Universidade do Minho
Escola de Engenharia

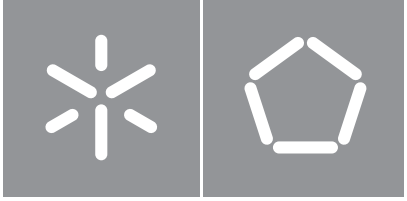
José Pedro Rodrigues Manso

Decoding Activities of Daily Living and Incipient Falls using Brain Signals

Decoding Activities of Daily Living and Incipient Falls using Brain Signals
José Pedro Rodrigues Manso

UMinho | 2023

October 2023



Universidade do Minho
Escola de Engenharia

José Pedro Rodrigues Manso

Decoding Activities of Daily Living and Incipient Falls using Brain Signals

Master's Dissertation
Master's in Informatics Engineering

Work supervised by
Professor Dr. Cristina P. Santos
Professor Dr. Nuno F. Ribeiro

October, 2023

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the Universidade do Minho.

Acknowledgements

I would like to start by thanking Professor Cristina P. Santos, who was indispensable for more than just scientific oversight, but also in terms of support, advice, encouragement, and suggestions. Undoubtedly, Professor Cristina enabled the materialisation of the project's ideas through a methodical approach.

To Nuno, who was always present, available to help and offer valuable recommendations to enrich the quality of the study. I am grateful for all the discussions that, ultimately, aided in the development of this achievement. Thank you for all the moments and for all the encouragement, especially at the most difficult times. I am also grateful for consistently guiding me towards achieving my highest potential.

I would also like to express my gratitude to the entire BiRDLab team for their warm welcome and the supportive work environment they have fostered, especially José Miguel, André and Ricardo. I want to give a special thanks to Luís for his invaluable assistance. Thank you for the countless hours dedicated to overcoming the challenges encountered on the journey. I would also like to extend special thanks to those who participated in the the slip-like perturbations protocol, and those who made the two protocols possible. To Raimundo, for his patience and for helping me whenever I needed.

My appreciation also goes to all my colleagues and friends who embarked on this journey with me, with special recognition to Rafael and João Pedro. I take pride in the moments we have shared and the way we have collectively tackled unexpected challenges along the way. I will never forget the moments we went through even in tough times.

To my primary school teacher, Odília Santos. Thank you for never giving up on me.

I am thankful for all my friends, with a special dedication to Marta, André, Ana, Afonso, Luís, Gonçalo, Hugo, Sara, Rita, Leonardo and Sofia. Your companionship has provided invaluable moments of much-needed diversion that greatly contributed to my well-being. I cannot express enough how grateful and proud I am to have you alongside me.

Finally, I would like to thank the people that represent the reason why I am what I am today. To my family, in particular my sister, my mother and my father, this dissertation is dedicated to you. It is simply a privilege to grow alongside and because of you. You have made me a better person all my life, and I cannot thank you enough. I owe you a lot. To my grandparents, who have taught me and gave me so much. All of you, together, symbolise everything I want to be, both as a person and as a professional.

To all of you, my absolute most sincere thank you!

"The chances you take, the people you meet, the people you love, the faith that you have. That's what's going to define you."

— Denzel Washington

Abstract

Decoding Activities of Daily Living and Incipient Falls using Brain Signals

Falls represent one of the biggest causes of deaths related to unintentional injuries. The increasing number of occurrences is associated a continuously expanding elderly population, along with its detrimental effects on the survival and well-being of those aged 65 and above, has turned the issue of falls into a global public health concern. It is estimated that 684,000 people worldwide lose their lives due to falls, which happen approximately 37.3 million times annually. As a result, the financial expenses associated with hospitalisations are significant and present a complex challenge.

The Electroencephalogram (EEG) technique is widely utilised to assess brain electrical activity and detect indicators of balance disruptions, such as Perturbation Evoked Potentials (PEPs), in brain signals. This is possible because EEG data provides insights into motor planning and intention, making it a valuable tool for monitoring both falls and Activities of Daily Living (ADLs). Accordingly, this dissertation will establish two experimental protocols: one for simulating slip-like incidents and another for ADLs, with the aim of collecting EEG data. The primary goal of this dissertation is to leverage Artificial Intelligence (AI)-based systems to identify slip-like perturbations and various ADLs using the data from both protocols. The ultimate objective is to integrate these algorithms into assistive robotic devices, e.g. exoskeletons.

In the context of the methods employed, the PEP components were identified within a time frame of 75–137 ms after the external perturbation onset. To analyse the pre-processed EEG data, four distinct artificial neural networks were evaluated, each with varying network architecture parameters. Among these architectures, the Convolutional Neural Network (CNN)-Long Short-Term Memory (LSTM) model, trained to predict EEG perturbations, exhibited superior classification performance, achieving an accuracy rate of 86% when using a short time window of 100 ms. In contrast, for classifying ADL, the best result obtained was 53% accuracy, and this was also achieved using the CNN-LSTM architecture.

Keywords: EEG, ADL, Falls, Artificial Intelligence, Slip-like perturbations

Resumo

Descodificação de atividades do dia a dia e quedas iminentes com recurso a sinais cerebrais

As quedas representam uma das maiores causas de morte relacionadas com lesões não intencionais. O número crescente de ocorrências está associado a uma população idosa em constante expansão, juntamente com os seus efeitos prejudiciais na sobrevivência e no bem-estar das pessoas com 65 anos ou mais, o que transformou a questão das quedas num problema de saúde pública global. Estima-se que 684.000 pessoas em todo o mundo perdem a vida devido a quedas. Consequentemente, as despesas financeiras são significativas e representam um desafio complexo.

A técnica eletroencefalograma (EEG) é amplamente utilizada para avaliar a atividade eléctrica cerebral e detetar potenciais evocados por perturbação (PEPs) e Atividades Diárias (ADLs). Assim, esta dissertação irá estabelecer dois protocolos experimentais: um para simular incidentes do tipo escorregar e outro para ADLs, com o objetivo de recolher dados EEG. O objetivo principal desta dissertação é utilizar sistemas baseados em Inteligência Artificial (IA) para identificar perturbações do tipo escorregar e várias ADLs utilizando os dados de ambos os protocolos. O objetivo final é integrar estes algoritmos em dispositivos robóticos de assistência, por exemplo, exoesqueletos.

No contexto dos métodos utilizados, os componentes PEP foram identificados num período de tempo de 75-137 ms após o início da perturbação externa. Para analisar os dados EEG, foram avaliadas quatro redes neurais artificiais distintas, cada uma com parâmetros de arquitetura de rede variáveis. Entre estas arquitecturas, o modelo de Rede Neuronal Convolucional (CNN)-Long Short-Term Memory (LSTM) apresentou um desempenho de classificação superior, alcançando uma taxa de precisão de 86% quando se utilizou uma janela de tempo curta de 100 ms. Em contraste, para a classificação de ADL, o melhor resultado obtido foi uma precisão de 53%, o que também foi conseguido utilizando a arquitetura CNN-LSTM.

Palavras-chave: EEG, ADL, Quedas, Inteligência Artificial, Perturbações do tipo escorregar

Contents

List of Figures

List of Tables

Acronyms

ACC Anterior Cingulate Cortex

ADL Activities of daily living

AI Artificial Intelligence

AM Amplitude Modulated

ANN Artificial Neural Network

APA Anticipatory Postural Adjustment

ASR Artifacts Subspace Reconstruction

B Batch size

BCI Brain Computer Interface

BF Biceps Femoris

BiRDLab Biomedical Robotic Devices Laboratory

BMI Body Mass Index

CAR Common Average Reference

CCA Canonical Correlation Analysis

CDP Computerized Dynamic Posturography

CLDA Closed-loop Decoder Adaptation

CMEMS Center of Micro Electro-Mechanical Systems

CNN Convolutional Neural Network

CPA Compensatory Postural Adjustment

D Dropout

DL Deep Learning

E Epochs

ECG Electrocardiogram

EEG Electroencephalogram

EEMD-CCA Ensemble Empirical Mode Decomposition with Canonical Correlation Analysis

EMG Electromyography

EOG Electrooculography

ERD Event-Related Desynchronisation

ERP Event-related Potential

ERS Event-Related Synchronisation

FAC Functional Ambulation Categories

FES-I Falls Efficacy Scale – International

FIR Finite Impulse Response

GDS Geriatric Depression Scale

GL Gastrocnemius Lateralis

GM Gastrocnemius Medialis

GND Ground

HO Hold-out

HW Hardware

IC Independent Component

ICA Independent Component Analysis

IIR Infinite Impulse Response

IMU Inertial Measurement Units

IPL Inferior Parietal Lobe

KPI Key Performance Indicator

LOB Loss of Balance

LOO Leave-one-out

LSTM Long Short-Term Memory

MCC Matthews Correlation Coefficient

ME Motor Execution

MMP Movement Monitoring Potential

MMSE Mini Mental State Examination

MRCP Movement Related Cortical Potentials

MT Muscle Trigger

O Overlap

OS Operative System

PEP Perturbation Evoked Potential

PPC Posterior Parietal Cortex

pre-SMA pre-Supplementary Motor Area

PREP Processing Pipeline

PRISMA Preferred Reporting Items for Systematic Review and Meta-Analysis

rASR Riemannian Artifact Subspace Reconstruction

REF Reference

RELICA Reliable Independent Component Analysis

RF Rectus Femoris

RP Readiness Potential

SCP Slow Cortical Potential

SL Sequence Length

STG Superior Temporal Gyrus

TA Tibialis Anterior

UKF Unscented Kalman Filter

VL Vastus Lateralis

VM Vastus Medialis

VR Virtual Reality

WHO World Health Organization

Chapter 1

Introduction

In this document, it is presented the work developed during the second year of the Master's in Informatics Engineering during the academic year of 2022/23. All the investigation work documented was developed at Biomedical Robotic Devices Laboratory (BiRDLab) located at the Center of Micro Electro-Mechanical Systems (CMEMS), University of Minho, Braga, Portugal. The aim of this project is to create an Artificial Intelligence (AI)-driven strategy capable of identifying early signs of falls and monitoring Activities of daily livings (ADLs). This technology has the potential to activate preventative measures in auxiliary robotic devices to mitigate falls in the future. To achieve this goal, it is imperative to gain a comprehensive understanding of the neuronal mechanisms triggered during falls or specific activities like walking or standing. Such insights can empower auxiliary robotic devices to assess the risk of a fall and initiate preventive actions to safeguard the user.

1.1 Motivation and Problem Statement

Falls are a leading cause of injury and the second leading cause of death related to unintentional injuries. Its psychological and social repercussions lead to a lower quality of life. The World Health Organization (WHO) reported that between 28% and 35% of people aged at least 65 years fall each year and the rate increases to 32% - 42% for people aged over 70 years [14]. Based on a retrospective review of the causes and conditions of significant falls, the U.S. Public Health Service concluded that two-thirds of fall-related deaths may be avoidable [65]. Obviously, there are costs associated with these statistics. In 2016, for example, the value was $36.4 billion and it was expected to increase throughout the years. For that reason, there is an urgent need to create prediction and detection technologies to mitigate this problem.

Conventional therapeutic modalities, encompassing activities such as physical exercise, medication

management, regular vision assessments, and home safety enhancements, have demonstrated considerable success in reducing and preventing falls. However, a significant segment of the population afflicted by severe illnesses or injuries remains incapable of engaging in routine ADLs or adhering to standard care protocols, as indicated by the findings of studies regarding Perturbation Evoked Potentials (PEPs) and Electroencephalogram (EEG) decoding [62]. In response to the imperative need to enhance the rehabilitation of individuals with locomotor impairments, a range of innovative robotic-assisted devices have been devised, including prosthetic limbs, electrical stimulation systems, and exoskeletons. Moreover, these technological advancements have found applications in fall prevention strategies [10].

Near-fall detection holds promise for the early identification of older individuals at risk of falling, offering the opportunity for proactive intervention before actual falls occur. Near falls encompass instances of momentary balance loss during trips, slips, and missteps, where swift corrective actions prevent a full fall. Notably, near falls occur more frequently than actual falls, serving as valuable precursors for assessing future fall risk [29]. To facilitate a natural and effective recovery, it is essential for patients to have full control over devices like exoskeletons, prosthetic limbs, orthotic aids, and robotic canes. Empowering patients to actively manage and utilise these devices enhances mobility, restores independence, and lowers the risk of falls, thereby contributing to a safer and more efficient rehabilitation process.

Exoskeletons find widespread application in rehabilitation, offering active support for sitting, standing, and walking, as well as potential utility in fall risk reduction and prevention. However, it is essential to acknowledge the significant risk of falls while using exoskeletons. Current exoskeletons approved by the U.S. Food and Drug Administration (FDA) incorporate various safety measures to prevent falls, yet the efficiency of these strategies remains unexplored. These devices are typically intended for use in conjunction with a trained partner. Further research is needed to assess the effectiveness of these safety mechanisms and enhance the safety and usability of exoskeletons in rehabilitation and fall prevention [62].

This dissertation aims to develop a strategy capable of deciphering potential falls, with a view to activating fall prevention mechanisms in robotic assistive devices in the future. Consequently, the primary objectives of this research are to delve into brain signals during slip-like perturbations - since slips represent the most prevalent type of falls — and explore ADLs recognition. To achieve these aims, two meticulous data collection protocols will be executed, involving diverse subjects, young and elderly individuals, walking in both controlled indoor settings and on a treadmill during simulated slip-like perturbations. Leveraging EEG data, advanced AI models will be employed to discern the subjects' activities and precisely identify instances of slip-like perturbations.

1.2 Goals and Research Questions

1. **Goal 1**: The primary objective of this dissertation is to conduct a comprehensive survey and interpretative analysis of the existing research related to slip-like perturbations while utilising EEG in the

context of walking and maintaining a stable stance. This investigation seeks to shed light on several key aspects, including: i) the purpose of the study; (ii) the specific Hardware (HW) equipment employed in the experiments; (iii) a detailed examination of the ADLs and perturbation scenarios used in the research; (iv) an exploration of the experimental protocols and methodologies employed; (v) a thorough assessment of the primary findings, inherent limitations, and persistent challenges encountered in these studies. Key Performance Indicators (KPIs): i) requirements, functional and technical specifications; ii) a submitted journal article on this topic.

2. **Goal 2**: Creation of a public dataset with multivariate labelled data from daily activities and slip-like perturbations. Beyond sensory data, it will be collected anthropometric and demographic data, as well as data from clinical scales. KPIs: (i) definition of an experimental protocol for data acquisition regarding slip-like perturbations; (ii) definition of an experimental protocol for data acquisition regarding ADLs; (iii) comprehensive free-living senior ADL and slip-like perturbations public database.

3. **Goal 3**: The objective is to develop an AI-based system for identifying slip-like perturbations. This can be achieved by designing and implementing robust AI-based models trained on relevant data, and optimising the models hyper-parameters. KPIs: (i) Artificial Neural Network (ANN) architectures applied to decode brain signals from slip-like perturbations; (ii) obtain accuracy above 73.5% [73], benchmark to scientific literature; (iii) identify EEG channels where PEPs can be found.

4. **Goal 4**: The goal is to create an AI-driven system for the recognition of various ADLs. This objective can be realised through the development and deployment of AI models trained on pertinent datasets, and the fine-tuning of model hyper-parameters to optimise their performance in classifying diverse ADLs. (i) ANN architectures applied to decode brain signal; (ii) obtain accuracy above 90% [72], benchmark to scientific literature.

The following Research Questions (RQs) are anticipated to be addressed by the dissertation's conclusion:

1. **RQ1**: How has EEG been utilised in existing research to capture and identify daily activities and slip-like perturbations during walking and stance, and what are the key insights gained from these studies? This RQ is related to Goal 1 and is answered in Chapters 2 and 3.

2. **RQ2**: To what extent can current AI models accurately detect slip-like perturbations in healthy subjects using EEG? And what influence does artifacts have on the model's performance? This RQ is related to Goals 2 and 3 and is answered in Chapter 5.

3. **RQ3**: Can the developed EEG recognition system effectively identify daily activities in elderly subjects, and to what extent does artifacts impact its performance? This RQ is related to Goals 2 and 4 and is answered in Chapter 6.

1.3 Contribution to knowledge

The following are the primary contributions that the work documented in this dissertation aims to make:

1. A thorough literature review of physical and virtual perturbations with the objective of identifying and understanding the perturbation methods applied, as well as the activity purpose, EEG data collection and processing, sensors and equipment used and the participants information. Additionally, the experimental procedures used, key findings, limitations, and upcoming challenges are described;
2. Optimised slip-like perturbation recognition AI model using EEG;
3. Enhanced ADL recognition AI model leveraging Brain Computer Interface (BCI)-Informed Deep Learning;
4. Comprehensive public dataset with multivariate labelled data from ADLs and slip-like perturbations. It will contain sensory, anthropometric and demographic data, as well as data from clinical scales.

1.4 MsC Thesis Outline

Chapter 2 serves as one of the main foundation for this thesis, providing a comprehensive review of existing experimental protocols involving physical and/or virtual perturbations while utilising EEG. This chapter not only summarises the literature but critically examines the methodologies employed to induce perturbations, the sensor technologies incorporated, and the specifics of the EEG systems used. By doing so, it lays the groundwork for understanding the state-of-the-art techniques and technologies in the field.

In Chapter 3, our focus pivots toward the exploration of ADLs. The central objective of this chapter is to undertake a comprehensive literature review, shedding light on the intricate process of decoding various modes of locomotion such as walking and stair climbing/descending. The primary aim is to harness this knowledge to devise a solution for mitigating imbalance events within the context of human gait, leveraging the power of EEG. Through an examination of the existing literature, our intention is to glean insights into the diverse methodologies employed for understanding and quantifying ADLs.

In Chapter 4, we dive into the pressing necessity to tackle the challenges associated with falls among individuals. In response to this imperative, our thesis project has forged a solution that meticulously addresses these challenges. Within this chapter, we elucidate each phase of our project, emphasising the critical aspects, with particular focus on the data collection process. This phase is pivotal as it encompasses the design and execution of experimental protocols tailored to capture EEG data during slip-like perturbations and the performance of ADLs. The thorough explanation of this data collection process is instrumental in comprehending how we acquire the crucial information needed for our research and lays the groundwork for subsequent analyses and insights into fall prevention.

In Chapter 5, it is described the utilisation of advanced AI-based models with the primary goal of detecting perturbations within EEG trials. Leveraging the datasets gathered from slip-like perturbations, our focus turns to training AI-based models. These models are developed and fine-tuned to classify the presence of any perturbations within the brain signal that might signify a balance impairment during the walking process. To assess the efficacy of these AI-based models, this chapter also includes a comprehensive data analysis. This analysis entails the evaluation of model performance using a multitude of metrics, shedding light on the models' accuracy, precision, recall, and other relevant indicators.

Chapter 6 extends our exploration of deep learning models, this time applied to ADLs. We focus on evaluating these models' performance in decoding EEG signals during various ADLs. In this chapter, we also detail the organisation of our datasets, ensuring they align with specific ADL scenarios, and describe our data labelling methods.

Finally, Chapter 7 represents a critical juncture in our thesis journey. It synthesises the conclusions drawn from the work accomplished up to the point of writing this document. It provides a reflection on the key findings, highlights the contributions made, and discusses the implications of our research. This chapter also sets the stage for future work and explores potential avenues for further research and development in the field of EEG-based perturbation studies and daily activity monitoring.

Chapter 2

Review: Decoding Brain Signals for Fall Prevention

2.1 Introduction

EEG has been one of the most used techniques to measure brain electrical activity, since its data enables the detection of motor planning and intention. Non-invasive EEG is a powerful technique that allows medical professionals and researchers to measure brain activity with high temporal resolution. By placing non-invasive electrodes on the scalp, EEG is able to capture the electrical signals that represent the dynamics of voltage fluctuations in the brain, providing unparalleled insight into the workings of the human brain. Through EEG, medical professionals and researchers can gain further insight into a variety of neurological diseases and disorders, as well as gain a better understanding of normal brain activity [64]. Not only does it offer a secure and portable technology, but it also allows early detection so that robotic devices can prevent a fall from happening [60].

Some advances in signal processing have allowed source-resolved EEG dynamics during walking. Wittenberg et al. [86] have shown evidence of an increase in brain activation during the performance of balance control tasks by healthy adults, regardless of mechanical, cognitive, or sensory challenges they experienced. Also, Jiang et al. [33] attempted to establish a generic model for cross-task fall detection based on neural response features of ipsilateral postural perturbation events, which can also detect contralateral perturbation events. In order to provide additional insights into the mechanisms involving balance control, others sensors combine with EEG. A large amount of studies, such as Annese et al. [9] which implemented a real-time algorithm for falling risk prediction, use Electromyography (EMG) combined with EEG, in order to collect both muscular and brain activity.

The detection of falls has been developed with the use of wearable sensors. Pang et al. [51] intended to summarise and analyse the evidence regarding this detection, not only slip-like perturbations but also

trips, stumbles, missteps, incorrect weight transfer, or temporary loss of balance using wearable devices. The authors concluded that the use of a single lightweight sensor that has the capacity to distinguish not only different types of falls but also ADLs is a promising low-cost technology and also clinical tool for long-term continuous monitoring of older people and clinical populations at risk of falls.

Despite the numerous benefits of their studies, there are some limitations to consider. Primarily, their studies are conducted with young adults in a laboratory setting and do not accurately reflect the complexities of falling among people who are at high risk, such as the elderly. By not accounting for the distinct characteristics of the elderly, they are unable to assess the true effectiveness of their research in this vulnerable population. In fact, only one study aimed to investigate the detection of falls in high-risk population [27]. Additionally, it was challenging to evaluate the reported performance variances between the various methods. It is still unclear if wearable devices can accurately detect near falls in older adults or other high-risk populations while at home or other community settings.

Currently, a variety of sensors are used to detect or forecast falls from ADLs. One innovative trend to increase system performance is the sensor-fusion method, which merges numerous data streams from the relevant sensors. Ren et al. [63] presented a comprehensive review among the latest studies on fall-related technologies. Although they came to the conclusion that sensor-fusion method is a novel trend that can improve the performance of the system, the authors also agreed that by integrating many data sources, there is a raise of the computing complexity and system cost, as well as redundant information.

The lack of real-world fall data is a major challenge in fall-related systems. This is due to a variety of issues, including costs, information redundancy, the lack of a general evaluation framework, and the high rate of false alarms. Most studies only collect data on their own simulated activities in unique environments, making it difficult to replicate the results. Furthermore, the subjects in these studies typically have varying features, which further complicates the replication of results. As a result, the lack of real-world fall data is a significant issue that must be addressed in order to improve the efficacy of fall-related systems. From the analytical algorithm point of view, they concluded that threshold-based method is a classical and more basic approach when detecting falls, while machine learning-based method has been widely researched to improve accuracy of the system.

Siragy et al. [71] assessed on the reactive recovery responses between over-ground and split-belt treadmill trips and slips as well as the effect of aging on these responses, since there is a variety of studies that have examined recovery responses from falls during walking, specifically trips and slips. In order to simulate real world falls/perturbations, those studies included treadmills and/or over-ground methods in their experiments. The authors in [71] found that, although perturbations on the treadmill and over-ground cause the center of mass to become unstable, the recovery response to these perturbations was different on the treadmill. Particularly, treadmill trips and slips resulted in more predictable recovery responses. Moreover, treadmills may be limited in their ability to study the diversity of aging effects on perturbation recovery responses because older people have trouble scaling perturbation recovery. One of

the main limitations pointed out was that the falls induced within the studies were manipulated during the respective protocol without a follow-up assessment for prospective falls. Thus, the authors found unclear how accurately experimentally created perturbations mimic the fall mechanisms that occur in real-world settings.

Ultimately, fall detection is crucial to create assistive robotic devices such as prostheses or exoskeletons. Studies [48, 80] showed that there is potential of building controlled wearable devices to assist elderly or disabled subjects in order to improve their quality of life. The authors in [48], for instance, built a device with a real-time detection system of slip-like perturbations who counteracted torques at the hips to assist balance recovery.

The main goal of this study is to review the literature to understand what kind of perturbations are performed using EEG, how they are performed and how the EEG data is processed. The state-of-the-art is examined in this chapter, taking into consideration a review process and eligibility standards created prior to the research and motivated by the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) flowchart, as a method of addressing the goals present in Section 1.2. Thus, all articles that have been identified in the literature are analysed, with the following topics being addressed: (i) "What type of participants were included in the studies?"; (ii) "What are the main sensor systems used to collect data during perturbation-based protocols?"; (iii) "What are the characteristics of the EEG system during data collection (number of channels, sampling frequency)?"; (iv) "What methods are used to perform perturbations?"; (v) "Which methods are used to process EEG data?" and (vi) "Is it common the use of ecologically valid virtual environments to produce visual perturbations?".

2.2 Review Strategy

A systematic search was accomplished in Scopus, IEEE, WebOfScience and PubMed databases. This search was carried out until the October 24th of 2022 using the set of keywords: "provoked falls"AND "eeg"OR "erp"AND "classification"AND "perturbation"; ("balance"OR "posture") AND "eeg"AND "perturbation"AND "falls"; "eeg"AND ("fall"OR "balance"AND "loss"); "eeg"AND ("fall"OR "balance"OR "posture") AND "perturbation"). The keywords used were based on previous systematic reviews.

A total of 1281 articles were collected from the mentioned databases and 1041 remained after duplicates removal. The remaining articles were then screened based on their title and abstract, which resulted in the removal of 792 articles. Following that, the remaining articles were screened using the following inclusion criteria: i) perturbations were performed exclusively by an external factor; ii) the paper was not a review or journal; iii) the purpose of the study did not look to address any problem related to a particular disease; iv) the paper mentioned the use of EEG in its experimental protocol. The reason for the removal of reviews and journals is related to the purpose of this review, since it is about finding studies that describes experimental protocols for external perturbations. The age, like the equipment used, was not used as an

inclusion criteria in order to achieve a wider set of methods used to perform perturbations. In some cases, it was not possible to ascertain if the papers fulfilled the eligibility criteria mentioned previously only by reading the title and abstract. For that reason, those papers were carefully read and analysed so that they could be included or not. With that, a group of 24 articles was then obtained through the filtering procedure. Finally, the references of each remaining article were read in order to verify if there was any new paper that could also be included. The inclusion process for the references was identical to the one mentioned before. With that, a final group of 29 articles was obtained. Figure 1 presents the PRISMA flowchart which illustrates the literature search previously described.

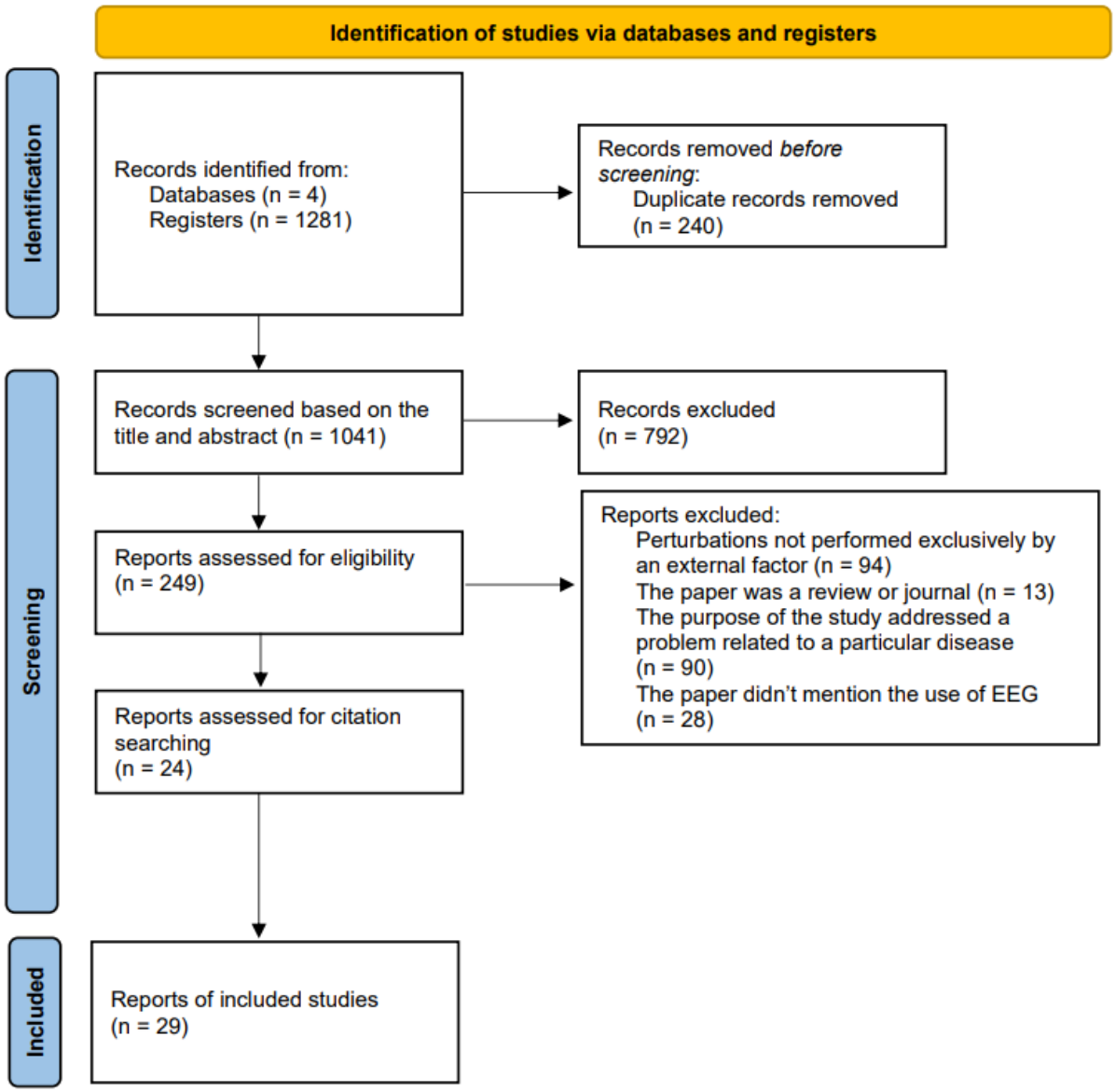


Figure 1: PRISMA Flowchart - Brain Signals for Fall Prevention.

Table 1: Physical and virtual perturbations in the group of 29 articles

Perturbation Condition (No. of studies)	Type of perturbation	No. of studies	Articles	
Physical (25)	Stance	19	[4, 5, 11, 15, 16, 23, 31, 33, 42, 43, 45, 46, 53, 54, 58, 68, 73, 74, 82]	
	Walking	6	Over-ground	[13]
			Treadmill	[40, 41, 56, 70, 84]
Virtual (5)	Stance	3	[55, 75, 85]	
	Walking	3	Treadmill [55, 56, 57]	

2.3 Reviewed Articles

From the 29 included studies, physical perturbations (25 studies) were more prevalent than virtual perturbations (5 studies). In cases such as [55], it was considered both a stance and walking perturbations study. In addition, these two categories were divided into stance and walking perturbations, in order to provide a more complete information on the different types of perturbations that can occur. Also, the walking perturbations were divided again into over-ground and treadmill, depending on the conditions that each study implemented. A separate analysis was conducted for both physical and virtual perturbations since they have different approaches on how to perform perturbations. To demonstrate this strategy, Table 1 depicts the articles collected in the literature that deliver physical and virtual perturbations using EEG.

2.3.1 Physical Perturbations

2.3.1.1 Stance Perturbations

Purpose of the studies

Some studies preferred to analyse human balance control, either for identifying cortical characteristics during certain activities, proving that the brain engages more cortical activity when the balance is challenged, or both purposes [15, 23, 33, 42, 43, 46, 58]. Mierau et al. [42, 43], for instance, analysed what cortical regions became activated during a wide range of repetitive balance tasks (such as balancing on a stable surface) differing in difficulty by changing the base of support, surface stability, or both. Before heading into the experiment, the authors hypothesised that cortical networks are a massive contribution to an optimisation of balance control. For that reason, they chose to work with EEG. On the other hand, Payne et al. [53] aimed to predict that people with lower balance ability would exhibit larger cortical N1 responses during balance perturbations.

Table 2: Physical stance perturbations articles

Study	**Purpose**	**Participants (Number/Age)**	**EEG Characteristics and Sensors Location**	**Perturbation Method**	**Equiment**	**Performance**
Mierau et al. [42]	Balance control analysis	(39/24.7 avg)	EEG (32 channels, 1000 Hz)	Anchor release	Anchor, surface	N/A
Saadat et al. [68]	Examine external perturbations effects	(19 young and 20 older/24.25 and 65.55)	EEG (32 channels, 500 Hz)	Force belt	Force belt, load, cable	N/A
Jiang et al. [33]	Balance control analysis	(15/23-28)	EEG (60 channels, 1000 Hz), Handle force (hands)	Handle force	Handle force, air bag	Maximum Cross-Task Detection Rate - 94.2%; Latency - 314.4 ms
Adkin et al. [4]	Investigate the effects of postural threat on the cortical response	(10/22-28)	EEG (500 Hz), EMG (legs)	Haulotte, Crane	Haulotte, crane	N/A
Solis-Escalante et al. [73]	Present a new perspective on the cortical representation of postural stability	(9/26.3 avg)	EEG (126 electrodes, 2048 Hz), Radboud Falls Simulator (feet)	Movable platform	Safety harness equipment, ankle brace, platform	Accuracy - 73.5%

Ravindran et al. [74]	Emphasise the importance of of EEG for fall prevention	(7/18-32)	EEG (64 channels, 250 Hz), EMG, Neurocom (ear lobes and legs)	Balance platform	Exoskeleton, platform	F-score - 75.0 ± 4.3%
Varghese et al. [82]	Examine external perturbations influence	(14/19-31)	EEG (64 channels, 1000 Hz), EMG (legs)	Force plate	Platform, cable	N/A
Mierau et al. [43]	Balance control analysis	(37/24.7 avg)	EEG (32 channels, 1000 Hz), EOG and EMG (eyes, legs)	Platform	Platform, anchor, steel cables, gaze fixation cross	N/A
Goel et al. [23]	Balance control analysis	(10/25.3 avg)	EEG (64 channels, 1000 Hz), CDP (feet)	Force plate	CDP force platform	N/A
Mochizuki et al. [46]	Balance control analysis	(8/22-44)	EEG (64 channels, 1000 Hz), EMG (trunk, TA, AD and biceps)	Force plate	SynAmps2 amplifier, belt, force plate	N/A
Ditz et al. [16]	Classify PEPs from EEG	(15/26.7 avg)	EEG (29 channels, 500 Hz), EOG (eyes)	Custom-built tilting chair	Custom-built tilting chair, amplifier	Accuracy - 87.6 ± 8.0%
Quant et al. [58]	Balance control analysis	(8/18-42)	EEG (500 Hz), EMG (legs)	Inverted pendulum	Foot platform, inverted pendulum	N/A

Payne et al. [54]	Test cognitive set shifting ability expressivity in balance recovery	(19/71 avg)	EEG (32 electrodes, 1000 Hz), EOG and EMG (eyes, legs)	Support surface	Support surface and Vicon Nexus	N/A
Dietz et al. [15]	Balance control analysis	(10/N/A)	EEG, EOG and EMG (eyes and legs)	Treadmill	Treadmil, computer system (Sirius)	N/A
Bogost et al. [11]	Emphasise the importance of cortical areas in balance control	(15/20.3 avg)	EEG (256 electrodes, 1000 Hz)	Surface platform	Surface platform	N/A
Jacobs et al. [31]	Examine external perturbations influence	(12/27 avg)	EEG, EOG and EMG (TA, MG, RA, ES)	Force plate	Support surface	N/A
Adkin et al. [5]	Examine external perturbations effects	(8/17-32)	EEG (500 Hz), EMG (AP, MG)	Padded customized device	Padded customized device	N/A

Mochizuki et al. [45]	Examine cortical activity that precedes the onset of perturbations	(23/29.1 for the 1st and 26.1 for the 2nd experiment)	EEG (64 channels, 1000 Hz), EMG (trunk and legs)	Force plate	Cable, load, platform, belt	N/A
Payne et al. [53]	Influence of balance on cortical activity	(20/26)	EEG (32 electrodes, 1000 Hz), EOG (eyes)	Support surface	Support surface, camera	N/A

This hypothesised that challenging balance recovery can interfere with cortical activity. The effects of postural set on the cortical response evoked by, specifically, an external perturbation was also mentioned in some studies [5, 31, 68, 82]. Each study included a single method or a group of methods to perform external perturbations, such as forward trunk perturbations or load releasing. A few studies also performed unpredictable external perturbations, in order to understand the relation between predictability and cortical activity. Adkin et al. [5], for example, found that unpredictable perturbations evoked a large negative potential, which was not discernible for perturbations with predictable onset timing and direction in spite of the presence of significant compensatory balance reactions. This conclusion was also found in Payne et al. [53], whose purpose was to prove that people engage more cortical activity when balance is challenged, both in predictable and unpredictable perturbations. In addition to the predictability of perturbations, the authors asserted that larger cortical amplitudes are also related to greater difficulty recovering balance, either due to more difficult perturbations or to lower intrinsic balance ability.

