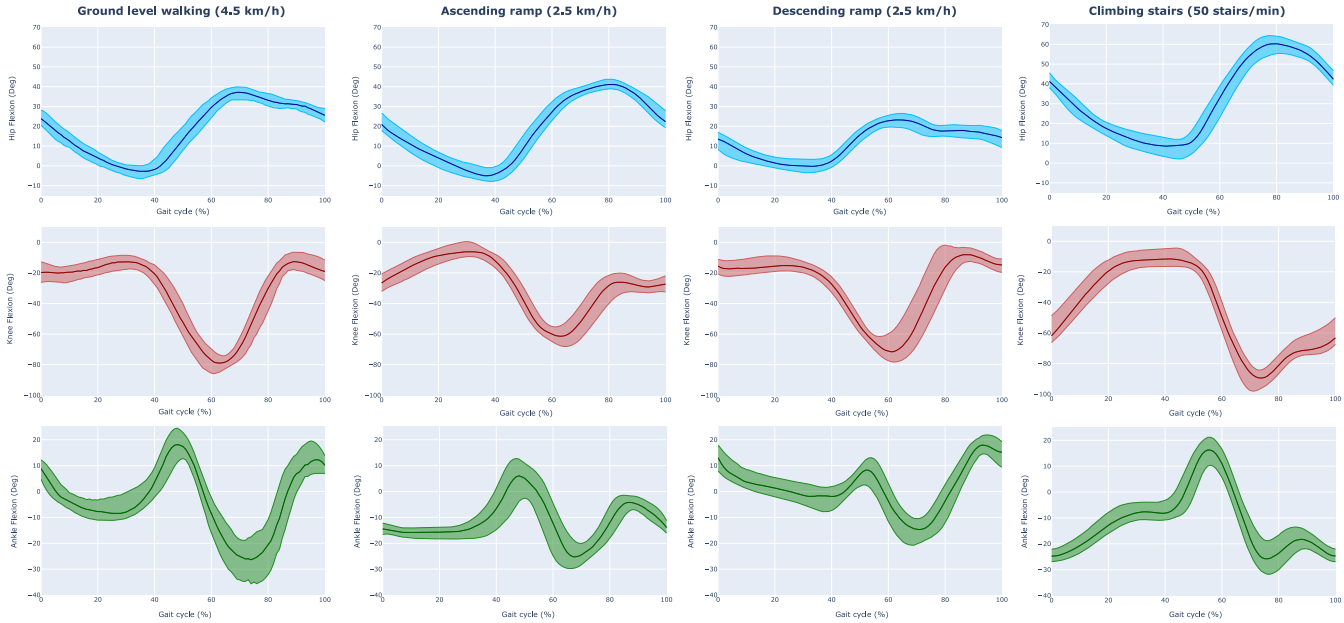


Fig. 3. Collected kinematics data during the experiment. The flexion/extension angles of the hip, knee, and ankle joints (top, middle, and bottom row) for different types of walking (flat surface, ascent ramp, descent ramp, and stairs) are shown, and normalized by the gait cycle. The figures above show the median of the twelve participants, and the shaded areas indicate the range between the first and third quartiles.

used to evaluate the model with the intra-participant and inter-participant data.

The proposed 1D-CNN model has been trained with the Keras Python Library. The model was trained for 50 epochs and a batch size of 8. To adjust the weights of the network we used the Adam algorithm as the optimizer, and the Categorical Cross-Entropy (CE) was used as the loss function, defined as:

$$CE = -\sum_{i}^{C} y_i \cdot log(\hat{y}_i) \tag{1}$$

where y_i is the actual class, $\hat{y}_i$ is the CNN score for each class, and C is the number of classes.

E. Model explainability

It is common to think of machine learning models as black boxes since insights into decision-making are mostly opaque to humans. However, understanding decision-making can be critical in areas such as healthcare. In our particular case, it may be interesting to know how the trained CNN model discriminates between the four ADLs performed during the experimental session.

Some methods have been proposed to explain model predictions to solve the black-box problem. We used the Gradient-weighted Class Activation-Map (Grad-CAM) algorithm [45] to examine the class activation map [46]. The Grad-CAM method has been used to explain deep learning models in image analysis by noting the impact of the pixels on the prediction. Extending this method to time series analysis with a CNN, we can use Grad-CAM to highlight segments of the joint trajectories to understand how our model distinguishes each gait activity. In the next sections, the contribution of the hip, knee, and ankle flexion/extension trajectories will be analyzed.

IV. Results

A. Model evaluation

In Fig. 3 the hip, knee, and ankle flexion/extension angles have been represented during the ADLs performed during the experimental session.

We evaluated the performance of the CNN model trained with data from the intra-participants and inter-participants. The inference process was made in a PC with an Intel(R) Core(TM) i7-1165G7 @ 2.80GHz with 16 GB of RAM, and each inference took 1 $\pm$ 0.37 milliseconds. In addition, the performance of the proposed model is compared with four state-of-the-art models proposed in different studies: ResNet [47], t-LeNet [48], MCDCNN [49], and Time-CNN [50]. Table I collects the accuracy and the CE with the intra-participant and inter-participant data with the proposed model and the state-of-the-art models. The results show that our classifier achieves an accuracy of 99.65% and a CE of 0.0054 with

TABLE I
ACCURACY AND CATEGORICAL CROSS-ENTROPY (CE) OBTAINED WITH THE INTRA-PARTICIPANTS AND INTER-PARTICIPANTS.

Model	Dataset	Accuracy(%)	CE
Proposed 1D-CNN	Intra-participants	99.65	0.0054
	Inter-participants	99.56	0.0025
ResNet [47]	Intra-participants	100.00	0.0004
	Inter-participants	99.46	0.0082
t-LeNet [48]	Intra-participants	99.98	0.0132
	Inter-participants	99.21	0.0287
MCDCNN [49]	Intra-participants	99.98	0.0036
	Inter-participants	99.46	0.0039
Time-CNN [50]	Intra-participants	99.98	0.0003
	Inter-participants	98.09	0.0102

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3400296

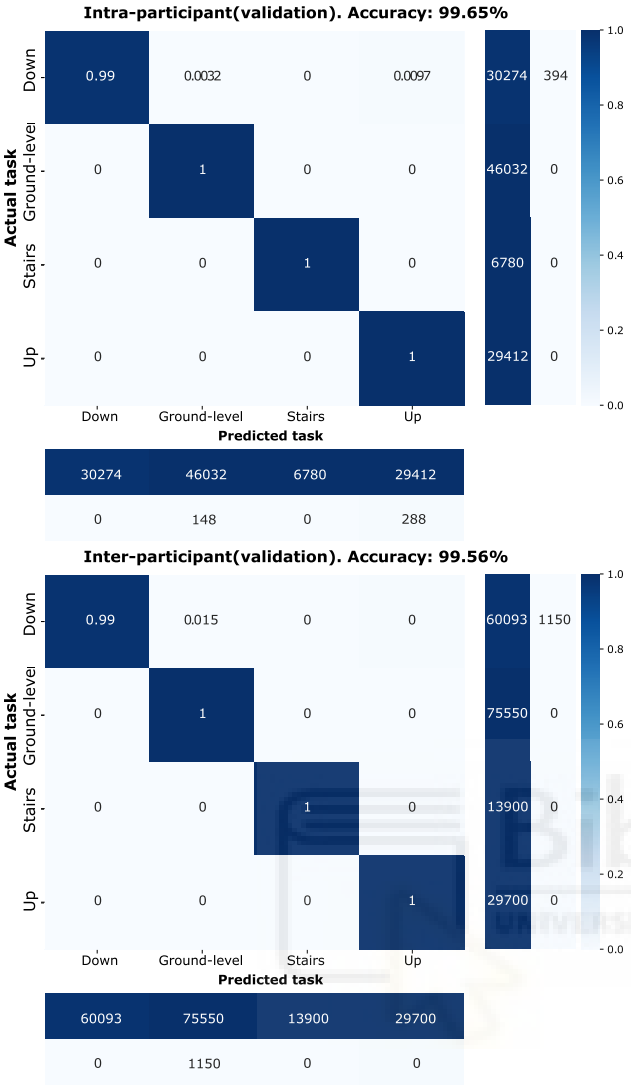


Fig. 4. Confusion matrices obtained with the proposed CNN classifier for intra-participants and inter-participants. The rows of the matrix show the ADL that was performed, and the columns show the ADL that was predicted. The matrices include the four gait ADLs that were done in the experiment: walking on a flat surface, going up a ramp (Up), going down a ramp (Down), and climbing stairs.

intra-participant data, and an accuracy of 99.56% and a CE of 0.0025 with inter-participant data. For a better comprehension of the classification results, confusion matrices are represented in Fig. 4 with the intra-participant and inter-participant data. The results show that CNN could confuse ground-level walking, ascending a ramp, and descending a ramp. For the intra-participants, 99% of the descent ramp steps are classified correctly, where 0.32% of the steps are confused with ground-level walking, and 0.97% is confused with ascending a ramp. The inter-participant results show similar behavior since 1.5% of the descent ramp steps are classified incorrectly as ground-level walking.

B. Model interpretability

We applied the t-Distributed Stochastic Neighbor Embedding (t-SNE) to project the 64-dimensional feature map obtained by the convolutional and global average pooling layers into a 2-dimensional graph [51]. Fig. 5 shows the representation of the generated feature maps for 2 intra-participants and 2 inter-participants, using perplexity and early exaggeration values of 50. This graph shows diverse groups of points, also known as clusters, where each point represents the feature map obtained by the CNN for each step, and clusters represent a group of similar steps.

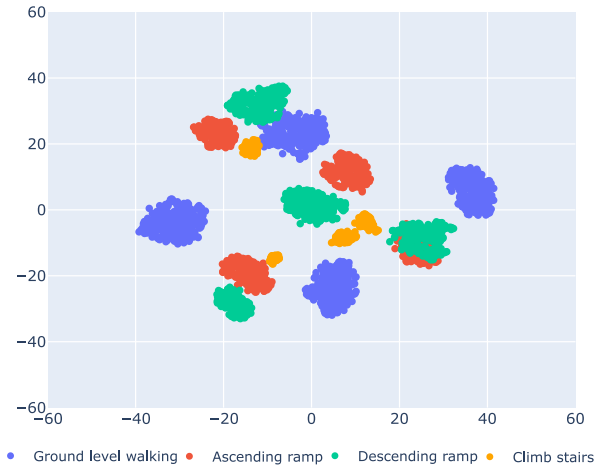


Fig. 5. Two-dimensional representation of the feature map generated by the CNN model proposed for 2 intra-participants and 2 inter-participants. The projection was obtained using the t-SNE embedding. Each gait ADL is represented by a color: blue for ground-level walking, red for ascending a ramp, green for descending a ramp, and yellow for climbing stairs.