The importance of EEG to decode neural activity was addressed in [74]. In particular, the authors claimed that EEG data contain short-latency brain information about an incipient fall, which can contribute to build BCI systems for fall prevention.

Participants Information

In the broader context, it is worth noting that the majority of participants involved in various experimental protocols were selected due to their youth and overall healthy physical condition. This criterion encompassed their body weight, ensuring that the subjects were in a good physical shape. However, it is important to highlight an outlier in this trend, as exemplified in the study conducted by Solis-Escalante et al. [73]. This particular research project featured a relatively small group of nine subjects. Out of these, six were individuals in robust health, aligning with the typical criteria established across the experimental

protocols. Notably, the remaining three participants in this study presented a contrasting profile, as they were individuals dealing with chronic hemiparetic stroke.

Sensors

As mentioned previously, every study used EEG as one of its main sensors. Additionally, EMG was also present in studies [5, 15, 31, 43, 45, 46, 54, 58, 74, 82], since it is possible to combine muscular with brain activity, which provides more complete data. Some studies like [5, 15, 16, 31, 43, 53, 54] also added Electrooculography (EOG) to study eye movements.

Perturbation Method

The study [42] aimed to identify cortical regions that become activated during a wide range of continuous standing balance tasks differing in difficulty by changing the base of support, surface stability, or both. Mierau et al. [43], on the other hand, wanted to show that transient balance perturbations are associated with characteristic changes in electroencephalographic. Both studies, for the experimental protocol, tried to provoke a perturbation by releasing an anchor that would unbalance a platform that the subjects were balancing on. Thus, subjects had to position themselves in order to stabilise their posture, while their dominant leg was fixed in the center of the platform.

Saadat et al. [68] experimented a protocol that had predictable and unpredictable perturbations, in order to investigate the effects of predictable and unpredictable external perturbations on cortical activity in healthy young and older adults. The perturbation consisted of releasing a load that was attached via a cable to a belt worn by the subjects at their sternum level. The subjects stood barefoot with their feet 24 cm apart and were asked to maintain their postural balance after the load was released on the belt. Varghese et al. [82] and [46] used a similar lean and release cable system to perform a perturbation.

Jiang et al. [33] simulated two types of falling-risk events on the subjects without movement intention which caused by unpredictably random erroneous motion execution of the external equipment. They tested the neural and behavioural responses of those rapidly unpredictable postural perturbation. The experiment equipment consisted of, mainly, an inflatable airbag, two handles that moved up and down in response to the airbag's inflation and deflation, and a force sensor attached to both the handle and the airbag. The airbag was fully inflated during the resting condition, and the subjects maintained a stable stance while leaning forward by around 20 degrees, shifting their center of gravity in the direction of the handles. When the system was unstable, one of the two handles would quickly descend, mimicking falling as a result of the airbag deflating (which takes maximum to 300 ms). The support force of one of the upper limbs suddenly diminished, causing the subject to lose balance.

Adkin et al. [4] investigated the effects of postural threat on the cortical response associated with postural reactions to predictable and unpredictable perturbations to upright stance. When standing with their feet shoulder-width apart and their eyes closed during the experimental protocol, the subjects encountered

a series of predictable and unpredictable trunk perturbations both at ground level and at the edge of a platform that was 3.2 meters above the ground. The perturbations were delivered to the trunk by the experimenter using a padded hand-held bar. This device was instrumented with a load transducer to provide online monitoring of perturbation magnitudes. With that, they discovered that unpredictable perturbations generated a large negative potential (N1) which was increased by 84% when standing at the high compared to low surface height. The magnitude of change in this potential was related to the magnitude of change in balance perceptions, such as confidence and fear. According to the authors, predictable perturbations resulted in a potential associated with anticipation before the perturbation. Threat from postural position had no effect on the brain activity that was seen in response to predicted perturbations.

As described previously, some studies included platforms as one of the main equipment to deliver perturbations [11, 18, 31, 53, 54, 73, 74]. Solis-Escalante et al. [73], for instance, studied cortical dynamics involved in top-down regulation of postural responses carry information about directional postural changes imposed by sudden perturbations to standing balance. In this particular case, the sudden perturbation consisted on a support surface horizontal translation, in which the subjects were perturbed with movements of the support surface in four different directions. The support surface translations caused the body to sway in the opposite direction of the surface movement. Ravindran et al. [74] studied a similar protocol, but the platform generated forward, backward or tilted perturbations in a random order.

Dietz et al. [15] used a treadmill as the main equipment to perform a perturbation. The acceleration unit of the treadmill allowed rapid changes of its speed. This applied, for an adult subject, a step change of treadmill speed from either 0 to 6 km/h or from 3 to 9 km/h, over a 70 ms period during gait. Ditz et al. [16] innovated by using a custom-built chair that could tilt either to the left or right (random generator). The perturbation onset was randomly set within the first 4 seconds of each trial for each subject.

EEG Data Processing

Saadat et al. [68] opted to process the EEG data by converting it into ASCII format and importing into a software called NeuroGuide. The recorded signals were pre-processed by an EEG expert and denoised to eliminate signals disturbed by eye movements or muscle artifacts. The difference with z-score absolute power was computed by fast Fourier transform based on the NeuroGuide qEEG normative database. The exact moment of load release (T0 = 0) was marked on the EEG tracing using a foot pedal triggered by the examiner. The T0 time point was considered the reference to which all EEG signals were compared and analysed. Data were analysed in three time intervals; the Anticipatory Postural Adjustment (APA) from −1000 to −500 ms (APA1), the APA from −500 to 0 ms (APA2), and the early Compensatory Postural Adjustment (CPA) from 0 to +500 ms (CPA1). The acquired EEG signals upon each set of 15 trials were averaged, and the mean absolute power and coherence were calculated for alpha (8–12 Hz) and beta (12.5–25 Hz) frequency bands.

For motion artifact caused by the movement of the electrode during the experiment, Jiang et al. [33]

used a bandage to fix the EEG cap, and the 0.1–10 Hz bandpass filter removed the EMG artifacts in EEG (main frequency is 500–2000 Hz) and partly unnecessary EEG components of the most sensorimotor rhythm in order to avoid the interference of active motion intention by Event-Related Synchronisation (ERS) or Event-Related Desynchronisation (ERD) phenomenon and Slow Cortical Potential (SCP). The raw EEG signals were bandpass filtered by Butterworth filter of 0.1-10 Hz and downsampled to 50 Hz (5 times of the effective frequency band, preserving the original characteristics of the signal while improving the processing speed). The continuous EEG was segmented into epochs from t = −2 to 4 s relative to the perturbation event onset.

Solis-Escalante et al. [73] also processed EEG data with a bandpass filter between 2 and 200 Hz (4th order Butterworth Infinite Impulse Response (IIR) filter, zerophase shift), and then downsampled to 512 Hz. Highly contaminated channels were identified by visual inspection and removed from the recordings. On average, 126 channels remained for analysis. The remaining channels were re-referenced to the common average. To improve the signal-to-noise ratio of the cortical activity contained in the EEG, artifact correction was performed in four stages. The Cleanline plugin in EEGLAB was used in the initial step to eliminate line noise (50 Hz) and its harmonics (100 Hz and 150 Hz), which employs frequency-domain regression to estimate and decrease sinusoidal distortions. In the second stage, non-stationary high-amplitude artifacts, such as muscle artifacts, eye movements, and electrode movements, were reduced with the clean_rawdata plugin also in EEGLAB, which implements the Artifacts Subspace Reconstruction (ASR) method. In the third stage, the EEG was segmented into epochs from −2 to 3 s relative to perturbation onset.

Goel et al. [23] downsampled raw EEG signals to 100 Hz and then also high-pass filtered at 0.1 Hz using a 4th order Butterworth filter. To apply robust common referencing methods and remove artifactual EEG channels, a standardized EEG Processing Pipeline (PREP) was used. ASR was applied to reconstruct any artifactual section in the EEG data. Next, to compensate for the rank deficiency in the data, interpolated channels during the PREP pipeline were removed before running the Independent Component Analysis (ICA) using adaptive mixture ICA. The ICA was performed on EEG data from all 18 trials for each subject.

Ditz et al. [16] considered three different statistical parameters for the rejection of artifact-tainted trials. First, the authors performed an amplitude threshold rejection removing all trials with an amplitude that exceeded ±125 μV. Afterwards, they tested trials for an abnormal joint probability and an abnormal kurtosis. The rejection threshold was four times the standard deviation for both tests.

2.3.1.2 Walking Perturbations

Purpose of the studies

A group of studies, as mentioned in 2.3.1.1, had as a purpose to analyse and identify cortical regions associated with loss of balance, whether the perturbations are expected or not [41, 56, 70]. Mezzina et al. [41] investigated changes in the cortical involvement when their subjects were managing slippages. The authors developed a strategy for cortical dynamics information extraction to separate features linked to reactive balance recovery from those associated to steady walking. The study [56] aimed to analyse the varying responses of cortical and muscle connection patterns to temporary sensory disruptions affecting balance during both walking and standing..

Finally, a few studies proposed novel architectures for pre-impact fall detection [40, 84]. In [84], for instance, the architecture consists on a multi-sensor system that simultaneously examines cortical and muscle involvement when sudden slippages happen while walking steadily. Alternatively, [13] proposed and validated a system for the reliable recognition of losses of balance situations, addressing the high specificity and system robustness against the ADLs.

Table 3: Physical walking perturbations articles

Study	**Purpose**	**Participants (Number/Age)**	**EEG Characteristics and Sensors Location**	**Perturbation Method**	**Equiment**	**Performance**
Mezzina et al. [41]	Balance control analysis	(4/26.3 avg)	EEG (32 channels, 500 Hz), EMG (AT, LG, VM, RF, BF and LL)	Platform acceleration	SENLY	N/A
De Venuto et al. [84]	Architecture for pre-impact fall detection	(6/N/A)	EEG (32 channels, 500 Hz), EMG (trunk, legs)	Platform acceleration	SENLY	Accuracy - 96.21%
De Venuto et al. [13]	Architecture for fall detection	(9/26.3 avg)	EEG (32 channels, 500 Hz), EMG (legs)	Platform acceleration	SENLY	Accuracy - 98.91%; Latency - 370.62 ± 60.85 ms

Mezzina et al. [40]	Architecture for pre-impact fall detection	(6/26.3 avg)	EEG (32 channels, 500 Hz), EMG, 3D Kinematics (lower limbs, feet)	Platform acceleration	SENLY	Accuracy - 93.33%
Peterson et al. [56]	Balance control analysis	(30/22.5 avg)	EEG (136 channels, 512 Hz), EMG (legs, waist)	Treadmill-mounted balance beam	Treadmill, VR headset, body-support harness	N/A
Sipp et al. [70]	Balance control analysis	(26/23 avg)	EEG (256 channels, 512 Hz), EMG (lower-limb)	Balance beam	Treadmill-mounted balance beam, orthopaedic shoes	N/A

Participants Information

All included subjects were in a healthy condition, demonstrating no significant medical concerns or pre-existing conditions that could potentially confound the results of the study. The stringent health screening processes ensured the reliability and validity of the data collected throughout the study.

Sensors

Once again, EEG is used in all studies. Most studies place the EMG sensors on the legs, since subjects walk in the respective experimental setups. Other than that, there does not seem to be any particular difference to stance studies in terms of the sensors used.

Perturbation Method

The majority of studies included force belts to induce a physical perturbation while subjects are walking. Studies [13, 40, 41, 84] used SENLY, which consists of a mechatronic platform that aims to perform slip-like perturbations in steady conditions, like walking or stance. Peterson et al. [56] used a treadmill since it is more practical for the subjects to walk, even though the authors also included a Virtual Reality (VR) headset as one of the equipment. The reason why this study was also included in the physical perturbations cluster is because there were physical perturbations induced. In particular, subjects were pulled at the waist by a cable while wearing the VR headset.

EEG Data Processing

For Mezzina et al. [41], the impacts of noise and artifacts were first removed from the raw data using pre-processing. In particular, low-pass filtering data (zero-lag, 4th order Butterworth low-pass filter) with a cutoff of 10 Hz reduced high-frequency associated noise from digitised coordinates. Raw ground reaction force signals were also bandpass-filtered (Butterworth filter, 4th order, cutoff at 0.5–10 Hz). The EEG data were progressively band-filtered between 1 Hz and 40 Hz by using an 8th order Butterworth filter before the transmission. The onset of the Vastus Medialis (VM), namely Muscle Trigger (MT), was then used to enable the analysis of EEG signals for each gait cycle (both right and left sides). The EEG post-processing was activated once the rising edge of the VM (both sides) MT was discovered. This subset of EEG data underwent an online Riemannian Artifact Subspace Reconstruction (rASR). Mezzina et al. [40] used a bandpass frequency span of 1 Hz - 40 Hz through a 8th order Butterworth online filter.

Peterson et al. [56] decided to, firstly, downsample EEG data to 256 Hz and then apply a 1 Hz high pass filter. Then, they applied a cleanline 60 Hz noise reduction, removed bad channels and used ASR. After that, they used a selective low-pass filter, following with bad channel interpolation and dipole fitting. In particular, they discarded dipoles with residual variance $> 15\%$. Manually, they retained cortical dipoles and grouped dipole clustering, retaining clusters with more than 15 subjects.

2.3.2 Virtual Perturbations

2.3.2.1 Stance Perturbations

Purpose of the studies

Sun et al. [75] aimed to analyse the interaction between postural control and anxiety during balance-demanding height and perturbation shifts. The authors also applied machine learning methods to validate the effectiveness of using EEG-based measures of anxiety in order to classify between subjects with low or high balance function. In [55], the main goal was to determine the possibility to identify common cortical responses to visual and physical balance perturbations. Finally, [85] presented a system that enables the simultaneous recording of behavioural, physiological, and neurological data in an immersive virtual reality environment.

Participants Information

All studies were conducted with young subjects who were in good physical and mental conditions. These participants were carefully selected to represent a vibrant and healthy demographic, ensuring that the research findings are applicable to this specific age group.

Table 4: Virtual stance perturbations articles

Study	**Purpose**	**Participants (Number/Age)**	**EEG Characteristics and Sensors Location**	**Perturbation Method**	**Equiment**	**Performance**
Sun et al. [75]	Examine interaction between postural control, neurological feedback and anxiety	(6/21 avg)	EEG (64 channels), EMG, EKG (chest, legs, feet)	Virtual platform, HTC Vive	Safety harness, force plate	N/A
Peterson et al. [55]	Examine cortical responses to visual and physical perturbations using EEG	(30/22.5)	EEG (136 channels, 512 Hz), tensile load cells, motion capture camera, EMG (neck, head, feet and sacrum)	Webcam rotation, motors placed on treadmill	Safety harness, treadmill	N/A
Widdowson et al. [85]	Architecture to collect neural, physiological, and behavioural data in VR environment	N/A	EEG (64 channels), optoelectronic sensors, EOG, HMD, EMG, inertial systems (TA, MG)	Virtual platform, stepper-motor driven system	Safety harness	N/A

Sensors

All studies used both EEG and EMG. Sun et al. [75] decided to add an electrocardiogram, since its purpose was to analyse the interaction between balance function and anxiety during balance-demanding height and perturbation shifts, and this sensor can help measuring the rhythm of the heart beats. Peterson et al. [55] used tensile load cells to measure the forces applied by the subjects. Widdowson et al. [85] decided to use inertial sensors, EOG and EMG to monitor changes in movement.

Perturbation Method

Every study included both virtual and physical perturbations. Studies [75, 85], for instance, used a virtual platform that went up or down. For the physical perturbation, [75] applied a 10-degree toe-down perturbation, while [85] used sliding and tilting force plates. Peterson et al. [55] performed virtual perturbations by rotating the virtual field. As for the physical perturbation, the authors used a steel cable attached to an electromechanical motor (with two motors placed on each side of a treadmill). A steel cable that was attached to the subject's safety harness was pulled when the motor was instructed to spin an attached bar 90 degrees away from the person.

Virtual Environment

Sun et al. [75] created a virtual platform went up to 2.5 m, 5 m and 7.5 m, or the pit around platform can drop to 2.5 m, 5 m and 7.5 m. In [55], there is no description of the virtual environment itself, but it is described that subjects could walk and jumped around the environment using body gestures. As for Widdowson et al. [85], the VR environment consisted of a large room with two platforms: A and B. Platform A raised (and lowered) from 0 to 8 m above ground, while platform B consisted of the surrounding floor and lowered (and raised) from 0 to 8 m below ground, via operator's defined inputs.

EEG Data Processing

In Sun et al. [75], signals epoched and filtered via a bandpass Finite Impulse Response (FIR) filter (1-45Hz), ICA to reject muscular, cardiac, eye or other equipment related artifacts - Independent Component (IC) scalp topography heat map, Event-related Potential (ERP) and power dispersion plot. Peterson et al. [55] processed all EEG signals using custom EEGLAB scripts. EEG data were downsampled to 256 Hz, 1 Hz high-pass filtered, merged across all conditions, and referenced to the common median of all channels. They reduced 60 Hz line noise using Cleanline. Afterwards, rejected bad channels that had high standard deviations or were uncorrelated for $> 1\%$ of the time. Finally, they also performed selective low-pass filtering using ensemble empirical mode decomposition and canonical correlation analysis.

2.3.2.2 Walking Perturbations

Purpose of the studies

Peterson et al. [57] aimed to ascertain whether brief visual perturbations induced by immersive VR affect subjects' brain activity and behavioural responses as they practice a novel balancing task. In [55], the main goal was to determine the possibility to identify common cortical responses to visual and physical balance perturbations, whether by using EEG and ICA. Finally, [56] aimed to compare how different cortical and muscle connection patterns responded to transient sensory perturbations that interfered with postural control when walking and standing.

Table 5: Virtual walking perturbations articles

Study	**Purpose**	**Participants (Number/Age)**	**EEG Characteristics and Sensors Location**	**Perturbation Method**	**Equiment**	**Performance**
Peterson et al. [57]	Examine transient visual perturbations influence cortical and behavioural activity in an immersive VR environment	(30/22.6 avg)	EEG (136 channels, 512 Hz), Microsoft Kinect, HR monitor, eye movements and mastoids electrodes, EMG, skin conductance (scalp, feet, sacrum, sternum, neck, head, eyes)	Webcam rotation, balance beam	Body support harness for safety, with extended support straps to increase unrestricted mediolateral movement, treadmill-mounted balance beam	N/A

Peterson et al. [55]	Examine cortical responses to visual and physical perturbations using EEG	(30/22.5)	EEG (136 channels, 512 Hz), tensile load cells, motion capture camera, EMG (neck, head, feet and sacrum)	Webcam rotation, motors placed on treadmill	Safety harness, treadmill	N/A
Peterson et al. [56]	Balance control analysis	(30/22.5 avg)	EEG (136 channels, 512 Hz), EMG (legs, waist)	Treadmill-mounted balance beam	Treadmill, VR headset, body-support harness	N/A

Participants Information

All studies involved the inclusion of young subjects who were in optimal health and well-being. The research consistently emphasised the participation of individuals within the youthful age demographic, ensuring that the findings would be directly applicable to this specific group.

Sensors

Once again, both articles employ the simultaneous use of EEG and EMG to gather data. Additionally, for the purpose of motion detection, [57] opted for the utilisation of the Microsoft Kinect system. This approach allows for a comprehensive and multifaceted analysis, combining neural activity data from EEG and muscular responses from EMG with motion tracking information provided by Microsoft Kinect.

Perturbation Method

Peterson et al. [57] performed a virtual perturbation by rotating its virtual point of view, adding 1 second of delay to make it more difficult to anticipate. As for the physical perturbation, the authors used a balance beam. Also, study [55] performed virtual perturbations by rotating the virtual field. As for the physical perturbation, the authors used a steel cable attached to an electromechanical motor (with two motors placed on each side of a treadmill). When the motor was commanded, it rotated an attached bar 90° away from the subject, which pulled on a steel cable connected to the subject's safety harness.

The authors in [56] performed a 20-degree field-of-view rotation using a VR headset with an attached

webcam. The webcam view was displayed on the VR headset, providing subjects a pass-through VR experience. The authors performed perturbations by digitally rotating the VR view 20 clockwise or counterclockwise. This rotation occurred immediately from 0 to 20, lasting for a half second before the VR view returned to its starting position. They also performed a physical perturbation described in 2.3.2.2.

Virtual Environment

Studies [56, 57] conspicuously lacked any detailed description of the virtual environment employed in their respective research. However, it is noteworthy that in [55], the authors did provide an insightful portrayal of the virtual environment. In this specific study, participants were granted the freedom to navigate and interact with the digital landscape through body gestures, enabling them to walk and engage in dynamic movements within the virtual space.

EEG Data Processing

Peterson et al. [57] processed EEG data using custom scripts in EEGLAB. Firstly, they downsampled the data to 256 Hz, highpass filtered the EEG data at 1 Hz, and merged all conditions, referencing all channels to a common median reference to reduce the impact of potential outlier channels. Used Cleanline to reduce 60 Hz electrical noise. Then, rejected channels that had an abnormally high standard deviation from all other electrodes, had a kurtosis above five standard deviations, or were uncorrelated with other channels for more than 1% of the time; multiple denoising procedures to reduce prominent artifacts in the data: first, ASR with a standard deviation of 20 to remove large artifacts. Next, Ensemble Empirical Mode Decomposition with Canonical Correlation Analysis (EEMD-CCA) method performed selective low-pass filtering, specifically targeting large high-frequency activity with low auto correlation such as muscle activity and line noise.

Peterson et al. [55] also processed all EEG signals using custom EEGLAB scripts. EEG data were downsampled to 256 Hz, 1 Hz highpass filtered, merged across all conditions, and referenced to the common median of all channels. They reduced 60 Hz line noise using Cleanline. Then, rejected bad channels that had high SDs, had kurtosis $>$ 5 SDs, or were uncorrelated for $>$ 1% of the time. Further, denoised the remaining EEG channels. To remove large mechanical artifacts, they used ASR with a threshold of 20 standard deviations. The authors also performed selective low-pass filtering using ensemble empirical mode decomposition and canonical correlation analysis.

Peterson et al. [56] decided to, firstly, downsample EEG data to 256 Hz and then apply a 1 Hz high pass filter. Then, they applied a cleanline 60 Hz noise reduction, removed bad channels and used ASR. After that, they used a selective low-pass filter, following with bad channel interpolation and dipole fitting. In particular, they discarded dipoles with residual variance $>$ 15% . Manually, they retained cortical dipoles and grouped dipole clustering, retaining clusters with more than 15 subjects.

2.4 Discussion

(i) Which methods and walking conditions are used to provoke physical and virtual perturbations?

In order to perform a perturbation, there has to be a stimulus that can make the subject change his posture. This stimulus can happen in different ways, both in physical and virtual perturbations. For physical perturbations, it was observed that these methods varied according to the type of perturbation. In the stance condition, [42, 43] used an anchor that, when released, unbalanced a platform that a subject was standing on. Also, load releasing was experimented by [68], where subjects had to maintain postural balance while the load was being released on a belt that they worn at their sternum. It was discovered that [33] used an inflatable airbag and two handles that moved up or down depending on the airbag's inflation level to cause perturbations. Specifically, when the airbag was deflating, one of the handles would rapidly descend, which caused the subjects to lose balance. The authors in [4] experienced a series of predictable and unpredictable perturbations delivered to the trunk using a padded hand-held bar, both at ground level and at the edge of a platform that was 3.2 meters above the ground. Dietz et al. [15] decided to include a treadmill for their experimental protocol because it permitted quick changes of speed. In a stance situation, the treadmill would change its speed from 0 to 6 km/h, causing the subjects to lose balance. Finally, [16] used a specially constructed chair that could tilt to the left or right based on a random generator. In this manner, each subject's perturbation moment was selected at random to occur during the first four seconds of each trial.

A large amount of physical stance perturbations studies [11, 18, 31, 53, 54, 73, 74] caused losses of balance with the inclusion of a platform as one of the main equipment to create loss of balance. However, some studies such as [73, 74] only needed the platform itself to perform perturbations through horizontal translations or forward, backward or tilted perturbations in a random order.

For the physical walking perturbation studies, it was observed that all articles included treadmills to perform perturbations, which is understandable since it is logistically practical. In fact, the majority of studies used the SENLY platform [13, 40, 41, 84], which consists of a mechatronic platform designed to produce slip-like perturbations and is made up of a set of equipment, including treadmills and force belts. With that, while subjects were walking, perturbations were induced in the force belts.

Studies that included VR, in stance or walking conditions, performed virtual perturbations by sudden translation and rotation of the virtual environment that subjects were on. While experiencing these perturbations, subjects could also be physically perturbed. In [55], for instance, when the motor was ordered to spin an attached bar 90 degrees away from the subject, a steel cable that was fastened to the subject's safety harness was yanked.

Regarding walking conditions, only one study [13] applied over-ground walking in their experimental protocol, while the remaining eight worked with treadmills [13, 40, 41, 55, 56, 57, 70, 84]. Although it

is more complicated and laborious to work with over-ground walking especially in virtual perturbations, it would be beneficial to minimise this disproportion of articles, mainly taking into account that the over-ground analysis would bring the possibility of obtaining data closer to reality.

(ii) What type of subjects were included in the studies?

Generally, all studies included young healthy subjects in whom height and weight were within the average. Two (2) studies divided subjects in groups according to their age and health condition, respectively [68, 73]. Solis-Escalante et al. [73] included 9 subjects: 6 in a healthy condition and 3 with chronic hemiparetic stroke. Saadat et al. [68] decided to divide the subjects in two groups: younger and older.

(iii) Is it preferable to deliver perturbations during stance or walking?

The preference to perform perturbations between stance and walking may depend on the purpose of the study. The current literature, as it can be observed in Table 1, tends to study more stance perturbations compared to walking. However, not every fall happens in a stance condition. Slipping falls, for example, occur during walking, so it is also important to explore these type of perturbations. The use of EEG during walking not only requires more signal processing, since the subject is constantly moving, but it also can lead to more artifacts. These are some of the possible reasons why there is a trend towards stance perturbations, since it is more convenient to analyse and test cortical activity when subjects are not moving.

(iv) What procedures are implemented to maintain responses to perturbations unbiased?

Experimental protocols must rigorously apply unbiased perturbations in order to closely simulate real-world scenarios. As a result, it is imperative to establish and enforce procedures designed to enhance the ability of all participants to maintain a high level of focus and to abstract themselves from external distractions to the maximum extent achievable. These measures are crucial for preserving the integrity and validity of the experimental data, ensuring that the subjects' responses are not unduly influenced by extraneous factors and that the results accurately reflect the intended parameters of the study. Consequently, the stringent implementation of these procedures is pivotal for maintaining the scientific rigour and credibility of the research endeavour.

Some studies decided to start with an adaptation phase so that every subject could feel acclimatised during the protocol. In stance studies such as [74], subjects typically asked to stand comfortably for a small period of time (in this case, it was 2 minutes) to acquire eyes open resting-state activity. For walking studies and since most of them used a treadmill, the acclimatisation period consisted on free walking during a certain amount of time. Mezzina et al. [41], for example, performed about 5 minutes of free walking to discover their favourite speed and give them time to become used to the environment. Also, in [4], subjects rated their confidence, fear and stability levels using scales from 0 to 100. These procedures not also help

the subjects to feel as relaxed as possible but can also help EEG data processing by understanding how each subject behaves before a perturbation is performed.

During the perturbations, most studies described that subjects were asked to maintain posture. Mierau et al. [43] used a platform in their protocol and instructed subjects to, following a perturbation, keep its oscillations to a minimum. They asked subjects to focus their vision on a fixed location at eye-level, which is one of the main procedures used in the current literature to maintain responses to perturbations unbiased. Studies [55, 57] specified instructions to look straight ahead and avoid looking down. Also, since subjects were on a balance beam, they were asked to move only their hips side-to-side to maintain posture, avoiding rotation across the longitudinal axis of their body. Some studies even added a minimum distance that a subject should focus their vision. In [54], the authors defined this distance as 4.5 m. Virtual perturbation studies did not differ from the physical perturbation studies in this respect.

(v) What are the main sensor systems used to collect data during perturbation-based protocols?

In order to collect cortical activity and since the main purpose of the review includes this sensor as a mandatory one, EEG is considered one of the main sensors used to collect data. Its data are fundamental to understanding the mechanisms that trigger a perturbation at brain level, not only at the perturbation onset but also in the time instants that follow and precede it. Also, many studies opted for the addition of the muscle activity data collection, using specifically the EMG sensor [4, 5, 13, 15, 31, 40, 41, 43, 45, 46, 54, 55, 56, 58, 61, 70, 75, 82, 84, 85]. The combination of cortical and muscular activity is beneficial because it provides more complete information about the mechanisms that are triggered not only at brain level but also at muscle level.

EOG is primarily used to detect eye movements. Some studies included this sensor [5, 15, 16, 31, 43, 53, 54, 85] and combined it with EEG. Payne et al. [53], for instance, aimed to study cortical activity associated with predictable and unpredictable activities. For that reason, the authors decided to analyse eye movements and observe its differences in both predictable and unpredictable activities. Finally, only one study included inertial sensors systems to monitor changes in movement strategies [85]. Also, one study included Electrocardiogram (ECG) to measure electrical signals produced by the heart [75]. Specifically in the virtual perturbation articles, there were no significant changes in the main sensors that were included.

(vi) What was the participants' walking speed during the trials?

Most of the articles described walking speeds that allowed the subjects to walk successfully and without the risk of falling. It was observed that studies [55, 56, 57, 70] performed a speed of 0.22 m/s. Alternatively, [40, 84] experimented the speed of 1.10 $\pm$ 0.07 m/s and [41] 1.07 $\pm$ 0.07 m/s. With that, it can be concluded studies that included VR did not change the walking speed in their experimental protocols.

(vii) Is the number of channels a factor influencing the EEG decoding performance?

From the articles that have presented performance metrics, it is not clear whether there is a relation between the number of channels and the EEG decoding performance. On one hand, in [74], only 8 channels were required for the decoder and the average accuracy using raw data was 78%. Studies [13, 40, 84] have shown that, using 13 channels, it is possible to obtain accuracy levels above 93%. On the other hand, [73] used 126 channels and the best accuracy obtained was 73.5%. However, there are some variables that can influence these results, such as signal processing capacity or the number of active electrodes that may not be mentioned in the articles.

There was one article that tested different layouts in order to clarify the possible influence of the number of channels in the EEG decoding performance. Specifically, [16] examined the effects of channel reduction on classification performance. The four different layouts were: the 29 recorded channels in their entirety, a medium-sized layout with 15 channels, a tiny layout with 5, and a minimum configuration with 1 channel. Based on these four layouts, the individual participant peak accuracy was, respectively, 95.1, 94.4, 93.1 and 87.6% and the peak accuracy for the grand average was 83.5, 82.7, 81.2 and 71.5%. These findings show that, although it is not be linear, there is a correlation between the number of channels of the EEG sensor and the system's performance.

(viii) What methods were used to process EEG data?

There are several procedures to process EEG data in the current literature, which can differ on the purpose of each study in particular. Ditz et al. [16] performed an amplitude threshold rejection, which removed all trials with an amplitude that exceeded ±125 μV. In [56], they used a selective low-pass filter, following with bad channel interpolation and dipole fitting, discarding dipoles with residual variance $>$ 15% . For post-processing, [55] performed selective low-pass filtering using ensemble empirical mode decomposition and canonical correlation analysis. However, one of the most common first sequence of steps is downsampling of a given frequency, followed by bandpass filter, especially the Butterworth filter, which has a certain range of frequency depending on the study. In [33], for instance, raw EEG data were bandpass filtered by Butterworth filter of 0.1-10 Hz and downsampled to 50 Hz (5 times of the effective frequency band, preserving the original characteristics of the signal while improving the processing speed). Another example is [73], which also processed EEG data with a bandpass filter between 2 and 200 Hz and then downsampled to 512 Hz. Also, in order to eliminate line noise and its harmonics, the authors used the cleanline plugin in EEGLAB.

Another methods used in the literature are the ASR [23, 55, 56, 57, 73] and the rASR [41], which corrects artifacts that can happen in eye, muscle or electrode movements. This method is also implemented in the clean_rawdata plugin in EEGLAB [73]. ICA is another method that has the ability to remove/subtract data artifacts without deleting the impacted data sections, being performed in [23, 57, 75]. Associated with these methods, is the removal of bad channels, which is a technique that can be applied using EEGLAB.

Peterson et al. [57], for instance, rejected channels that had an abnormally high standard deviation from all other electrodes, had a kurtosis above five standard deviations, or were uncorrelated with other channels for more than 1% of the time. The authors also applied the EEMD-CCA method, performing selective low-pass filtering, specifically targeting large high-frequency activity with low auto correlation such as muscle activity and line noise. In [55], the authors went through similar procedures, but also performed selective low-pass filtering using ensemble empirical mode decomposition and canonical correlation analysis.

In reference to the various perturbation types and conditions, it is noted that no discernible correlation existed between these factors and the methodologies employed for processing EEG data. This lack of correlation can be elucidated by observing that the articles within each respective group adhered to comparable procedures during the processing of EEG data.

(ix) Is it common the use of ecologically valid virtual environments to produce visual perturbations?

Virtual perturbation articles, in general, implemented simple virtual environments that served only to fulfil the purpose of a visual perturbation and did not focus on being visually close to reality. This can be observed in [75], the authors created a virtual platform that could make vertical translations. Widdowson et al. [85] created a large room with two platforms that could also be raised or lowered. In [55], it is described that subjects could walk and jump around using body gestures. Thus, the virtual environments found in the literature are not ecologically valid, at least from a visual point of view. However, the visual perturbations applied to the subjects were close to reality, since the subjects' posture was generally challenged at the moment they were visually perturbed.

2.5 Conclusion

The current state-of-art of incipient falls using brain signals analysed in this review showed that the current literature is more focused on physical rather than virtual perturbations. Specifically in the physical group, most studies experiment perturbations in stance rather than walking. Apart from EEG, the main sensor included is EMG. The combination of cortical and muscular activity provides more data about the mechanisms that are triggered when a perturbation occurs. It was shown that most participants included in the studies were in a healthy condition and some even decided to divide them in groups, taking into account their age, for instance. There is a common group of methods to process EEG data, such as ASR, bad channel removal and bandpass filters.

Concerning virtual perturbation studies, the lack of ecologically valid virtual environments is an important aspect to point out, since all studies opted to implement simple environments that aimed only to perform perturbations without focusing on realistic details. In general, there is a proportionally higher variety of methods to perform perturbation in stance studies compared to walking. In the latter type, all

studies used treadmills as one of the main equipment in their protocols. For that reason and comparing the number of stance and walking studies, there is a lack of methods to perform perturbations regarding particularly in walking. The resolution of this problem would benefit the performance of a perturbation decoding system.

Chapter 3

Review: Decoding Brain Signals for ADL Recognition

3.1 Introduction

Despite the advancements in devices and technology aimed at facilitating natural and efficient recovery, it remains essential for individuals to have complete control over the devices they utilise. In this context, BCI has been developed to decode motor or cognitive goals from recordings of brain activity and convert these predictions into instructions for computer programs or robotic equipment. [44]. This part of the BCI which deciphers the user's motor intent from recorded brain activity is typically referred to as a neural decoder. To ensure a wider applicability of individualised brain mapping, BCI technologies, based on motion decoders from brain recordings through non-invasive technologies like scalp EEG signals, are intriguing – controlled rehabilitation devices. The scalp-mounted electrodes record signals with amplitudes in the micro-volt range in real-time. In the development of EEG-based BCIs and in order to achieve online signal decoding, numerous training trials are necessary to generate individually customised parameters for characterising signals [36].

The main goal of this study is review the literature to understand the process of decode different modes of locomotion (walking, climbing/descending stairs, ascending/descending on a ramp) which are classified as ADLs and develop a solution for imbalance events of human gait through the use of EEG which function is capture and register the brain signals [3, 67]. Based on recommended procedures identified in the scientific literature, a protocol for data collecting will be put into place [3]. It is intended that the group of subjects with determined healthy subjects using at least a 16 channel EEG system. These sensor data will be gathered while several subjects are walking in a controlled environment. In order to employ AI models to detect the subject's locomotor intents, EEG data must be processed to provide meaningful features through feature engineering [37]. The most important brain rhythms must be identified, together

with tasks and algorithms for pre-processing and eliminating the most often utilised artifacts, such as the ICA and ASR. Therefore, it is intended to evaluate the classification of AI-based models, and compare them in order to have an optimised benchmark model for real-time use with low computational cost. It is expected that the best model will be able to forecast when a loss of balance will occur.

3.2 Methods

3.2.1 Research Strategy

A systematic review was conducted and is a way of exploring existing literature. It starts with choosing a set of keywords, along with Boolean operators, to try and extract only the most pertinent papers in literature. The keywords selected should be relevant to the study of studies that decoded locomotion modes by recording signals from EEG and using sophisticated machine learning techniques for lower limb gait analysis. They should be general enough not to miss applicable literature and encompass the varying techniques, sensors, devices and terminology researchers use in their research but also restrictive to eliminate irrelevant research. The aforementioned keywords have been used to extract papers available in three databases: WEB OF SCIENCE, PUBMED and SCOPUS. The sets of keywords were: (”EEG”OR ”Electroencephalogram”) AND (”ADL”OR ”Walking”OR ”Climb”OR ”Descend”OR ”Sit-to-Stand”OR ”Stand-to-Sit”) AND (”BCI”) AND (”Decoding”) AND (”Classification”) AND (”Machine Learning”OR ”Deep Learning”).

The results have also been further filtered using inclusion and exclusion criteria in addition to the keywords. Only conference and journal papers were included in the results. A few papers were excluded because the full-text version was unavailable to us or was otherwise inaccessible. There were 151 papers in all that were discovered using these inclusion and exclusion criteria. After that, duplicate publications found in databases were eliminated, along with papers published before 2015, bringing the final number of papers down to 98. These publications have only been examined based on their abstracts, with less important or irrelevant studies manually removed (i.e., their subject matter has no bearing on our review).

The use of upper-limb robotics rather than lower-limb robotics, motor imagery rather than gait analysis, and the absence of electroencephalography were a few of the factors that led to the rejection of studies. The results were also restricted to journal articles and not conference papers. This review contained 16 of the 54 full-text papers that we evaluated for eligibility. The complete text of these publications has been examined with an emphasis on the variables the researchers are taking into account, the algorithms utilised, the sensing modalities, the types of subjects used to test/train their algorithms on, and the general effectiveness of their systems. This process is illustrated in Figure 2.

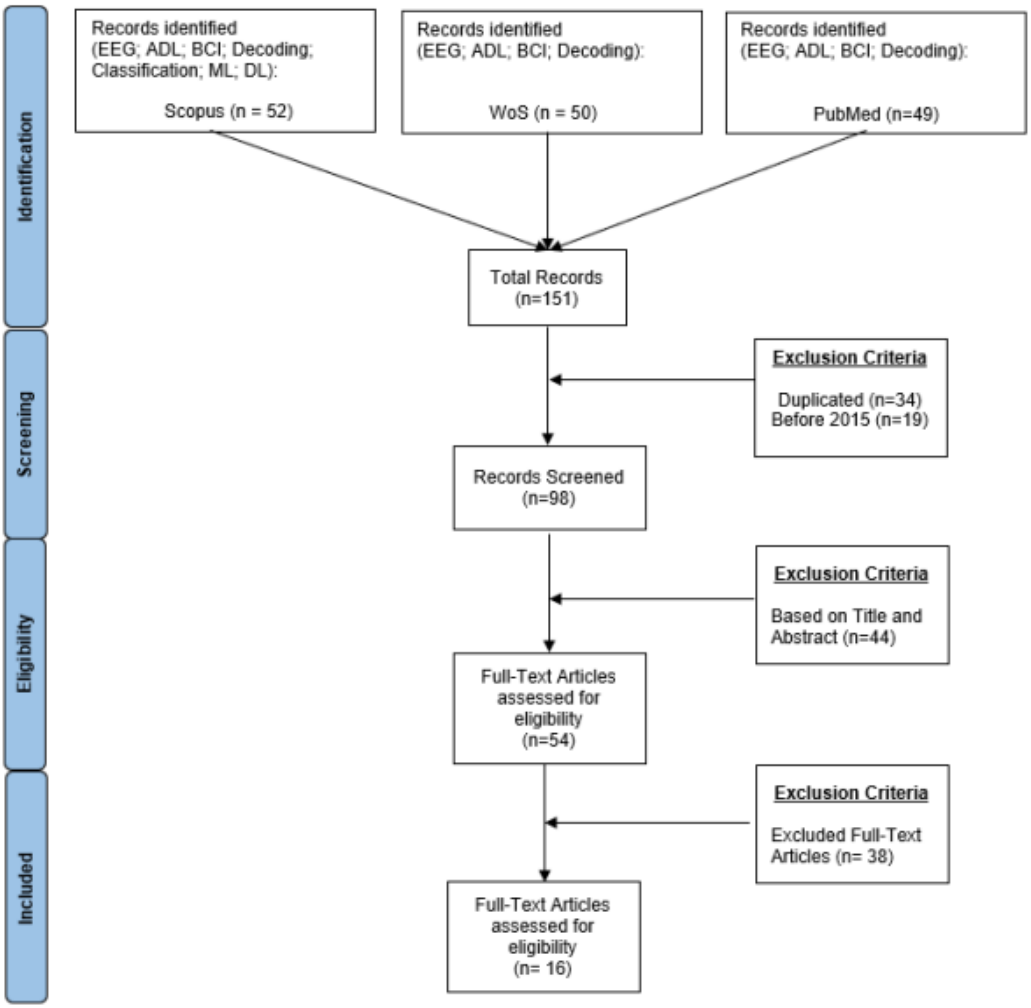


Figure 2: PRISMA Flowchart - ADLs using brain signals.

3.3 Reviewed Articles

Gait decoding was delivered in 16 studies, with seven papers conducting these studies on a treadmill at speeds ranging from 1.5 km/h to 3.5 km/h. Each session in 3 papers started with the subject standing still and upright for 2 minutes (mins). The treadmill's speed was slowly increased to 1.6 km/h in [77] experiment and the subject walked normally at this gait speed for 5 mins. When the subjects had completed the baseline phase of the experiment, they were directed to walk on a treadmill at a fixed slow speed of 1.5 km/h [72]. In order to calibrate the decoder for the following BCI control, individuals were instructed to walk steadily for 15 minutes during the task [78]. Chisari et al. [79] and [67] experiment the participants to walk on a treadmill at two distinct velocities, 2.5 km/h and 3.5 km/h. Two acquisitions lasting ten minutes each were made for each walking speed. Healthy participants in [22] study attended one experimental session in which they were asked to either walk in a passive or in an active mode at a speed of 1.5Km/h

wearing an exoskeleton. Subjects walked on the treadmill for 49 seconds, at which point it was shut off. The treadmill took at least 7 seconds to accelerate to a steady speed and 7 seconds to slacken and halt.

Table 6: ADLs Experimental Protocol - Procedure

Article	**Participants**	**Age (avg)**	**Task**	**Treadmill Speed (km/h)**	**Trials**	**Time per trial (min)**
[72], [78]	8 Healthy	24.5	sit-to-stand; stand-to-sit	1.6	3	20
[12]	8 Healthy	24.5	walking	1.6	3	20
[79], [67]	11 Healthy	30	walking	2.5 / 3.5	2 for each speed	10 for each trial
[32]	10 Healthy	25	resting; walking intention; walking	N/A	50	0.5
[22]	10 Healthy	32	walking; stand	1.5	7 active walking; 7 passive walking	1
[34]	13 Healthy	24	stand-to-sit; sit-to-stand; walking; step up; side step; back step;	N/A	50 for each task	180
[66]	27 Healthy	24	walking	N/A	N/A	N/A
[77]	5 Healthy	24.5	walking	1.6	N/A	N/A
[76]	5 Healthy	8	stand-to-sit; sit-to-stand; walking; stop;	1.6	20	N/A
[87]	1 Healthy; 1 Non-Healthy	N/A	walking; turn left; turn right; stop	N/A	20	N/A
[52]	3 Healthy; 8 Non-Healthy	N/A	walking	N/A	9	1

[19]	4 Healthy	23	stand-to-sit; sit-to-stand	N/A	78	2
[39]	20 Healthy	22	step forward; step up; step back	N/A	10	4
[26]	7 Healthy; 2 Non-Healthy	32.6	walking; stop	N/A	145	175

The 16 manuscripts presented in Table 6 included six that depicted overground mobility through gait. The experiment paradigm in [32] consisted of three states: a resting state of approximately 10 s, an intention state, and 10 s exoskeleton walking. Prasad et al. [87] is composed by 2 tasks: the first challenge involved four classes in a single session and had the subjects performing various moves (i.e., walking forward, turning right, turning left and stop, following the marked path on the ground, and the second task, subjects only executed walking and stop motions according to audible beep instructions).

Eguren et al. [76] performed four different types of locomotion tasks (i.e., sitting and standing, start and stop overground walking). The participants stood still for 15 seconds at the beginning of the session and in the sitting and standing task completed 20 sit-to-stand and stand-to-sit transitions that were triggered by a visual cue that was presented to them. They started the self-initiated sit-to-stand and stand-to-sit transitions after the calm standing time. The transitions were separated by an average of 10 seconds of waiting. When 20 transitions were finished, the participants were verbally told. They finished 20 walk-to-stand and stand-to-walk transitions throughout the start and stop overground walking exercise. Then, they were asked to remain motionless for 15 seconds to start the session. Prior to the transition from standing to walking, the participants maintained their standing position for an interval of 10 to 15 seconds that varied at random (indicated by a green light). After, they were told to stop walking when a red light started to flash on a 10-meter boardwalk. One attempt of the walk-to-stop transition was finished, and the individual then walked back to the starting point. In total, 20 experiments were conducted during the process.

Jochumsen et al. [34] performed 6 tasks which were: (1) stand-to-sit; the subject was standing in front of a chair (height of seat: 45 cm) and had to sit on that, (2) sit-to-stand; the subject was sitting on the chair and had to stand up, (3) walking; the subject had to walk three strides (starting with the right leg), (4) step up; the subject had to step up to a plateau (height: 16 cm) starting with the right foot, (5) side step; the subject took one step to the right side, and (6) back step; the subject took one step back starting with the right foot. Each run had six blocks, each of which contained ten motions of the same type (i.e., after each run 10 movement trials were performed of each movement type). The subjects stood calmly and gazed on a point on the wall four meters away for the two-minute resting recording that was conducted after each run (a total of five recordings). The individual was required to start the movement assignment once

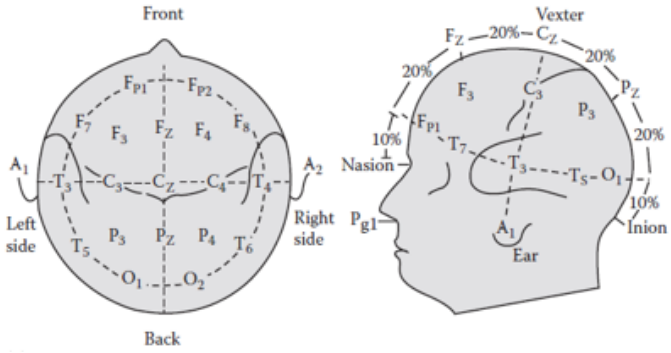


Figure 3: The 10/20 international system of electrode positions for EEG. Taken from [47].

three seconds had passed on the clock. There were 15 seconds between each movement trial. During the 3-second countdown and while completing the movement task, the subjects were told not to blink or make any other facial movements. The experiment lasted around three hours and was conducted in an electrically protected chamber.

The single trial performed in [52] experiment is composed of trial start, auditory cue, start walking, stop walking and trial end. Each trial took roughly a minute. Following the cue sign, each subject was instructed to enter the room. In order to reduce ocular and head movement artifacts, they strolled along at their own natural rate while gazing at the designated dot. In Tan Abdullah Al-Mamun et al. [66], the subjects had to perform free walking along an approximately 21 meter corridor.

Lastly, in the experiment [12] for sit-to-stand and stand-to-sit, a video stimulus that lasted for 4s to 5s and showed either the sit-to-stand or stand-to-sit video task, was presented to guide the participants to avoid the ambiguity of the instructions. The program started with a sitting position and included alternate sit-to-stand and stand-to-sit exercises for 5 times each.

3.3.1 EEG Setup

From all the articles reviewed in this study, the EEG data were collected and labelled according to the extended 10-20 international system (example in Figure 3). Under this system, the skull is divided into six areas from nasion to inion with interval rates of 10%, 20%, 20%, 20%, 20%, and 10% (Fp: frontopolar, F: frontal, C: central, P: parietal, and O: occipital, respectively), and also divided into the same ratios from left to right preauricular points (T3: temporal, C3: central, Cz, C4, and T5, respectively).

Four (4) studies describe the usage of a 64-channel active EEG electrode setup, four of which were employed as EOG sensors to record and filter out eye-related distortions. In three of the four publications cited, the sampling frequency was set to 100 Hz [72, 77, 78]. Ground (GND) and Reference (REF) channels were placed on the left and right earlobe (A1 and A2), respectively. The T7 and T8 channels were respectively shifted to FCz and AFz. As EOG, FT9, FT10, TP9, and TP10 were utilised to record eye blinks

and movements. This change seeks to increase decoding accuracy for two key reasons: 1) GND and REF channels in the standard setup were close to the motor cortex and 2) EOG sensors were required in the process of artifact removal [72, 77]. In Eguren et al. [76] study, a 64-channel was used to record wirelessly at 1000 Hz from the face and scalp. In order to record blinks and eye movements, channels TP9, PO9, PO10, and TP10 were taken out of the cap and used for EOG. Furthermore, these data were eliminated from all analyses in this experiment.

EEG data were captured using a 64-channel EEG amplifier with a custom signal-preamplifying active electrode cap at a sampling rate of 2048 Hz/channel (bandwidth DC - 1024 Hz). However, EEG was re-sampled at 1024 Hz prior to further pre-processing [67, 79]. Jochumsen et al. [34] also recorded EEG using a 64-channel sampled with 1200 Hz whereas in [87], the sample frequency was 100 Hz. Electrical signals from 62 electrodes were recorded by Nienhuis et al. for their paper on [22], while Jeong et al. [32] and Park et al. [52] used a 32-channel wireless EEG data system. However, in the first paper, the authors digitised the frequency at a sampling of 1000 Hz, but in contrast with rest of articles presented in this study, [87] EEG (64 channels) was recorded by combining two 32-channel amplifiers (100 Hz).

From all studies, only [12] obtained EEG signals using 11 passive electrodes using a sampling rate of 1200 Hz, and the reference and ground electrodes placed on the left and the right earlobes, respectively. Two passive electrodes placed beneath and close to the right eye's outer can were used to collect EOG signals. The technology utilised in this work and the other 5 papers was designed to simultaneously record EMG signals throughout the experiment. These signals are frequently used to modify gait pattern and in some applications, in order to determine the commencement of movement. EMG electrodes were placed on Tibialis Anterior (TA), VM and Biceps Femoris (BF) of each leg [67, 79], Rectus Femoris (RF), TA, and Gastrocnemius Lateralis (GL) of two lower limbs [12], Gastrocnemius Medialis (GM), Semintendinosus and Vastus Lateralis (VL) [22], two bipolar Ag/AgCl electrodes on the TA and BF muscles of the right leg (these muscles are known as fast activation muscles for walking) [32]. In [26], the reference electrode was positioned on the bony surface of the right knee, and the EMG channel was positioned at the mid-belly of the right leg's TA muscle. The muscle TA was selected because it is one of the first to contract throughout a gait cycle.

Table 7: Experimental Protocol - Sensor Detail

Article	EEG Channels	EEG Frequency	Additional Sensors
[72], [78], [77]	64	100Hz	EOG
[79], [67]	64	2048Hz	EMG
[34]	64	1200Hz	
[76]	64	1000Hz	EOG
[66]	62	1000Hz	
[22]	62	500Hz	EMG
[32]	32	1000Hz	EMG
[52]	32	500Hz	
[87]	64	100Hz	
[12]	11	1200Hz	EOG and EMG
[19]	8	250Hz	EMG
[39]	19	200Hz	
[26]	8	500Hz	EMG

3.4 EEG Data Acquisition and Processing

The EEG signals are recorded from active electrodes, amplified by a EEG amplifier, and then digitalised. The EEG signals typically have an amplitude of micro-volts. The signals were initially filtered using the subsequent filter types: i) 4th order Butterworth in 5 papers, ii) 2nd order Butterworth filter in 4 papers, iii) FIR in 2 papers, and iv) Zero phase 24th Chebyshev type II in 1 paper. Additionally, different filter classes were used: i) high-pass filter from 0.1 Hz to 10 Hz were used, based on [26, 67], to remove DC power supply bias from the signal, a value of 1 Hz is applied, and in accordance with [32] value of 10 Hz is applied to remove delta (1-4 Hz), theta (4-8 Hz) band from the signals; ii) most articles used a bandpass filter between 0.1 Hz and 50 Hz (on average), which is the frequency range best suited for gait decoding and allows the division of the EEG spectrum, as shown in Figure 4 into delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz) and gamma (30-50 Hz) bands [66].

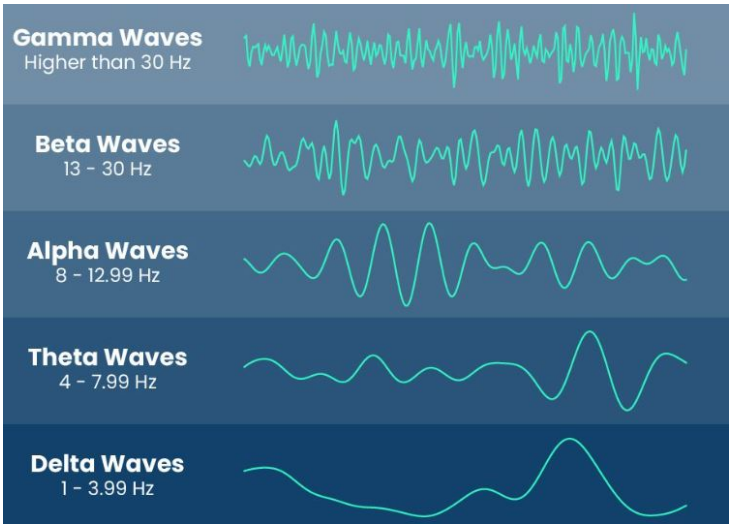


Figure 4: Brain Waves Frequencies. Adapted from [21].

In Prasad et al. [87], the obtained data were normalised by channel by summing the mean and dividing by the standard deviation, and then filtered in the 0.1-2Hz band using a second order Butterworth filter (z-score) [76, 77, 87]. EEG data during walking may typically contain movement and other artifacts, both physiological and non-physiological, despite all experimental measures used to minimise interference from movement artifacts. Gait-locked artifacts in particular, which coincide in timing and frequency with brain activity [30].

As a result, particular care should be made to avoid using task-related artifacts instead of neural correlates to decode the task itself when decoding gait activity from EEG signals. To reduce this danger, a number of pre-processing techniques to reject movement artifacts have been presented in the literature, the majority of which are based on ICA [35]. ICA was employed to remove the independent components (Figure 5) representing artifacts by visual inspection [52] to project the data from the scalp channels domain to the IC domain. The interference from the power lines was eliminated using a specially designed 50 Hz comb notch filter without actual poles [79].

The EEGLAB toolbox in Matlab [2] can also automatically reject bad channels based on the probability and Kurtosis statistics of the distribution of the entire signal in each channel using a threshold of 90% [79]. Bad channels are indicated as standard deviation greater than 1000 V or kurtosis of more than five standard deviations from the mean [78]. The residual signals are then subjected to a Common Average Reference (CAR) spatial filtering to increase the signal-to-noise ratio. A CAR filter is applied to the chunk of EEG data inside each window and the band (1-8 Hz) was extracted with a 4th-order zerolag Butterworth filter [67]. As a result of the averaging procedure, CAR reduces the uncorrelated random noise with a zero mean for each sample period by deducting the mean value of all electrodes [36]. By visually inspecting the data, epochs with high amplitude artifacts (over 100 V) and irregularities were manually deleted [67, 79]. However, epochs were rejected only if they exceeded +/- 150 μV [34].

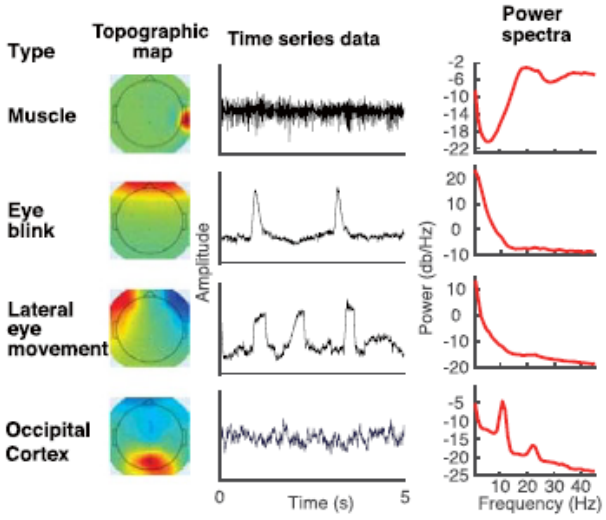


Figure 5: Examples of artifact and brain ICs. Taken from [25].

High amplitude artifacts were eliminated using an ASR (e.g. eye blinks, muscle burst) [76]. Principal component analysis is used by ASR to identify channels that significantly deviate from baseline data with little movement artifact in sliding window EEG data [78]. Prasad et al. [87] evaluated the classification accuracy of ASR-based and non-ASR-based classifications to evaluate the possible impact of motion artifacts.

Then, Reliable Independent Component Analysis (RELICA) was used to make sure that movement-related artifacts did not have an impact on classification performance. For each individual, the extracted ICs were then clustered across RELICA repeats based on how similar they were to one another into a set of clusters that matched the number of EEG channels. Each IC in the RELICA architecture can have a quality index assigned to it based on how compact the cluster is to which it is connected [79]. The EEG signals from Step I were subjected to the ICA weights, which projected the data into the ICs domain. By back-projecting the EEG signals to the original domain while only using the components linked to brain activity, stereotyped artifacts (such as neck muscles and eye movement) were discarded [67].

An overview of some of the techniques mentioned above can be found in Figure 6, where single-subject EEG data were, firstly, pre-processed (green) by going through two EEG preprocessing stages, Step I (cyan) and Step II (orange). The first, which is more conservative, removes noisy epochs, poor channels, and line noise. To increase the dependability of extracted ICs, the second method employs a more forceful rejection of artifacts (blue). Data prepared in accordance with the first pre-processing phase is subjected to ICs. Neural activity patterns in multiple frequency bands are associated with gait, and the most often used method relies on ERD/ERS, which were derived by dividing the power in the frequency of interest from active and passive walking by the appropriate baseline condition [22] in mu (8-13 Hz) and beta (14-20 Hz), i.e., sensory-motor rhythms [79].

Other works exploited Movement Related Cortical Potentials (MRCP). An MRCP comprises two main components: a Readiness Potential (RP) and a Movement Monitoring Potential (MMP). The RP is a negative-cortical potential that starts around 1-2 s before the onset of a voluntary movement and activated over the pre-Supplementary Motor Area (pre-SMA) or the contra-lateral primary motor cortex (M1). A slow positive deflection known as a MMP results from the motor process after the movement intention has been carried out. Therefore, a MRCP decodes the process of movement preparation/execution based on a single trial basis in delta (<3 Hz) or theta (4–8 Hz) [79] due to its properties such as a potential spontaneity and early detection of user intentions [32]. The EEG signals captured during the Motor Execution (ME) were used to extract the MRCP signals. Next, a 0.05 Hz high-pass filter was applied to the EEG readings (2nd order non-causal Butterworth filter). To filter out the electrical disturbances, the notch filter's frequency rate was set at 50 Hz. EEG signals were then down-sampled from 1200 to 250 Hz [12].

Two studies also used other techniques in their real-time operations in addition to the ones stated

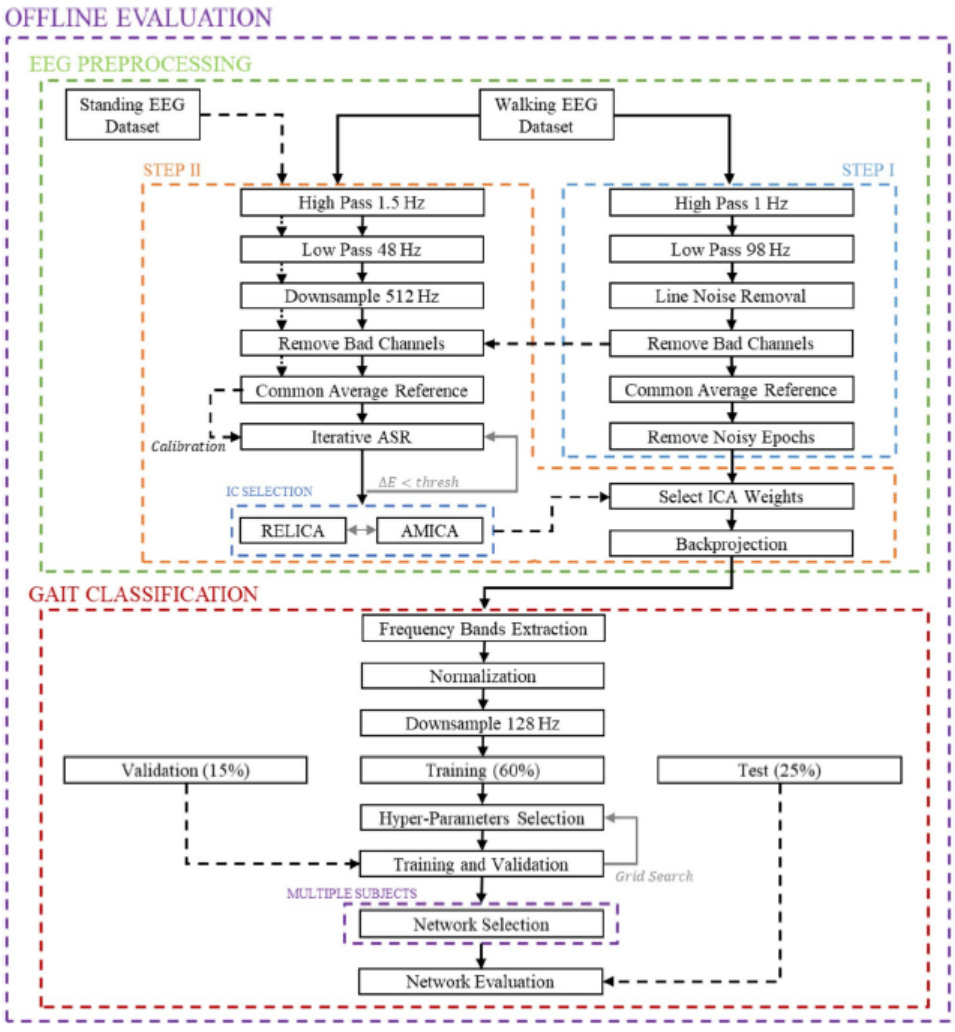


Figure 6: Schematic workflow of the offline EEG processing and classification. Taken from [79].

above: (1) Ravindran et al. [72] employed an H-infinity algorithm to simultaneously eliminate recording biases, amplitude drifts, eye blinks, and eye movements. The H-infinity algorithms' parameters remained constant with respect to real-time decoding. As peripheral channels often contain a lot of artificial components, (2) whereas Nakagame et al. [77] using data recorded from EOG channels, applied H-infinity filter to remove ocular artifacts and signal drifts, which may be the main sources of EEG contamination, from data taken from the EOG channels. The Unscented Kalman Filter (UKF) decoder was designed and validated using 50 channels that remained after the EOG and peripheral channels were eliminated [77] [72]. In order to enhance the performance in real-time, Closed-loop Decoder Adaptation (CLDA) was used to change the UKF's parameters during BCI operations [78].

The results of the majority of articles that trained and validated models to interpret EEG signals in order to classify the types of ADLs will be discussed in the following sections along with their separate findings. The percentage of data split and the models' accuracy as determined by the literature review are summarized in Table 8.

3.4.1 Gait Deconding from EEG with LSTM

Chisari et al. [79] employed a comprehensive offline pre-processing stage, depicted in Fig. 6, as a crucial step in their experimental methodology. This meticulous pre-processing was implemented with the specific aim of mitigating any potential influence of artifacts that might have been time-locked with the stepping frequency. The objective was to safeguard the integrity of their analysis by ensuring that these artifacts did not interfere with the optimisation of network hyper-parameters or the assessment of model performance.

Table 8: ADL Recognition - Dataset Split (Train, Validation, and Test) & Accuracy

Article	**Train**	**Validation**	**Test**	**Accuracy**
[72]	80%	20%	10%	90%
[79]	80%	20%	10%	90%
[67]	60%	15%	25%	80%
[32]	50%	shared with train	50%	86% multiple channels and 81% Cz Channel
[77]	4 segment of data	N/A	1 segment of data	43%
[12]	14 trials	N/A	1 trial	82.73%
[22]	90%	N/A	10%	89.9%
[52]	7 trials	N/A	1 trial	83.4%
[66]	N/A	N/A	N/A	CV = 77.8%
[19]	75	N/A	25	83.3%
[39]	70	N/A	30	86%
[26]	N/A	N/A	N/A	85.65%

In relation to cerebral correlates of gait, several EEG frequency bands (δ band (1-4 Hz), δ and θ bands (1-8 Hz), from δ to μ and low β bands (1-16 Hz), only μ and low β bands (8- 16 Hz), high β band (24-40 Hz)) were extracted as neurophysiological features to decode walking activity. A MRCP is a feature common to all bands, including low frequencies (1-3, 1-8, 1-16 Hz), which are characterised by their appearance before each swing phase. The ERD ranges, on the other hand, exhibit a desynchronization — smaller absolute signal amplitude — during the swing phase in comparison to the stance phase [79].

The impact of Long Short-Term Memory (LSTM) hyper-parameter selection on the decoding performance was evaluated during the offline evaluation. Since the 1-16 Hz frequency band is the widest band taken into consideration in this study, performance was measured in terms of accuracy and cross-entropy loss of the network trained on 60% of the entire dataset (training set) and used to predict 15% of the entire dataset (validation set). We investigated a network with two LSTM layers, 250 LSTM units in the first layer, and 100 LSTM units in the second layer. This network was trained using a constant learning rate of 10-3 (none learning rate schedule), achieving on average 92.8±3.1% of accuracy and 0.14±0.12 of loss in validation [79]. As a result, the three classifiers were trained for each individual using EEG signals that had been filtered in the 1–8 Hz frequency region, and their effectiveness was evaluated using the test set's 25% of wholly unknown data.

Chisari et al. [79] proposed in their paper a deep learning–based classifier for the decoding of gait events from scalp EEG signals. Their approach depends on the memory capacities of a particular kind of recurrent network, namely LSTM, to learn time correlations within the data. In fact, the brain activity that occurs before each leg's swing phase and the contro-lateral leg's stance phase coincide in time (Wagner et al 2014). This characteristic may make it difficult to distinguish clearly between the dynamics of each leg, as evidenced by memory-less classifiers' poor performance when decoding right and left gait events separately compared to decoding the two legs combined.

3.4.2 Gait Intention Using Spatio-Spectral CNN

The entire study was carried out online. The average of all test sets from k-fold cross-validation is used to determine accuracy. The total average is the average of all subjects, whereas the subaverage is the accuracy of each subject group on average. Gait/stand state recognition classification accuracy using the spatio-spectral Convolutional Neural Network (CNN) model was 83.4%. On the other hand, the gait and stand intention recognition accuracy achieved 77.3% and 77.7%, respectively. The results imply that while designing a CNN architecture, carefully choosing the properties of the data is essential. Park and Park [52] utilised only 19 participant EEG channels, which wasn't enough to reflect spatial aspects. Performance should improve using a deep learning model that incorporates the better spatial and temporal properties of the EEG data.

3.4.3 Real-Time EGG-based BCI

The paper results discussed below were processed using a combination of online pre-processing and offline decoding. Nakagome et al. [72] targeted to increase the lower limb EEG decoding accuracy and resilience from an algorithmic standpoint. For enhanced usability and system control of exoskeletons and neuroprosthetics, accurate lower limb decoding is crucial. UKF demonstrated its superior performance in early convergence with fewer samples and when assessed from the standpoint of the r-value. However, UKF demonstrated its weakness when measured by the R2 score and when a channel is disrupted. While alternative algorithms may outperform UKF decoders with bigger tap sizes, UKF decoders often perform better with fewer samples. This could be one of the benchmarks used to determine how many samples should be used for real-time decoding.

The impact of sampling frequency on performance was unknown before. Using a sample-by-sample decoding method, this study examined this problem and found that performance may be enhanced. Although data can be recorded at the greatest sampling frequency, if there is enough frequency range for reconstruction, the delta bandpassed characteristics can theoretically lower the sample size to 20 Hz. With this method, the future prediction time in [72] may theoretically be increased from 1 ms (when 100 Hz), to 5 ms (in 20 Hz), while still using the same decoding strategy. We also demonstrated that the performance was capable of being enhanced.

In particular, Luu et al. [12] found sustained α/μ suppression in the Posterior Parietal Cortex (PPC), and Inferior Parietal Lobe (IPL) regions and significant decreases (ERD) in the β band, indicating increased cortical involvement in the walking task. Additionally, the outcomes showed a significant rise in brain activity in the low frequency (Δ bands) in the Anterior Cingulate Cortex (ACC) region, illustrating the potential advantages of closed BCI, which makes use of cortical networks involved in motor learning and error monitoring. Additionally, there was low γ in the ACC and Superior Temporal Gyrus (STG) as well as β suppression in the ACC, PPC, and IPL. These results imply that a closed-loop BCI system promotes natural human gait control. With a closed-loop BCI system, the technology may potentially aid in a better understanding of cortical dynamics during walking.

3.4.4 Decoding Sensorimotor Rhythms

Both healthy participants and stroke patients had classification performances that distinguished walking from baseline that were higher than 93% and 89%, respectively [22]. According to the classifier weights, ERD in the mu rhythm, which was more bilaterally distributed, and ERD in the beta and low gamma bands, which were more centro-medially placed, were the major brain signals that contributed to this performance. Brain signals in the beta band centro-medially located (Cz electrode) were significantly different between passive walking and vigorous walking in comparison to baseline (with exoskeleton). Beta/gamma proved to be more medially concentrated over the foot sections, whilst the mu-band effect looked to be

more lateralised over the hand parts. This could be as a result of cerebral mu-cancellation brought on by implicit hand activity or volume conduction effects. It is commented in article [22] that some studies have measured cerebral activity during actual gait, showing that walking increases cerebral activity bilaterally in the medial primary sensorimotor cortices, the supplementary motor area and the prefrontal cortex.