In Fig. 6 we have represented the mean joint trajectories of the lower limbs during the four gait ADLs over the class activation maps obtained with Grad-CAM for the last convolutional layer. The activation map is represented as a heatmap that highlights the important parts of the inputs. Blue parts of the graph can be seen as low attention parts of the inputs, and red values as high attention parts. The ground-level walking activation map shows high activation during the terminal stance phase (45%-65%), the swing phase and the initial contact; the ascending ramp activation map shows strong activation during the loading response and mid stance phase (around 10%-50%) and the terminal swing phase; the descending ramp activation class shows high activation during the terminal swing phase (75%-90%); and the climbing stairs activation map has an strong activation during the swing phase (70%-90%).

In addition, the Grad-CAM method could also help us to understand why some steps are classified incorrectly. Fig. 7 collects two steps classified wrongly as descending a ramp. The joint angles and the class activation maps are represented, and the softmax values for the actual class and the predicted class are specified.

C. Model adaption

Several models have been trained with the same proposed architecture, but changing the number of inputs of the models to check the feasibility of adapting the CNN model presented. With this purpose, we have trained seven 1D-CNN models for

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3400296

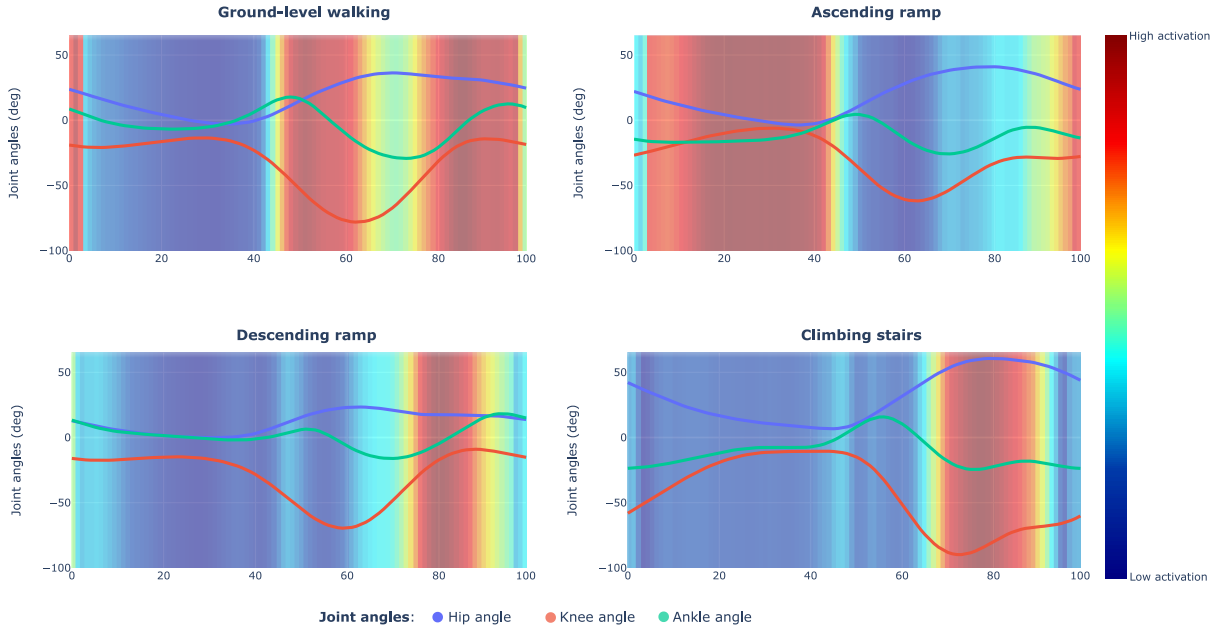


Fig. 6. Mean lower limb joint trajectories during ground-level walking, ascending a ramp, descending a ramp, and climbing stairs overlaid with the mean class activation maps. The class activation maps have been obtained with the Grad-CAM method, where blue parts represent low activation, and reds represent high activation.

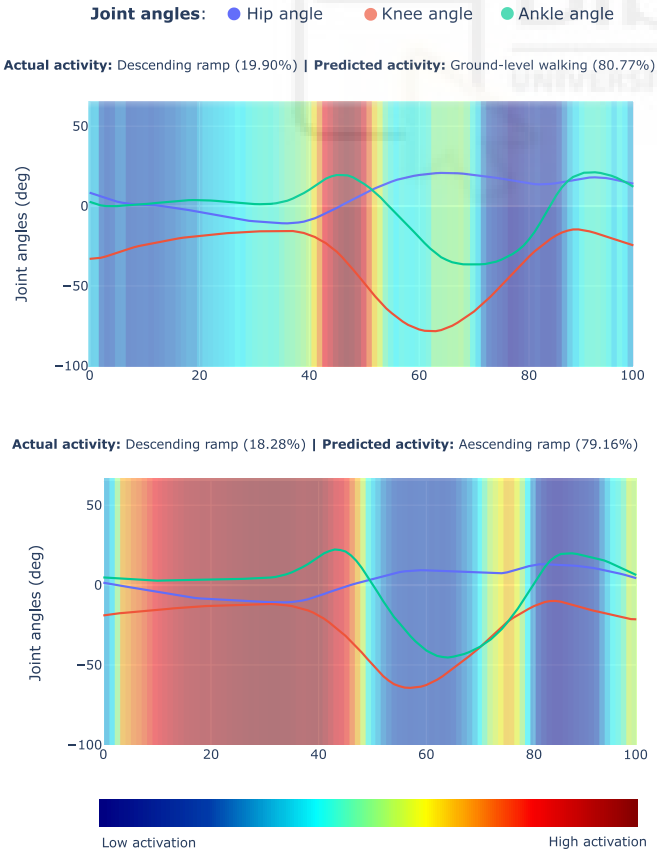


Fig. 7. Descending ramp steps classified incorrectly as ground-level walking and ascending ramp steps. The class activation maps and the joint trajectories are represented, as well the softmax values for the predicted and actual classes are collected.

a different number of epochs, all of them with the architecture proposed.

In Table II and Table III the accuracy of the models is shown according to the number of inputs and training epochs.

With the intra-participant data, when the CNN model is trained with the hip angle, the accuracy ranges from 26.91% (25 epochs) to 97.82% (200 epochs); with the knee angle, the accuracy ranges from 38.59% (200 epochs) to 93.21% (250 epochs); with the ankle angle, the accuracy ranges from 43.64% (10 epochs) to 99.91% (250 epochs); with the hip and knee angles the accuracy ranges from 38.41% (10 epochs) to 98.87% (250 epochs); with the hip and ankle angles the accuracy ranges from 82.32% (75 epochs) to 100% (100, 150, 200, and 250 epochs); and with the knee and ankle angles, the accuracy ranges from 30.49% (25 epochs) to 99.91% (200 epochs).

With the inter-participant data, when the CNN model is trained with the hip angle, the accuracy ranges from 25.71% (25 epochs) to 98.28% (200 epochs); with the knee angle it ranges from 35.23% (200 epochs) to 85.97% (250 epochs); with the ankle angle it ranges from 37.24% (10 epochs) to 98.18% (150 epochs); with the hip and knee it ranges from 37.63% (10 epochs) to 89.94% (250 epochs); with the hip and ankle angles it ranges from 63.54% (75 epochs) to 99.95% (150 epochs); and with the knee and ankle angles, the accuracy ranges from 29.83% (150 epochs) to 90.72% (200 epochs).

As detailed in previous sections, the results show that the model trained with all the angles of the lower limb joints achieves accuracy close to 100% for all the range of epochs with the intra-participant and inter-participant data.

TABLE II
INTRA-PARTICIPANT ACCURACY (%) ACCORDING TO EPOCHS NUMBER DURING TRAINING

	Epochs							
	10	**25**	**50**	**75**	**100**	**150**	**200**	**250**
Hip, knee, and ankle	100.00	100.00	100.00	99.91	100.00	100.00	99.65	100.00
Hip	37.54	26.91	70.73	76.05	46.34	93.73	97.82	96.52
Knee	47.30	65.85	70.12	51.74	45.12	71.52	38.59	93.21
Ankle	43.64	57.49	81.10	86.24	60.28	96.43	44.16	99.91
Hip and knee	38.41	45.56	91.46	67.07	72.30	99.39	72.83	98.87
Hip and ankle	92.42	99.65	99.91	82.32	100.00	100.00	100.00	100.00
Knee and ankle	41.20	30.49	69.60	47.91	69.33	31.97	99.91	84.84

TABLE III
INTER-PARTICIPANT ACCURACY (%) ACCORDING TO EPOCHS NUMBER DURING TRAINING

	Epochs							
	10	**25**	**50**	**75**	**100**	**150**	**200**	**250**
Hip, knee, and ankle	100.00	99.80	99.90	100.00	100.00	100.00	99.90	100.00
Hip	37.63	25.71	70.17	55.30	39.99	68.45	98.28	85.97
Knee	37.44	60.84	81.31	39.60	44.46	67.22	35.23	85.77
Ankle	37.24	46.37	89.11	85.33	59.67	98.18	34.44	97.55
Hip and knee	37.63	43.67	79.00	57.85	62.71	92.89	75.47	89.94
Hip and ankle	86.16	96.42	99.46	63.54	99.71	99.95	99.26	99.31
Knee and ankle	44.50	31.99	69.92	43.92	60.94	29.83	85.57	90.72

V. DISCUSSION

We trained a CNN to recognize four gait ADLs: ground-level walking, ramp ascent, ramp descent, and stair climbing. Several authors have proposed using accelerometer and gyroscope data from IMUs to recognize different gait activities [20], [25]–[27], [29], [30], [33], [34]. We propose a different approach from other studies since we used multiple IMUs to capture the joint angles of the lower limb in order to classify the gait activities. This way, we can do a biomechanical analysis and also adjust the assistance given by lower limb exoskeletons. Hence, the hip, knee, and ankle flexion/extension angles were introduced as inputs to a 1D-CNN. Our method, based on the of T. Seel et al. [36] approach, has the advantage of estimating joint angles without depending on the position and orientation of the IMUs. This means that variations in sensor placement do not affect the error in gait activity classification. Furthermore, the calibration method of this algorithm is simple, and it does not require accurate movements, as it only requires walking and performing random leg movements for a short period of time.