A Canonical Correlation Analysis (CCA) has shorter computational time and can be used experimentally. Also, because of that, CCA can be easily implemented during online BCI. Even though the primary objective of the work is not to create a BCI for gait rehabilitation, it is important to comprehend and gather all the data necessary to first assess the feasibility of decoding walking intention from brain signals during lower limb rehabilitation by a robot because it might be useful in situations and experiments where the subjects are not in good health.

When stroke patients are immersed in a robot-gait training system, [22] offline classification findings reveal that it is possible to discriminate between resting and walking with high accuracy. García-Cossio et al. [22] showed that the online application of this approach (i.e., using a logistic regression classifier to distinguish walking intention from resting) can be used to accurately control the on/off state of a treadmill using EEG signals from healthy subjects. These results showed that developing a BCI for gait rehabilitation was feasible.

3.4.5 Decoding Movement-Related Cortical Potentials

Jeong et al. [32] performed single-channel and multiple-channel decoding. The statistics show that there was only a little (less than 5%) performance difference between the multi- and single-channel approaches. This shows that a BCI would be able to detect movements using just a single EEG channel, enabling the use of assistive technology to provide afferent input. Furthermore, it was established that the Cz channel performed better in terms of discriminant performance than other nearby channels. One of the causes for the Cz channel activation in this study is their experimental results, which relate to lower-limb movement (i.e., walking). Additionally, the motor cortex reflects each electrode's increasing cortical density during a bodily movement according to the location of the body. The motor cortex's central region, which is close to the Cz channel, has a close relationship with the lower limb. The single-channel BCI technique with the best sites uses the fewest possible EEG feature vectors to make a final decision, potentially cutting down on computing time for model training and real-time scenarios.

Jochumsen and Niazi [34] identified the movements could be classified with respect to the idle activity with accuracies in the range of 80%–90%. The stand-to-sit and sit-to-stand actions were the most discriminable movement types, accounting for 54% of the movements that were properly classified based on movement type. The findings show that it is possible to identify and classify movement intents associated with functional motions. This information could be used to build a BCI that introduces task diversity in neurorehabilitation.

By classifying movement intentions and idle activities, the detection of movement intents was measured. In a real-time BCI decoder, this information won't be available because the epochs were retrieved with knowledge of when the movements occurred; as a result, fewer movements may be successfully detected. It is possible to alter the detection algorithm to be less cautious, but this raises the possibility of false positives. Building a cue-based BCI system in which the detector is only active during particular time frames is one potential solution [34].

3.4.6 Neural Decoding of Gait in Developing Children

A minimum of six out of the thirteen publications analysed in this study demonstrated that walking causes significant EEG alterations at low frequencies (< 10 Hz) [66, 72, 76, 77, 78, 87]. EEG in the delta band (0.1 – 3 Hz) was utilised to neurally decode walking, so the EEG features match time-domain delta band Amplitude Modulated (AM) potentials [72, 76, 78]. It has been shown that EEG waves in this range can provide important information about human gait. The results were comparable to invasive BCI methods when slow cortical potentials in the delta band EEG were combined with a linear decoder to understand human gait while walking on a treadmill. These research also showed that a real-time closed loop BCI for human treadmill walking has a good decoding accuracy from an EEG signal in the delta band.

EEG signals need to be carefully monitored since they convey both multi-sensory afferent feedback and efferent motor control during human walking. To understand human gait, it is challenging to identify pertinent neural properties from scalp EEG data. To deal with this, [76] employed a causal model that predicts current kinematics based on EEG signals before the current motor command. Afferent feedback changed in future EEG readings, which are not included as inputs in the specified causal model, also includes delays relevant to the current motor command in the feedback signals.

Finally, the construction of a reliable neural decoder depends on the dynamic nature of brain signals. Neural decoders are often trained offline by mapping real motions to neural signals. While switching from open-loop to closed-loop BCI, this technique, on the other hand, disregards participants' brain dynamics, which frequently leads to lower online performance.

3.5 Discussion

The comprehensive examination of the existing literature on BCIs for controlling lower-limb robotic systems revealed a common trend in the experiments conducted. These experiments frequently involved the classification of discrete state directives, such as walking, stopping, turning, standing, and sitting. Notably, the review identified a limited application of EEG denoising techniques, or where applied, these techniques often lacked thorough testing.

Neural classification, on the other hand, encompassed the utilisation of numerous neural features and decoding methods. On an overarching level, the performance of these BCI systems appears promising. However, they still remain some distance away from practical real-world applications. Several factors contribute to this gap, including the relatively small sample sizes used in the experiments, potential safety concerns, and various technical challenges. Furthermore, it's worth highlighting that none of the reviewed papers addressed experiments related to balance loss in elderly individuals, which is a central focus of this study.

A Deep Learning (DL)-based model for decoding gait has shown the potential to utilise non-invasive brain activity recordings effectively for the detection of walking events, including phases such as swing and stance, in healthy individuals. Tortora et al.'s study [79] indicate that a memory-based deep learning classifier may offer a promising avenue for enhancing brain-controlled rehabilitation or restoration devices, particularly when employed in real-time scenarios with the capability to control each leg individually.

In the work conducted by Park et al. [52], a spatio-spectral CNN model was created, resulting in a notable gait state identification accuracy of 83.4%. Additionally, the study achieved accuracies of 77.3% and 77.7% in identifying intentions related to gait and standing, respectively. These outcomes demonstrate the ability of this study to successfully decode the intended gait actions of individuals, including those with subacute or chronic strokes, as well as healthy subjects.

Four separate studies [72, 76, 77, 78] collectively affirm that motion artifacts can be considered negligible within the delta band frequencies, especially when considering the gait speeds employed in this research (specifically, 1.6 km/h). The study outlined in [22] successfully decoded the intention to walk from cortical patterns within the sensorimotor strip. This was achieved during robot-assisted gait training involving both healthy subjects and stroke patients with minor lower limb disabilities. Notably, the research identified that the modulation of low gamma activity in the central midline regions was correlated with the stages of the gait cycle in healthy subjects. However, this correlation was not observed in stroke patients, suggesting a difference in the neural patterns between the two groups.

Lower extremity functional movements have been shown to generate MRCPs that are discernible in EEG recordings. Interestingly, the study identified that the MRCP with the most pronounced amplitude was elicited during the stand-to-sit exercise, while the MRCP associated with walking had the least amplitude. The research demonstrated that these MRCPs could be recognised even during periods of inactivity, and different individual movement tasks could be effectively categorised [87]. Although direct comparisons with previous studies are challenging due to variations in experimental procedures and methodologies, the detection of MRCPs achieved a remarkable high grand-average decoding performance of 86% across all patients in a single session [32].

Chapter 4

Project Conceptual Design

4.1 Introductory Insight

In light of the pressing need to address the challenges associated with falls among individuals, the development of prediction and detection technologies has emerged as an urgent imperative. This necessity has driven the creation and application of assistive devices, including prostheses, exoskeletons, and various technological solutions, all aimed at preventing falls [60, 86]. However, a significant portion of these devices still relies on reactive measures based on past events, underscoring the crucial need for faster and more intuitive interventions. Moreover, recognising the importance of ADLs recognition is paramount, as it plays a pivotal role in enhancing the quality of life for individuals, particularly the elderly.

To confront this challenge effectively, it is crucial to gain a deep understanding of the mechanisms triggered during the onset of a fall and ADLs, encompassing brain-level processes. This understanding can pave the way for rapid and timely responses, affording individuals more reaction time in critical moments, as well as assess the risk of fall through the identification of risky ADLs [60, 86]. The central objective of this dissertation is to tackle this issue by formulating a strategy for decoding incipient falls and ADLs, which can subsequently activate fall prevention mechanisms in auxiliary robotic devices or alert notifications [48, 80]. The central emphasis will be on decoding brain signals during slip-like perturbations, given that slips constitute the most prevalent type of falls, followed by the recognition of ADLs.

To accomplish this objective, two comprehensive data collection protocols are implemented. One involves the observation of healthy subjects on a treadmill within a controlled environment as they simulate slip-like perturbations, as the second engages elderly in various ADLs. Leveraging EEG data and processing it through tools such as EEGLAB, AI-based systems are employed to recognise ADLs and to decode slip-like perturbations.

This dissertation endeavours to make a significant contribution to the exploration of brain signal responses to slip-like perturbations and ADLs by analysing previously collected EEG experimental data. Through a thorough analysis of this experimental data, it aims to gain an understanding of the brain's signal responses and to identify specific patterns indicative of balance loss - PEPs - or regular motion during ADLs. This knowledge can subsequently inform the development of technological solutions aimed at addressing the challenge of fall prevention and ADL recognition, respectively.

By extracting quantifiable data from the examination of slip-like perturbations and ADLs, which can serve as inputs for training AI-based models, this research significantly enhances the potential for the development of BCI systems dedicated to these scopes. Moreover, this study goes beyond the selection of the most suitable AI-based algorithms; it also seeks to provide valuable insights into the design of fall prevention devices.

4.2 Research Hypothesis

The research endeavours conducted as part of this project are founded on the following assumptions:

- EEG has the potential to expedite the detection of perturbations when they occur [60, 64];
- The individual neural parameters of users enable the timely detection and prediction of PEPs, which can serve as an effective method for detecting slip-like perturbations during walking [62, 83];
- Daily activities are associated with distinct patterns of cortical activity [12, 22, 53];
- Sudden muscular movements have the capacity to introduce movement artifacts into EEG data, which can significantly impact the experimental protocol [33, 68];
- ANN exhibit the ability to classify slip-like perturbations and ADLs based on EEG data [49, 79];
- The use of a treadmill for subjects' walking, as opposed to overground walking, may introduce variations in EEG data due to the more repetitive nature of the activity [71].

4.3 Project Solution

Considering the research hypothesis, this chapter outlines the project's progression through three pivotal phases (Figure 7). The initial phase involved comprehensive data collection, employing two distinct experimental protocols that enlisted the participation of ten and six subjects, respectively. This phase laid the foundation for the subsequent step, which is data processing, where EEG data is meticulously labelled and processed using EEGLAB, setting the stage for the comparative analysis phase. This comprehensive

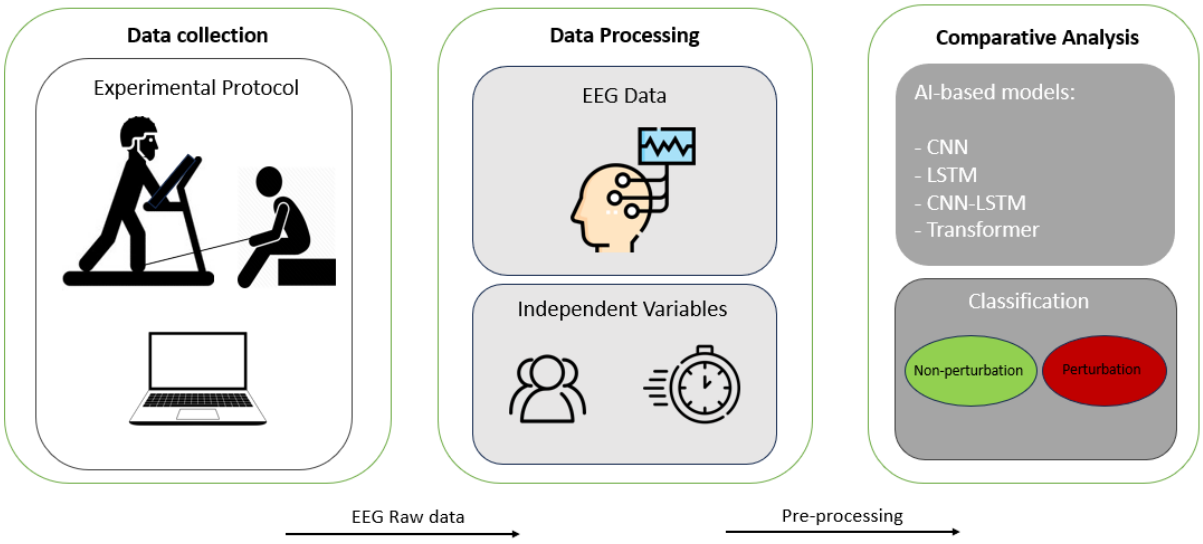


Figure 7: Schematic of slip-like perturbation protocol phases.

analysis not only elucidates the methodologies adopted but also delves into an insightful comparison of their performances.

4.3.1 Data Collection

Phases 2 and 3 will only be possible with experimental protocols that allow the collection of data from several sensors. Two experimental protocols were designed considering the good practices of the scientific literature. The first consists of carrying out slip-like perturbations during walking, while the second is related to ADLs recognition.

4.3.1.1 Slip-like perturbations protocol

The first experimental protocol, meticulously orchestrated within the distinguished confines of BiRDLab at University of Minho, was fundamentally driven by the overarching ambition to unravel the intricacies of cerebral responses in the face of slip-like perturbations. To proceed with the study, certain criteria were set for the subjects. They needed to exhibit: i) normal and healthy walking; ii) complete postural stability; iii) be over 18 years old; and iv) have a body mass below 135 kg. Individuals with any disease, particularly neurological disorders, that affected their ability to walk were excluded, as well as those who had recently undergone surgeries that impaired mobility. Considering these requirements, ten healthy adult (5 males, 5 females) were selected for the experiment, with an average age of 23 ± 1.48, height of 1.7 ± 0.12 m, and weight of 64.6 ± 3.26 kg. Prior to participating in the experimental trials, all subjects provided written informed consent and willingly agreed to take part. Each participant completed the Waterloo Footedness Questionnaire to assess their preferred foot in a qualitative manner.

For this study, two sensor systems were utilised. The first one was a wireless EEG headset called g.NAUTILUS PRO, which consisted of 16 channels equipped with active wet electrodes. This headset was employed to record brain activity in a controlled environment. The second system employed an Inertial Measurement Units (IMU) on the rope. This IMU was responsible for recording the acceleration that occurred when a rope was pulled to simulate a slip-like perturbation while the subject was walking on a treadmill. The EEG system recorded the brain activity data at a rate of 500Hz, and the IMU recorded the acceleration caused by the rope pulling action at a rate of approximately 500Hz.

During the experiments, the subjects were equipped with a safety harness device to prevent falls in the case of an irreversible Loss of Balance (LOB). This device consisted of a vest that was attached to a rope connected to a ceiling structure. The length of the rope was adjusted so that there was a minimum of 15 cm distance between the subjects' knees and the treadmill belt. To ensure proper weight distribution, the participants were instructed to lift their feet, causing their entire body weight to be supported by the harness system (Figure 8).

In order to gather synchronised data, a software called Sync Lab Desktop was employed. This software, developed by BiRDLab, was specifically designed for Windows Operative System (OS) and allowed for simultaneous initiation/termination of data collection from the sensors mentioned earlier. The Sync Lab program facilitated the transmission of electrical trigger signals, which were delivered through a HW interface called Sync Box. This was connected to a PC via a USB connection and was able to send a pulse at the start/stop of each trial so that both the EEG device and the IMU would start/stop simultaneously.

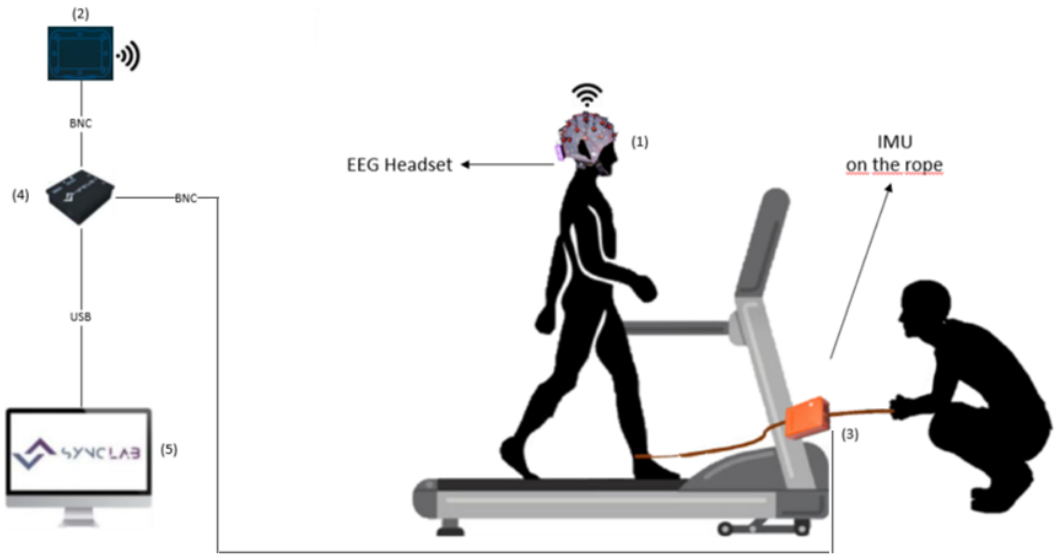


Figure 8: Experimental setup for balance loss data collection. (1) EEG headset. (2) EEG base station to capture the signals from the headset and send to SyncLab. (3) IMU to record the acceleration regarding to pulling rope. (4) Sync Box to synchronise the base station and the IMU. (5) PC with SyncLab to collect and process the sync data.

The experimental protocol was carried out involving two different treadmill speeds: 1.6 km/h and 2.5 km/h. The subjects walked on the treadmill, and the protocol was repeated with each subject, only changing the speed. As an attempt to avoid any distraction, the subjects wore noise reducing earplugs and had to look in a specific object while maintaining focus. They also wore a neck brace to minimise the artifacts that head movements provoke (Figure 9).

The subjects were given instructions to handle unexpected slip-like perturbations while walking on the treadmill (flat – 0°). They were also required to walk on the treadmill for a duration of 20 to 30 minutes at each speed. To avoid any prior bias in their brain activity response, they were not informed about the specific protocol. During the experiments, a hidden rope was attached to the subject's shoe at the moment their heel made contact with the treadmill belt. The subjects were unable to expect perturbations as they could not see the trained operator while walking. A second person manually marked these perturbations, which were later used to synchronise the data with the IMU recordings. Since the rope remained connected to one of the subject's feet throughout the two trials, the subjects were unaware of whether a perturbation would occur or not.

During each experimental session, which included both speeds, perturbations were introduced by the operator. In the beginning, subjects walked for about 5 minutes to get used to the movement. After that, 40 perturbations were randomly delivered at various moments while the subjects were walking. These perturbations were interspersed with non-perturbation phases, and each phase had a duration ranging from 20 to 40 seconds to increase the level of unpredictability.

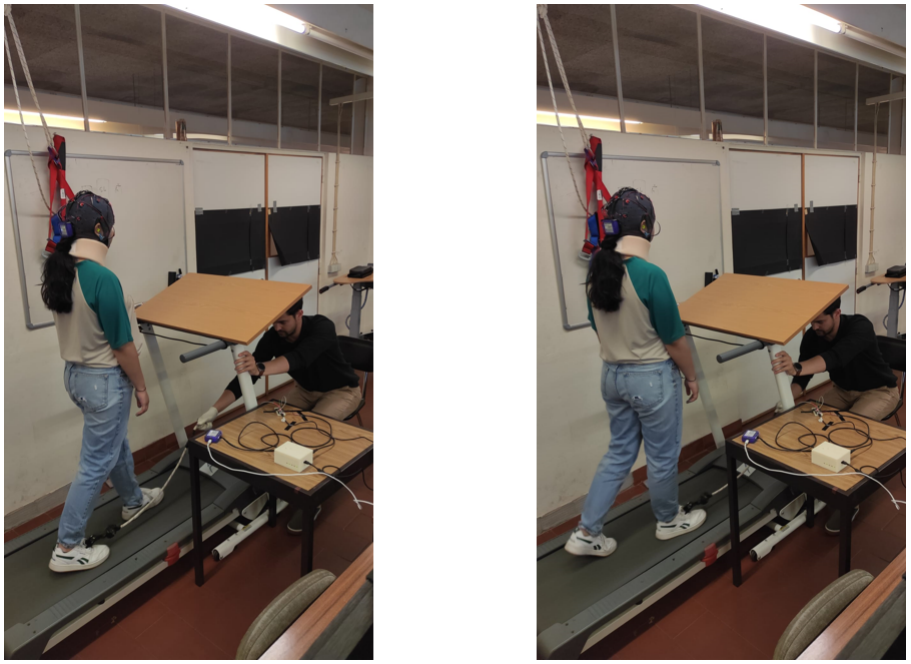


Figure 9: Experimental protocol - slip-like perturbations.

4.3.1.2 ADLs protocol

The development of the second data acquisition experimental protocol was a collaborative effort involving a team of experts specialising in the field of geriatrics. Notable figures in this collaborative endeavour included Dr. Gorjão Clara, an accomplished doctor with extensive expertise in gerontology, and Marta Sousa, who holds administrative and care-giving roles at the *Fundo Social de Braga* nursing home. The primary objective of this collaborative effort was to intricately define a set of inclusion and exclusion criteria, specify the ADLs to be undertaken by the selected participants, and meticulously delineate the conditions under which data acquisition would transpire.

In accordance with the outlined exclusion criteria, certain groups of individuals were excluded from participation. This included seniors grappling with conditions such as dementia (Mini Mental State Examination (MMSE) $\leq$ 15), severe depression (Geriatric Depression Scale (GDS) $\geq$ 11), substantial physical limitations that rendered them unable to walk unassisted (Functional Ambulation Categories (FAC) $<$ 4), orthopedic, cardiac, or respiratory disorders with an impact on locomotion, morbid obesity (Body Mass Index (BMI) $\geq$ 30), and a moderate to high level of fear regarding falling (Short Falls Efficacy Scale – International (FES-I) $\geq$ 14). The selection of activities, as presented in Figure 11, was thoughtfully curated to align with the specific needs and physical constraints of the elderly target audience.

Subsequently, several activities were meticulously executed (further described in Chapter 3), and inertial data was acquired under strictly controlled conditions to ensure the utmost safety and well-being of all subjects. Throughout the execution of these activities, each subject was accompanied by dedicated caregivers and the individual responsible for data collection. A comprehensive array of tests and clinical assessments, as mentioned earlier, was administered by a team comprising clinicians, nurses, and caregivers from nursing homes. This multifaceted approach aided in the identification of potential subjects suitable for data collection.

Following the demographic data collection process, the initial step in inertial data acquisition involved the precise placement of a waistband and the EEG headset on the subjects. The waistband was meticulously positioned to ensure the alignment of the inertial sensor with the spinal axis, located in the central region of the subjects' waist (Figure 10).

Each subject embarked on a circuit comprising activities outlined in Figure 11 and detailed in Table 9. This circuit was executed three times by each subject at a pace commensurate with their comfort level. Each individual activity within the circuit was sustained for a minimum duration of 2 minutes, with exceptions made for specific transitions such as Sit-to-Stand, Stand-to-Sit, and Lay Down.

Figure 10: Experimental protocol with an elderly performing walking.

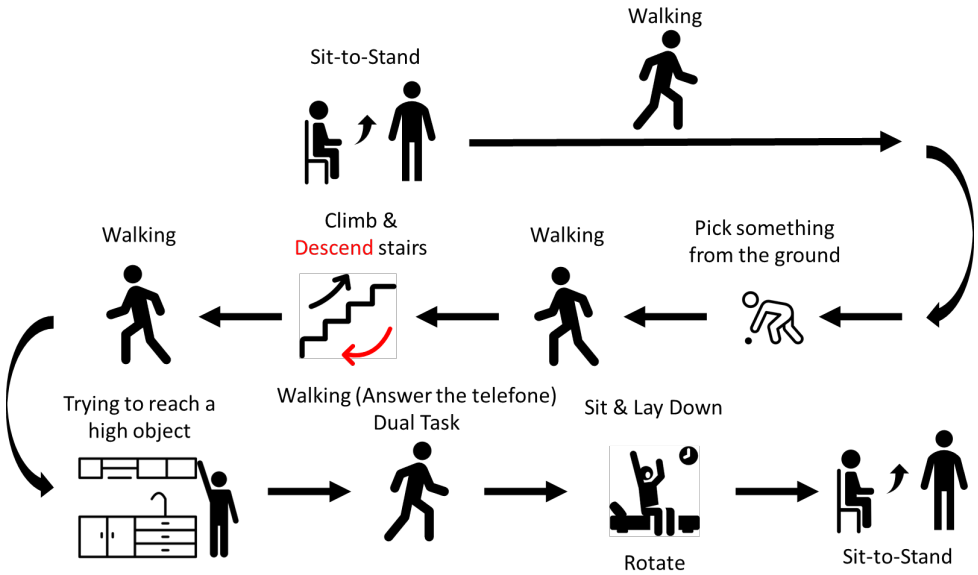


Figure 11: Experimental protocol circuit to collect ADLs.

Table 9: ADLs that elderly subjects performed under controlled trials

Code	Activity
ADL001	Walking
ADL002	Standing
ADL003	Walking upstairs
ADL004	Walking downstairs
ADL005	Sit in a regular height chair, wait a moment, and up
ADL006	Sitting a moment, lying, wait a moment, and sit again
ADL007	Being on one's back change to lateral position, wait a moment, and and change to one's back (Lying)
ADL008	Standing, pick something from the ground, and getting up
ADL009	Get into and out of bed without sitting

Throughout the circuit, vocal commands were issued by the test administrator to prompt subjects to transition between activities, with each activity's start time meticulously recorded within the mobile application to capture timestamps. Furthermore, if any subject displayed signs of fatigue, the test was promptly paused to allow for sufficient recovery and ensure the safety and well-being of the subjects.

4.3.2 Data Processing

4.3.2.1 Slip-like perturbations protocol

The slip-like perturbation events were automatically recorded by an IMU, and the magnitude of the acceleration vector was utilised to indicate the occurrence of provoked perturbations at both speeds. To identify and eliminate any spikes that could represent erroneous situations, such as dropping the rope or the rope hitting the treadmill, the manual recording of events using SyncLab was suggested as a solution to make sure that the labelling process is correct and it will not indicate misleading data. Since the clicks are manual, there is a certain delay compared to the beginning of a perturbation. For that matter, a MATLAB script was developed to help synchronise these events, enabling their adjustment and validation before initiating the actual data processing stage. Also, it was essential to cross-reference the time recorded in the dataset with the time recorded by IMU and make necessary adjustments, including re-sampling the IMU data since its frequecy is not exactly the same as the EEG device.

An example of the result from one trial is depicted in Figure 12. The orange lines represent the manual recordings of a perturbation. The yellow crosses at the bottom of the graph represent the exact moment of a perturbation, which corresponds to a local minimum. On the other hand, the orange crosses indicate the existence of a perturbation through the local maximum detected.

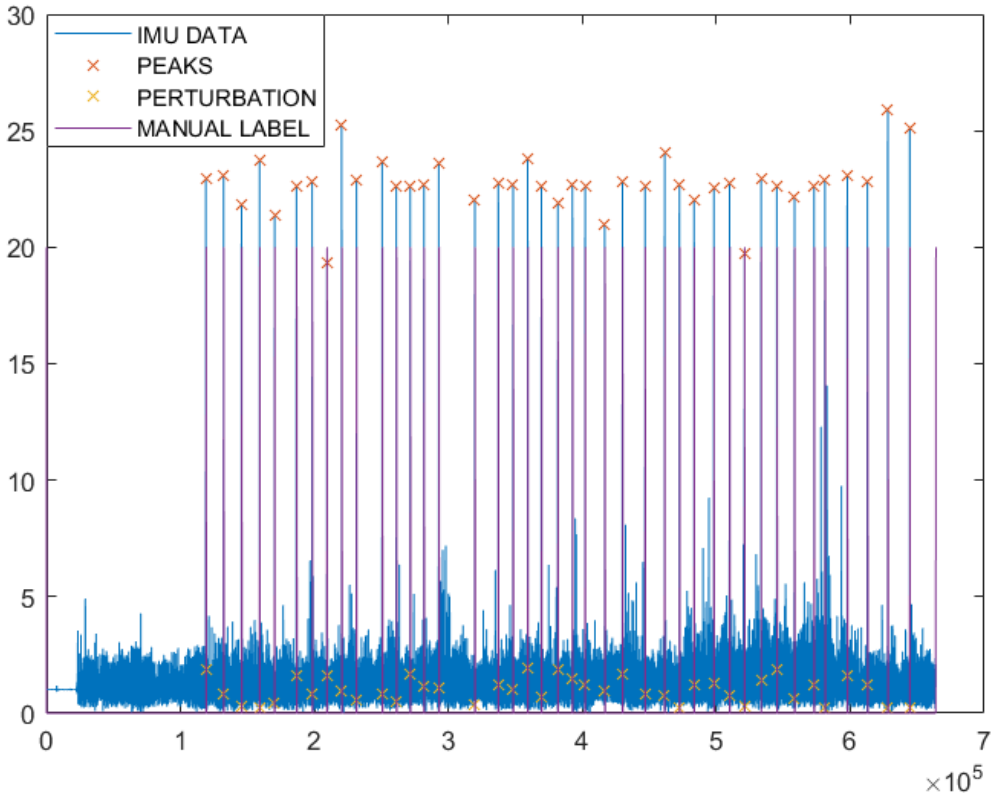


Figure 12: Manual labels and IMU acceleration magnitude for label correction using MATLAB.

Subsequent to the data acquisition, the EEG data underwent processing using custom-made scripts within the EEGLAB toolbox, executed in MATLAB. The EEG dataset comprised 21 rows: row 1 represented time (with a resolution of 2 ms per register), rows 2-17 contained EEG channels, row 18 was left blank as it presents the battery level, row 19 encompassed manual events related to perturbations, row 20 was related to actual events related to perturbations, and row 21 contained the peak magnitudes related to perturbations. Each column in the dataset corresponded to the values obtained from brain signals captured by the electrodes, directly correlating with the trial's duration.

Due to the susceptibility of EEG signals to noise and their non-stationary nature, EEG recordings can be viewed as a combination of independent cerebral sources and non-cerebral sources known as artifacts (e.g., ocular, muscular, etc.). These artifacts are considered disruptions in brain electrical signals, necessitating a pre-processing procedure to eliminate them in order to facilitate the extraction of pertinent event-related information. Thus, the initial processing step involved removing irrelevant attributes (rows 1, 18, 19, and 21), as 2-17 were selected for processing, while the registers in row 20 were imported as events. Figure 13 resumes the steps made to process EEG data with EEGLAB.

To eliminate low-frequency drift, a 4th-order Butterworth filter with a high-pass cutoff frequency of 0.2 Hz was applied [19, 67, 76, 77]. Additionally, to attenuate line noise at 50 Hz, a notch filter was utilised

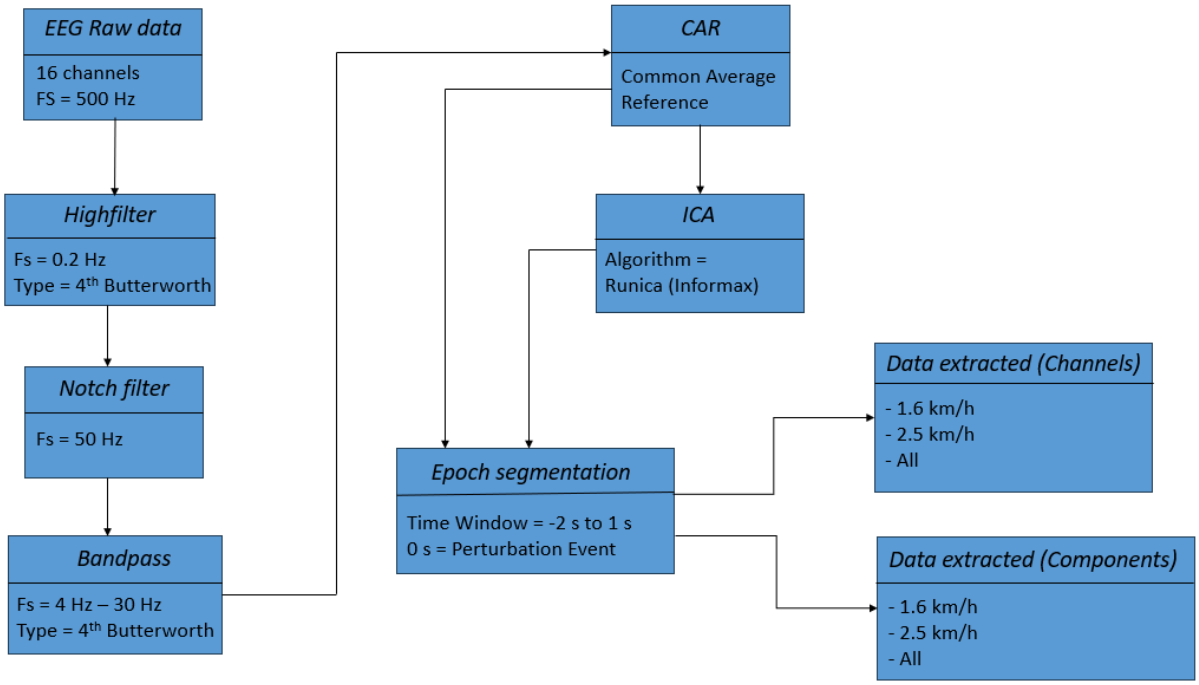


Figure 13: EEGLAB processing pipeline.

[12]. However, even after implementing these filtering steps, it was not possible to identify patterns related to PEPs. As a result, a band-pass filter ranging from 4 Hz to 30 Hz was employed to remove low frequencies associated with motor imagery or stable movements and frequencies above 30 Hz that could introduce muscular artifacts related to natural walking and postural recovery following a loss of balance [62].

Next, as part of the pre-processing procedure, the processed EEG signal was re-referenced using CAR. This method is commonly used in EEG to enhance the identification of small signal sources in highly noisy recordings [1]. The re-referenced EEG signal was then prepared for ICA, a technique used to extract independent components representing distinct EEG source signals. Following the application of the Infomax [7] algorithm, both channels obtained after CAR re-referencing and components extracted after ICA were further segregated into separate datasets for subsequent analysis.

Previous studies [17, 50, 62, 81, 83] indicated that components centered around Cz and distributed over frontal, central, and parietal areas (FCz, C3, C4, and CPz) exhibit a strong correlation with balance loss detection. However, since the EEG headset used did not include frontocentral and centroparietal electrodes, and the electrodes were fixed, four different approaches were further tested in Chapter 5: i) using all channels, ii) removing all channels except Fz, C3, Cz, C4, Pz (which are closer to the central area), iii) using Fz and Cz, as previous trials showed similar responses related to provoked perturbations as reported in the literature [17, 50, 62, 81, 83], and iv) using only Cz.

To study the event-related dynamics of continuously recorded EEG data, data epochs time-locked to events of interest were extracted. In the case of perturbation events, the EEG signals were segmented into

epochs ranging from -2 seconds to +1 second relative to the perturbation onset, with 0 ms included in the window, representing the first sample at the perturbation onset. Additionally, due to poor data quality, one subject was excluded from the 1.6 km/h dataset analysis.

4.3.2.2 ADLs protocol

The labelling of ADLs was facilitated through the utilisation of a smartphone application, which had been developed by members of the BiRDLab team. This application, presented in Figure 14, was purposefully designed to streamline the data labelling process by enabling annotators to effortlessly label activities through a series of simple clicks. During the data acquisition process, this application would save the initiation time of various ADLs into a text file. Importantly, this text file's name was customised, allowing users to adapt it to their specific needs or preferences.

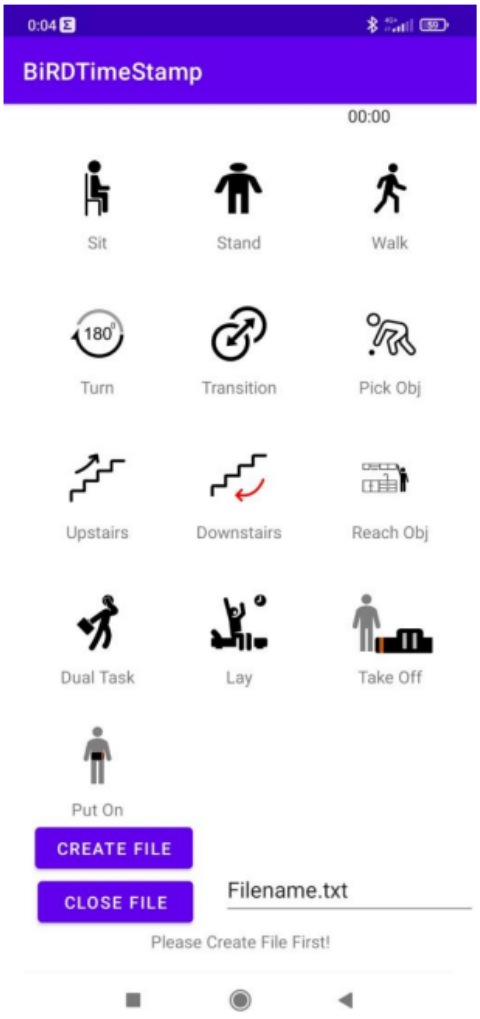


Figure 14: ADLs labelling smartphone application.

Additionally, the application featured a built-in timer functionality, which served as a valuable tool for the person responsible for data collection. This timer function allowed them to precisely monitor and record the duration of each activity, ensuring that the timing information was both accurate and reliable. The integration of such features within the application contributed to the efficiency and accuracy of the data labelling process, ultimately aiming to enhance the quality of the labelled dataset.

The following processing steps, which involved the use of EEGLAB, closely mirrored the procedures outlined in 4.3.2.1, with one notable exception. In this case, the epoch segmentation step was deemed unnecessary due to the specific characteristics of the data. As a result, two subsequent datasets were constructed by extracting the CAR and ICA data from each subject. This streamlined approach ensured that the data preparation remained consistent with the previous chapter, with due consideration for the unique requirements of this particular dataset. Following this, all datasets underwent normalisation using the minmax method. The normalisation procedure was tailored to each individual subject. Class weights were also applied since the created datasets were imbalanced.

4.4 Comparative Analysis

In this sub-chapter, the objective is to introduce the deep learning models and their associated hyper-parameters, setting the stage for a comprehensive comparative analysis. Our primary aim is to determine the most effective model and identify the optimal hyper-parameter configurations through a meticulous evaluation. To achieve this, we will employ two distinct methods within our evaluation strategy. The first phase of the evaluation strategy was designed to fine-tune the models and its associated hyper-parameters, ensuring optimal performance on each dataset. The process involved a utilisation of the Hold-out (HO) method, which facilitated the exploration of diverse hyper-parameter combinations while maintaining a keen focus on performance metrics [8]. To achieve this, the dataset was partitioned into three distinct subsets: training, validation and testing.

With this partitioning in place, each machine learning model was executed following the predefined hyper-parameters, which were previously established through a thorough exploration process. This rigorous approach ensured that each model variant was trained on a consistent subset of subjects, thereby eliminating any potential bias that could skew the performance assessment. Having optimised our machine learning model through the preceding phase, we progressed to the Leave-one-out (LOO) Phase, a robust evaluation approach that validated the generalisation capabilities of our model across a diverse set of subjects [20]. This phase allowed us to scrutinise the model's performance while systematically varying the testing subject, offering a more comprehensive understanding of its predictive prowess.

This method consisted on the execution of ten iterations, each corresponding to a unique testing subject. For each iteration, the model was trained on the amalgamation of subjects designated for training

and validation, thereby incorporating the knowledge distilled from diverse subsets of the dataset. Subsequently, the model was evaluated against the subject earmarked for testing. This process yielded ten distinct assessments, effectively encompassing the entire dataset as the testing subject in a controlled and non-repetitive manner. By iteratively excluding one subject at a time and employing the previously optimised hyper-parameters, we garnered a holistic perspective on the model's ability to generalise and adapt to new subjects, thereby enhancing its credibility and reliability.

To ensure a rigorous and equitable comparison between the results obtained from each method, we relied on a comprehensive set of performance metrics. These metrics provide us with a clear and objective basis for evaluating the performance of our models, allowing us to draw meaningful conclusions from the comparative analysis of both protocols.

4.4.1 Classification Models and Architecture Parameters

The rapid progress and advancements in DL techniques have opened up promising opportunities in the analysis of gait for controlling or potentially controlling active lower limb exoskeletons and orthoses. DL algorithms have shown great potential in handling complex and large-scale data, which is crucial in gait analysis due to the multifaceted nature of human movement. By leveraging deep learning, researchers and engineers can extract meaningful patterns and features from gait-related data, enabling more accurate and robust control of assistive devices.

These technologies offer the potential to enhance mobility and independence for individuals with mobility impairments or those undergoing rehabilitation. As deep learning continues to advance, the future of active lower limb exoskeletons and orthoses holds great promise in contributing to improved mobility and quality of life for a diverse range of users [59, 72]. Using the literature analysis carried out in Chapters 2 and 3, four distinct neural network architectures were chosen and adapted to the problem of this dissertation. Thus, architectures of CNNs [62] and LSTMs [24, 25, 79] were built. A study also used Transformers to presents an activity recognition system [69], supplying a foundational framework for training models to recognise activities. For that reason, an architecture of Transformers were added to the deep learning model set. In addition, even not mentioned at the state-of-the-art, a CNN-LSTM architecture that involves layers of CNN for feature extraction on input data combined with LSTMs to support sequence prediction was employed and tested during the development of this thesis [6].

The data described in 5.2.1 allowed the extraction of features that was accomplished with the sliding window method, where a signal is split into a determined number of equal-sized windows (subgroups), from which different attributes can be evaluated. The window size depends on the resolution of activity recognition. In order to analyse long-term activity patterns, a stream of activity data may be processed as one by employing a large window size. However, if numerous actions are crammed into one window, it becomes unable to detect balance loss events. When dealing with perturbations that occur in a relatively

short period of time, a smaller window size may be more appropriate. The hyper-parameters are values that control the learning process of neural networks [28]. These parameters are set before the training process begins and can significantly impact the performance of the model.

Below are the main hyper-parameters and their descriptions: i) The batch size determines the number of samples considered for updating the internal model parameters (weights and biases) before performing a gradient descent update; ii) The number of epochs refers to the number of times the learning algorithm runs on the entire training dataset. It controls how many times the neural network's weights are updated during training. Using an appropriate number of epochs is crucial to avoid underfitting or overfitting; iii) The optimisation algorithm is responsible for adjusting the network's weights and biases during training to minimise the loss function. The loss function quantifies how well the model's predictions match the actual targets, and the optimiser's goal is to minimise this loss; iv) Dropout is a regularisation technique used during training to randomly deactivate a portion of neurons in a neural network. This helps prevent overfitting by encouraging the network to rely on more robust features and avoids over-reliance on specific neurons; v) The number of neurons in a hidden layer determines the capacity and complexity of the neural network. More neurons can increase the network's representation power, but it also introduces the risk of overfitting; vi) Cross-entropy loss, also known as log loss, is commonly used in classification problems where the model's output is a probability value between 0 and 1. It quantifies the difference between the predicted probabilities and the actual labels, and as this difference increases, the cross-entropy loss increases as well. Minimising the cross-entropy loss during training improves the model's accuracy.

In some studies, ANNs can work with raw data as input, but other studies opt to use previously extracted features as inputs [17, 62, 79]. To conduct a comparative research analysis on how different neural network architectures respond to the same type of data, four feature subsets were used as inputs during the training and testing of the constructed architectures.

Throughout all operations, test specifications such as the loss function, number of epochs, optimiser, number of hidden layers, dropout rate, and learning rate were kept constant for all architectures based on the findings of Ravindran et al [62]. However, different batch sizes were used to train the models [62, 79]. The following table provides a summary of these characteristics and their respective values.

The neural networks used in this study were implemented with the underlying architecture without employing internal layer optimisations. This approach allowed for a comparison of the network's performance while exploring the impact of different input feature sets and batch sizes.

4.4.2 Model Evaluation Metrics

Evaluating a model is a crucial task in building an effective AI-based model. The metrics used for evaluation measure the quality and performance of the model. In the case of binary classification tasks, such as the ones used in the slip-like perturbations study, the models can only produce two possible results. The

supervised deep learning approach used in this work follows these steps: i) Fit a model on the training data, where the model learns to make predictions based on the input data and the corresponding labels. ii) Test the model on the testing data, where the model's predictions are compared to the true values (correct labels) of the observations. These metrics allow for a comprehensive assessment of the model's performance and can be used to calculate important metrics such as the ones that follow, offering insights into the model's effectiveness in handling the binary classification task.

To measure the effectiveness of the classification model, the generated classification outcomes are compared with the actual values of the provided observations. This comparison is done using a confusion matrix. The confusion matrix is a table that summarises the classification model's performance in predicting instances from different classes. It provides an organised way to map the predictions to the original classes to which the data belong. This matrix has two axes: one for the expected label (actual class) and one for the predicted label (model's predicted class): i) True Positives (TP): The number of instances correctly predicted as positive (perturbation); ii) True Negatives (TN): The number of instances correctly predicted as negative (non-perturbation); iii) False Positives (FP): The number of instances incorrectly predicted as positive when they are actually negative; iv) False Negatives (FN): The number of instances incorrectly predicted as negative when they are actually positive.

4.4.2.1 Accuracy

Accuracy is a foundational metric in machine learning that measures the correctness of predictions made by a model. It plays a vital role in model evaluation, selection, and hyper-parameter tuning, making it a critical aspect of building successful machine learning solutions. However, it is essential to consider the context and potential limitations, particularly when dealing with imbalanced datasets, to ensure a well-rounded evaluation of model performance. Accuracy measures how often the model correctly predicts both positive and negative instances in the dataset. The accuracy is calculated using the formula:

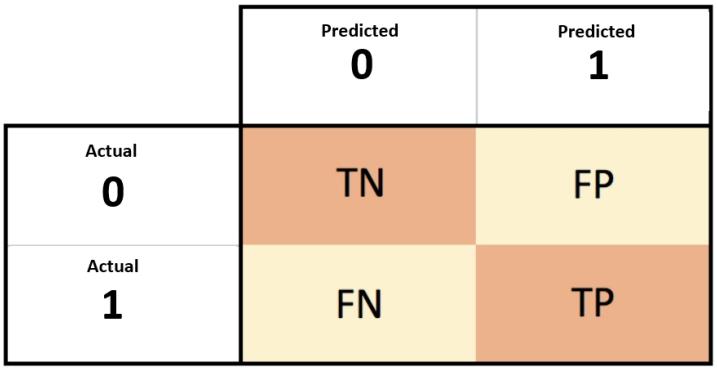


Figure 15: Confusion Matrix. Obtained from https://subscription.packtpub.com/book/data/9781838555078/6/ch06lvl1sec34/confusion-matrix.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

However, this may not be the most suitable metric for imbalanced datasets, where one class significantly outweighs the other, as it can be biased towards the majority class.

4.4.2.2 Recall (Sensitivity)

Recall (Sensitivity) is an essential performance metric in machine learning, particularly in situations where correctly identifying positive instances is of paramount importance. It plays a crucial role in applications dealing with imbalanced datasets, medical diagnostics, and anomaly detection. By focusing on recall, it is possible to build models that are more sensitive to positive instances, reducing the risk of missing critical cases and ultimately enhancing the reliability and effectiveness of the system. However, it is crucial to consider the trade-offs between recall and other metrics, such as precision, depending on the specific requirements of the task at hand. While high recall is desirable in certain scenarios, it may come at the cost of increased false positives. In cases where false positives can have severe consequences or significant resource implications.

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

4.4.2.3 Specificity

Specificity (true negative rate) is a vital performance metric, especially in scenarios where avoiding false positives or false alarms is crucial. It complements recall (sensitivity) in providing a comprehensive evaluation of the model's ability to correctly identify both positive and negative instances. By understanding specificity, it is possible to build models that are more specific in detecting negative cases, reducing the risk of false positives and enhancing the model's overall reliability and effectiveness. However, it is essential to strike a balance between specificity and recall, considering the specific needs and priorities.

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

4.4.2.4 Precision

Precision is a fundamental metric, particularly in scenarios where avoiding false positives is critical. It complements other evaluation metrics like recall and specificity, providing a comprehensive understanding of a model's predictive capabilities. With the help of precision, models can be more precise in identifying positive cases, reducing the risk of false positives and enhancing the model's overall reliability and effectiveness. A high precision value close to 100% indicates that the model has a low rate of false positives. It

means that when the model predicts an instance as positive, it is highly likely to be correct. This implies that the model is precise in identifying positive cases. Conversely, a low precision score suggests that the model is making a significant number of false positive predictions, leading to a lack of precision in positive classifications.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

4.4.2.5 F1-Score

The F1-score is mostly applied in scenarios with imbalanced datasets and applications where both false positives and false negatives are of significance. By taking into account both precision and recall, the F1-score offers a balanced evaluation of a model's performance. Machine learning models can utilise the F1-score to select appropriate models, understand the trade-offs between precision and recall, and build models that effectively handle various real-world challenges. The F1-score ranges between 0 and 1, where 1 represents a perfect F1-score, indicating both high precision and high recall. A higher F1-score signifies a model that strikes a good balance between precision and recall, effectively handling both false positives and false negatives. Conversely, a lower F1-score suggests that the model is biased towards either precision or recall, or it is not performing well in correctly classifying positive instances.

$$F1 - Score = \frac{2 \times (Recall \times Precision)}{Recall + Precision} \tag{5}$$

4.4.2.6 MCC

The Matthews Correlation Coefficient (MCC) is a valuable performance metric especially in binary classification tasks with imbalanced datasets. It provides a balanced evaluation of a model's performance by considering all aspects of the confusion matrix. It can be used to compare models, especially in scenarios where precision and recall are equally important. However, it is crucial to consider the specific requirements and priorities of the application when selecting an appropriate evaluation metric for the task at hand. The MCC is defined as the correlation between the observed and predicted binary classifications. Mathematically, it can be represented using the confusion matrix elements:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{6}$$

The MCC ranges between -1 and +1, where +1 represents a perfect classifier, 0 indicates random predictions, and -1 indicates a classifier performing entirely opposite to the true classifications.

4.5 Outcomes

The global aging population is leading to a higher prevalence of neurological disorders such as dementia, Parkinson's disease, and cerebrovascular accidents. These conditions not only have various implications but also contribute to reduced mobility and an increased risk of falls, which often result in injuries. The impact of these challenges on healthcare systems worldwide is significant. Moreover, with the growing importance of ADLs in the lives of elderly individuals, understanding their interaction with neurological conditions is crucial. In light of these challenges and the need for alternative strategies to prevent falls, this dissertation aims to gather quantitative information by analysing EEG data. The goal is to train AI-based models to identify slip-like perturbations and ADLs and classify them, thus capturing the natural human response to these events. This research endeavours to contribute to a deeper understanding of the interplay between ADLs and falls, ultimately paving the way for innovative strategies to enhance the well-being of the aging population.

This work focuses on three main phases: (i) the analysis of slip-induced experimental data, which involves the application of best practices for collecting EEG data and pre-processing techniques to remove artifacts and select relevant features as discussed in the previous chapter. The objective is (i) to detect the loss of balance during treadmill walking, which will be further discussed in Chapter 5; (ii) the decoding of perturbation events for fall prevention robotic devices; (iii) the recognition of ADLs, which will be further discussed in Chapter 6; (iv) a comparative analysis from the outcomes of both studies. The AI framework developed by the BiRDLab team was adapted for this project, and the selection of DL architectures was based on the literature review presented in Chapters 2 and 3, where DL-based models and their computational performance will be evaluated and discussed in Chapters 5 and 6.

4.6 Conclusions

The outcomes of the analysis of experimental data is a crucial step in identifying the signature of loss of balance and a certain ADL in EEG signals, shedding light on the brain's response to slip-like perturbations and ADLs. This analysis will provide a deeper understanding of how the brain signals respond to slip-like perturbations across various subjects, gait speeds, and ADL scenarios. It aims to address the gaps identified in the existing literature, where, for instance, studies on loss of balance predominantly focused on experiments involving subjects standing and experiencing postural reactions. This research pursuits to bridge the knowledge gap by investigating the neural correlates of falls and ADLs, offering valuable insights into fall prevention and ADL recognition. This dissertation aims to bridge this gap and compare the effectiveness of AI-based models while considering their computational performance. The outcomes of this evaluation will be discussed in Chapters 5 and 6, as well as an associated comprehensive analysis.

Chapter 5

Slip-like Perturbations Detection using Brain Signals

5.1 Introductory Insight

This chapter focuses on the utilisation of AI-based models to detect the presence of perturbations in EEG data collected in the first experimental protocol described in Chapter 4, including all the pre-processing steps that led to the creation of six datasets. The objective is to decode the brain signals in response to slip events and determine if there was a loss of balance during walking. By utilising previously collected fall data and specifying the characteristics of slip-like perturbations, the AI-based models are trained to classify the presence of a perturbation in the brain signal.

It is worth noting that studies examining perturbations using an EEG paradigm, as well as the temporal relationship between different signal modalities, are scarce. Moreover, based on the literature review conducted in Chapter 2, there is are few studies that have evaluated the influence of balance perturbations on EEG specifically during walking. Furthermore, only very few studies have developed AI-based models to identify instances of balance loss. This highlights the novelty and significance of this research in addressing these gaps in the existing literature.

5.2 Methods

In this section, we will discuss the methods used to validate and recognise the LOB induced by perturbations, which are the focus of this study. The following methods will be covered: i) labelling the data into two classes, perturbation and non-perturbation, based on the segmentation of data mentioned in the literature review presented in Chapter 4; ii) applying AI-based models to the datasets and specifying the parameters used for training these models; iii) describing the strategies employed to split the data into training,

validation, and test datasets; iv) discussing the metrics used to assess the quality and performance of the developed models. By covering these methods, we aim to provide a comprehensive understanding of the validation and recognition process of LOB caused by perturbations in this study.

5.2.1 Data Labelling

After data segmentation, the first step was the labelling process. Each sample obtained from the segmented data was assigned a label to identify it as either a perturbation or a non-perturbation event. Following the approach suggested by Ravindran et al. [62], Class 0 (non-perturbation) consisted of individual trial windows during the baseline period, specifically from 1200 ms to 500 ms prior to the onset of the perturbation. On the other hand, Class 1 encompassed EEG segments ranging from -200 ms to +500 ms post perturbation onset. These defined classes allowed for the differentiation of EEG data corresponding to perturbation and non-perturbation events. Figure 16 depicts an epoch example from an EEG trial where the two classes are evident, demonstrating the difference between them in a cortical level.

In order to label the samples obtained, a Python script was written and every sample that was not within the segments of Class 0 or 1 was discarded. The final datasets were extracted according to the speed and filters used. For each filter (ICA and CAR), there was a dataset for 1.6 km/h, 2.5 km/h and both speeds, resulting in a total of 6 datasets (Table 10). Every dataset contained all channels from every subject depending on the speed. The only exception is subject 4, whose 1.6 km/h trial was excluded.

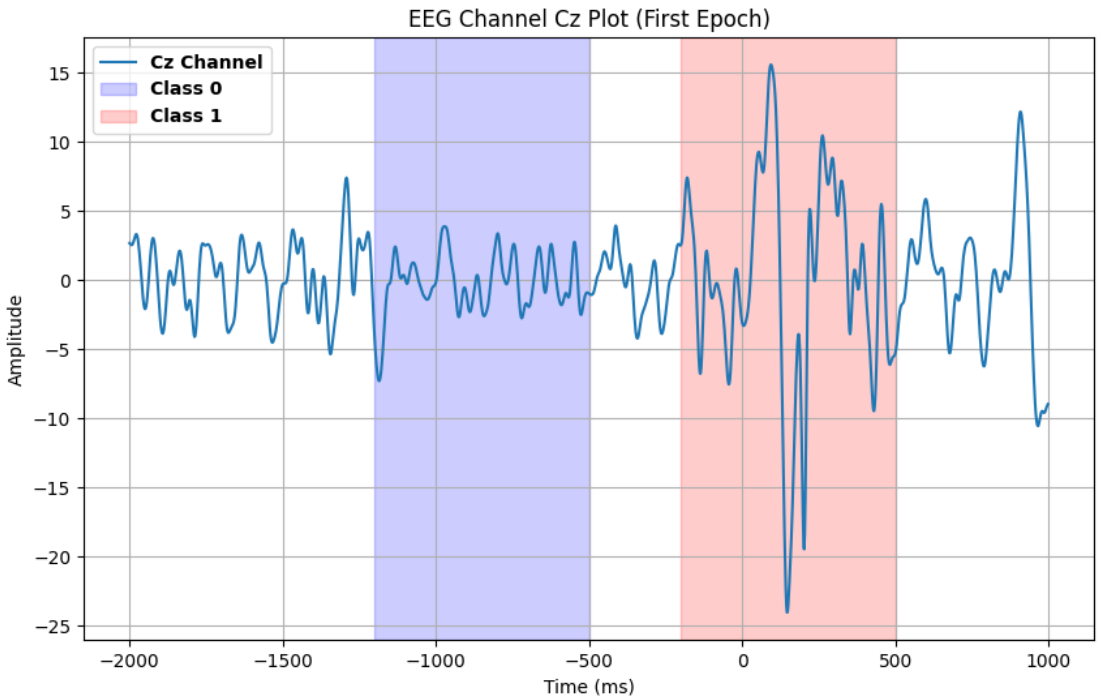


Figure 16: Epoch of Cz channel from EEG trial demonstrating classification process.

Table 10: Datasets generated from data labelling

Dataset	Treadmill Speed (km/h)	Channels/ICA
1_6_CAR	1.6	Channels
2_5_CAR	2.5	Channels
ALL_CAR	1.6 and 2.5	Channels
1_6_ICA	1.6	ICA
2_5_ICA	2.5	ICA
ALL_ICA	1.6 and 2.5	ICA

5.2.2 Architecture Parameters

Expanding on the hyper-parameters discussed in Chapter 4, which played a crucial role in configuring our deep learning models, Table 11 serves as an exhaustive reference point for the range of values that underwent extensive testing in the HO method. This comprehensive examination of hyper-parameter values was undertaken with the utmost care, as these parameters are instrumental in shaping the models' performance. Also, Table 12 depicts the division made for each dataset and the respective DL models used. CAR datasets have four different subsets, while ICA datasets only have one.

The choice of hyper-parameters for our study was guided by the need to adequately prepare our model for optimal performance in analysing new testing data. Consequently, the maximum number of training epochs was set at 100, ensuring that the model received sufficient training. Additionally, we considered the temporal characteristics of the data. The study [62] revealed that PEP components emerge 75–137 ms after a perturbation occurs. Consequently, the window size or Sequence Length (SL) was set at 100 ms to align with these findings. It is worth noting that future investigations could explore different window sizes, as [62] previously used a window of 200 ms.

To maintain consistency and comparability with prior research [62, 79], certain test specifications were held constant across all model architectures. This included the choice of loss function, optimiser, and learning rate. Conversely, variability was used in batch sizes, dropout rates, and overlap values during model training. This variation allowed the assessment of their effects on model performance and the gaining of a deeper understanding of how these hyper-parameters influence the study's outcomes.

5.2.3 Model Building and Evaluation

Regarding the first HO iteration, the split was configured as follows: 70% of the data for training, 20% for validation, and the remaining 10% for testing. This partitioning facilitated a balanced allocation of subjects, wherein seven subjects were earmarked for training, two for validation, and one for testing. The specific distribution entailed subjects 1, 2, 3, 4, 5, 6, and 10 for training, subjects 8 and 9 for validation, and subject 7 for testing. In the case of LOO, each subject was systematically left out as part of the testing set

Table 11: Hyper-parameters for Deep Learning models - Slip-like perturbations

Hyper-parameters	Value(s)
Number of Epochs	100
Window Size	50
Batch Size	32, 64, 128, 256
Dropout	20%, 50%
Overlap	50%, 80%
Optimiser	Adam
Loss Function	Cross Entropy Loss

Table 12: Subset of features per dataset and Deep Learning Models - slip-like perturbations protocol

Dataset Type	Channels/ICs	DL Models
CAR	All Channels Fz, C3, Cz, C4, Pz Fz, Cz Cz	CNN LSTM CNN-LSTM Transformer
ICA	All Components	CNN LSTM CNN-LSTM Transformer

in one iteration to ensure that all subjects were thoroughly tested, resulting in a total of 10 iterations per execution. This methodology allowed for a comprehensive evaluation of the model's performance across the entire subject population.

After completing the initial LOO iteration and assessing the models' performance across all subjects, we observed that the first three subjects were recorded using the gTec gel, while the remaining subjects were recorded with an alternative gel named Spectra 360. Recognising the potential impact of this change, we decided to conduct a second iteration of the evaluation process. In this second iteration, the focus was exclusively on the model variant that demonstrated the best performance during the initial LOO Phase. The purpose of this secondary evaluation was to isolate the effect of the gel type on the model's predictive prowess. Additionally, we made adjustments to expedite the evaluation process by reducing the hyper-parameters combinations by half, ensuring faster execution while still maintaining the analysis integrity.

The second iteration of both the HO and the LOO method followed a similar structure to the initial phase, with a respective partitioning of the dataset into training, validation, and testing subsets. However, this time, we excluded subjects 1, 2, and 3 from both training, validation, and testing due to the unique use of the gTec gel in these individuals. The objective was to isolate and evaluate the specific influence of the gTec gel in comparison to the new gel's performance. Subjects 4 and onwards continued to be allocated as previously described. For the second HO iteration, the data split was arranged as follows: Subjects 5, 6, and 10 were included in the training set, Subjects 4, 7, and 8 formed the validation set, and Subject 9 was

designated for the testing set. By performing this secondary iteration and adjusting the hyper-parameter values, we aimed to gain clear insights into the impact of gel type on the model's performances. This additional step in our evaluation process allowed us to refine our understanding of the model's behaviour and enhance the credibility of the results.

5.3 Results

This sub-chapter presents the final results obtained using the four deep learning architectures described in Section 4.4. The results from different iterations and methodologies are detailed in the appendices: A (First Iteration of Hold-Out), B (First Iteration of Leave-One-Out), C (Second Iteration of Hold-Out), and D (Second Iteration of Leave-One-Out). In the first iteration of HO, the focus is on the top-performing five combinations, presenting their results. For the first iteration of LOO, we provided the mean and standard deviation for each model's performance. Regarding the second iteration of HO, we narrowed our analysis to the two best runs of the most promising deep learning model from the first HO iteration. Lastly, in the second iteration of LOO, we report the mean and standard deviation, along with the performance of each subject using the best model identified in the first LOO iteration.

5.3.1 Hold-out - first iteration

5.3.1.1 CAR

In the initial series of experiments conducted at a speed of 1.6 km/h, utilising all 16 channels, the best performance results were obtained with the CNN-LSTM model (Accuracy: 0.73, F1-Score: 0.72, Precision: 0.73, MCC: 0.46, Recall: 0.72, Specificity: 0.72). The best combination of hyper-parameters was Epochs (E)=100, Batch size (B)=128, Overlap (O)=80 and Dropout (D)=0.5. Generally, all models had similar behaviour in terms of performance, although CNN-LSTM stood out both on the best run overall.

Reducing the number of channels to 5, it is possible to observe higher results, as this time CNN presented the best results (Accuracy: 0.7613, F1-Score: 0.7521, Precision: 0.7954, MCC: 0.5643, Recall: 0.7708, Specificity: 0.7708). The best combination of hyper-parameters was B=32, O=50 and D=0.2. The CNN-LSTM model also improved its results, whilst both LSTM and Transformer kept presenting accuracy values below 70%, as observed with 16 channels.

In a further reduction of channels to 2, the results presented slightly lower accuracy values. The best run was executed using the CNN model again (Accuracy: 0.7574, F1-Score: 0.7454, Precision: 0.8119, MCC: 0.5807, Recall: 0.7723, Specificity: 0.7723). The best hyper-parameter combination was, once again, B=32, O=50 and D=0.2. The similarity between CNN and CNN-LSTM were still very evident. The results in the remaining 2 models were relatively maintained comparing to 16 and 5 channels.

With the use of the Cz channel only, since it exhibits a strong correlation with balance loss detection, the results were surprising, as they reached approximately the same performance as using 5 channels. By a small margin comparing to CNN, this time the CNN-LSTM model managed to achieve the best results (Accuracy: 0.7671, F1-Score: 0.7572, Precision: 0.8267, MCC: 0.6034, Recall: 0.7805, Specificity: 0.7805), with the best parameters combination B=128, O=80, D=0.2. This helps to demonstrate the veracity of the current literature [17, 50, 62, 81, 83] about the influence of mainly the Cz channel on the balance loss detection. Thus, not only these channels have a strong correlation and cause a positive impact on the overall performance but also the remaining channels may even influence negatively, as it has been demonstrated that increasing the number of channels does not improve the final results. The differences can be noted from Table 20 to 23.

In the initial 1.6 km/h experiments, the use of 16 channels showed similar results comparing to the equivalent number of channels in the 1.6 km/h speed. CNN-LSTM was the model with the best results (Accuracy: 0.7392, F1-Score: 0.7385, Precision: 0.747, MCC: 0.4922, Recall: 0.7452, Specificity: 0.7452). All models achieved accuracy values of at least 70%, with CNN-LSTM and CNN continuing to stand out. With 5 channels, the results were relatively similar, with CNN-LSTM being once more the model with the best results (Accuracy: 0.7366, F1-Score: 0.7360, Precision: 0.7480, MCC: 0.4909, Recall: 0.7429, Specificity: 0.7429), with the parameters B=64, O=80 and D=0.2. Comparing with the same number of channels in the 1.6 km/h speed, these results could be lower because of the higher speed that subjects were walking at. However, analysing the overall performance from all models, there was a better similarity between them.

Observing the results of the utilisation of 2 and 1 channels, the performance metrics values are maintained even when reducing the number of channels. This demonstrates the consistency of the models CNN-LSTM and CNN, mainly. The 1.6 and 2.5 km/h executions were distinct in terms of overall performance as the higher speed presented lower results. Combining the two previously mentioned speeds did not result in a decrease in performance. In fact, the values remained relatively the same. Figure 17 shows that some metrics like precision were even higher with only one channel comparing to all 16.

In short, the results were expected to be better considering the goal that was set. This under-performance may be due to the gel used in the first three subjects, as mentioned in 6.2.3. Also, the set of training, validation and test may not be optimal, as the first three subjects were part of the training data, which could have decreased the final results. On the other hand, all datasets have shown a consistent level of performance throughout the decrease of the number of channels, which indicates the strong influence of channels like Cz and Fz. This can help improve the model's performance by eliminating unnecessary data and thus reducing execution time and memory spent.

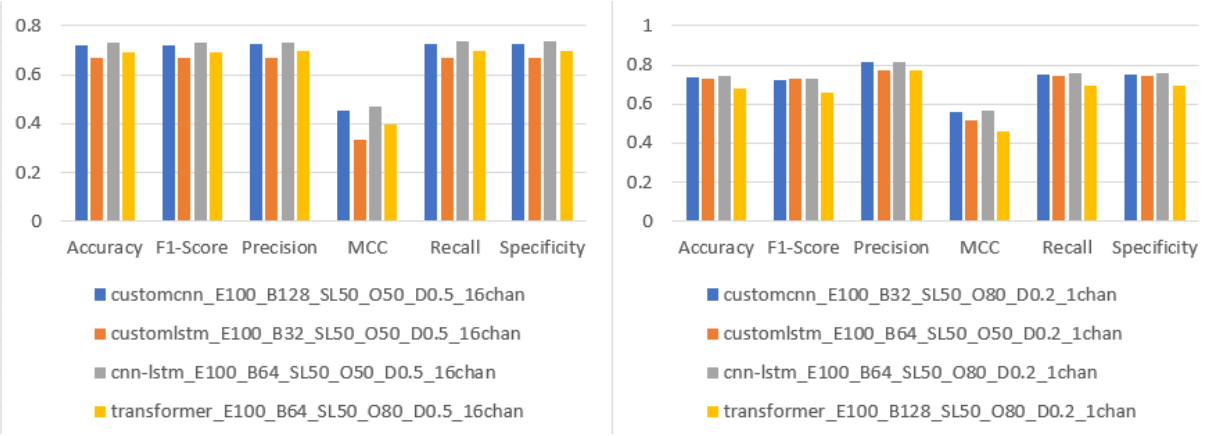


Figure 17: Comparative analysis of EEG channels: Hold-Out (1st iteration) with CAR datasets at different speeds (1.6 km/h and 2.5 km/h) using 16 (Left) and 1 (Right) EEG channels.

5.3.1.2 ICA

For the 1.6 km/h speed, the model that demonstrated the best behaviour was Transformer (Accuracy: 0.664, F1-Score: 0.658, Precision: 0.663, MCC: 0.321, Recall: 0.658, Specificity: 0.658), with the parameters being B=32, O=80 and D=0.5. The same happened for 2.5 km/h (Accuracy: 0.538, F1-Score: 0.35, Precision: 0.269, MCC: 0, Recall: 0.5, Specificity: 0.5), whereas this time the best parameters were B=32, O=50 and D=0.2. The combination of both speeds was similar (Accuracy: 0.53, F1-Score: 0.347, Precision: 0.265, MCC: 0, Recall: 0.5, Specificity: 0.5), where the parameters were the same as 2.5 km/h with the exception of the overlap value, which was 80.

The results suggest that ICA may not be recommended to this purpose, not only because it presented lower results than any of the CAR datasets despite requiring more processing resources, but also because the values obtained are far from intended. Even so, the gel used in the first 3 subjects may cause a significant influence in the overall results. Thus, ICA datasets were not discarded for the second iteration.

5.3.2 Leave-One-Out - first iteration

5.3.2.1 CAR

The models' accuracy levels demonstrated remarkable consistency, with variations not exceeding 4% between the highest and lowest values in each dataset (1.6 km/h, 2.5 km/h and the combination of both. The best results were observed with the use of 2 channels by CNN-LSTM, with an accuracy of 68%. Notably, the same underlying pattern observed during the initial hold-out iteration persisted throughout all channel combinations. This pattern indicated that reducing the number of channels consistently led to improved final results. However, it is worth noting that in comparison to the results discussed in 5.3.1.1, the results in this method exhibited a slight decline in accuracy. This difference may be attributed to the utilisation of distinct sets of training, validation, and test subjects.

In the context of the 2.5 km/h dataset, the accuracy values exhibited a range between 60% and 70%, approximately. Notably, all models displayed an enhancement in their performance as the number of channels decreased. Significantly, the CNN-LSTM model achieved its highest accuracy of 70.7% with 2 channels used. When compared to the results from the first dataset, the accuracy figures were notably higher. This observation underscores the idea that a speed of 2.5 km/h does not introduce significant motion artifacts that hinder the models' classification accuracy. Furthermore, the amalgamation of both speed levels within the same dataset did not pose a hindrance, as the results remained relatively consistent with those of the previous datasets.

5.3.2.2 ICA

In the context of the three datasets analysed, the accuracy values yielded remained consistently below the 59% threshold. This comparative evaluation highlights a notable disparity when juxtaposed with the results achieved through the use of the CAR technique. The observations suggest a departure from optimal performance, but intriguingly, the root cause does not appear to be rooted in the gTec gel application itself. A closer examination of the data indicates that the initial three subjects do not exhibit the lowest accuracy values, effectively ruling out the gel as a primary source of concern. As elucidated in 5.3.1.2, the underlying issue appears to be intricately linked to the efficiency of ICA when applied to this specific context. The somewhat sub-optimal outcomes underscore the inherent complexities associated with signal separation and noise reduction, particularly within the domain of EEG signal processing.