We can extract features describing the gait using the angles of the lower limb joints to distinguish between activities. However, we decided to use a CNN model since Deep Learning methods can extract features autonomously from the input data [18], also known as feature maps. The 1D-CNN proposed shows promising results, with 99.65% accuracy for intra-participants, and 99.56% accuracy for inter-participants (Fig. 4). In addition, the results collected in Table I show that our model outperforms other state-of-the-art architectures with the inter-participant data. However, it must be noted that the accuracy and the CE are in a similar range for all the trained models. These results suggest that introducing the lower limb joint trajectories is a robust method for gait ADL recognition, outperforming other approaches proposed previously (section II). Furthermore, the computational time results with our model are around 1 ms for each inference on a laptop, suggesting that our method can be employed to perform HAR in real time. Despite our CNN model achieved high accuracy with both intra-participants and inter-participants, in the confusion matrices can be observed that ground-level and walking on a ramp are sometimes classified wrong. The t-SNE and Grad-CAM methods could be the basis to clarify this behavior.

In Fig. 5, we have represented the projection of the 64-dimensional feature maps extracted with the proposed CNN model with data from four users: two intra-participants and two inter-participants. The resulting graph shows diverse groups of points (clusters), where each point represents the feature map obtained by the CNN for each step, and clusters represent a group of similar steps. The feature maps generated by our CNN model for climbing stairs are grouped in four different clusters, while the rest of the clusters could be composed of one or two ADLs. This could be explained by the similarity in the joint trajectories during the ground-level walking and walking on a ramp (Fig. 3). In addition, although there are clusters formed with two ADLs, in the graph can be observed that the ADLs are joined into four different groups.

The number of groups coincides with the number of users in the projection. This could mean that each participant has a unique gait pattern, so the feature maps generated by the CNN would not be equal for all the users. However, one might wonder why a good classification result is achieved when t-SNE does not show a clear separation of the extracted features. This may be because neural networks can recognize nonlinear boundaries so that ADL can be separated in the 64-dimensional space.

The t-SNE embedding method can help us understand how the features extracted by the CNN are grouped. However, it cannot verify which parts of the joint trajectories are important in determining the user's gait ADL. To examine which parts of the joint trajectories are crucial in classifying each gait ADL, we can use the Grad-CAM method. In Figure 6, the class activation maps for each gait ADL over the hip, knee, and ankle flexion/extension angles are represented. For instance, in the case of climbing stairs, the highest activation is produced during 70%-90% of the gait cycle. When we compare the hip, knee, and ankle trajectories during those periods with the rest of the gait ADLs, there is a notable difference. However, the t-SNE graph and confusion matrices show that wrong classifications occur more often during ground-level walking, ascending a ramp, and descending a ramp. Specifically, certain steps descending a ramp are misclassified as ground-level walking or ascending a ramp.

The Grad-CAM method can be a powerful tool to understand why our CNN model makes incorrect predictions. To this end, we can use joint trajectories and activation maps for two incorrect classifications, as shown in Figure 7. The left graph represents a step performed while descending a ramp but classified as ground-level walking. The class activation map shows high activation between 40% to 50%. This activation occurs around the peak of the ankle angle, which suggests that the gait pattern is more similar to flat walking. This is because, at this instant, the ankle angle is greater than the hip angle, which does not occur when walking on a descent ramp. In the case of the step classified as ascending a ramp, a high activation can be observed during the loading and stance phase, which is from 10% to 50%. This activation corresponds to the activation map of the ascending ramp steps and could be due to the similarity of the knee trajectory during this section of the gait cycle.

The feasibility of reducing the number of inputs to the 1D CNN was also examined. The results collected in Table II and Table III show that high accuracy can also be achieved by reducing the number of inputs. It should be noted that by introducing a single input of the 1D CNN, the model achieves an accuracy of 97.82% (intra-participants) and 98.28% (inter-participants) when using the hip angle, and an accuracy of 93. 21% (intra-participants) and 85.77% (inter-participants), and an accuracy of 99.91% (intra-participants) and 97.55% (inter-participants) for the ankle angle trajectory. Although the accuracy is reduced compared to the introduction of hip, knee, and ankle flexion/extension, these results suggest the feasibility of adapting the model to different situations. Thus, the 1D-CNN classifier could be adapted to analyze the biomechanics of a single lower limb joint, as well as to detect gait activities for different types of lower limb exoskeletons.

Although the results obtained by adapting the model suggest that our method can be used with exoskeletons, it must be noted that we intend to further develop and test the proposed method. Furthermore, we intend to integrate this HAR system to adapt the assistance of active lower limb exoskeletons and to provide adequate assistance on different types of surfaces.

VI. Conclusions

This work presents a deep learning model to perform HAR during gait. The classifier is based on a 1D-CNN and uses the flexion/extension angles of the hip, knee, and ankle measured with IMUs. It is worth noting that the method used in this study to measure the lower limb angles does not depend on the position or location of the sensors, making the method very reliable.

The CNN classifier proposed achieves high accuracy with intra-participants (99.65%) and inter-participants (99.56%), so it can be assumed that the proposed model presents a high generalization. The t-SNE and Grad-CAM algorithms were used to understand this behavior, and it can be concluded that this is due to anomalous steps or a high similarity of joint trajectories between activities.

Furthermore, we have trained the 1D-CNN varying the number of inputs. The results suggest that the 1D-CNN architecture proposed could be adapted to analyze the biomechanics of a single lower limb joint, as well as to detect gait activities for different types of lower limb exoskeletons.

References

[1] M. Roberts, D. Mongeon, and F. Prince, "Biomechanical parameters for gait analysis: a systematic review of healthy human gait," *Phys Ther Rehabil*, vol. 4, no. 1, p. 6, 2017.

[2] K. Hyodo, T. Masuda, J. Aizawa, T. Jinno, and S. Morita, "Hip, knee, and ankle kinematics during activities of daily living: a cross-sectional study," *Brazilian Journal of Physical Therapy*, vol. 21, no. 3, pp. 159–166, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S141335551730045X

[3] R. Baker, "Gait analysis methods in rehabilitation," *Journal of neuroengineering and rehabilitation*, vol. 3, pp. 1–10, 2006.

[4] S. Nadeau, M. Betschart, and F. Bethoux, "Gait analysis for poststroke rehabilitation: the relevance of biomechanical analysis and the impact of gait speed," *Physical Medicine and Rehabilitation Clinics*, vol. 24, no. 2, pp. 265–276, 2013.

[5] 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, no. 2, pp. 3362–3394, 2014.

[6] D. S. Pamungkas, W. Caesarendra, H. Soebakti, R. Analia, and S. Susanto, "Overview: Types of lower limb exoskeletons," *Electronics*, vol. 8, no. 11, p. 1283, 2019.

[7] A. M. Dollar and H. Herr, "Lower extremity exoskeletons and active orthoses: challenges and state-of-the-art," *IEEE Transactions on robotics*, vol. 24, no. 1, pp. 144–158, 2008.

[8] W. Huo, S. Mohammed, J. C. Moreno, and Y. Amirat, "Lower limb wearable robots for assistance and rehabilitation: A state of the art," *IEEE systems Journal*, vol. 10, no. 3, pp. 1068–1081, 2014.

[9] R. Baud, A. R. Manzoori, A. Ijspeert, and M. Bouri, "Review of control strategies for lower-limb exoskeletons to assist gait," *Journal of NeuroEngineering and Rehabilitation*, vol. 18, no. 1, pp. 1–34, 2021.

[10] T. Xue, Z. Wang, T. Zhang, and M. Zhang, "Adaptive oscillator-based robust control for flexible hip assistive exoskeleton," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3318–3323, 2019.

[11] W. Huo, S. Mohammed, Y. Amirat, and K. Kong, "Fast gait mode detection and assistive torque control of an exoskeletal robotic orthosis for walking assistance," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1035–1052, 2018.

[12] R. Auberger, M. F. Russold, R. Riener, and H. Dietl, "Patient motion using a computerized leg brace in everyday locomotion tasks," *IEEE Transactions on Medical Robotics and Bionics*, vol. 1, no. 2, pp. 106–114, 2019.

[13] X. Zhao, W. Chen, X. Yan, J. Wang, and X. Wu, "Real-time stairs geometric parameters estimation for lower limb rehabilitation exoskeleton," in *2018 Chinese Control And Decision Conference (CCDC)*, 2018, pp. 5018–5023.

[14] V. G. Santos, L. B. P. Nascimento, D. H. S. Fernandes, D. S. Pereira, P. J. Alsina, and M. V. Araújo, "Step modeling and safe path planning for a lower limb exoskeleton," in *2019 19th International Conference on Advanced Robotics (ICAR)*, 2019, pp. 560–565.

[15] S. Carvalho, J. Figueiredo, and C. P. Santos, "Environment-aware locomotion mode transition prediction system," in *2019 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, 2019, pp. 1–6.

[16] R. Begg, M. Palaniswami, and B. Owen, "Support vector machines for automated gait classification," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 5, pp. 828–838, 2005.

[17] A. Mannini, D. Trojaniello, A. Cereatti, and A. M. Sabatini, "A machine learning framework for gait classification using inertial sensors: Application to elderly, post-stroke and huntington's disease patients," *Sensors*, vol. 16, no. 1, 2016. [Online]. Available: https://www.mdpi.com/1424-8220/16/1/134

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[19] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 6999–7019, 2022.

[20] I. H. Lopez-Nava, L. M. Valentín-Coronado, M. Garcia-Constantino, and J. Favela, "Gait activity classification on unbalanced data from inertial sensors using shallow and deep learning," *Sensors*, vol. 20, no. 17, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/17/4756

[21] S.-M. Lee, S. M. Yoon, and H. Cho, "Human activity recognition from accelerometer data using convolutional neural network," in *2017 IEEE International Conference on Big Data and Smart Computing (BigComp)*, 2017, pp. 131–134.

[22] S. K. Yadav, K. Tiwari, H. M. Pandey, and S. A. Akbar, "A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions," *Knowledge-Based Systems*, vol. 223, p. 106970, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705121002331

[23] H. A. Varol and Y. Massalin, "A feasibility study of depth image based intent recognition for lower limb prostheses," in *2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2016, pp. 5055–5058.

[24] B. Laschowski, W. McNally, A. Wong, and J. McPhee, "Preliminary design of an environment recognition system for controlling robotic lower-limb prostheses and exoskeletons," in *2019 IEEE 16th international conference on rehabilitation robotics (ICORR)*. IEEE, 2019, pp. 868–873.