5.3.3 Hold-out - second iteration

5.3.3.1 CAR

The use of 16 and 5 channels in 1.6 and 2.5 km/h datasets, separately, displayed approximately the same results or even lower. With 16 channels, the accuracy improved by just 1% (73%), whereas with 5 the value was decreased by more than 10%. The main difference was with the use of 1 and 2 channels, where the accuracy was increased by more than 10%. The use of the Cz channel only, in 2.5 km/h, for instance, went from 73% to 85%. This means that the same pattern observed in the first iteration was verified in the second as well. The channels reduction resulted in an increase in the performance.

The combination of both speeds performed the best results, with an accuracy of 86% with 1 channel. Not only the accuracy was higher the smaller the number of channels used, but the combination of both speeds was not a detrimental factor in the models' performance. This could be a negative influence since the cortical activity does not behave in the exactly the same manner when subjects are walking in different speeds. However, it is noted that the accuracy value was the metric that stood out compared to the remaining metrics. Figure 18 demonstrates that, even though the accuracy was increased in the second iteration, the remaining metrics were substantially lower.

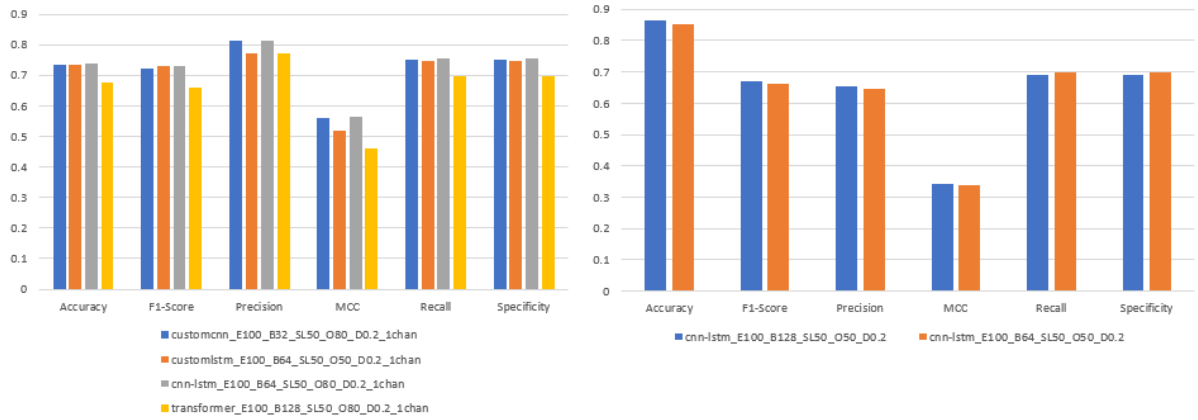


Figure 18: Comparing 1st and 2nd Hold-Out iterations: CAR datasets at different speeds with 1 EEG channel.

This case was observed in the three datasets, which suggests that the removed subjects, although they were not beneficial in optimising the models' accuracy, contributed to its robustness and improved its ability to classify unseen data. Ideally, those subjects should be replaced with new subjects that have used the new gel, contributing to data coherence and to lower the rate of false positives and false negatives.

5.3.3.2 ICA

Overall, the results were even lower than the first iteration. The best result in this iteration was performed in the 1.6 km/h speed, with an accuracy of 53%. This shows that the second iteration was not effective in ICA datasets, so the existence of the first three subjects' data ended up being beneficial. Even so, looking at the results of the first iteration, it can be concluded that ICA is not an optimal choice in this context of binary classification. For this reason, the most valid result to consider should be that of the first iteration, as the exclusion of the first three subjects did not lead to an improvement in accuracy.

5.3.4 Leave-One-Out - second iteration

5.3.4.1 CAR

When compared to the initial iteration and after removing the first three subjects, the mean accuracy level across all datasets remained remarkably consistent throughout the different numbers of channels, suggesting that the exclusion of those subjects did not notably impact the overall data quality. The best result was performed in the 2.5 km/h dataset with 1 EEG channel, achieving an accuracy value of 70%. Thus, the best result to consider is in the first iteration. It is important to emphasise that these results reflect the average performance across all subjects. However, it is essential to acknowledge that the variation between each iteration is an important factor to consider. Figure 19 illustrates the extent of performance variation observed in individual subjects. Despite greater consistency among various metrics in this method, the

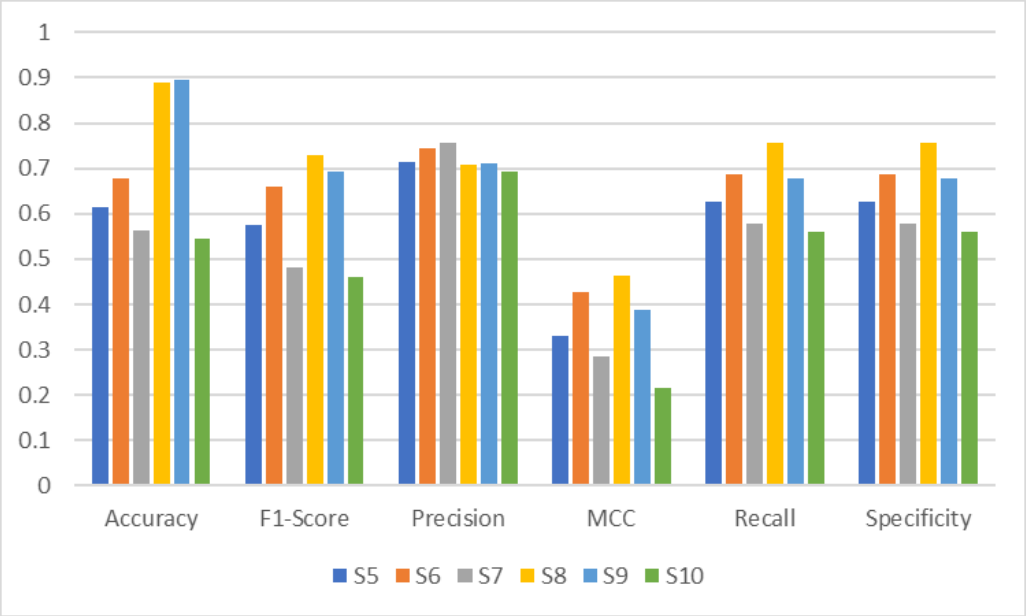


Figure 19: Difference between subjects' performance in the second LOO iteration.

HO approach yields higher accuracy results than the LOO method. Hold-Out achieves an accuracy of 86% as the best result (attained in the second iteration), while LOO achieves approximately 71% (attained in the first iteration).

5.3.4.2 ICA

As it has been consistently observed throughout previous chapters, ICA datasets failed to yield improved results, with the obtained values still falling considerably short of the intended target. While the best accuracy value obtained in the first iteration was approximately 60%, in the second iteration the best value was 55%. This persistent discrepancy between the actual outcomes and the desired benchmarks underscores the limitations and challenges associated with the application of this kind of processing method in this research or the need for further optimisation. The same can be concluded about the difference between the CAR and ICA results, which were consistently presenting divergent results.

5.3.5 Models and Parameters Evaluation

In this section, the ICA results were excluded due to their substantially lower performance compared to the CAR results. The analysis is based on tables in Appendix A, B, C, and D. Tables 13 and 15 display the count of individual executions that surpassed 75% of accuracy across various hyper-parameter sets. In Tables 14 and 16, the same principle is applied, but this time we calculated the number of executions where the average accuracy of all subjects exceeded 65%. Among all datasets, only the 1.6 km/h dataset achieved values of at least 75% accuracy, as others could not reach this threshold. Turning to the table 13,

Table 13: Number of runs from CAR datasets (Accuracy $\geqslant$ Target) - Hold-out first iteration

Dataset	DL Models		EEG Channels		Hyper-parameters					
	Model	ACC $\geqslant$ 75%	Channels	ACC $\geqslant$ 75%	Batch Size	ACC $\geqslant$ 75%	Overlap	ACC $\geqslant$ 75%	Dropout	ACC $\geqslant$ 75%
1.6 km/h	CNN	11	16	✗	32	9	50%	9	20%	21
	LSTM	5	5	✓	64	6				
	CNN-LSTM	10	2	✓	128	6	80%	17	50%	5
	Transformer	0	1	✓	256	5				
2.5 km/h	CNN	0	16	✗	32	0	50%	0	20%	0
	LSTM	0	5	✗	64	0				
	CNN-LSTM	0	2	✗	128	0	80%	0	50%	0
	Transformer	0	1	✗	256	0				
ALL	CNN	0	16	✗	32	0	50%	0	20%	0
	LSTM	0	5	✗	64	0				
	CNN-LSTM	0	2	✗	128	0	80%	0	50%	0
	Transformer	0	1	✗	256	0				

Table 14: Number of trials from CAR datasets (Accuracy $\geqslant$ Target) - Leave-one-out first iteration

Dataset	DL Models		EEG Channels		Hyper-parameters					
	Model	ACC $\geqslant$ 65%	Channels	ACC $\geqslant$ 65%	Batch Size	ACC $\geqslant$ 65%	Overlap	ACC $\geqslant$ 65%	Dropout	ACC $\geqslant$ 65%
1.6 km/h	CNN	4	16	✓	32	3	50%	3	20%	5
	LSTM	0	5	✓	64	1				
	CNN-LSTM	4	2	✓	128	4	80%	6	50%	4
	Transformer	1	1	✓	256	1				
2.5 km/h	CNN	3	16	✗	32	2	50%	3	20%	6
	LSTM	1	5	✓	64	3				
	CNN-LSTM	3	2	✓	128	2	80%	4	50%	1
	Transformer	0	1	✓	256	0				
ALL	CNN	3	16	✗	32	3	50%	3	20%	6
	LSTM	1	5	✓	64	4				
	CNN-LSTM	3	2	✓	128	0	80%	4	50%	1
	Transformer	0	1	✓	256	0				

Table 15: Number of runs from CAR datasets (Accuracy $\geqslant$ Target) - Hold-out second iteration

Dataset	DL Models		EEG Channels		Hyper-parameters					
	Model	ACC $\geqslant$ 75%	Channels	ACC $\geqslant$ 75%	Batch Size	ACC $\geqslant$ 75%	Overlap	ACC $\geqslant$ 75%	Dropout	ACC $\geqslant$ 75%
1.6 km/h	CNN-LSTM	4	16	✗	32	0	50%	2	20%	2
			5	✗	64	2				
			2	✓	128	2	80%	0	50%	0
			1	✓	256	0				
2.5 km/h	CNN-LSTM	6	16	✗	32	0	50%	6	20%	6
			5	✓	64	3				
			2	✓	128	3	80%	0	50%	0
			1	✓	256	0				
ALL	CNN-LSTM	4	16	✗	32	0	50%	2	20%	2
			5	✗	64	2				
			2	✓	128	2	80%	0	50%	0
			1	✓	256	0				

focusing on deep learning models, both CNN and CNN-LSTM consistently achieved accuracy above 75%. These results were only attained with 5 EEG channels or fewer. Batch size trends towards values of 32, 64, and 128. The preferred overlap was 80%, indicating a preference for more windows, even if it means training with repeated data. Finally, a clear preference for a 20% dropout rate was observed.

Regarding Table 14, it reaffirms the similarity of results between CNN and CNN-LSTM. Like Table 13, achieving higher accuracy was consistent with 5 or fewer channels. Similarly, the preference for batch size

Table 16: Number of trials from CAR datasets (Accuracy $\geqslant$ Target) - Leave-one-out second iteration

Dataset	DL Models		EEG Channels		Hyper-parameters					
	Model	ACC $\geqslant$ 65%	Channels	ACC $\geqslant$ 65%	Batch Size	ACC $\geqslant$ 65%	Overlap	ACC $\geqslant$ 65%	Dropout	ACC $\geqslant$ 65%
1.6 km/h	CNN	3	16	✗	32	3	50%	2	20%	2
			5	✓	64	0				
			2	✓	128	0	80%	1	50%	1
			1	✓	256	0				
2.5 km/h	CNN	3	16	✗	32	3	50%	2	20%	3
			5	✓	64	0				
			2	✓	128	0	80%	1	50%	0
			1	✓	256	0				
ALL	CNN	3	16	✗	32	3	50%	3	20%	3
			5	✓	64	0				
			2	✓	128	0	80%	0	50%	0
			1	✓	256	0				

values and dropout rates remains evident. In addition, the speed factor did not significantly influence the number of executions with average accuracy levels above 65% in the three datasets.

Comparing table 15 to 13, despite having only 2 executions per number of channels' combination, it is evident that even in the ALL dataset, there were 4 executions that met the required accuracy. This suggests that all datasets performed relatively similarly, with consistent batch size, overlap, and dropout patterns observed. Considering Table 16, it is noteworthy that it also reflects an identical number of executions that succeeded in achieving accuracy above 65%. In fact, across the three datasets, at least one execution reached at least 70% accuracy. This observation suggests that there is not a substantial difference in performance between the first and second LOO iterations, along with a slight improvement in accuracy.

5.4 Discussion

In the context of this research, several AI-based architectures were systematically constructed with the primary objective of decoding instances of balance loss induced by slip-like perturbations. The overarching aim was to identify the optimal combination of architectural components that could exhibit superior performance within the realm of binary classification, distinguishing between perturbation and non-perturbation events. This comprehensive approach was designed with careful consideration of the existing literature, ensuring that this study aligns with the current state-of-the-art methodologies and techniques.

Following this approach, two strategies were implemented: hold-out and leave-one-out. In the first, each dataset was separated in training, validation and test sets and the main goal was to set the best hyper-parameters combination in order to use them in leave-one-out. This method, for each combination of number of channels, had as many iterations as the number of subjects. Every iteration used a different subject as part of the testing set while the remaining were considered as the training set.

Regarding the models' performance, it is noted that, through a direct observation on the results presented in Appendix A, B, C and D, CNN-LSTM and CNN achieved the best results compared to the remaining models. Transformer, on the other hand, obtained the worst results in most executions. Despite the difference in behaviour between the two methods, the results are generally consistent when compared between HO and LOO.

The main purpose of hold-out was to tune the hyper-parameters while testing models' performance on unseen data. As it was described in 5.3.3.1, the accuracy was up to 86%. Although it shows that the performance did not decrease when training and testing with different speeds, the remaining metrics were not as good as expected. However, by observing the results obtained in the first iteration, it is noted that this discrepancy did not exist, which may suggest that the removal of the first three subjects, even though it increased the accuracy values, caused a loss of information, which reduces the diversity in the datasets and models' capacity to classify unseen data. The smaller the datasets, the more sensitive the metrics can be when it comes to outliers or noise. Another possibility is the hold-out split, which was modified due to the changing of the number of subjects. The split in the second iteration may have been more favourable by chance, resulting in better accuracy values, but it does not necessarily indicate a better model performance. In order to minimise this problem, it should be performed multiple iterations with different random splits to evaluate the final robustness.

Figure 19 illustrates the differences in subjects' accuracy values. It is important to note that these disparities may not be entirely accurate, as they can be significantly influenced by the initial hold-out split. For instance, subjects 8 and 9 were assigned to the validation set during the first hold-out iteration, and based on that split, the best hyper-parameter combination was replicated for use in both the first and second LOO iterations. This factor could potentially explain the distinctions observed in these subjects compared to the others, and it implies that similar variations could occur with different subjects if the HO split were changed. Nevertheless,the difference between the subjects still plays a pivotal role in influencing the final results.

Leave-one-out cross-validation did not accomplished the expected results, especially when compared to hold-out. In fact, HO ended up being more positive in terms of accuracy values than the first. However, this technique should still be part of the global strategy because it provides a robust evaluation of the performance by iteratively training and testing each model using the available data. While in HO there was a specific set of training, validation and testing data, LOO trains and tests on all individual data points. Thus, the models were not able to generalise as favourably as in HO, at least when it comes to accuracy.

CAR achieved considerably higher results than ICA, both in HO and LOO methods. One important aspect to be pointed out is the fact that these are prepossessing techniques with different purposes. CAR primarily aims to reduce noise and standardise EEG references, while ICA is employed for separating sources, removing artifacts and extracting features. In this context, CAR is more efficient, which can be due to the characteristics of EEG signals and the type of perturbations detected. A possibility of the inefficiency

of ICA can be related to the lack of complexity or independent sources in the data. An alternative would be testing with less independent components, following the same procedure as in CAR with different channel combinations, but in this case taking to account which components are more related to loss of balance.

The hold-out executions, for each dataset, had better results the smaller the number of channels. In fact, as observed in 5.3.3.1, the best result was obtained using only 1 channel. This contributes to demonstrate that PEPs are broadly dispersed across the fronto-centro-parietal regions. For that reason, the channels that are not part of this region can be discarded from the datasets that the model processes since they do not add valuable information, requiring less memory and increasing efficiency.

In comparison with the current literature, [73] achieved accuracy levels above 73% by performing support surface translations in standing subjects. Since our work involves subjects in walking conditions rather than stationary, it introduces additional complexities in EEG data analysis, particularly in artifact removal, owing to the inherent movement artifacts associated with walking. Consequently, achieving higher levels of performance becomes inherently more challenging. Despite these added challenges, our study successfully surpassed the previously reported accuracy levels.

While some studies have opted to enhance EEG data analysis by integrating additional sensors, achieving commendable results, it is important to note that our study pursued a more focused approach. Mezzina et al. [40], for instance, reached an impressive sensitivity level exceeding 93% by implementing a multi-sensor architecture aimed at joint analysis of muscular and cortical activity, showcasing the potential benefits of sensor fusion. However, our study deliberately relied solely on EEG data, eschewing the integration of other sensors. This focused approach can be considered a limitation in the context of achieving even higher results. By exclusively utilising EEG data, it is acknowledged that we may not fully harness the potential synergies that can arise from the integration of multiple sensor modalities. Despite this, our study contributes valuable insights into the effectiveness of EEG data in isolation, providing a baseline for future investigations seeking to optimise performance while considering the trade-offs associated with sensor selection and integration.

Some other papers obtained higher accuracy values than this work [22, 72, 79, 84]. Tortora et al. [79], for instance, achieved 90% accuracy while also using EEG and EMG data combined. Nakagome et al. [72] performed a 15 minute walk in its experimental protocol with a speed of approximately 1.6 km/h using a 64-channel EEG electrode system, while our study performed two isolated speeds and then combined them to analyse the potential influence on the models' behaviour when given as an input data from more than one speed. The lack of subjects also seems to be an issue in the current literature, and this study sought to innovate by including ten subjects equally distributed by gender. De Venuto et al. [84] constructed a classification system that was also provided with EEG and EMG data. Although the authors kept an accuracy of 96%, the system was tested in just 6 subjects.

Chapter 6

ADL Recognition using Brain Signals

6.1 Introductory Insight

This thesis chapter delves into the realm of utilising deep learning models for the classification and interpretation of ADLs. Focused on the second experimental protocol detailed in Chapter 4, our goal is to decode the intricate fabric of ADLs. At the core of this exploration lies the essential pre-processing of raw data, encompassing data labelling, cleansing, and transformation, facilitated by the versatile EEG signal processing tool, EEGLAB. The ability to not only classify but also compare the results from different models holds profound implications across domains such as healthcare, psychology, and human-computer interaction. Throughout this chapter, we bridge the divide between raw data and meaningful insights, while highlighting the pivotal role of deep learning in recognising daily activities. This chapter commences with the pre-processing steps, which lay the groundwork for the subsequent deployment of deep learning models, culminating in a comprehensive comparison of diverse outcomes.

6.2 Methods

In this section, we will delve into the methodologies utilised for the validation and identification of various ADLs. The ensuing approaches will be addressed: i) categorising the data into eight distinct classes, which is based on the data segmentation outlined in the literature review provided in Chapter 3; ii) implementing AI-based models on the datasets, with a detailed account of the specific parameters employed during the model training process; iii) elucidating the techniques used for partitioning the data into training, validation, and test datasets; iv) a comprehensive discussion on the metrics employed to evaluate the efficacy and performance of the models developed.

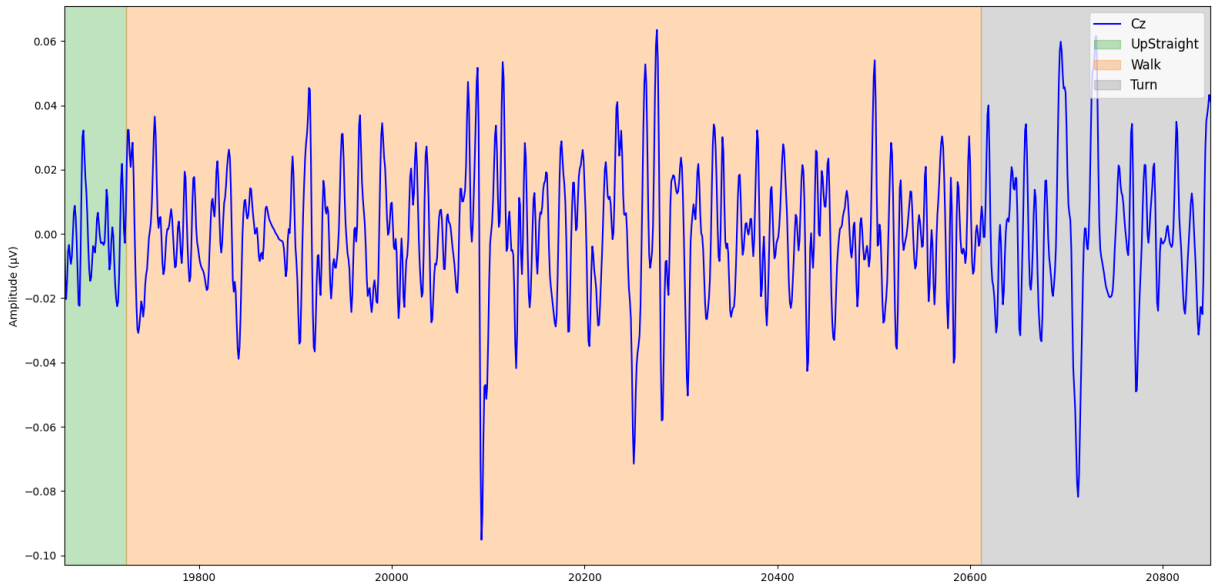


Figure 20: Segment of data of Cz channel including three classes from ADLs experimental protocol: 'Upstraight', 'Walk' and 'Turn'.

6.2.1 Data Labelling

The labelled epochs were combined into two separate datasets using a Python script. One dataset compiled the CAR epochs, while the other dataset compiled the ICA epochs, both of which were extracted from EEGLAB. Of paramount significance is the fact that, as a result of this process, each dataset became a repository of eight labelled classes. These classes, which form the foundation of the subsequent classification endeavours, encompass a spectrum of eight activities: 'Transition,' 'Walk,' 'Upstraight,' 'Lying,' 'Upstairs,' 'Downstairs,' 'Pick,' and 'Turn.' Figure 20 illustrates a segment of data where each sub-segment is labelled by three of the eight mentioned classes.

To minimise the textual content within the datasets, a label conversion process was implemented, transitioning from string-based labels to a numerical format. This conversion assigns a unique integer identifier to each label, ranging from 0 to 7. The resulting label identifiers are detailed in Table 17, with each identifier associated with its respective class. This conversion not only enhances the dataset's overall organisation and its data processing, but also streamlines the labelling system, enabling more efficient handling and analysis of the dataset.

6.2.2 Architecture Parameters

In the selection of our models, a thoughtful deliberation was undertaken, informed by the findings detailed in Chapter 5. Additionally, it was drawn upon the insights garnered from the extensive literature review in Chapter 3. This process led us to opt for the adoption of two top-performing models, namely the CNN and the CNN-LSTM architectures. Table 18 depicts the values that underwent extensive testing for each

Table 17: ADLs Labels Identifiers

Label Identifier	Description
0	Transition
1	Walk
2	Upstraight
3	Lying
4	Upstairs
5	Downstairs
6	Pick
7	Turn

hyper-parameter with H0, while Table 19 presents the two datasets created along with the two mentioned deep learning models architectures. The model that presents the best results along with the best hyper-parameter combination is used in L0O.

The hyper-parameter values that were ultimately chosen were carefully selected, taking into consideration the performance of values explored in Chapter 5, where more favourable results were observed. This decision was influenced by the valuable insights derived from the study documented in [38], particularly in relation to the determination of appropriate window size values.

6.2.3 Model Building and Evaluation

The evaluation strategy employed in this study encompassed multiple phases, each designed to assess the performance of our chosen classification models. The first phase involved an exhaustive optimisation process to fine-tune the models' parameters and enhance their performance. This optimisation was executed using the hold-out method. In this phase, out of the six subjects in our dataset, four were allocated to the training subset, one to the validation subset, and the remaining subject to the testing subset. In this case, subjects 1, 2, 3 and 4 were part of the training subset, 5 was part of the validation subset and 6 was part of the test subset.

Table 18: Hyper-parameters for Deep Learning models - ADLs protocol

Hyper-parameters	Value(s)
Number of Epochs	50
Window Size	150, 750
Regularisation	L1
Batch Size	32, 128
Dropout	20%
Overlap	50%
Optimiser	Adam
Loss Function	Cross Entropy Loss

Table 19: Subset of features per dataset and Deep Learning Models - ADLs protocol

Dataset Type	Channels/ICs	DL Models
CAR	All Channels	CNN CNN-LSTM
ICA	All Components	CNN CNN-LSTM

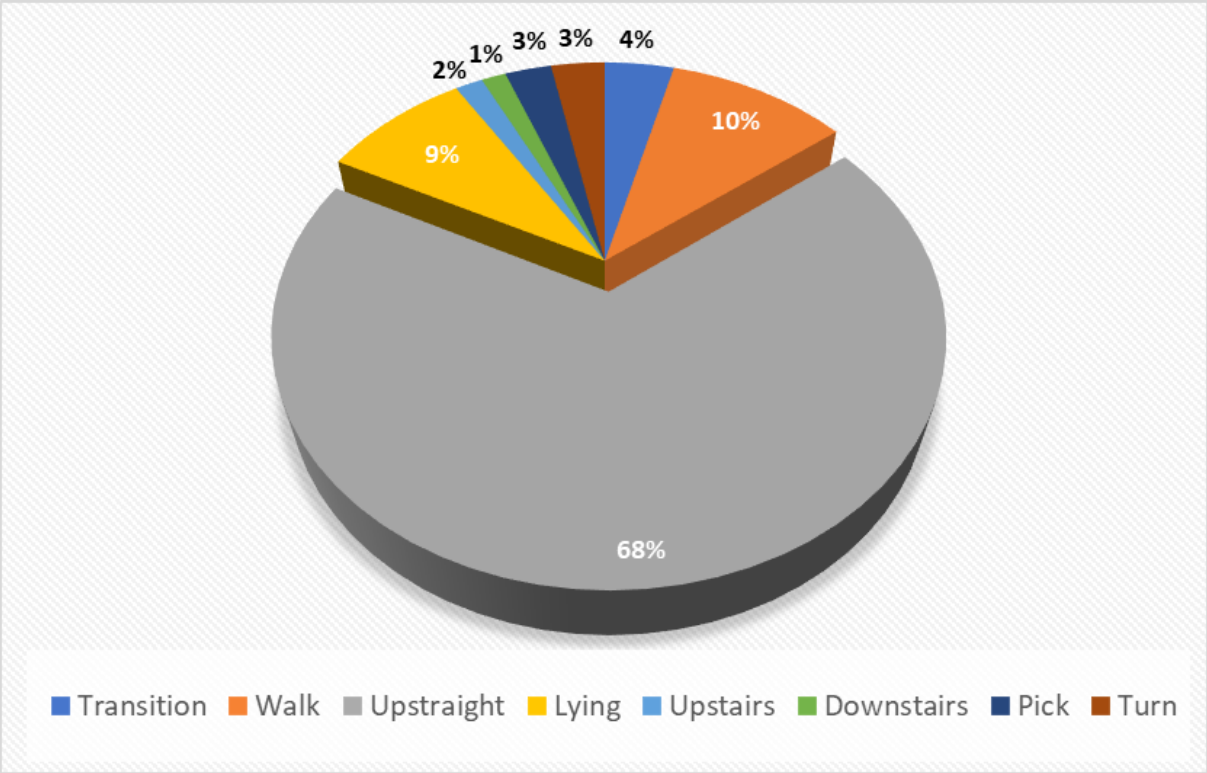


Figure 21: Class distribution in ADLs datasets.

Subsequently, we leveraged the LOO method for our evaluation, which entails systematically designating one subject as the testing set while using the remaining subjects for training. This process was iteratively repeated six times, corresponding to the number of subjects in our dataset. In each execution, we meticulously selected the two best hyper-parameter combinations from Chapter 5, based on the achieved accuracy levels, to ensure the models' optimal performance. This comprehensive evaluation strategy allowed the exploration of the models' capabilities and parameter configurations, providing a robust foundation for the subsequent analysis and discussion of the final results.

Addressing dataset imbalance was a primary concern throughout the experimental protocol, primarily due to the inherent variability in the duration of each collected activity. As a result, the final datasets exhibited an imbalance in the distribution of activity labels, as it can be observed in Figure 21. Balancing the dataset by discarding samples was an undesirable option as it would result in the loss of valuable information. Consequently, we opted to maintain the dataset in its original form while implementing class weights.

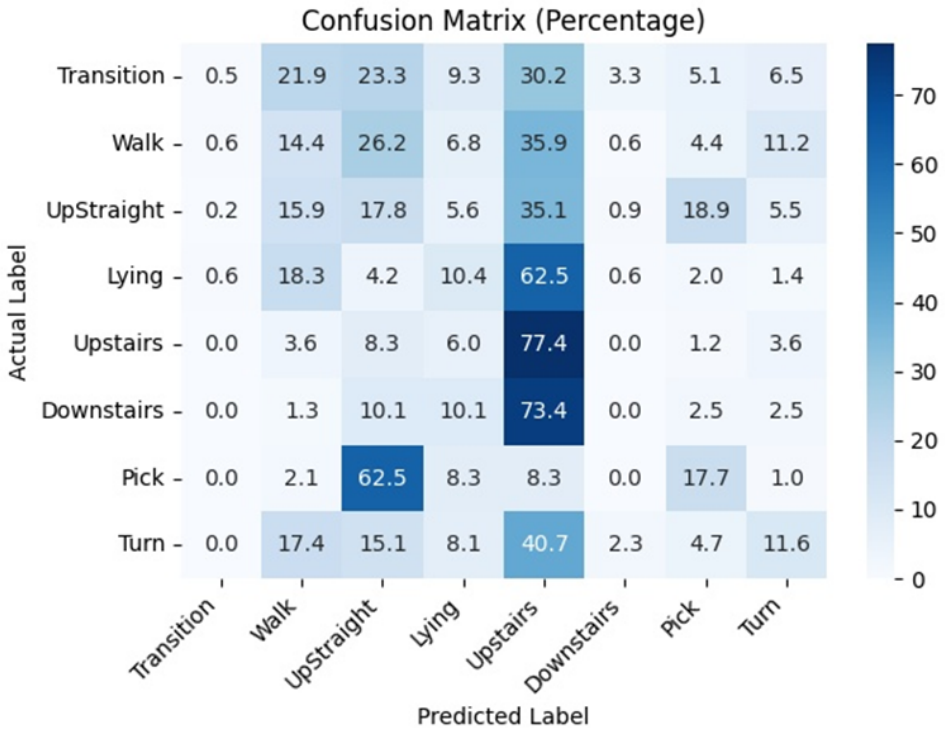


Figure 22: Confusion Matrix from ADLs protocol CAR results (HO) using the CNN model, 128 batch size and 750 window length.

This strategic use of class weights ensures that no class is unduly favored due to the varying quantities of available data, preserving the integrity and comprehensiveness of the dataset. This approach not only maintains dataset completeness but also can significantly enhance model performance by addressing the imbalance challenge.

6.3 Results

Based on the information extracted from Table 19, the pre-processed inputs were bifurcated into two distinct datasets. This chapter is dedicated to a comprehensive analysis of the conclusive outcomes obtained by considering the various phases of our meticulously crafted evaluation strategy. The results, detailed and tabulated, can be found in Appendix E and Appendix F.

6.3.1 Hold-out

6.3.1.1 CAR

The CNN-LSTM architecture emerged as the standout performer, attaining a test accuracy of 26%. Notably, the CNN model delivered similar results, although both models fell slightly short of the benchmark set by

current literature in the field. It is worth highlighting that the experiments unveiled a pattern where higher batch sizes seemed to adversely impact test accuracy. This suggests that there might be an optimal batch size, and exceeding it could lead to performance degradation. On a positive note, we observed a significant enhancement in test accuracy when we increased the window size from 150 to 750. This underscores the role of the temporal context captured by the window size in improving the model's predictive capabilities.

Diving deeper into the results and exploring the confusion matrix, a notable achievement was the 77.4% accuracy in classifying the 'Upstairs' activity, as presented in Figure 22. In contrast, other activities involve more complex cortical activity as the subject engages in motion, making them more challenging to distinguish accurately. This insight underscores the nuanced nature of activity recognition and highlights the importance of contextual understanding for precise classification.

6.3.1.2 ICA

In the ICA dataset, a notable discrepancy emerged when examining the testing accuracy of the models. The CNN model, while robust, achieved a testing accuracy of approximately 24%, whereas the CNN-LSTM architecture exhibited a significant performance improvement with an approximate accuracy of 45%. However, it is essential to note that these results, while promising, still fell short of the benchmarks set by previous research in the field.

Similar to the results in the CAR dataset, the ICA dataset showcased a comparable trend in terms of batch size and window size adjustments. In particular, we observed that the increase in batch size did not yield improved testing accuracy, indicating that there may be a threshold beyond which larger batch sizes do not contribute positively to model performance. On the other hand, an increase in window size, as noted previously, appeared to enhance the model's predictive capabilities, further reinforcing the importance of temporal context in activity recognition. When examining the confusion matrix for the ICA dataset, the highest accuracy value observed was 62.4% for the 'Upstraight' class, as depicted in Figure 23.

6.3.2 Leave-one-out

6.3.2.1 CAR

Given that CNN-LSTM was the sole architecture employed in conjunction with a single set of hyperparameters, the highest achieved accuracy reached approximately 53%. This accuracy surpasses the results attained in the HO methodology for both the CAR and ICA approaches. It is noted that within this top-performing result, an interesting pattern emerged: the training accuracy exhibited a slight decrement compared to the validation accuracy and notably trailed behind the testing accuracy. This observation highlights the CNN-LSTM model's robustness, as it exhibits a better generalisation to unseen data during testing comparing to the HO results, as well as the remaining LOO outcomes. Notably, the model's testing accuracy surpasses its training accuracy, suggesting its ability to adapt to new data patterns.

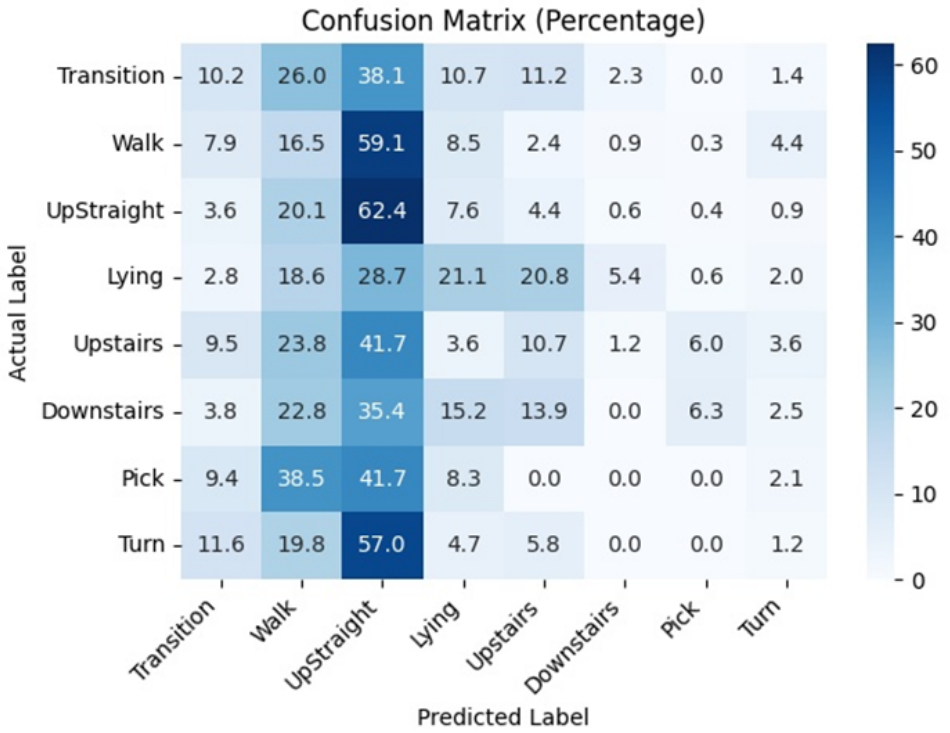


Figure 23: Confusion Matrix from ADLs protocol ICA results (HO) using the CNN-LSTM model, 32 batch size and 750 window length.