[25] C. Pham, N. N. Diep, and T. M. Phuong, "e-shoes: Smart shoes for unobtrusive human activity recognition," in *2017 9th International Conference on Knowledge and Systems Engineering (KSE)*. IEEE, 2017, pp. 269–274.

[26] C. A. Ronao and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert systems with applications*, vol. 59, pp. 235–244, 2016.

[27] S.-M. Lee, S. M. Yoon, and H. Cho, "Human activity recognition from accelerometer data using convolutional neural network," in *2017 ieee international conference on big data and smart computing (bigcomp)*. IEEE, 2017, pp. 131–134.

[28] M.-K. Yi, W.-K. Lee, and S. O. Hwang, "A human activity recognition method based on lightweight feature extraction combined with pruned and quantized cnn for wearable device," *IEEE Transactions on Consumer Electronics*, 2023.

[29] A. A. Badawi, A. Al-Kabbany, and H. Shaban, "Multimodal human activity recognition from wearable inertial sensors using machine learning," in *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*. IEEE, 2018, pp. 402–407.

[30] R. Chereshnev and A. Kertész-Farkas, "Hugadb: Human gait database for activity recognition from wearable inertial sensor networks," in *Analysis of Images, Social Networks and Texts: 6th International Conference, AIST 2017, Moscow, Russia, July 27–29, 2017, Revised Selected Papers 6*. Springer, 2018, pp. 131–141.

[31] Y. Makihara, H. Mannami, A. Tsuji, M. A. Hossain, K. Sugiura, A. Mori, and Y. Yagi, "The ou-isir gait database comprising the treadmill dataset," *IPSJ Transactions on Computer Vision and Applications*, vol. 4, pp. 53–62, 2012.

[32] V. B. Semwal, A. Gupta, and P. Lalwani, "An optimized hybrid deep learning model using ensemble learning approach for human walking activities recognition," *The Journal of Supercomputing*, vol. 77, no. 11, pp. 12 256–12 279, 2021.

[33] M. Liu, D. Wang, and H. Huang, "Development of an environment-aware locomotion mode recognition system for powered lower limb prostheses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 4, pp. 434–443, 2015.

[34] P. T. Chinimilli, S. Redkar, and W. Zhang, "Human activity recognition using inertial measurement units and smart shoes," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 1462–1467.

[35] J. Slemenšek, I. Fister, J. Geršak, B. Bratina, V. M. van Midden, Z. Pirtošek, and R. Šafarič, "Human gait activity recognition machine learning methods," *Sensors*, vol. 23, no. 2, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/2/745

[36] T. Seel, J. Raisch, and T. Schauer, "Imu-based joint angle measurement for gait analysis," *Sensors*, vol. 14, no. 4, pp. 6891–6909, 2014. [Online]. Available: https://www.mdpi.com/1424-8220/14/4/6891

[37] A. G. Cutti, A. Ferrari, P. Garofalo, M. Raggi, A. Cappello, and A. Ferrari, "'outwalk': a protocol for clinical gait analysis based on inertial and magnetic sensors," *Medical & biological engineering & computing*, vol. 48, pp. 17–25, 2010.

[38] S. Kumar, K. Gopinath, L. Rocchi, P. T. Sukumar, S. Kulkarni, and J. Sampath, "Towards a portable human gait analysis monitoring system," in *2018 International Conference on Signals and Systems (ICSigSys)*, 2018, pp. 174–180.

[39] D. Martínez-Pascual, J. M. Catalán, A. Blanco-Ivorra, M. Sanchís, F. Arán-Ais, and N. García-Aracil, "Estimating vertical ground reaction forces during gait from lower limb kinematics and vertical acceleration using wearable inertial sensors," *Frontiers in Bioengineering and Biotechnology*, vol. 11, 2023.

[40] D. Martínez-Pascual, J. M. Catalán, J. V. García-Pérez, M. Sanchís, F. Arán-Ais, and N. García-Aracil, "Activity classification with inertial sensors to perform gait analysis," in *International Symposium on Distributed Computing and Artificial Intelligence*. Springer, 2023, pp. 74–82.

[41] W. Zijlstra and A. L. Hof, "Assessment of spatio-temporal gait parameters from trunk accelerations during human walking," *Gait & Posture*, vol. 18, no. 2, pp. 1–10, 2003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S096663620200190X

[42] F. Gustafsson, "Determining the initial states in forward-backward filtering," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 988–992, 1996.

[43] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, "Convolutional neural networks for time series classification," *Journal of Systems Engineering and Electronics*, vol. 28, no. 1, pp. 162–169, 2017.

[44] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv preprint arXiv:1312.4400*, 2013.

[45] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[46] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[47] Z. Wang, W. Yan, and T. Oates, "Time series classification from scratch with deep neural networks: A strong baseline," in *2017 International joint conference on neural networks (IJCNN)*. IEEE, 2017, pp. 1578–1585.

[48] A. Le Guennec, S. Malinowski, and R. Tavenard, "Data augmentation for time series classification using convolutional neural networks," in *ECML/PKDD workshop on advanced analytics and learning on temporal data*, 2016.

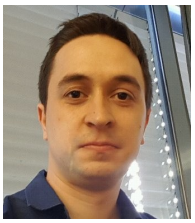
[49] Y. Zheng, Q. Liu, E. Chen, Y. Ge, and J. L. Zhao, "Exploiting multi-channels deep convolutional neural networks for multivariate time series classification," *Frontiers of Computer Science*, vol. 10, pp. 96–112, 2016.

[50] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, "Convolutional neural networks for time series classification," *Journal of Systems Engineering and Electronics*, vol. 28, no. 1, pp. 162–169, 2017.

[51] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.

David Martínez-Pascual received a B.S. degree in Industrial Electronics and Automation Engineering from Miguel Hernández University, Elche, Spain, in 2019 and an M.S. degree in robotics from Miguel Hernández University, Elche, Spain, in 2020. He is currently pursuing a Ph.D. degree in bioengineering at Miguel Hernández University, Elche, Spain.

FFrom 2019 to 2020, he was a Research Assistant in the Rehabilitation and Assistive Unit inside the Biomedical Neuroengineering Research Group at the Bioengineering Institute of Miguel Hernández University, Elche, Spain. Since 2020 he has been a Predoctoral Researcher in the Robotics & Artificial Intelligence Group at the Bioengineering Institute of Miguel Hernández University, Elche, Spain. He is a co-author of more than 10 research publications between ISI/Scopus papers and conference papers and is co-inventor of two patent applications. His main research interests include rehabilitation and assistive robotics, user modeling, and human-robot interaction.



José M. Catalán was born in Alicante, Alicante, Spain in 1991. He holds a M.Sc. in telecommunications technology engineering in electronic systems by the Miguel Hernández University, Elche, Spain, in 2014, a M.S. degree in telecommunications engineering by the Miguel Hernández University, Elche, Spain, in 2016, and a PhD in industrial and telecommunications technologies at Miguel Hernández University, Elche, Spain, in 2021. Since 2014, he has been a researcher in the Robotics and Artificial Intelligence research group at the Bioengineering Institute of Miguel Hernández University, Elche, Spain. Dr. José M. Catalán is co-author of more than 61 research publications, approximately 17 are ISI/Scopus articles, 8 are ISI/Scopus Book Chapters and the rest in national and international conferences. He is co-inventor of 6 patents/patent applications. He has actively participated in many national and international research projects, some of which include 4 international projects financed by the European Union (one ECHORD++ (FP7) and one European H2020-ICT-AIDE project). Dr. Jose M. Catalán is also co-director of a research chair in advanced technologies for rehabilitation.

Andrea Blanco-Ivorra received the B.S. degree in Mechanical Engineering from Miguel Hernández University, Elche, Spain, in 2014, a M.S. degree in Robotics from Miguel Hernández University, Elche, Spain, in 2017, and a Ph.D. in Industrial and Telecommunications Technologies at Miguel Hernández University, Elche, Spain, in 2022. From 2015 to 2017, she was a Research Assistant in Rehabilitation and Assistive Unit inside Biomedical Neuroengineering Research Group at the Bioengineering Institute of Miguel Hernández University, Elche, Spain. From 2018 to 2020 she was a Researcher in Training in the Systems Engineering and Automation Department of Miguel Hernández University, Elche, Spain. Since 2021, she has been Research Staff in Robotics and Artificial Intelligence research group at the Bioengineering Institute of Miguel Hernández University, Elche, Spain. She is co-author of more than 30 research publications, 13 are ISI/Scopus articles, 12 are ISI/Scopus Book Chapters and the rest in national and international conferences. She is co-inventor of 6 patent applications. She has about 7 years of research experience and has participated as a member of the research team in 2 international projects (ECHORD++ (FP7) and the European H2020-ICT-AIDE project) and more than 5 national projects.

Mónica Sanchís received a degree in Chemistry and a master's degree in Prevention of Occupational Risks, Safety, Industrial Hygiene, and Ergonomics Evaluations applied to working places, from the University of Alicante, Spain, in 1997 and 1999, respectively. She started 2008 her professional career at INESCOP as a researcher in Quality Department. From 2017 she was in charge of the Footwear Functional Analysis Department and work in the R&D INESCOP Project Office. She has extensive experience in the analysis of footwear and the preparation of reports on the measurement of feet and lasts, fit, dimensional evaluation of footwear, wear and comfort tests, as well as functional analysis using techniques of applied bioengineering such as internal and external baropodometry, surface electromyography, infrared thermography, or motion analysis. She has several publications on the slip resistance of footwear and she actively participates in the delivery of training courses at professional training and expert levels. She currently teaches classes in the "Master in Design, New Technologies, and Innovation Management in the footwear sector" at the University of Alicante and has been a speaker at different congresses in which footwear-related issues have been discussed.

Francisca Arán-Ais received the M. Sc degree in chemistry in 1996, and the PhD degree in chemistry science in 2000, both from the University of Alicante. With more than 25 years' experience in footwear sector, she is currently R&D coordinator at INESCOP, the Spanish Footwear Technological Center. Previously, she was responsible of the Technology Area including 4 different labs: Advanced Materials and Adhesives, Biotechnology, Circular Economy and Sustainability, and Footwear Functional Analysis. Therefore, her current research focuses on different fields related to footwear, such as adhesion and adhesives, advanced and multifunctional materials to improve comfort and footwear functionality, circular economy principles implementation to improve footwear sustainability such as debonded adhesives, recovery and recycling, biodegradation, composting, biorefineries, etc. She actively participates and coordinates European, National and Regional funded projects related to the mentioned topics. She is the author or co-author of more than 150 publications and more than 200 communications to national and international conferences. She is co-inventor of 2 patent applications. She is member of the Footwear Chair of the Miguel Hernández University.

Nicolas Garcia-Aracil (Member, IEEE) received the M.Sc. degree in control engineering and the master's degree in design, robotics and industrial automation from the University of Murcia, Spain, in 1996 and 1997, respectively, and the Ph.D. degree in control engineering from Miguel Hernández University, Elche, Spain.

He has been appointed as a Visiting Scientist at the Institut National de Recherche en Informatique et en Automatique (INRIA), Sophia Antipolis, France, in 2003, the Institute of Robotics and Mechatronics, DLR, Oberpfaffenhofen-Wessling, Germany, in 2006, and the Universita Campus Biomedico, Rome, Italy, in 2009. He is currently a Full Professor of control and systems engineering at Miguel Hernández University. He is also the Head of the Robotics and Artificial Intelligence Unit, Biomedical Neuroengineering Research Group, Bioengineering Institute, Miguel Hernández University.

He was the Project Coordinator of the European H2020-ICT-AIDE Project. He is the author or coauthor of a broad range of research publications. He is the coauthor of more than 50 ISI/Scopus papers and the co-inventor of more than six patents/patent applications. He served as the Program Co-Chair for the 2012 IEEE BIOROB.

BIBLIOGRAPHY

Ahmed, S.; Mayo, N. E.; Higgins, J.; Salbach, N. M.; Finch, L. and Wood-Dauphinee, S. L. (2003), «The Stroke Rehabilitation Assessment of Movement (STREAM): A Comparison With Other Measures Used to Evaluate Effects of Stroke and Rehabilitation,» *Physical Therapy* **83**(7), pp. 617–630, DOI: 10.1093/ptj/83.7.617 (cit. on p. 109).

Ai, Q.; Liu, Z.; Meng, W.; Liu, Q. and Xie, S. Q. (2023), «Machine Learning in Robot-Assisted Upper Limb Rehabilitation: A Focused Review,» *IEEE Transactions on Cognitive and Developmental Systems* **15** (4), pp. 2053–2063, DOI: 10.1109/TCDS.2021.3098350 (cit. on pp. 2, 18).

Alessandro, L.; Olmos, L. E.; Bonamico, L.; Muzio, D. M.; Ahumada, M. H.; Russo, M. J.; Allegri, R. F.; Gianella, M. G.; Campora, H.; Delorme, R., et al. (2020), «Multidisciplinary rehabilitation for adult patients with stroke,» *Medicina* **80**(1), pp. 54–68 (cit. on p. 1).

Altmann, A.; Toloşi, L.; Sander, O. and Lengauer, T. (2010), «Permutation importance: a corrected feature importance measure,» *Bioinformatics* **26**(10), pp. 1340–1347, DOI: 10.1093/bioinformatics/btq134, eprint: https://academic.oup.com/bioinformatics/article-pdf/26/10/1340/48851160/bioinformatics_26_10_1340.pdf (cit. on p. 49).

Andrade, K. d. O.; Pasqual, T. B.; Caurin, G. A. and Crocomo, M. K. (2016), «Dynamic difficulty adjustment with Evolutionary Algorithm in games for rehabilitation robotics,» in: *2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH)*, IEEE, pp. 1–8 (cit. on p. 17).

Andrade, K. O.; Joaquim, R. C.; Caurin, G. A. and Crocomo, M. K. (2018), «Evolutionary algorithms for a better gaming experience in rehabilitation robotics,» *Computers in Entertainment (CIE)* **16**(2), pp. 1–15 (cit. on p. 17).

Andrychowicz, M.; Denil, M.; Gomez, S.; Hoffman, M. W.; Pfau, D.; Schaul, T.; Shillingford, B. and De Freitas, N. (2016), «Learning to learn by gradient descent by gradient descent,» *Advances in neural information processing systems* **29** (cit. on p. 46).

Bäck, T. and Schütz, M. (1996), «Intelligent mutation rate control in canonical genetic algorithms,» in: *International Symposium on Methodologies for Intelligent Systems*, Springer, pp. 158–167 (cit. on p. 93).

Badawi, A. A.; Al-Kabbany, A. and Shaban, H. (2018), «Multimodal human activity recognition from wearable inertial sensors using machine learning,» in: *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, IEEE, pp. 402–407 (cit. on p. 14).

Badesa, F. J.; Morales, R.; Garcia-Aracil, N.; Sabater, J. M.; Casals, A. and Zollo, L. (2014), «Auto-adaptive robot-aided therapy using machine learning techniques,» *Computer methods and programs in biomedicine* **116**(2), pp. 123–130 (cit. on pp. xx, 16, 17).

Badesa, F. J.; Morales, R.; Garcia-Aracil, N. M.; Sabater, J. M.; Zollo, L.; Papaleo, E. and Guglielmelli, E. (2016), «Dynamic Adaptive System for Robot-Assisted Motion Rehabilitation,» *IEEE Systems Journal* **10**(3), pp. 984–991, DOI: 10.1109/JSYST.2014.2318594 (cit. on p. 16).

Beghriche, T.; Attallah, B.; Brik, Y. and Djerioui, M. (2023), «A multi-level fine-tuned deep learning based approach for binary classification of diabetic retinopathy,» *Chemometrics and Intelligent Laboratory Systems* **237**, p. 104820 (cit. on p. 115).

Berger, V. W. and Zhou, Y. (2014), «Kolmogorov–smirnov test: Overview,» *Wiley statsref: Statistics reference online* (cit. on p. 49).

Blank, A. A.; French, J. A.; Pehlivan, A. U. and O'Malley, M. K. (2014), «Current trends in robot-assisted upper-limb stroke rehabilitation: promoting patient engagement in therapy,» *Current physical medicine and rehabilitation reports* **2**, pp. 184–195 (cit. on p. 2).

Bradley, M. M. and Lang, P. J. (1994), «Measuring emotion: the self-assessment manikin and the semantic differential,» *Journal of behavior therapy and experimental psychiatry* **25**(1), pp. 49–59 (cit. on p. 96).

Byrd, R. H.; Schnabel, R. B. and Shultz, G. A. (1988), «Approximate solution of the trust region problem by minimization over two-dimensional subspaces,» *Mathematical programming* **40**(1), pp. 247–263 (cit. on p. 30).

Byrne, E. A. and Parasuraman, R. (1996), «Psychophysiology and adaptive automation,» *Biological psychology* **42**(3), pp. 249–268 (cit. on p. 16).

Canbek, G. (2023), «BenchMetrics Prob: benchmarking of probabilistic error/loss performance evaluation instruments for binary classification problems,» *International Journal of Machine Learning and Cybernetics* **14**(9), pp. 3161–3191 (cit. on p. 114).

Catalán, J. M.; Blanco-Ivorra, A.; García-Pérez, J. V.; Vales, Y.; Martínez-Pascual, D.; Ezquerro, S.; Garrote, A.; Costa, T.; Lledó, L. D. and García-Aracil, N. (2023a), «Patients' physiological reactions to competitive rehabilitation therapies assisted

by robotic devices,» *Journal of NeuroEngineering and Rehabilitation* **20**(1), p. 41 (cit. on pp. 3, 85, 121).

Catalán, J. M.; Blanco-Ivorra, A.; García-Pérez, J. V.; Vales, Y.; Martínez-Pascual, D.; Ezquerro, S.; Garrote, A.; Costa, T.; Lledó, L. D. and García-Aracil, N. (2023b), «Patients' physiological reactions to competitive rehabilitation therapies assisted by robotic devices,» *Journal of NeuroEngineering and Rehabilitation* **20**(1), p. 41 (cit. on p. 87).

Catalán, J. M.; García-Pérez, J. V.; Blanco, A.; Ezquerro, S.; Garrote, A.; Costa, T.; Bertomeu-Motos, A.; Díaz, I. and García-Aracil, N. (2021), «Tele-rehabilitation versus local rehabilitation therapies assisted by robotic devices: a pilot study with patients,» *Applied Sciences* **11**(14), p. 6259 (cit. on pp. xxvi, xxvii, 107, 109, 110).

Catalán, J. M.; García-Pérez, J. V.; Blanco, A.; Martínez, D.; Lledó, L. D. and García-Aracil, N. (2021), «Differences in physiological reactions due to a competitive rehabilitation game modality,» *Sensors* **21**(11), p. 3681 (cit. on p. 87).

Chereshnev, R. and Kertész-Farkas, A. (2018), «Hugadb: Human gait database for activity recognition from wearable inertial sensor networks,» in: *Analysis of Images, Social Networks and Texts: 6th International Conference, AIST 2017, Moscow, Russia, July 27–29, 2017, Revised Selected Papers 6*, Springer, pp. 131–141 (cit. on p. 14).

Choi, A.; Lee, J.-M. and Mun, J. H. (2013), «Ground reaction forces predicted by using artificial neural network during asymmetric movements,» *International Journal of Precision Engineering and Manufacturing* **14**(3), pp. 475–483 (cit. on pp. 11, 58).

Clauser, C. E.; Mc Conville, J. T. and Young, J. W. (1969), *Weight, volume, and center of mass of segments of the human body*, tech. rep. (cit. on pp. 34, 35).

Corke, P. I. (2007), «A simple and systematic approach to assigning Denavit–Hartenberg parameters,» *IEEE transactions on robotics* **23**(3), pp. 590–594 (cit. on p. 32).

Csikszentmihalyi, M. (2000), *Beyond boredom and anxiety.* Jossey-bass (cit. on p. 16).

Czikszentmihalyi, M. (1990), *Flow: The psychology of optimal experience* (cit. on p. 16).

Darzi, A. and Novak, D. (2021), «Automated affect classification and task difficulty adaptation in a competitive scenario based on physiological linkage: An exploratory study,» *International Journal of Human-Computer Studies* **153**, p. 102673 (cit. on p. 16).

Delp, S. L.; Anderson, F. C.; Arnold, A. S.; Loan, P.; Habib, A.; John, C. T.; Guendelman, E. and Thelen, D. G. (2007), «OpenSim: open-source software to create and analyze dynamic simulations of movement,» *IEEE transactions on biomedical engineering* **54**(11), pp. 1940–1950 (cit. on p. 10).