This deviation in performance between training and testing phases does not necessarily indicate overfitting. In an overfitting scenario, a model excels in training but struggles with testing data. In contrast, this model demonstrates better generalisation to the testing data. The suggestion of further fine-tuning is aimed at optimising the model's performance and not necessarily addressing overfitting, given the model's testing performance. Concerning the performance within each class, the analysis of confusion matrices revealed an interesting pattern, particularly in the 'Upstraight' class. The accuracy associated with this class were notably distinct from those of the remaining classes, as it can be observed in Figure 24.

6.3.2.2 ICA

The outcomes exhibited a noticeable disparity when compared to the CAR method, which is in contrast to the findings in HO. In the LOO approach, ICA produced lower results than CAR. In this context, the highest recorded accuracy value reached approximately 46%, which is in relative proximity to the 53% achieved in CAR. Nevertheless, both methodologies continue to fall short of the desired performance target. When examining the results obtained from ICA in the HO method, a consistent pattern emerges with the LOO method for this approach. LOO marginally improved the best accuracy, only by 2%. This suggests that this method did not significantly alter the performance outcomes compared to HO in the case of ICA. It is noteworthy that throughout these analyses, the training accuracy consistently exceeded

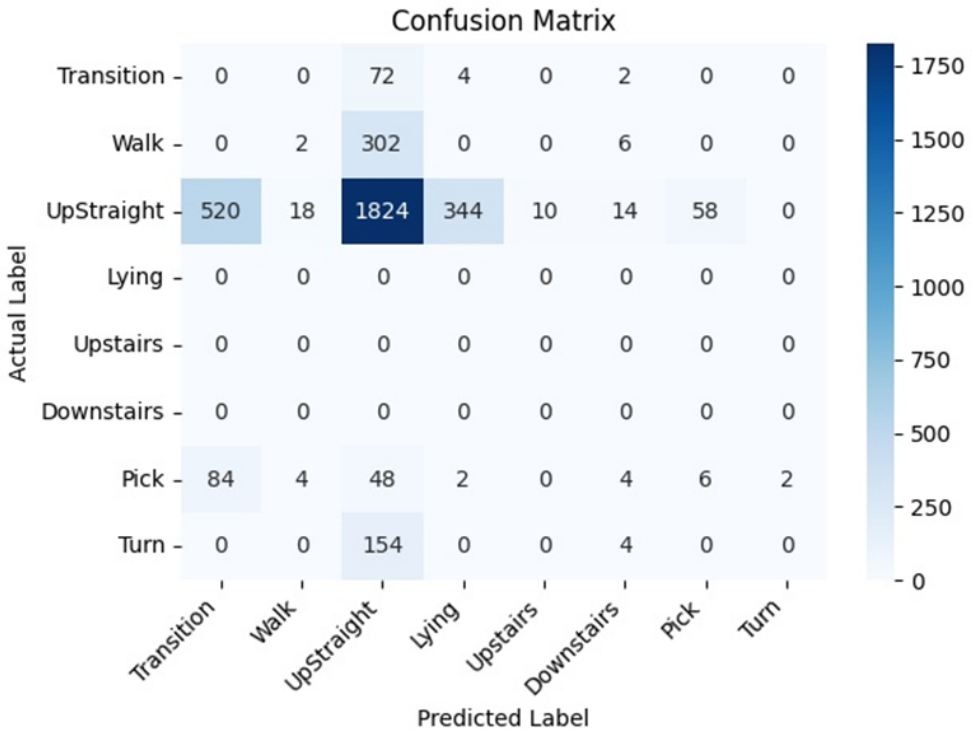


Figure 24: Confusion Matrix from ADLs protocol CAR results (LOO) using the CNN-LSTM model.

the testing accuracy, implying a potential issue with overfitting that could be adversely affecting the model's performance. This pattern should be addressed to optimise the model's performance.

In terms of the performance within individual classes, a comparable pattern emerged, particularly with regards to the 'Upstraight' class, akin to the observations made in CAR. However, in this instance, the model did not exhibit the same level of consistency in showing this pattern, as there were more instances where the predicted class did not align with the actual class label. For instance, in Figure 25, there were multiple occurrences where the predicted class was 'Lying' while the true label was 'Upstraight'.

6.4 Discussion

In the course of this investigation, two distinct AI-based architectures were meticulously developed with the primary objective of decoding various ADLs. The intention behind this endeavour was to identify the most optimal combination, thereby ascertaining superior performance in the domain of classification, while considering the existing studies in the literature.

Pursuing this methodical approach, two strategies were instituted: the hold-out and leave-one-out techniques. In the former, each dataset underwent a segregation process, resulting in the creation of distinct training, validation, and test sets. The principal aim here was to discern the most effective hyper-parameter combinations, subsequently intended for deployment in the LOO method, akin to the process delineated

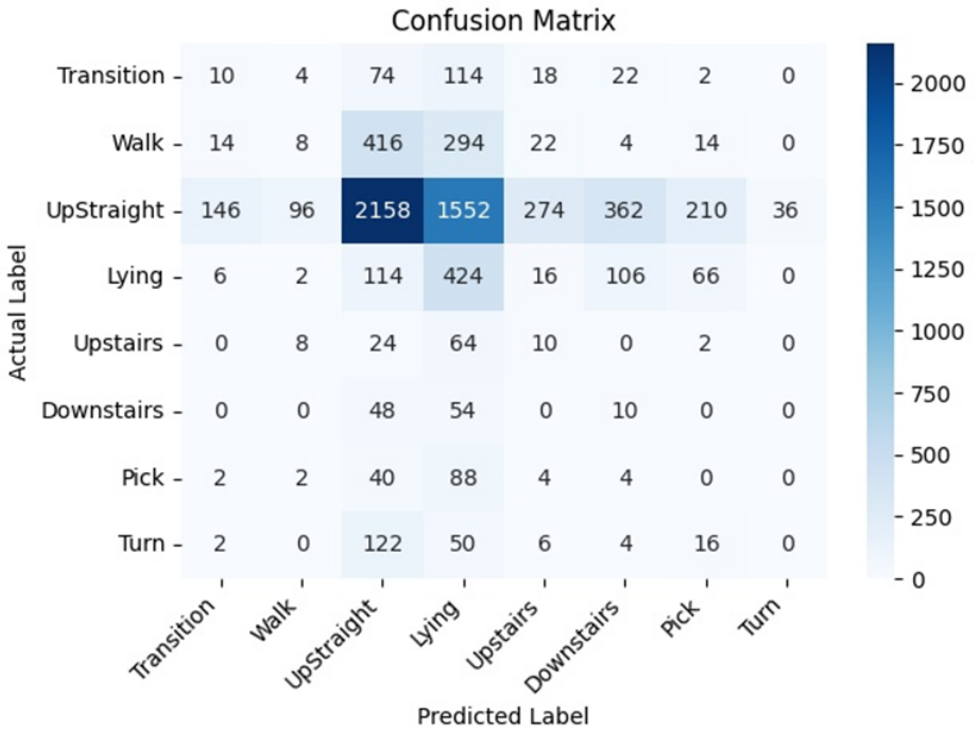


Figure 25: Confusion Matrix from ADLs protocol ICA results (LOO) using the CNN-LSTM model.

in Chapter 5. This method entailed iterations equivalent to the number of individual subjects involved in the study. Each iteration featured a distinct subject within the testing set, while the remaining subjects constituted the training set. However, differing from the methodology expounded in Chapter 5, there was no segmentation of the dataset with regard to diverse combinations of EEG channels.

As for the outcomes derived from the HO technique, it is discernible that the CNN-LSTM architecture yielded the most favourable results, attaining an accuracy of 44% as its highest achievement. Nonetheless, when juxtaposed with this particular outcome, both models exhibited a comparable level of performance. In the broader perspective, these findings, while indicative of progress, remain notably distant from the initially envisaged benchmarks. It is noted that these results, particularly in the context of the HO technique, may be influenced by the specific training, validation, and test set split, which could potentially introduce biases that impact the true performance of the models. Moreover, an intriguing observation arises when comparing the CAR dataset to the ICA dataset. Unlike the outcomes presented in Chapter 5, the ICA dataset displayed superior results in this instance, whereas the CAR dataset achieved an accuracy of only approximately 27%.

Regarding the LOO outcomes, there is a substantial difference between this method and HO when it comes to the CAR and ICA datasets. Here, CAR exhibited better performance compared to ICA, in contrast to what occurred in the HO scenario. For that reason, it is challenging to determine the best-performing dataset in terms of results. CAR achieved a maximum accuracy of 53%, while ICA attained 46%. This means

that LOO managed to obtain a better accuracy than HO, which did not happen in the slip-like perturbation outcomes in Chapter 5. Also, no clear pattern emerges regarding the best testing subjects, as the best results for each dataset were obtained from different subjects.

In 6.2.3, we delved into the intricate issue of dataset imbalance and discussed the process of weight distribution as a strategy to mitigate this challenge. While this approach represents a critical step in rectifying class imbalances, it has come to light that further measures may be necessary to effectively address the issue. By analysing the various confusion matrices in this chapter, our models' predictions revealed a recurring pattern: a predominant number of accurate predictions pertained to a certain class. In case of LOO, for instance, it was the 'Upstraight' class. This phenomenon can be attributed to the fact that 'Upstraight' boasts the highest number of samples within the datasets. Consequently, this disparity highlights the need for additional optimisation initiatives geared towards rectifying the imbalance in favour of the remaining classes. It is imperative to intensify the efforts in re-balancing the dataset, ensuring that no class remains disproportionately under-represented. By doing so, it is possible to, ultimately, improve the overall performance of the DL models.

In short, the attained results fall significantly short of the benchmarks set forth in the current literature [12, 32, 77]. For instance, in [32], a commendable accuracy rate of 86% was achieved. However, it is paramount to consider that the majority of studies have tended to limit the scope of ADLs tested, unlike our comprehensive study which encompasses a spectrum of eight distinct classes. The intricacies arise from the increased complexity of the task due to the higher number of transitions between various ADLs. Moreover, it is noteworthy that certain activities, such as sitting and standing, may exhibit similarities in terms of the cortical activity they invoke, given that subjects maintain a stance condition in both scenarios, further adding to the intricacies of the classification task.

Chapter 7

Conclusion

Falls constitute a predominant cause of unintentional fatalities on a global scale. The gravest impact is borne by the elderly, who confront an elevated risk of falling due to the cognitive, physical, and sensory impairments inherent in the aging process. Even in cases where a fall does not culminate in physical injury, it can engender a profound fear of falling, thereby adversely affecting both mental and physical well-being. Extensive scholarly inquiry has spotlighted slip-like perturbations as primary precursors to falls.

In addressing this issue, the investigation of cerebral signals in response to slip-like perturbations emerges as an essential preliminary endeavour. Consequently, the initial stage of this study involved the cultivation of theoretical acumen in the domains of EEG signals, BCI applications, and ANNs, alongside a comprehensive exploration of the most contemporary research pertinent to the decoding of brain signals related to ADLs. Moreover, it was imperative to develop systems and algorithms for the prediction and detection of falls. Owing to the infrequency of fall incidents, these algorithms are frequently constructed using data derived from simulated falls within a controlled laboratory setting. Regrettably, a dearth of authentic fall datasets prevails, with only scant experimental data from inertial sensors in existence.

This dissertation delineates the process of decoding cerebral responses to slip-like perturbations and ADLs. The decoding strategy was partitioned into two experimental protocols: one encompassing slip-like perturbations involving young, healthy individuals during treadmill ambulation, and the other comprising a battery of ADLs performed by elderly subjects. EEG data collected from both protocols was subjected to decoding via ANN, with the objective of predicting the corresponding labels.

In conducting an exhaustive review of slip-like perturbations employing EEG, an exploration of the methods used for inducing perturbations using EEG, EEG data processing techniques, the principal sensor systems employed, and EEG system characteristics was executed. Notably, most studies confine their experimentation to stance rather than walking conditions. The amalgamation of muscular sensors such as

EMG with EEG was a recurrent occurrence. Nevertheless, there is a conspicuous dearth of experiments in walking conditions, especially during over-ground ambulation, as most studies opt for treadmill walking.

Subsequently, a second review was conducted pertaining to ADLs, with the aim of comprehending the process of decoding various modes of locomotion classifiable as ADLs and devising a solution for gait imbalance events via EEG. This review serves as an extension of a prior study carried out at the BiRDLab at the University of Minho. Within this review, it was observed that the performance of these systems holds promise. However, practical application remains a distant prospect due to the limited sample sizes used, potential safety concerns, and other impediments. Moreover, none of the reviewed studies addressed balance loss in the elderly.

Given the foregoing limitations identified in the current literature, the present dissertation aims to provide a comprehensive strategy to decode and recognise the cerebral signal responses to slip-like perturbations and ADLs. This analysis draws upon data previously collected at the BiRDLab and the *Fundo Social de Braga* nursing home. In the context of deep learning results achieved in this dissertation, our approach was structured into two distinct methods: hold-out and leave-one-out. Particularly in the slip-perturbations study, each comprising two iterations. This facilitated a comprehensive comparative analysis involving various deep learning models and hyper-parameters. Thus, it allows us to discern the most pertinent combination for classifying either perturbation and non-perturbation events or various ADLs.

Regarding the slip-like perturbations study, for the hold-out method, the CNN-LSTM model delivered the best performance, including an accuracy of 86.69%, an F1-score of 67.59%, a precision rate of 66%, a MCC of 0.35, a recall rate of 69.87%, and a specificity of 69.87%. In contrast, the leave-one-out method, while producing commendable results, exhibited lower performance metrics, with an accuracy of 70.89%, an F1-score of 62.24%, a precision rate of 70.03%, an MCC of 0.35, a recall rate of 65.23%, and a specificity of 65.23%. These findings underscore the robustness of the CNN-LSTM classifier in the context of the hold-out method. In the LOO approach, despite the slightly reduced performance metrics compared to CNN-LSTM, the CNN model remains a valuable contender for perturbation recognition. These results contribute to the ongoing pursuit of enhancing fall prevention systems, fostering a safer and more secure environment for the elderly population.

In contrast to the slip-like perturbations study, the study on ADL recognition exhibited more favourable outcomes with the LOO cross-validation method than with the HO method. Specifically, within the LOO approach, the highest achieved accuracy was 53%, surpassing the 46% achieved in HO, both executed with the CNN-LSTM architecture. Additionally, in this study, ICA demonstrated competitive levels of accuracy when compared to CAR in both the HO and LOO scenarios. These results, while falling short of expectations, can be attributed to the challenges posed by the extensive number of classes to classify and the similarities between some of them, aligning with the complexities observed in the existing literature.

The outcomes of this investigation underscore the latent promise of constructing deep learning architectures, and they beckon towards further refinement and optimisation. These advancements hold the

potential to yield even higher performance levels, paving the way for their integration into real-world applications. Moreover, these results furnish valuable insights for designing future studies tailored to diverse populations, thus facilitating the expansion of our understanding and the applicability of these models.

In essence, these findings lay the groundwork for a compelling proposition: EEG signals encompass concise neural information with short latency, intricately linked to the anticipation of an impending fall. Such revelations not only propel us closer to the development BCI systems for fall prevention and ADL recognition but also accentuate the transformative potential of this research in enhancing safety and well-being in various settings, particularly for the elderly.

The work undertaken in this study has facilitated the fulfilment of the RQs delineated in Chapter 1:

- **RQ1**: How has EEG been utilised in existing research to capture and identify daily activities and slip-like perturbations during walking and stance, and what are the key insights gained from these studies?

  This question is addressed in both Chapter 2 and 3. In pursuit of expanding the scope of individualised brain mapping, BCI technologies have gained prominence and have the ability to leverage motion decoders that can be derived from EEG signals to control rehabilitation devices. Thus, the development of EEG-based BCIs necessitates numerous training trials to establish customised parameters for signal characterisation and online decoding [36].

  Among the studies examined, some have employed EEG to analyse and pinpoint cortical regions associated with balance loss [41, 56, 70]. A significant body of research has opted to investigate human balance control through EEG, aiming to identify cortical features during daily activities with minimal latency. These studies collectively reveal that the brain exhibits increased cortical activity when confronted with balance challenges, serving both the purpose of understanding cortical engagement during daily activities and underlining the significance of this increased activity in balance control [15, 23, 33, 42, 43, 46, 58]. In terms of predictability, the study [53] stands out by demonstrating that individuals exhibit heightened cortical activity in response to balance challenges, whether those challenges are predictable or unpredictable.

- **RQ2**: To what extent can current AI models accurately detect slip-like perturbations in healthy subjects using EEG? And what influence does artifacts have on the models' performance?

  Prior studies documented in the literature [40, 72, 79] have reported higher accuracy levels, consistently exceeding 90%. Notably, the study [79] stands out as it leveraged a LSTM architecture and incorporated data from both EEG and EMG sensors. In our study, focusing solely on EEG data in the context of binary classification, we accomplished a maximum accuracy of 86% presented in Chapter 5. This finding suggests that there is room for further optimisation, potentially through the

integration of muscular sensors like EMG, to enhance the accuracy and effectiveness of AI-based models for detecting slip-like perturbations.

Furthermore, Chapter 2 addresses the absence of studies that have introduced perturbations during walking. Even among the limited studies that have explored this aspect, such as those authored by Mezzina et al. [41] and Venuto et al. [84], it is noteworthy that the majority have employed different equipment to induce perturbations. This disparity in the equipment utilised for perturbation induction suggests that making direct comparisons with the existing literature is not feasible, given that these studies have not adhered to the same experimental protocol procedures.

In terms of the impact of artifacts, they can indeed influence the system's performance as they do not represent the actual cortical activity in a given instant of time, so it becomes more difficult to classify. It is crucial to minimise the presence of artifacts, particularly within the experimental protocol, by employing techniques that mitigate head or muscular movements that may introduce motion artifacts. By doing so, the collected data aligns more closely with the underlying cortical activity. While some artifacts may still be present, they can be addressed through the use of EEGLAB techniques as outlined in Chapter 4. However, it is important to recognise that this approach is not ideal, as it does not fully replicate real-world scenarios.

- **RQ3**: Can the developed EEG recognition system effectively identify daily activities in elderly subjects, and to what extent does artifacts impact its performance?

  The results presented in Chapter 6 indicate that the performance of the developed EEG recognition system fell short of the predefined target. A significant contributing factor to this discrepancy is the system's high number of classes, which inherently poses greater challenges for accurate prediction. Most studies that included ADLs [22, 72] do not have that many classes as our study. In fact, the current literature presents studies that are more focused on a small number of ADLs, therefore achieving higher results, since it is easier for the models to predict based on a small number of transitions of ADLs (i.e., walking to stand).

  In our investigation, we have extended the boundaries of this domain by encompassing eight distinct classes. Consequently, it is anticipated that our study may yield comparatively diminished results. The best results were obtained using a CNN-LSTM architecture, achieving an accuracy of approximately 53%. This can be attributed to the intricacies associated with a higher number of classes, which, from a cortical perspective, often exhibit inherent similarities, as exemplified by activities such as Sitting and Laying, where subjects adopt analogous stances. Consequently, it is of paramount importance to rationalise the intricacies of our system to achieve a more efficacious and precise solution, thereby rendering it suitable for practical deployment in real-world scenarios. Furthermore, in contradistinction to prior studies, our research encompasses a broader spectrum of ADLs, potentially offering novel insights for future enhancements in this field.

As it was described in RQ2, artifacts can notably impact the model's performance, particularly in ADL recognition, since it contains various movements in the different types of ADLs. Minimising their influence is crucial, particularly in the experimental setup. Employing techniques to reduce head and muscular movements that generate motion artifacts is vital for accurate data collection. While EEGLAB methods can address some artifacts, they may not fully replicate real-world scenarios.

7.1 Future work

As the pursuit of enhancing the performance and applicability of EEG-based slip-like perturbation and ADLs detection continues, several avenues for future research emerge. Firstly, regarding the slip-like perturbations study, there is a need for balancing the models in terms of the number of windows generated for each class. Despite the balance between perturbation and non-perturbation samples within each dataset, there remains variation in the number of windows generated for each class, sometimes being quite discrepant. In this regard, future work should prioritise adapting class weights to ensure a more balanced model training process.

In Chapter 5, we discussed the sub-optimal results obtained with the Transformer architecture in the two methods implemented. Consequently, it is essential to explore different architectures for each deep learning model to assess if performance can be improved while maintaining or even enhancing the trade-off between performance and execution time. Also, it is imperative to redirect our attention towards the refinement of models through the application of the leave-one-out cross-validation method. Nevertheless, it is necessary to reserve a subset of subjects during this fine-tuning process for subsequent validation, thereby ensuring the availability of independent testing data for subsequent assessments. This allows for a more robust evaluation and validation of the models' performance.

The integration of EEG with other sensors, such as EMG, presents an avenue for enhancing the system's capabilities and performance. Furthermore, the application of normalisation techniques, which has been only preliminarily tested, should be implemented more extensively in the created datasets. By normalising each EEG channel, it is plausible that the combination of EEG data from different subjects can yield improved results through deep learning models.

Finally, regarding the decoding of ADLs, there is a clear need to streamline the classification. Reducing the number of classes, potentially transforming it into binary classification (stance and movement), could significantly enhance the obtained results. This adjustment would involve aggregating previous classes related to movement and standing into the respective new classes. These enhancements are expected to advance the field of EEG-based ADL recognition.

Bibliography

[1] L. K. A, M. R. M, L. N. B, J. M. D, A. D. J, and K. D. R. "Using a Common Average Reference to Improve Cortical Neuron Recordings From Microelectrode Arrays." In: *The Journal of Neurophysiology* 101 (2009). doi: 10.1152/jn.90989.2008. url: https://journals.physiology.org/doi/full/10.1152/jn.90989.2008.

[2] S. M. A. Delorme. "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis." In: *Journal of Neuroscience Methods* (2003). doi: 10.1016/j.jneumeth.2003.10.009.

[3] L. W. F. A. Presacco and t. D. J. L. Contreras-Vidal". In: *IEEE Xplore* 20.2 (2012), pp. 1–3. doi: 10.1109/TNSRE.2012.2188304.

[4] A. L. Adkin, A. D. Campbell, R. Chua, and M. G. Carpenter. "The influence of postural threat on the cortical response to unpredictable and predictable postural perturbations." eng. In: *Neuroscience letters* 435.2 (2008), pp. 120–125. issn: 0304-3940 (Print). doi: 10.1016/j.neulet.2008.02.018.

[5] A. L. Adkin, S. Quant, B. E. Maki, and W. E. McIlroy. "Cortical responses associated with predictable and unpredictable compensatory balance reactions." eng. In: *Experimental brain research* 172.1 (2006), pp. 85–93. issn: 0014-4819 (Print). doi: 10.1007/s00221-005-0310-9.

[6] A. Agga, A. Abbou, M. Labbadi, Y. E. Houm, and I. H. Ou Ali. "CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production." In: *Electric Power Systems Research* 208 (2022), p. 107908. issn: 0378-7796. doi: https://doi.org/10.1016/j.epsr.2022.107908. url: https://www.sciencedirect.com/science/article/pii/S0378779622001389.

[7] D. Allexandre, A. Hoxha, H. V. Shenoy, S. Saleh, S. E. Selvan, and G. H. Yue. "Altered Cortical and Postural Response to Balance Perturbation in Traumatic Brain Injury – An EEG Pilot Study." In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2019), pp. 1543–1546. doi: 10.1109/EMBC.2019.8856645.

[8] E. Allibhai. *Holdout vs. Cross-validation in Machine Learning*. en. Oct. 2018. url: https://medium.com/@eijaz/holdout-vs-cross-validation-in-machine-learning-7637112d3f8f (visited on 03/25/2022).

[9] V. F. Annese, M Crepaldi, D Demarchi, and D De Venuto. "A digital processor architecture for combined EEG/EMG falling risk prediction." In: *Proceedings of the 2016 Design, Automation and Test in Europe Conference and Exhibition, DATE 2016*. 2016, pp. 714–719. url: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84973645498{\&}partnerID=40{\&}md5=213b543d831eca402a9a2ee201164d29.

[10] A. S. E. S. M. I. B. S. Rupal S. Rafique and G. S. Virk. "Lower-limb exoskeletons: Research trends and regulatory guidelines in medical and non-medical applications." In: *International Journal of Advanced Robotic Systems* (2017), pp. 1–2. doi: 10.3390/s18103342.

[11] M. D. Bogost, P. I. Burgos, C. E. Little, M. H. Woollacott, and B. H. Dalton. "Electrocortical Sources Related to Whole-Body Surface Translations during a Single- and Dual-Task Paradigm." eng. In: *Frontiers in human neuroscience* 10 (2016), p. 524. issn: 1662-5161 (Print). doi: 10.3389/fnhum.2016.00524.

[12] R. Chaisaen, N. M. P. Autthasan, P. Leelaarporn, N. Kunaseth, S. Tammajarung, P. Manoonpong, S. C. Mukhopadhyay, and T. Wilaiprasitporn. "Decoding EEG Rhythms During Action Observation, Motor Imagery, and Execution for Standing and Sitting." In: *IEEE Sensors Journal* 20 (2020). doi: 10.1109/JSEN.2020.3005968.

[13] D De Venuto and G Mezzina. "High-specificity digital architecture for real-time recognition of loss of balance inducing fall." In: *Sensors (Switzerland)* 20.3 (2020). doi: 10.3390/s20030769.

[14] D. De Venuto, V. F. Annese, M. Ruta, E. Di Sciascio, and A. L. Sangiovanni Vincentelli. "Designing a Cyber-Physical System for Fall Prevention by Cortico-Muscular Coupling Detection." In: *IEEE Design and Test* 33.3 (2016), pp. 66–76. issn: 21682356. doi: 10.1109/MDAT.2015.2480707.

[15] V Dietz, J Quintern, and W Berger. "Cerebral evoked potentials associated with the compensatory reactions following stance and gait perturbation." eng. In: *Neuroscience letters* 50.1-3 (1984), pp. 181–186. issn: 0304-3940 (Print). doi: 10.1016/0304-3940(84)90483-x.

[16] J. C. Ditz, A. Schwarz, and G. R. Müller-Putz. "Perturbation-evoked potentials can be classified from single-trial EEG." eng. In: *Journal of neural engineering* 17.3 (2020), p. 36008. issn: 1741-2552 (Electronic). doi: 10.1088/1741-2552/ab89fb.

[17] J. C. Ditz, A. Schwarz, and G. R. Müller-Putz. "Perturbation-evoked potentials can be classified from single-trial EEG." In: *Journal of Neural Engineering* 17.3 (2020), p. 036008. doi: 10.1088/1741-2552/ab89fb. url: https://doi.org/10.1088/1741-2552/ab89fb.

[18] R. B. Duckrow, K Abu-Hasaballah, R Whipple, and L Wolfson. “Stance perturbation-evoked potentials in old people with poor gait and balance.” eng. In: *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* 110.12 (1999), pp. 2026–2032. issn: 1388-2457 (Print). doi: 10.1016/s1388-2457(99)00195-9.

[19] Z. Fang, N. Darren, J. Wang, Y. Zheng, and L. Y. Y. H. “An Integrated Fall Prevention System with Single-Channel EEG and EMG Sensor.” In: *2021 4th International Conference on Circuits, Systems and Simulation (ICCSS)* (2021), pp. 183–189. doi: 10.1109/ICCSS51193.2021.9464220.

[20] J. Figueiredo, C. P. Santos, and J. C. Moreno. “Automatic recognition of gait patterns in human motor disorders using machine learning: A review.” In: *Medical Engineering and Physics* 53 (2018), pp. 1–12. issn: 18734030. doi: 10.1016/j.medengphy.2017.12.006. url: https://doi.org/10.1016/j.medengphy.2017.12.006.

[21] S. Foundation. *Alpha Waves and Sleep.* Accessed: 01.04.2023.

[22] E. García-Cossio1, M. Severens, B. Nienhuis, J. Duysens, P. Desain, N. Keijsers, and J. Farquhar. “Decoding Sensorimotor Rhythms during Robotic-Assisted Treadmill Walking for Brain Computer Interface (BCI) Applications.” In: *Plos One* (2016). doi: 10.1109/TNSRE.2020.2966826.

[23] R. Goel, R. A. Ozdemir, S. Nakagome, J. L. Contreras-Vidal, W. H. Paloski, and P. J. Parikh. “Effects of speed and direction of perturbation on electroencephalographic and balance responses.” eng. In: *Experimental brain research* 236.7 (2018), pp. 2073–2083. issn: 1432-1106 (Electronic). doi: 10.1007/s00221-018-5284-5.

[24] V. H. Greg, M. Carlos, and N. Gonzalo. “A Review on the Long Short-Term Memory Model.” In: *Artificial Intelligence Review* 53 (Dec. 2020). doi: 10.1007/s10462-020-09838-1.

[25] K. W. H. Yokoyama N. Kaneko and K. Nakazawa. “Neural decoding of gait phases during motor imagery and improvement of the decoding accuracy by concurrent action observation.” In: *Journal of Neural Engineering* (2021). doi: 10.1088/1741-2552/ac07bd.

[26] S. M. S. Hasan, M. R. Siddiquee, R. Atri, R. Ramon, J. S. Marquez, and O. Bai. “Prediction of gait intention from pre-movement EEG signals: a feasibility study.” In: *Journal of NeuroEngineering and Rehabilitation* 17 (2020). doi: https://doi.org/10.1186/s12984-020-00675-5.

[27] T. Iluz, E. Gazit, T. Herman, E. Sprecher, M. Brozgol, N. Giladi, A. Mirelman, and J. M. Hausdorff. “Automated detection of missteps during community ambulation in patients with Parkinson’s disease: A new approach for quantifying fall risk in the community setting.” In: *Journal of NeuroEngineering and Rehabilitation* 11.1 (2014), pp. 1–9. issn: 17430003. doi: 10.1186/1743-0003-11-48.

[28] P. Ishaani and C. C. a. "A novel LSTM–CNN–grid search-based deep neural network for sentiment analysis." In: *The Journal of Supercomputing* 17 (2021). doi: 10.1007/s11227-021-03838-w. url: https://doi.org/10.1007/s11227-021-03838-w.

[29] P. Ivan, O. Yoshiro, S. Daina, L. S. R., and B. M. A. "Detection of Near Falls Using Wearable Devices: A Systematic Reviews." In: *Journal of Geriatric Physical Therapy* 42 (2019), pp. 48–56. doi: 10.1519/JPT.0000000000000181. url: https://journals.lww.com/jgpt/Fulltext/2019/01000/Detection_of_Near_Falls_Using_Wearable_Devices__A.5.aspx.

[30] K. L. S. J. E. Kline H. J. Huang and D. P. Ferris. "Isolating gait-related movement artifacts in electroencephalography during human walking." In: *Journal of Neuro Engineering* (2015). doi: 10.1088/1741-2560/12/4/046022.

[31] J. V. Jacobs, K. Fujiwara, H. Tomita, N. Furune, K. Kunita, and F. B. Horak. "Changes in the activity of the cerebral cortex relate to postural response modification when warned of a perturbation." eng. In: *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* 119.6 (2008), pp. 1431–1442. issn: 1388-2457 (Print). doi: 10.1016/j.clinph.2008.02.015.

[32] J. H. Jeong, N. S. Kwak, C. Guan, and S. W. Lee. "Decoding Movement-Related Cortical Potentials Based on Subject-Dependent and Section-Wise Spectral Filtering." In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28.3 (2020). doi: 10.1109/TNSRE.2020.2966826.

[33] S Jiang, H Qi, J Zhang, S Zhang, R Xu, Y Liu, L Meng, and D Ming. "A pilot study on falling-risk detection method based on postural perturbation evoked potential features." In: *Sensors (Switzerland)* 19.24 (2019). doi: 10.3390/s19245554. url: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85076737048{\&}doi=10.3390{\%}2Fs19245554{\&}partnerID=40{\&}md5=b07d249c17ecd785132b17b94a9bc574.

[34] M. Jochumsen and I. K. Niazi. "Detection and classification of single-trial movement-related cortical potentials associated with functional lower limb movements." In: *Journal of Neural Engineering* (2020). doi: 10.1088/1741-2552/ab9a99.

[35] H. J. H. K. L. Snyder J. E. Kline and D. P. Ferris. "Independent Component Analysis of Gait-Related Movement Artifact Recorded using EEG Electrodes during Treadmill Walking." In: *Frontiers in Neuroscience* (2015). doi: 10.3389/fnhum.2015.00639.

[36] L. P. R. C. M. J. d. R. K. Lee D. Liu. "A brain-controlled exoskeleton with cascaded event-related desynchronization classifiers." In: *Robotics and Autonomous Systems* 90 (2017), pp. 15–23. doi: 10.1016/j.robot.2016.10.005.

[37] A. P. P. A. M. R. M. M. A. F. A. N. B. S. B. M. Rashid N. Sulaiman and S. Khatun. "Current Status, Challenges, and Possible Solutions of EEG-Based Brain-Computer Interface: A Comprehensive Review." In: *Frontiers in Neurorobotics* (2020). doi: 10.3389/fnbot.2020.00025.

[38] L. M. Martins, N. F. Ribeiro, F. Soares, and C. P. Santos. "Inertial Data-Based AI Approaches for ADL and Fall Recognition." In: *Sensors* 22.11 (2022), pp. 1–20. issn: 14248220. doi: 10.3390/s22114028.

[39] L. Mercado, L. Alvarado, G. Quiroz-Compean, R. Romo-Vazquez, H. Vélez-Pérez, M. Platas-Garza, A. A. González-Garrido, J. Gómez-Correa, J. A. Morales, A. Rodriguez-Liñan, L. Torres-Treviño, and J. M. Azorín. "Decoding the torque of lower limb joints from EEG recordings of pre-gait movements using a machine learning scheme." In: *Neurocomputing* 446 (2021), pp. 118–129. issn: 0925-2312. doi: https://doi.org/10.1016/j.neucom.2021.03.038. url: https://www.sciencedirect.com/science/article/pii/S0925231221004069.

[40] G Mezzina, F Aprigliano, S Micera, V Monaco, and D De Venuto. "EEG/EMG based Architecture for the Early Detection of Slip-induced Lack of Balance." In: *Proceedings - 2019 8th International Workshop on Advances in Sensors and Interfaces, IWASI 2019*. 2019, pp. 9–14. doi: 10.1109/IWASI.2019.8791252. url: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85071426274{\&}doi=10.1109{\%}2FIWASI.2019.8791252{\&}partnerID=40{\&}md5=73413b6e8000b0780de24589da71ced9.

[41] G Mezzina, F Aprigliano, S Micera, V Monaco, and D. D. Venuto. "Cortical reactive balance responses to unexpected slippages while walking: a pilot study." In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019, pp. 6868–6871. isbn: 1558-4615 VO -. doi: 10.1109/EMBC.2019.8856925.

[42] A Mierau, B Pester, T Hulsdunker, K Schiecke, H. K. Struder, and H Witte. "Cortical Correlates of Human Balance Control." In: *BRAIN TOPOGRAPHY* 30.4 (2017), pp. 434–446. issn: 0896-0267. doi: 10.1007/s10548-017-0567-x.