Dempster, W. T. and Gaughran, G. R. (1967), «Properties of body segments based on size and weight,» *American journal of anatomy* **120**(1), pp. 33–54 (cit. on pp. 34, 35).

Dietterich, T. (1995), «Overfitting and undercomputing in machine learning,» *ACM computing surveys (CSUR)* **27**(3), pp. 326–327 (cit. on p. 56).

Drillis, R.; Contini, R. and Bluestein, M. (1964), «Body segment parameters,» *Artificial limbs* **8**(1), pp. 44–66 (cit. on p. 34).

DuBois, D. (1916), «A formula to estimate the approx-imate surface area if height and weight be known,» *Arch intern med* **17**, pp. 863–871 (cit. on p. 34).

Dziedzic, D. (2016), «Dynamic difficulty adjustment systems for various game genres,» *Homo Ludens* **9**(1), pp. 35–51 (cit. on p. 15).

Dziedzic, D. and Włodarczyk, W. (2018), «Approaches to measuring the difficulty of games in dynamic difficulty adjustment systems,» *International Journal of Human–Computer Interaction* **34**(8), pp. 707–715 (cit. on p. 15).

Esser, P.; Dawes, H.; Collett, J. and Howells, K. (2009), «IMU: Inertial sensing of vertical CoM movement,» *Journal of Biomechanics* **42**(10), pp. 1578–1581, DOI: https://doi.org/10.1016/j.jbiomech.2009.03.049 (cit. on p. 43).

Faria, A. L.; Pinho, M. S. and Bermúdez i Badia, S. (2020), «A comparison of two personalization and adaptive cognitive rehabilitation approaches: a randomized controlled trial with chronic stroke patients,» *Journal of neuroengineering and rehabilitation* **17**, pp. 1–15 (cit. on pp. 2, 85).

Fogarty, T. (1989), «Varying the probability of mutation in the genetic algorithm,» in: *Proc. 3rd Int'l Conf. on Genetic Algorithms,(1989)*, Morgan Kaufmann, pp. 104–109 (cit. on p. 93).

Fong, J.; Ocampo, R.; Gross, D. P. and Tavakoli, M. (2020a), «Intelligent robotics incorporating machine learning algorithms for improving functional capacity evaluation and occupational rehabilitation,» *Journal of occupational rehabilitation* **30**, pp. 362–370 (cit. on p. 19).

Fong, J.; Ocampo, R.; Gross, D. P. and Tavakoli, M. (2020b), *Intelligent Robotics Incorporating Machine Learning Algorithms for Improving Functional Capacity Evaluation and Occupational Rehabilitation*, DOI: 10.1007/s10926-020-09888-w (cit. on p. 2).

Friedrich, M.; Gittler, G.; Halberstadt, Y.; Cermak, T. and Heiller, I. (1998), «Combined exercise and motivation program: effect on the compliance and level of disability of patients with chronic low back pain: a randomized controlled trial,» *Archives of physical medicine and rehabilitation* **79**(5), pp. 475–487, DOI: 10.1016/S0003-9993(98)90059-4 (cit. on p. 107).

Gilgen-Ammann, R.; Schweizer, T. and Wyss, T. (2019), «RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise,» *European journal of applied physiology* **119**(7), pp. 1525–1532 (cit. on p. 97).

Goršič, M.; Cikajlo, I.; Goljar, N. and Novak, D. (2017), «A multisession evaluation of an adaptive competitive arm rehabilitation game,» *Journal of neuroengineering and rehabilitation* **14**(1), pp. 1–15 (cit. on p. 16).

Gulli, A. and Pal, S. (2017a), *Deep learning with Keras*, Packt Publishing Ltd (cit. on pp. 46, 64).

Gulli, A. and Pal, S. (2017b), *Deep learning with Keras*, Packt Publishing Ltd (cit. on p. 114).

Gustafsson, F. (1996), «Determining the initial states in forward-backward filtering,» *IEEE Transactions on Signal Processing* **44**(4), pp. 988–992, DOI: 10.1109/78.492552 (cit. on p. 43).

Hasan, S. and Dhingra, A. K. (2022), «Biomechanical design and control of an eight DOF human lower extremity rehabilitation exoskeleton robot,» *Results in Control and Optimization* **7**, p. 100107 (cit. on p. 2).

Hassanat, A.; Almohammadi, K.; Alkafaween, E.; Abunawas, E.; Hammouri, A. and Prasath, V. S. (2019), «Choosing mutation and crossover ratios for genetic algorithms—a review with a new dynamic approach,» *Information* **10**(12), p. 390 (cit. on p. 93).

He, K.; Zhang, X.; Ren, S. and Sun, J. (2015), «Delving deep into rectifiers: Surpassing human-level performance on imagenet classification,» in: *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034 (cit. on p. 114).

Heins, S.; Dehem, S.; Montedoro, V.; Dehez, B.; Edwards, M.; Stoquart, G.; Rocca, F.; De Deken, P.-H.; Mancas, M. and Lejeune, T. (2017), «Robotic-assisted serious game for motor and cognitive post-stroke rehabilitation,» in: *2017 IEEE 5th international conference on serious games and applications for health (SeGAH)*, IEEE, pp. 1–8 (cit. on pp. 3, 85).

Hocine, N.; Gouaich, A.; Cerri, S. A.; Mottet, D.; Froger, J. and Laffont, I. (2015), «Adaptation in serious games for upper-limb rehabilitation: an approach to improve training outcomes,» *User Modeling and User-Adapted Interaction* **25**, pp. 65–98 (cit. on pp. xx, 17, 18).

Hu, S.; Mendonca, R.; Johnson, M. J. and Kuchenbecker, K. J. (2021), «Robotics for occupational therapy: Learning upper-limb exercises from demonstrations,» *IEEE Robotics and Automation Letters* **6** (4), pp. 7781–7788, DOI: 10.1109/LRA.2021.3098945 (cit. on p. 19).

iDRhA (n.d.), *https://idrha.es/*, Retrieved from https://idrha.es/ (cit. on p. 86).

Jacobs, N.; Skorecki, J. and Charnley, J. (1972), «Analysis of the vertical component of force in normal and pathological gait,» *Journal of Biomechanics* **5**(1), pp. 11–34 (cit. on p. 41).

Jakob, V.; Küderle, A.; Kluge, F.; Klucken, J.; Eskofier, B. M.; Winkler, J.; Winterholler, M. and Gassner, H. (2021), «Validation of a sensor-based gait analysis system with a gold-standard motion capture system in patients with Parkinson's disease,» *Sensors* **21**(22), p. 7680 (cit. on p. 9).

Jiang, X.; Napier, C.; Hannigan, B.; Eng, J. J. and Menon, C. (2020), «Estimating Vertical Ground Reaction Force during Walking Using a Single Inertial Sensor,» *Sensors* **20**(15), DOI: 10.3390/s20154345 (cit. on pp. 12, 58).

Kalita, J.; Bharadwaz, M. P. and Aditi, A. (2023), «Prevalence, contributing factors, and economic implications of strokes among older adults: a study of North-East India,» *Scientific Reports* **13**(1), p. 16880 (cit. on p. 1).

Karafotias, G.; Hoogendoorn, M. and Eiben, Á. E. (2014), «Parameter control in evolutionary algorithms: Trends and challenges,» *IEEE Transactions on Evolutionary Computation* **19**(2), pp. 167–187 (cit. on p. 93).

Kesiktas, N.; Paker, N.; Erdogan, N.; Gülsen, G.; Biçki, D. and Yilmaz, H. (2004), «The use of hydrotherapy for the management of spasticity,» *Neurorehabilitation and neural repair* **18**(4), pp. 268–273 (cit. on pp. 3, 23).

Kim, H.-G.; Cheon, E.-J.; Bai, D.-S.; Lee, Y. H. and Koo, B.-H. (2018), «Stress and heart rate variability: a meta-analysis and review of the literature,» *Psychiatry investigation* **15**(3), p. 235 (cit. on p. 97).

Kingma, D. P. and Ba, J. (2014), «Adam: A method for stochastic optimization,» *arXiv preprint arXiv:1412.6980* (cit. on p. 47).

Kramer, O. and Kramer, O. (2017), *Genetic algorithms*, Springer (cit. on p. 92).

Kumar, S.; Gopinath, K.; Rocchi, L.; Sukumar, P. T.; Kulkarni, S. and Sampath, J. (2018), «Towards a portable human gait analysis & monitoring system,» in: *2018 International Conference on Signals and Systems (ICSigSys)*, pp. 174–180, DOI: 10.1109/ICSIGSYS.2018.8372660 (cit. on p. 32).

Kwakkel, G. (2006), «Impact of intensity of practice after stroke: issues for consideration,» *Disability and rehabilitation* **28**(13-14), pp. 823–830 (cit. on pp. 1, 103).

Laborde, S.; Mosley, E. and Thayer, J. F. (2017), «Heart rate variability and cardiac vagal tone in psychophysiological research–recommendations for experiment planning, data analysis, and data reporting,» *Frontiers in psychology* **8**, p. 213 (cit. on p. 97).

Lambora, A.; Gupta, K. and Chopra, K. (2019), «Genetic algorithm-A literature review,» in: *2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)*, IEEE, pp. 380–384 (cit. on p. 92).

Laschowski, B.; McNally, W.; Wong, A. and McPhee, J. (2022), «Environment classification for robotic leg prostheses and exoskeletons using deep convolutional neural networks,» *Frontiers in Neurorobotics* **15**, p. 730965 (cit. on pp. xx, 14, 15).

Lauer, J. (2023), «Video-driven simulation of lower limb mechanical loading during aquatic exercises,» *Journal of Biomechanics* **152**, DOI: 10.1016/j.jbiomech.2023.111576 (cit. on pp. xx, 10, 25).

Lauretti, C.; Cordella, F.; Guglielmelli, E. and Zollo, L. (2017), «Learning by Demonstration for Planning Activities of Daily Living in Rehabilitation and Assistive Robotics,» *IEEE Robotics and Automation Letters* **2** (3), pp. 1375–1382, DOI: 10.1109/LRA.2017.2669369 (cit. on p. 19).

Le Guennec, A.; Malinowski, S. and Tavenard, R. (2016), «Data augmentation for time series classification using convolutional neural networks,» in: *ECML/PKDD workshop on advanced analytics and learning on temporal data* (cit. on pp. 76, 77, 115).

Lin, M.; Chen, Q. and Yan, S. (2013), «Network in network,» *arXiv preprint arXiv:1312.4400* (cit. on p. 72).

Liu, C.; Lu, J.; Yang, H. and Guo, K. (2022), «Current state of robotics in hand rehabilitation after stroke: a systematic review,» *Applied Sciences* **12**(9), p. 4540 (cit. on p. 2).

Liu, D.; He, M.; Hou, M. and Ma, Y. (2022), «Deep learning based ground reaction force estimation for stair walking using kinematic data,» *Measurement* **198**, p. 111344, DOI: 10.1016/J.MEASUREMENT.2022.111344 (cit. on pp. xx, 13).

Liu, K.; Liu, T.; Shibata, K. and Inoue, Y. (2009), «Ambulatory measurement and analysis of the lower limb 3D posture using wearable sensor system,» in: *2009 International Conference on Mechatronics and Automation*, pp. 3065–3069, DOI: 10.1109/ICMA.2009.5245982 (cit. on p. 31).

Lopez-Nava, I. H.; Valentín-Coronado, L. M.; Garcia-Constantino, M. and Favela, J. (2020), «Gait Activity Classification on Unbalanced Data from Inertial Sensors Using Shallow and Deep Learning,» *Sensors* **20**(17), DOI: 10.3390/s20174756 (cit. on p. 14).

Losey, D. P.; McDonald, C. G.; Battaglia, E. and O'Malley, M. K. (2018), *A review of intent detection, arbitration, and communication aspects of shared control for physical human–robot interaction*, DOI: 10.1115/1.4039145 (cit. on p. 2).

Luo, L.; Peng, L.; Wang, C. and Hou, Z. G. (2019), «A Greedy Assist-as-Needed Controller for Upper Limb Rehabilitation,» *IEEE Transactions on Neural Networks and*

Learning Systems **30** (11), pp. 3433–3443, DOI: 10.1109/TNNLS.2019.2892157 (cit. on p. 18).

Maclean, N.; Pound, P.; Wolfe, C. and Rudd, A. (2000), «A critical review of the concept of patient motivation in the literature on physical rehabilitation,» *Social Science & Medicine* **50**(4), pp. 495–506, DOI: 10.1016/s0277-9536(99)00334-2 (cit. on p. 107).

Maier, M.; Ballester, B. R. and Verschure, P. F. (2019), «Principles of neurorehabilitation after stroke based on motor learning and brain plasticity mechanisms,» *Frontiers in systems neuroscience* **13**, p. 74 (cit. on pp. 18, 86, 103).

Makihara, Y.; Mannami, H.; Tsuji, A.; Hossain, M. A.; Sugiura, K.; Mori, A. and Yagi, Y. (2012), «The OU-ISIR gait database comprising the treadmill dataset,» *IPSJ Transactions on Computer Vision and Applications* **4**, pp. 53–62 (cit. on p. 15).

Makowski, D.; Pham, T.; Lau, Z. J.; Brammer, J. C.; Lespinasse, F.; Pham, H.; Schölzel, C. and Chen, S. H. A. (2021), «NeuroKit2: A Python toolbox for neurophysiological signal processing,» *Behavior Research Methods* **53**(4), pp. 1689–1696, DOI: 10.3758/s13428-020-01516-y (cit. on p. 96).

Marasović, T.; Cecić, M. and Zanchi, V. (2009), «Analysis and interpretation of ground reaction forces in normal gait,» *WSEAS Trans. Syst* **8**(9), pp. 1105–1114 (cit. on p. 41).

Marinho-Buzelli, A. R.; Bonnyman, A. M. and Verrier, M. C. (2015), «The effects of aquatic therapy on mobility of individuals with neurological diseases: a systematic review,» *Clinical rehabilitation* **29**(8), pp. 741–751 (cit. on p. 23).

Martinez, C. M.; Fong, J.; Atashzar, S. F. and Tavakoli, M. (2019), «Semi-autonomous robot-assisted cooperative therapy exercises for a therapist's interaction with a patient,» in: *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, IEEE, pp. 1–5 (cit. on p. 19).

Mirjalili, S. and Mirjalili, S. (2019), «Genetic algorithm,» *Evolutionary Algorithms and Neural Networks: Theory and Applications*, pp. 43–55 (cit. on p. 92).

Miyoshi, T.; Hiramatsu, K.; Yamamoto, S.-I.; Nakazawa, K. and Akai, M. (2008), «Robotic gait trainer in water: Development of an underwater gait-training orthosis,» *Disability and Rehabilitation* **30**(2), pp. 81–87 (cit. on p. 23).

Miyoshi, T.; Komatsu, F.; Takagi, M. and Kawashima, N. (2015), «Attempt toward a development of aquatic exercise device for gait disorders,» *Disability and Rehabilitation: Assistive Technology* **10**(6), pp. 501–507 (cit. on p. 24).

Monoli, C.; Fuentez-Perez, J. F.; Cau, N.; Capodaglio, P.; Galli, M. and Tuhtan, J. A. (2021), «Land and Underwater Gait Analysis Using Wearable IMU,» *IEEE Sensors Journal* **21** (9), pp. 11192–11202, DOI: 10.1109/JSEN.2021.3061623 (cit. on p. 9).

Morer, C.; Michan-Doña, A.; Alvarez-Badillo, A.; Zuluaga, P. and Maraver, F. (2020), «Evaluation of the feasibility of a two-week course of aquatic therapy and thalassotherapy in a mild post-stroke population,» *International Journal of Environmental Research and Public Health* **17**(21), p. 8163 (cit. on p. 32).

Najafi, M.; Rossa, C.; Adams, K. and Tavakoli, M. (2020), «Using Potential Field Function with a Velocity Field Controller to Learn and Reproduce the Therapist's Assistance in Robot-Assisted Rehabilitation,» *IEEE/ASME Transactions on Mechatronics* **25** (3), pp. 1622–1633, DOI: 10.1109/TMECH.2020.2981625 (cit. on pp. xxi, 19, 20).

Nas, K.; Yazmalar, L.; Şah, V.; Aydın, A. and Öneş, K. (2015), «Rehabilitation of spinal cord injuries,» *World journal of orthopedics* **6**(1), p. 8 (cit. on p. 23).

Nations, U. (2017), «Department of Economic and Social Affairs. Population Division. World population prospects: the 2017 revision: key findings and advance tables,» *Working Paper No. ESA/P/WP/248* (cit. on p. 1).

Niu, Z.; Zhong, G. and Yu, H. (2021), «A review on the attention mechanism of deep learning,» *Neurocomputing* **452**, pp. 48–62 (cit. on p. 114).

Oh, S. E.; Choi, A. and Mun, J. H. (2013), «Prediction of ground reaction forces during gait based on kinematics and a neural network model,» *Journal of Biomechanics* **46**(14), pp. 2372–2380, DOI: https://doi.org/10.1016/j.jbiomech.2013.07.036 (cit. on pp. 11, 58).

Oña, E.; Cano-de La Cuerda, R.; Sánchez-Herrera, P.; Balaguer, C. and Jardón, A. (2018), «A review of robotics in neurorehabilitation: Towards an automated process for upper limb,» *Journal of healthcare engineering* **2018** (cit. on p. 1).

Orselli, M. I. V. and Duarte, M. (2011), «Joint forces and torques when walking in shallow water,» *Journal of Biomechanics* **44** (6), pp. 1170–1175, DOI: 10.1016/j.jbiomech.2011.01.017 (cit. on p. 11).

Palumbo, A.; Vizza, P.; Calabrese, B. and Ielpo, N. (2021), «Biopotential Signal Monitoring Systems in Rehabilitation: A Review,» *Sensors* **21**(21), DOI: 10.3390/s21217172 (cit. on p. 88).

Pareek, S. and Kesavadas, T. (2020), «IART: Learning from Demonstration for Assisted Robotic Therapy Using LSTM,» *IEEE Robotics and Automation Letters* **5** (2), pp. 477–484, DOI: 10.1109/LRA.2019.2961845 (cit. on pp. xxi, 20, 21, 107, 115, 121).

Pareek, S.; Nisar, H. J. and Kesavadas, T. (2023), «AR3n: A Reinforcement Learning-Based Assist-as-Needed Controller for Robotic Rehabilitation,» *IEEE Robotics and Automation Magazine*, DOI: 10.1109/MRA.2023.3282434 (cit. on p. 19).

Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V., et al. (2011), «Scikit-learn: Machine

learning in Python,» *the Journal of machine Learning research* **12**, pp. 2825–2830 (cit. on p. 45).

Pham, C.; Diep, N. N. and Phuong, T. M. (2017), «e-Shoes: Smart shoes for unobtrusive human activity recognition,» in: *2017 9th International Conference on Knowledge and Systems Engineering (KSE)*, IEEE, pp. 269–274 (cit. on p. 14).

Pirovano, M.; Mainetti, R.; Baud-Bovy, G.; Lanzi, P. L. and Borghese, N. A. (2014), «Intelligent game engine for rehabilitation (IGER),» *IEEE Transactions on Computational Intelligence and AI in Games* **8**(1), pp. 43–55 (cit. on p. 16).

Pratiwi, H.; Windarto, A. P.; Susliansyah, S.; Aria, R. R.; Susilowati, S.; Rahayu, L. K.; Fitriani, Y.; Merdekawati, A. and Rahadjeng, I. R. (2020), «Sigmoid activation function in selecting the best model of artificial neural networks,» in: *Journal of Physics: Conference Series*, vol. 1471, 1, IOP Publishing, p. 012010 (cit. on p. 114).

Pratt, J. W. (1959), «Remarks on Zeros and Ties in the Wilcoxon Signed Rank Procedures,» *Journal of the American Statistical Association* **54**(287), pp. 655–667, DOI: 10.1080/01621459.1959.10501526, eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1959.10501526 (cit. on p. 49).

Qassim, H. M. and Wan Hasan, W. Z. (2020), «A Review on Upper Limb Rehabilitation Robots,» *Applied Sciences* **10**(19), DOI: 10.3390/app10196976 (cit. on p. 2).

Rajendra Acharya, U.; Paul Joseph, K.; Kannathal, N.; Lim, C. M. and Suri, J. S. (2006), «Heart rate variability: a review,» *Medical and biological engineering and computing* **44**, pp. 1031–1051 (cit. on p. 97).

Ravichandar, H.; Polydoros, A. S.; Chernova, S. and Billard, A. (2020), «Recent advances in robot learning from demonstration,» *Annual review of control, robotics, and autonomous systems* **3**, pp. 297–330 (cit. on p. 19).

Rekand, T.; Hagen, E. M. and Grønning, M. (2012), «Spasticity following spinal cord injury,» *Tidsskrift for Den norske legeforening* (cit. on p. 23).

Roper, J. A.; Bressel, E. and Tillman, M. D. (2013), «Acute aquatic treadmill exercise improves gait and pain in people with knee osteoarthritis,» *Archives of physical medicine and rehabilitation* **94**(3), pp. 419–425 (cit. on pp. 3, 23).

Seel, T.; Raisch, J. and Schauer, T. (2014), «IMU-Based Joint Angle Measurement for Gait Analysis,» *Sensors* **14**(4), pp. 6891–6909, DOI: 10.3390/s140406891 (cit. on pp. 28, 29, 31).

Seel, T.; Schauer, T. and Raisch, J. (2012), «Joint axis and position estimation from inertial measurement data by exploiting kinematic constraints,» in: *2012 IEEE International Conference on Control Applications*, pp. 45–49, DOI: 10.1109/CCA.2012.6402423 (cit. on p. 28).

Semwal, V. B.; Gupta, A. and Lalwani, P. (2021), «An optimized hybrid deep learning model using ensemble learning approach for human walking activities recognition,» *The Journal of Supercomputing* **77**(11), pp. 12256–12279 (cit. on p. 15).

Serra, J.; Pascual, S. and Karatzoglou, A. (2018), «Towards a Universal Neural Network Encoder for Time Series.,» in: *CCIA*, pp. 120–129 (cit. on pp. 113, 115).

Shahabpoor, E. and Pavic, A. (2018), «Estimation of vertical walking ground reaction force in real-life environments using single IMU sensor,» *Journal of Biomechanics* **79**, pp. 181–190, DOI: https://doi.org/10.1016/j.jbiomech.2018.08.015 (cit. on pp. 11, 42, 48).

Sharma, S.; Sharma, S. and Athaiya, A. (2017), «Activation functions in neural networks,» *towards data science* **6**(12), pp. 310–316 (cit. on p. 46).

Sheldon, M. R.; Fillyaw, M. J. and Thompson, W. D. (1996), «The use and interpretation of the Friedman test in the analysis of ordinal-scale data in repeated measures designs,» *Physiotherapy Research International* **1**(4), pp. 221–228 (cit. on p. 49).

Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I. and Salakhutdinov, R. (2014), «Dropout: a simple way to prevent neural networks from overfitting,» *The journal of machine learning research* **15**(1), pp. 1929–1958 (cit. on p. 47).

Svozil, D.; Kvasnicka, V. and Pospichal, J. (1997), «Introduction to multi-layer feed-forward neural networks,» *Chemometrics and Intelligent Laboratory Systems* **39**(1), pp. 43–62, DOI: https://doi.org/10.1016/S0169-7439(97)00061-0 (cit. on p. 46).

Thibaut, A.; Chatelle, C.; Ziegler, E.; Bruno, M.-A.; Laureys, S. and Gosseries, O. (2013), «Spasticity after stroke: physiology, assessment and treatment,» *Brain injury* **27**(10), pp. 1093–1105 (cit. on p. 23).

Too, E. C.; Yujian, L.; Njuki, S. and Yingchun, L. (2019), «A comparative study of fine-tuning deep learning models for plant disease identification,» *Computers and Electronics in Agriculture* **161**, pp. 272–279 (cit. on p. 115).

Tresser, S.; Kuflik, T.; Levin, I. and Weiss, P. L. (2021), «Personalized rehabilitation for children with cerebral palsy,» *User modeling and user-adapted interaction* **31**(4), pp. 829–865 (cit. on pp. xx, 17, 18).

Ulyanov, D.; Vedaldi, A. and Lempitsky, V. (2016), «Instance normalization: The missing ingredient for fast stylization,» *arXiv preprint arXiv:1607.08022* (cit. on p. 114).

Vales, Y.; Catal, J. M.; Lled, L. D.; Blanco, A.; Álvarez-Pastor, J.; Martínez-Pascual, D.; Garrote, A.; Costa, T.; Más, G. and García-Aracil, N. (2022), «Influence of session duration in robot-assisted neurorehabilitation therapies: a pilot study with patients,» in: *Iberian Robotics conference*, Springer, pp. 286–297 (cit. on p. 87).

Vales, Y.; Catalán, J. M.; Bertomeu-Motos, A.; García-Pérez, J. V.; Lledó, L. D.; Blanco-Ivorra, A.; Marzo, C. A.; Más, G. and García-Aracil, N. (2023), «Influence of Robotic Therapy on Severe Stroke Patients,» in: *2023 International Conference on Rehabilitation Robotics (ICORR)*, pp. 1–6, DOI: 10.1109/ICORR58425.2023.10304780 (cit. on p. 87).

Van der Maaten, L. and Hinton, G. (2008), «Visualizing data using t-SNE.,» *Journal of machine learning research* **9**(11) (cit. on p. 76).

Varol, H. A. and Massalin, Y. (2016), «A feasibility study of depth image based intent recognition for lower limb prostheses,» in: *2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, IEEE, pp. 5055–5058 (cit. on p. 14).

Veerbeek, J. M.; Langbroek-Amersfoort, A. C.; Van Wegen, E. E.; Meskers, C. G. and Kwakkel, G. (2017), «Effects of robot-assisted therapy for the upper limb after stroke: a systematic review and meta-analysis,» *Neurorehabilitation and neural repair* **31**(2), pp. 107–121 (cit. on pp. 2, 23).

Veldema, J. and Jansen, P. (2021), «Aquatic therapy in stroke rehabilitation: systematic review and meta-analysis,» *Acta Neurologica Scandinavica* **143**(3), pp. 221–241 (cit. on p. 33).

Verhulst, A.; Yamaguchi, T. and Richard, P. (2015), «Physiological-based Dynamic Difficulty Adaptation in a Theragame for Children with Cerebral Palsy.,» in: *PhyCS*, pp. 164–171 (cit. on p. 16).

Volpe, D.; Pavan, D.; Morris, M.; Guiotto, A.; Iansek, R.; Fortuna, S.; Frazzitta, G. and Sawacha, Z. (2017), «Underwater gait analysis in Parkinson's disease,» *Gait & Posture* **52**, pp. 87–94, DOI: 10.1016/J.GAITPOST.2016.11.019 (cit. on p. 9).

Wang, Z.; Yan, W. and Oates, T. (2017a), «Time series classification from scratch with deep neural networks: A strong baseline,» in: *2017 International joint conference on neural networks (IJCNN)*, IEEE, pp. 1578–1585 (cit. on pp. 76, 77, 115).

Wang, Z.; Yan, W. and Oates, T. (2017b), «Time series classification from scratch with deep neural networks: A strong baseline,» in: *2017 International joint conference on neural networks (IJCNN)*, IEEE, pp. 1578–1585 (cit. on p. 115).

Winstein, C. J.; Stein, J.; Arena, R.; Bates, B.; Cherney, L. R.; Cramer, S. C.; Deruyter, F.; Eng, J. J.; Fisher, B.; Harvey, R. L., et al. (2016), «Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association,» *Stroke* **47**(6), e98–e169 (cit. on pp. 1, 103).

Wouda, F. J.; Giuberti, M.; Bellusci, G.; Maartens, E.; Reenalda, J.; Beijnum, B.-J. F. van and Veltink, P. H. (2018), «Estimation of Vertical Ground Reaction Forces and Sagittal Knee Kinematics During Running Using Three Inertial Sensors,» *Frontiers in Physiology* **9**, DOI: 10.3389/fphys.2018.00218 (cit. on pp. xx, 12).

Xia, J.; Falconer, R. A.; Wang, Y. and Xiao, X. (2014), «New criterion for the stability of a human body in floodwaters,» *Journal of Hydraulic Research* **52**(1), pp. 93–104 (cit. on p. 34).

Xu, G.; Gao, X.; Pan, L.; Chen, S.; Wang, Q.; Zhu, B. and Li, J. (2018), «Anxiety detection and training task adaptation in robot-assisted active stroke rehabilitation,» *International Journal of Advanced Robotic Systems* **15**(6), p. 1729881418806433 (cit. on p. 16).

Yadav, S. K.; Tiwari, K.; Pandey, H. M. and Akbar, S. A. (2021), «A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions,» *Knowledge-Based Systems* **223**, p. 106970, DOI: https://doi.org/10.1016/j.knosys.2021.106970 (cit. on p. 14).

Young, A. (2009), «Comparison of Orientation Filter Algorithms for Realtime Wireless Inertial Posture Tracking,» in: *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*, pp. 59–64, DOI: 10.1109/BSN.2009.25 (cit. on p. 32).

Zadpoor, A. A. and Nikooyan, A. A. (2012), «The effects of lower-extremity muscle fatigue on the vertical ground reaction force: a meta-analysis,» *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* **226**(8), pp. 579–588 (cit. on p. 41).

Zhang, C. and Ma, Y. (2012), *Ensemble machine learning: methods and applications*, Springer (cit. on p. 45).

Zhao, B.; Lu, H.; Chen, S.; Liu, J. and Wu, D. (2017a), «Convolutional neural networks for time series classification,» *Journal of Systems Engineering and Electronics* **28**(1), pp. 162–169, DOI: 10.21629/JSEE.2017.01.18 (cit. on p. 72).

Zhao, B.; Lu, H.; Chen, S.; Liu, J. and Wu, D. (2017b), «Convolutional neural networks for time series classification,» *Journal of Systems Engineering and Electronics* **28**(1), pp. 162–169 (cit. on pp. 76, 77, 115).

Zheng, Y.; Liu, Q.; Chen, E.; Ge, Y. and Zhao, J. L. (2016), «Exploiting multi-channels deep convolutional neural networks for multivariate time series classification,» *Frontiers of Computer Science* **10**, pp. 96–112 (cit. on pp. 76, 77, 115).

Zhou, J.; Yang, S. and Xue, Q. (2021), «Lower limb rehabilitation exoskeleton robot: A review,» *Advances in Mechanical Engineering* **13**(4), p. 16878140211011862 (cit. on p. 2).

Zhou, L.; Li, Y. and Bai, S. (2017), «A human-centered design optimization approach for robotic exoskeletons through biomechanical simulation,» *Robotics and Autonomous Systems* **91**, pp. 337–347 (cit. on p. 2).

Zijlstra, W. and Hof, A. L. (2003), «Assessment of spatio-temporal gait parameters from trunk accelerations during human walking,» *Gait Posture* **18**(2), pp. 1–10, DOI: https://doi.org/10.1016/S0966-6362(02)00190-X (cit. on p. 45).