[43] A. Mierau, T. Hülsdünker, and H. K. Strüder. "Changes in cortical activity associated with adaptive behavior during repeated balance perturbation of unpredictable timing." eng. In: *Frontiers in behavioral neuroscience* 9 (2015), p. 272. issn: 1662-5153 (Print). doi: 10.3389/fnbeh.2015.00272.

[44] J. d. R. Millán. "The human-computer connection : an overview of brain-computer interfaces." In: *Metode Science Studies Journal-Annual Review* 9 (2018), p. 135. doi: 10.7203/metode.9.12639.

[45] G Mochizuki, K. M. Sibley, J. G. Esposito, J. M. Camilleri, and W. E. McIlroy. "Cortical responses associated with the preparation and reaction to full-body perturbations to upright stability." eng. In: *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* 119.7 (2008), pp. 1626–1637. issn: 1388-2457 (Print). doi: 10.1016/j.clinph.2008.03.020.

[46] G. Mochizuki, K. M. Sibley, H. J. Cheung, J. M. Camilleri, and W. E. McIlroy. "Generalizability of perturbation-evoked cortical potentials: Independence from sensory, motor and overall postural state." eng. In: *Neuroscience letters* 451.1 (2009), pp. 40–44. issn: 0304-3940 (Print). doi: 10.1016/j.neulet.2008.12.020.

[47] T. Mohammad. "DIMENSIONALITY REDUCTION AND CLASSIFICATION OF TIME EMBEDDED EEG SIGNALS." In: (Jan. 2023).

[48] V. Monaco, P. Tropea, F. Aprigliano, D. Martelli, A. Parri, M. Cortese, R. Molino-Lova, N. Vitiello, and S. Micera. "An ecologically-controlled exoskeleton can improve balance recovery after slippage." In: *Scientific Reports* 7.March (2017), pp. 1–10. issn: 20452322. doi: 10.1038/srep46721.

[49] B. A. Nasreddine, J. Nuraini, P. J. A., O. Sofia, and C. Chao. "Brain Computer Interfaces for Improving the Quality of Life of Older Adults and Elderly Patients." In: *Frontiers in Neuroscience* 14 (2020). doi: 10.3389/fnins.2020.00692. url: https://www.frontiersin.org/articles/10.3389/fnins.2020.00692.

[50] R. A. Ozdemir, J. L. Contreras-Vidal, and W. H. Paloski. "Cortical control of upright stance in elderly." In: *Mechanisms of Ageing and Development* 169 (2018), pp. 19–31. issn: 0047-6374. doi: https://doi.org/10.1016/j.mad.2017.12.004. url: https://www.sciencedirect.com/science/article/pii/S0047637417302440.

[51] I. Pang, Y. Okubo, D. Sturnieks, S. R. Lord, and M. A. Brodie. "Detection of Near Falls Using Wearable Devices: A Systematic Review." In: *Journal of Geriatric Physical Therapy* 42.1 (2019), pp. 48–56. issn: 21520895. doi: 10.1519/JPT.0000000000000181.

[52] S. Park, F. C. Park, J. Cho, and H. Kim. "EEG-based Gait State and Gait Intention Recognition Using Spatio-Spectral Convolutional Neural Network." In: *7th International Winter Conference on Brain-Computer Interface (BCI)* (2019). doi: 10.1109/IWW-BCI.2019.8737259.

[53] A. M. Payne and L. H. Ting. "Worse balance is associated with larger perturbation-evoked cortical responses in healthy young adults." In: *Gait and Posture* 80 (2020), pp. 324–330. doi: 10.1016/j.gaitpost.2020.06.018. url: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85086754838{\&}doi=10.1016{\%}2Fj.gaitpost.2020.06.018{\&}partnerID=40{\&}md5=779067250fb1b85a0e5d113423a04fd3.

[54] A. M. Payne, J. A. Palmer, J. L. McKay, and L. H. Ting. "Lower Cognitive Set Shifting Ability Is Associated With Stiffer Balance Recovery Behavior and Larger Perturbation-Evoked Cortical Responses in Older Adults." eng. In: *Frontiers in aging neuroscience* 13 (2021), p. 742243. issn: 1663-4365 (Print). doi: 10.3389/fnagi.2021.742243.

[55] S. M. Peterson and D. P. Ferris. "Differentiation in Theta and Beta Electrocortical Activity between Visual and Physical Perturbations to Walking and Standing Balance." In: *ENEURO* 5.4 (2018). issn: 2373-2822. doi: 10.1523/ENEURO.0207-18.2018.

[56] S. M. Peterson and D. P. Ferris. "Group-level cortical and muscular connectivity during perturbations to walking and standing balance." In: *NeuroImage* 198.January (2019), pp. 93–103. issn: 10959572. doi: 10.1016/j.neuroimage.2019.05.038. url: https://doi.org/10.1016/j.neuroimage.2019.05.038.

[57] S. M. Peterson, E. Rios, and D. P. Ferris. "Downloaded from www.physiology.org/journal/jn by ${individualUser.givenNames} ${individualUser.surname} (192.236.036.029) on." In: (2018). url: www.physiology.org/journal/jn.

[58] S Quant, A. L. Adkin, W. R. Staines, and W. E. McIlroy. "Cortical activation following a balance disturbance." eng. In: *Experimental brain research* 155.3 (2004), pp. 393–400. issn: 0014-4819 (Print). doi: 10.1007/s00221-003-1744-6.

[59] M. K. A.-H. R. Kolaghassi and K. Sirlantzis. "Systematic Review of Intelligent Algorithms in Gait Analysis and Prediction for Lower Limb Robotic Systems." In: *IEEE Access* 9 (2021). doi: 10.1109/ACCESS.2021.3104464.

[60] E. Ratti, S. Waninger, C. Berka, G. Ruffini, and A. Verma. "Comparison of medical and consumer wireless EEG systems for use in clinical trials." In: *Frontiers in Human Neuroscience* 11.August (2017). issn: 16625161. doi: 10.3389/fnhum.2017.00398.

[61] A. S. Ravindran, M Cestari, C Malaya, I John, G. E. Francisco, C Layne, and J. L. Contreras Vidal. "Interpretable Deep Learning Models for Single Trial Prediction of Balance Loss." In: *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*. Vol. 2020-Octob. 2020, pp. 268–273. doi: 10.1109/SMC42975.2020.9283206.

[62] A. S. Ravindran, C. A. Malaya, I. John, G. E. Francisco, C. Layne, and J. L. Contreras-Vidal. "Decoding neural activity preceding balance loss during standing with a lower-limb exoskeleton using an interpretable deep learning model." In: *Journal of Neural Engineering* 19.3 (2022), p. 036015. doi: 10.1088/1741-2552/ac6ca9. url: https://doi.org/10.1088/1741-2552/ac6ca9.

[63] L. Ren and Y. Peng. "Research of fall detection and fall prevention technologies: A systematic review." In: *IEEE Access* 7 (2019), pp. 77702–77722. issn: 21693536. doi: 10.1109/ACCESS.2019.2922708.

[64] M. Rubega, R. Di Marco, M. Zampini, E. Formaggio, E. Menegatti, P. Bonato, S. Masiero, and A. Del Felice. "Muscular and cortical activation during dynamic and static balance in the elderly: A scoping review." In: *Aging Brain* 1 (2021), p. 100013. issn: 25899589. doi: 10.1016/j.nbas.2021.100013. url: https://doi.org/10.1016/j.nbas.2021.100013.

[65] L. Z. Rubenstein. “Falls in older people: Epidemiology, risk factors and strategies for prevention.” In: *Age and Ageing* 35.SUPPL.2 (2006), pp. 37–41. issn: 00020729. doi: 10.1093/ageing/afl084.

[66] K. C. T. A. A.-M. N. T. A. B. S. K. Goh H. A. Abbass and J. Li. “Spatio–Spectral Representation Learning for Electroencephalographic Gait-Pattern Classification.” In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29.9 (2018). doi: 10.1109/TNSRE.2018.2864119.

[67] C. C. S. M. E. M. S. Tortora L. Tonin and F. Artoni. “Hybrid Human-Machine Interface for Gait Decoding Through Bayesian Fusion of EEG and EMG Classifiers.” In: *Frontiers in Neurorobotics* (2020), pp. 1–3. doi: 10.3389/fnbot.2020.582728.

[68] Z. Saadat, E. Sinaei, S. Pirouzi, M. Ghofrani, and M. Nami. “Cortical Activity During Postural Recovery in Response to Predictable and Unpredictable Perturbations in Healthy Young and Older Adults: A Quantitative EEG Assessment.” eng. In: *Basic and clinical neuroscience* 12.2 (2021), pp. 291–300. issn: 2008-126X (Print). doi: 10.32598/bcn.12.2.453.1.

[69] Y. Shavit and I. Klein. “Boosting Inertial-Based Human Activity Recognition with Transformers.” In: *IEEE Access* 9 (2021), pp. 53540–53547. issn: 21693536. doi: 10.1109/ACCESS.2021.3070646.

[70] A. R. Sipp, J. T. Gwin, S Makeig, and D. P. Ferris. “Loss of balance during balance beam walking elicits a multifocal theta band electrocortical response.” In: *Journal of Neurophysiology* 110.9 (2013), pp. 2050–2060. doi: 10.1152/jn.00744.2012. url: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84886998076{\&}doi=10.1152{\%}2Fjn.00744.2012{\&}partnerID=40{\&}md5=086b0f136de51a4902b0e080fc59e89f.

[71] T. Siragy, Y. Russo, W. Young, and S. E. Lamb. “Comparison of over-ground and treadmill perturbations for simulation of real-world slips and trips: A systematic review.” In: *Gait and Posture* 100.August 2022 (2023), pp. 201–209. issn: 18792219. doi: 10.1016/j.gaitpost.2022.12.015. url: https://doi.org/10.1016/j.gaitpost.2022.12.015.

[72] Y. H. A. S. R. S.Nakagome T.P. Luu and J. L. Contreras-Vidal. “An empirical comparison of neural networks and machine learning algorithms for EEG gait decoding.” In: *Nature* (2020). doi: 10.1038/s41598-020-60932-4.

[73] T Solis-Escalante, D. D. Kam, and V Weerdesteyn. “Classification of Rhythmic Cortical Activity Elicited by Whole-Body Balance Perturbations Suggests the Cortical Representation of Direction-Specific Changes in Postural Stability.” In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28.11 (2020), pp. 2566–2574. issn: 1558-0210 VO - 28. doi: 10.1109/TNSRE.2020.3028966.

[74] A. Sujatha Ravindran, C. A. Malaya, I. John, G. E. Francisco, C. Layne, and J. L. Contreras-Vidal. "Decoding neural activity preceding balance loss during standing with a lower-limb exoskeleton using an interpretable deep learning model." In: *Journal of Neural Engineering* 19.3 (2022). issn: 17412552. doi: 10.1088/1741-2552/ac6ca9.

[75] R. Y. Sun, R Kaur, L Ziegelman, S Yang, R Sowers, and M. E. Hernandez. *Using Virtual Reality to Examine the Correlation between Balance Function and Anxiety in Stance*. Ed. by I. H. Yoo, J. B. Bi, and X Hu. 2019.

[76] M. T. P. Luu, D. Eguren, M. Cestari, and J. L. Contreras-Vidal. "EEG-based Neural Decoding of Gait in Developing Children." In: *International Conference on Systems, Man and Cybernetics (SMC)* (2019). doi: 10.1109/SMC.2019.8914380.

[77] S. N. J. G. K. N. T. P. Luu Y. He and J. L. Contreras-Vidal. "Unscented Kalman Filter for Neural Decoding of Human Treadmill Walking from Non-invasive Electroencephalography." In: *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2016). doi: 10.1109/EMBC.2016.7591006.

[78] Y. H. T. P. Luu S. Nakagome and J. L. Contreras-Vidal. "Real-time EEG-based brain-computer interface to a virtual avatar enhances cortical involvement in human treadmill walking." In: *Nature* (2017). doi: 10.1038/s41598-017-09187-0.

[79] S. Tortora, S. Ghidoni, C. Chisari, S. Micera, and F. Artoni. "Deep learning-based BCI for gait decoding from EEG with LSTM recurrent neural network." In: *Journal of Neural Engineering* (2020). doi: 10.1088/1741-2552/ab9842.

[80] M. Trkov, S. Wu, K. Chen, J. Yi, T. Liu, and Q. Zhao. "Design of a Robotic Knee Assistive Device (ROKAD) for Slip-Induced Fall Prevention during Walking." In: *IFAC-PapersOnLine* 50.1 (2017), pp. 9802–9807. issn: 24058963. doi: 10.1016/j.ifacol.2017.08.887. url: https://doi.org/10.1016/j.ifacol.2017.08.887.

[81] J. P. Varghese, R. E. McIlroy, and M. Barnett-Cowan. "Perturbation-evoked potentials: Significance and application in balance control research." In: *Neuroscience & Biobehavioral Reviews* 83 (2017), pp. 267–280. issn: 0149-7634. doi: https://doi.org/10.1016/j.neubiorev.2017.10.022. url: https://www.sciencedirect.com/science/article/pii/S0149763417305699.

[82] J. P. Varghese, W. R. Staines, and W. E. McIlroy. "Activity in Functional Cortical Networks Temporally Associated with Postural Instability." eng. In: *Neuroscience* 401 (2019), pp. 43–58. issn: 1873-7544 (Electronic). doi: 10.1016/j.neuroscience.2019.01.008.

[83] J. P. Varghese, W. R. Staines, and W. E. McIlroy. "Activity in Functional Cortical Networks Temporally Associated with Postural Instability." In: *Neuroscience* 401 (2019), pp. 43–58. issn: 0306-4522. doi: https://doi.org/10.1016/j.neuroscience.2019.01.008. url: https://www.sciencedirect.com/science/article/pii/S0306452219300284.

[84] D. D. Venuto and G Mezzina. "Multisensing Architecture for the Balance Losses During Gait via Physiologic Signals Recognition." In: *IEEE Sensors Journal* 20.23 (2020), pp. 13959–13968. issn: 1558-1748 VO - 20. doi: 10.1109/JSEN.2020.2989823.

[85] C. Widdowson, J. Ganhotra, M. Faizal, M. Wilko, S. Parikh, Z. Adhami, and M. E. Hernandez. "Virtual reality applications in assessing the effect of anxiety on sensorimotor integration in human postural control." In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* 2016-October.August (2016), pp. 33–36. issn: 1557170X. doi: 10.1109/EMBC.2016.7590633.

[86] E. Wittenberg, J. Thompson, C. S. Nam, and J. R. Franz. "Neuroimaging of Human Balance Control: A Systematic Review." eng. In: *Frontiers in human neuroscience* 11 (2017), p. 170. issn: 1662-5161 (Print). doi: 10.3389/fnhum.2017.00170.

[87] S. P. A. K. Y. Zhang 2 and J. L. Contreras-Vidal. "Multiple Kernel Based Region Importance Learning for Neural Classification of Gait States from EEG Signals." In: *Frontiers in Neuroscience* (2017). doi: 10.3389/fnins.2017.00170.

Appendix A

Hold-out Results from Slip-like Perturbations Protocol - first iteration

The following tables are related to the results obtained in the first hold-out iteration, mentioned in Chapter 5. Each table contains the 5 best executions of each DL model. Every execution is identified by the following syntax: Bx_SL50_Oy_Dz.

Table 20: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: 1.6 km/h with CAR using 16 EEG channels

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B256_SL50_O80_D0.5	0.7100	0.7100	0.7200	0.4300	0.7100	0.7100
	B256_SL50_O50_D0.5	0.7200	0.7200	0.7300	0.4500	0.7200	0.7200
	B128_SL50_O50_D0.5	0.7200	0.7200	0.7400	0.4600	0.7300	0.7300
	B64_SL50_O80_D0.5	0.7100	0.7000	0.7200	0.4300	0.7100	0.7100
	B256_SL50_O80_D0.2	0.7100	0.7100	0.7200	0.4300	0.7100	0.7100
LSTM	B256_SL50_O50_D0.5	0.6700	0.6700	0.6700	0.3400	0.6700	0.6700
	B64_SL50_O80_D0.5	0.6700	0.6600	0.6800	0.3400	0.6600	0.6600
	B256_SL50_O50_D0.2	0.6700	0.6700	0.6700	0.3400	0.6700	0.6700
	B64_SL50_O80_D0.2	0.6700	0.6600	0.6800	0.3400	0.6600	0.6600
	B32_SL50_O50_D0.2	0.6700	0.6600	0.6800	0.3500	0.6700	0.6700
CNN-LSTM	B256_SL50_O80_D0.5	0.7200	0.7200	0.7200	0.4400	0.7200	0.7200
	B256_SL50_O50_D0.5	0.7200	0.7200	0.7300	0.4500	0.7200	0.7200
	B128_SL50_O80_D0.5	0.7300	0.7200	0.7300	0.4600	0.7200	0.7200
	B128_SL50_O50_D0.5	0.7300	0.7200	0.7300	0.4600	0.7300	0.7300
	B64_SL50_O50_D0.5	0.7200	0.7100	0.7200	0.4400	0.7200	0.7200
Transformer	B128_SL50_O80_D0.5	0.6900	0.6900	0.6900	0.3800	0.6900	0.6900
	B128_SL50_O50_D0.5	0.7000	0.6900	0.7000	0.3900	0.6900	0.6900
	B64_SL50_O80_D0.5	0.6900	0.6800	0.7000	0.3800	0.6900	0.6900
	B256_SL50_O80_D0.2	0.6900	0.6900	0.6900	0.3800	0.6900	0.6900
	B64_SL50_O50_D0.2	0.6900	0.6800	0.6900	0.3700	0.6800	0.6800

Table 21: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: 1.6 km/h with CAR using 5 EEG channels

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B32_SL50_O80_D0.5	0.7585	0.6960	0.7494	-	0.7440	0.7440
	B128_SL50_O80_D0.2	0.7607	0.7549	0.7838	0.5504	0.7676	0.7676
	B64_SL50_O80_D0.2	0.7603	0.7530	0.7722	0.5323	0.7606	0.7606
	B32_SL50_O80_D0.2	0.7607	0.6968	0.7373	-	0.7413	0.7413
	B32_SL50_O50_D0.2	0.7613	0.7521	0.7954	0.5643	0.7708	0.7708
LSTM	B128_SL50_O80_D0.5	0.6842	0.6780	0.6812	0.3605	0.6794	0.6794
	B64_SL50_O80_D0.5	0.6829	0.6723	0.6876	0.3637	0.6768	0.6768
	B256_SL50_O80_D0.2	0.6748	0.6689	0.6726	0.3433	0.6708	0.6708
	B128_SL50_O80_D0.2	0.6842	0.6780	0.6812	0.3605	0.6794	0.6794
	B64_SL50_O80_D0.2	0.6829	0.6723	0.6876	0.3637	0.6768	0.6768
CNN-LSTM	B128_SL50_O80_D0.5	0.7470	0.7376	0.7837	0.5366	0.7551	0.7551
	B32_SL50_O80_D0.5	0.7598	0.7053	0.7467	0	0.7519	0.7519
	B128_SL50_O80_D0.2	0.7491	0.7438	0.7638	0.5168	0.7537	0.7537
	B64_SL50_O80_D0.2	0.7585	0.7517	0.7774	0.5387	0.7624	0.7624
	B32_SL50_O80_D0.2	0.7598	0.6980	0.7156	0	0.7405	0.7405
Transformer	B128_SL50_O80_D0.5	0.6919	0.6872	0.6926	0.3823	0.6898	0.6898
	B128_SL50_O50_D0.5	0.6953	0.6925	0.6963	0.3908	0.6945	0.6945
	B64_SL50_O80_D0.5	0.6915	0.6828	0.6965	0.3838	0.6878	0.6878
	B256_SL50_O80_D0.2	0.6902	0.6850	0.6884	0.3754	0.6871	0.6871
	B64_SL50_O50_D0.2	0.6903	0.6764	0.6864	0.3674	0.6813	0.6813

Table 22: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: 1.6 km/h with CAR using 2 EEG channels

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B128_SL50_O80_D0.2	0.7474	0.7366	0.8062	0.5628	0.7602	0.7602
	B64_SL50_O80_D0.2	0.7338	0.7165	0.8062	0.5468	0.7479	0.7479
	B64_SL50_O50_D0.2	0.7327	0.7205	0.8006	0.5464	0.7497	0.7497
	B32_SL50_O80_D0.2	0.7457	0.6707	0.7212	-	0.7321	0.7313
	B32_SL50_O50_D0.2	0.7574	0.7454	0.8119	0.5807	0.7723	0.7723
LSTM	B256_SL50_O50_D0.5	0.6647	0.6594	0.6950	0.3719	0.6774	0.6774
	B128_SL50_O50_D0.5	0.6864	0.6849	0.6938	0.3852	0.6914	0.6914
	B32_SL50_O50_D0.5	0.6726	0.6643	0.6737	0.3450	0.6714	0.6714
	B256_SL50_O50_D0.2	0.6647	0.6594	0.6950	0.3719	0.6774	0.6774
	B128_SL50_O50_D0.2	0.6864	0.6849	0.6938	0.3852	0.6914	0.6914
CNN-LSTM	B256_SL50_O80_D0.2	0.7380	0.7288	0.7871	0.5346	0.7500	0.7500
	B128_SL50_O80_D0.2	0.7551	0.7459	0.8072	0.5715	0.7672	0.7672
	B64_SL50_O80_D0.2	0.7470	0.7335	0.8100	0.5655	0.7607	0.7607
	B32_SL50_O80_D0.2	0.7466	0.6761	0.7388	-	0.7350	0.7341
	B32_SL50_O50_D0.2	0.7535	0.7383	0.8116	0.5747	0.7672	0.7673
Transformer	B64_SL50_O50_D0.5	0.6716	0.6482	0.7741	0.4575	0.6926	0.6926
	B256_SL50_O80_D0.2	0.6538	0.6387	0.7006	0.3651	0.6673	0.6673
	B256_SL50_O50_D0.2	0.6617	0.6528	0.7071	0.3835	0.6777	0.6777
	B128_SL50_O80_D0.2	0.6684	0.6542	0.7154	0.3934	0.6814	0.6814
	B64_SL50_O50_D0.2	0.7101	0.6967	0.7774	0.5008	0.7270	0.7270

Table 23: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: 1.6 km/h with CAR using 1 EEG channel

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B256_SL50_O80_D0.2	0.7509	0.7402	0.8157	0.5769	0.7649	0.7649
	B128_SL50_O80_D0.2	0.7590	0.7480	0.8203	0.5885	0.7721	0.7721
	B64_SL50_O80_D0.2	0.7500	0.7352	0.8200	0.5776	0.7639	0.7639
	B64_SL50_O50_D0.2	0.7505	0.7397	0.8145	0.5776	0.7669	0.7669
	B32_SL50_O80_D0.2	0.7628	0.6880	0.7459	-	0.7459	0.7459
LSTM	B256_SL50_O80_D0.5	0.7513	0.7434	0.7951	0.5548	0.7616	0.7616
	B256_SL50_O50_D0.5	0.7554	0.7525	0.7915	0.5592	0.7683	0.7683
	B128_SL50_O50_D0.5	0.7525	0.7493	0.7858	0.5494	0.7642	0.7642
	B256_SL50_O50_D0.2	0.7554	0.7525	0.7915	0.5592	0.7683	0.7683
	B128_SL50_O50_D0.2	0.7525	0.7493	0.7858	0.5494	0.7642	0.7642
CNN-LSTM	B256_SL50_O80_D0.2	0.7628	0.7539	0.8241	0.5975	0.7765	0.7765
	B128_SL50_O80_D0.2	0.7671	0.7572	0.8267	0.6034	0.7805	0.7805
	B64_SL50_O80_D0.2	0.7551	0.7407	0.8246	0.5878	0.7695	0.7695
	B64_SL50_O50_D0.2	0.7604	0.7517	0.8170	0.5895	0.7753	0.7753
	B32_SL50_O80_D0.2	0.7556	0.6820	0.7392	-	0.7401	0.7401
Transformer	B256_SL50_O80_D0.2	0.6573	0.6304	0.7616	0.4279	0.6769	0.6769
	B256_SL50_O50_D0.2	0.6696	0.6543	0.7508	0.4363	0.6898	0.6898
	B128_SL50_O80_D0.2	0.6470	0.6129	0.7671	0.4184	0.6682	0.6682
	B128_SL50_O50_D0.2	0.7179	0.7094	0.7799	0.5122	0.7345	0.7345
	B64_SL50_O50_D0.2	0.7199	0.7055	0.7937	0.5250	0.7364	0.7364

Table 24: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: 2.5 km/h with CAR using 16 EEG channels

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B256_SL50_O50_D0.5	0.7158	0.7153	0.7254	0.4485	0.7231	0.7231
	B64_SL50_O80_D0.5	0.7049	0.6966	0.7089	-	0.7008	0.7008
	B32_SL50_O80_D0.5	0.7093	0.6371	0.6715	-	0.6672	0.6672
	B32_SL50_O50_D0.5	0.7064	0.6996	0.7103	-	0.7074	0.7027
	B32_SL50_O50_D0.2	0.7092	0.7028	0.7122	-	0.7101	0.7054
LSTM	B128_SL50_O50_D0.5	0.6717	0.6686	0.6720	0.3425	0.6706	0.6706
	B32_SL50_O80_D0.5	0.6691	0.5883	0.6281	0	0.6210	0.6210
	B128_SL50_O50_D0.2	0.6717	0.6686	0.6720	0.3425	0.6706	0.6706
	B32_SL50_O80_D0.2	0.6691	0.5583	0.6281	0	0.6210	0.6210
	B32_SL50_O50_D0.2	0.6679	0.6577	0.6673	0	0.6635	0.6588
CNN-LSTM	B256_SL50_O50_D0.5	0.7326	0.7320	0.7434	0.4835	0.7402	0.7402
	B128_SL50_O50_D0.5	0.7392	0.7385	0.7470	0.4922	0.7452	0.7452
	B64_SL50_O80_D0.5	0.7167	0.7065	0.7447	0	0.7208	0.7208
	B64_SL50_O50_D0.5	0.7186	0.7147	0.7214	0.4428	0.7215	0.7215
	B32_SL50_O50_D0.5	0.7139	0.7070	0.7304	0	0.7210	0.7163
Transformer	B256_SL50_O80_D0.5	0.6878	0.6864	0.6936	0.3854	0.6918	0.6918
	B128_SL50_O80_D0.5	0.6858	0.6800	0.6848	0	0.6826	0.6826
	B64_SL50_O50_D0.5	0.7073	0.6995	0.7308	0.4445	0.7149	0.7149
	B256_SL50_O80_D0.2	0.6939	0.6919	0.6947	0.3889	0.6942	0.6942
	B256_SL50_O50_D0.2	0.6979	0.6963	0.6985	0.3969	0.6985	0.6985

Table 25: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: 2.5 km/h with CAR using 5 EEG channels

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B32_SL50_O50_D0.2	0.7326	0.7303	0.7607	0.5047	0.7442	0.7442
	B32_SL50_O80_D0.5	0.7301	0.7244	0.7767	0.5185	0.7430	0.7430
	B32_SL50_O80_D0.2	0.7289	0.7265	0.7535	0.4916	0.7383	0.7383
	B64_SL50_O80_D0.2	0.7280	0.7262	0.7491	0.4858	0.7368	0.7368
	B128_SL50_O50_D0.2	0.7261	0.7230	0.7584	0.4965	0.7385	0.7385
LSTM	B32_SL50_O80_D0.2	0.6602	0.6596	0.6598	0.3203	0.6605	0.6605
	B32_SL50_O80_D0.5	0.6602	0.6596	0.6598	0.3203	0.6605	0.6605
	B64_SL50_O80_D0.2	0.6589	0.6573	0.6574	0.3146	0.6572	0.6572
	B64_SL50_O80_D0.5	0.6589	0.6573	0.6574	0.3146	0.6572	0.6572
	B256_SL50_O50_D0.2	0.6520	0.6510	0.6511	0.3029	0.6519	0.6519
CNN-LSTM	B64_SL50_O80_D0.2	0.7366	0.7360	0.7480	0.4909	0.7429	0.7429
	B128_SL50_O50_D0.2	0.7355	0.7333	0.7627	0.5093	0.7468	0.7468
	B32_SL50_O50_D0.2	0.7326	0.7319	0.7481	0.4892	0.7411	0.7411
	B64_SL50_O50_D0.2	0.7317	0.7307	0.7489	0.4896	0.7407	0.7407
	B64_SL50_O80_D0.5	0.7289	0.7246	0.7660	0.5058	0.7405	0.7405
Transformer	B64_SL50_O80_D0.2	0.7272	0.7252	0.7493	0.4854	0.7362	0.7362
	B64_SL50_O80_D0.5	0.7146	0.7112	0.7442	0.4690	0.7252	0.7252
	B64_SL50_O50_D0.5	0.7036	0.7008	0.7305	0.4454	0.7152	0.7152
	B64_SL50_O50_D0.2	0.6886	0.6886	0.6932	0.3861	0.6929	0.6929
	B256_SL50_O80_D0.2	0.6740	0.6740	0.6762	0.3527	0.6765	0.6765

Table 26: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: 2.5 km/h with CAR using 2 EEG channels

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B64_SL50_O50_D0.2	0.7355	0.7280	0.8010	0.5515	0.7526	0.7526
	B64_SL50_O80_D0.2	0.7325	0.7248	0.7932	0.5382	0.7470	0.7470
	B128_SL50_O80_D0.2	0.7293	0.7217	0.7878	0.5295	0.7436	0.7436
	B32_SL50_O80_D0.2	0.7232	0.7125	0.7997	0.5357	0.7393	0.7393
	B32_SL50_O50_D0.2	0.7205	0.7100	0.8004	0.5364	0.7394	0.7394
LSTM	B64_SL50_O80_D0.2	0.6398	0.6396	0.6403	0.2812	0.6409	0.6409
	B64_SL50_O80_D0.5	0.6398	0.6396	0.6403	0.2812	0.6409	0.6409
	B256_SL50_O50_D0.2	0.6313	0.6285	0.6505	0.2922	0.6418	0.6418
	B256_SL50_O50_D0.5	0.6313	0.6285	0.6505	0.2922	0.6418	0.6418
	B32_SL50_O80_D0.2	0.6285	0.6278	0.6279	0.2564	0.6285	0.6285
CNN-LSTM	B64_SL50_O80_D0.2	0.7289	0.7208	0.7901	0.5316	0.7435	0.7435
	B32_SL50_O80_D0.2	0.7272	0.7174	0.7999	0.5399	0.7430	0.7430
	B128_SL50_O80_D0.2	0.7211	0.7127	0.7822	0.5159	0.7358	0.7358
	B256_SL50_O80_D0.2	0.7203	0.7116	0.7828	0.5158	0.7352	0.7352
	B64_SL50_O50_D0.2	0.7186	0.7099	0.7857	0.5196	0.7362	0.7362
Transformer	B64_SL50_O50_D0.2	0.6745	0.6593	0.7579	0.4483	0.6948	0.6948
	B64_SL50_O50_D0.5	0.6492	0.6277	0.7474	0.4121	0.6716	0.6716
	B128_SL50_O50_D0.2	0.6445	0.6351	0.6898	0.3493	0.6607	0.6607
	B128_SL50_O80_D0.2	0.6386	0.6090	0.7571	0.4055	0.6599	0.6599
	B256_SL50_O50_D0.2	0.6360	0.6207	0.7003	0.3528	0.6553	0.6553

Table 27: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: 2.5 km/h with CAR using 1 EEG channel

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B128_SL50_O80_D0.2	0.7341	0.7240	0.8126	0.5594	0.7503	0.7503
	B64_SL50_O80_D0.2	0.7272	0.7159	0.8101	0.5499	0.7438	0.7438
	B32_SL50_O80_D0.2	0.7252	0.7136	0.8091	0.5470	0.7419	0.7419
	B64_SL50_O80_D0.5	0.7252	0.7134	0.8105	0.5483	0.7421	0.7421
	B32_SL50_O80_D0.5	0.7248	0.7130	0.8096	0.5470	0.7416	0.7416
LSTM	B128_SL50_O50_D0.2	0.7233	0.7211	0.7487	0.4828	0.7343	0.7343
	B128_SL50_O50_D0.5	0.7233	0.7211	0.7487	0.4828	0.7343	0.7343
	B64_SL50_O50_D0.2	0.7223	0.7196	0.7521	0.4861	0.7343	0.7343
	B64_SL50_O50_D0.5	0.7223	0.7196	0.7521	0.4861	0.7343	0.7343
	B64_SL50_O80_D0.2	0.7126	0.7079	0.7490	0.4726	0.7243	0.7243
CNN-LSTM	B64_SL50_O80_D0.2	0.7333	0.7237	0.8078	0.5538	0.7491	0.7491
	B32_SL50_O80_D0.2	0.7309	0.7206	0.8085	0.5521	0.7470	0.7470
	B128_SL50_O80_D0.2	0.7264	0.7155	0.8057	0.5449	0.7428	0.7428
	B32_SL50_O80_D0.5	0.7264	0.7154	0.8070	0.5461	0.7429	0.7429
	B128_SL50_O80_D0.5	0.7215	0.7096	0.8054	0.5396	0.7383	0.7383
Transformer	B256_SL50_O80_D0.2	0.6736	0.6544	0.7692	0.4551	0.6923	0.6923
	B64_SL50_O50_D0.2	0.6660	0.6442	0.7796	0.4600	0.6892	0.6892
	B256_SL50_O50_D0.2	0.6576	0.6372	0.7570	0.4300	0.6799	0.6799
	B128_SL50_O80_D0.2	0.6476	0.6207	0.7614	0.4194	0.6683	0.6683
	B128_SL50_O50_D0.2	0.6398	0.6135	0.7537	0.4076	0.6638	0.6638

Table 28: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 16 EEG channels

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B128_SL50_O50_D0.5	0.7195	0.7194	0.7278	0.4534	0.7256	0.7256
	B64_SL50_O50_D0.5	0.7195	0.7190	0.7322	0.4594	0.7272	0.7272
	B256_SL50_O50_D0.5	0.7176	0.7176	0.7200	0.4409	0.7208	0.7208
	B64_SL50_O80_D0.5	0.7130	0.7130	0.7175	0.4343	0.7168	0.7168
	B256_SL50_O50_D0.2	0.7120	0.7115	0.7119	0.4250	0.7131	0.7131
LSTM	B32_SL50_O50_D0.5	0.6707	0.6677	0.6685	0.3359	0.6674	0.6674
	B32_SL50_O50_D0.2	0.6698	0.6687	0.6685	0.3378	0.6693	0.6693
	B64_SL50_O50_D0.2	0.6557	0.6540	0.6539	0.3081	0.6542	0.6542
	B64_SL50_O50_D0.5	0.6557	0.6540	0.6539	0.3081	0.6542	0.6542
	B256_SL50_O50_D0.2	0.6538	0.6515	0.6517	0.3031	0.6514	0.6514
CNN-LSTM	B64_SL50_O50_D0.5	0.7317	0.7316	0.7332	0.4675	0.7343	0.7343
	B128_SL50_O50_D0.5	0.7298	0.7296	0.7392	0.4755	0.7364	0.7364
	B256_SL50_O50_D0.5	0.7298	0.7296	0.7392	0.4755	0.7364	0.7364
	B256_SL50_O80_D0.5	0.7280	0.7275	0.7392	0.4735	0.7343	0.7343
	B128_SL50_O80_D0.5	0.7195	0.7195	0.7233	0.4463	0.7230	0.7230
Transformer	B64_SL50_O80_D0.5	0.6915	0.6912	0.6990	0.3955	0.6966	0.6966
	B128_SL50_O50_D0.2	0.6914	0.6912	0.6929	0.3867	0.6938	0.6938
	B256_SL50_O50_D0.2	0.6895	0.6866	0.6876	0.3737	0.6861	0.6861
	B32_SL50_O50_D0.2	0.6886	0.6819	0.7339	0.4367	0.7038	0.7038
	B256_SL50_O80_D0.2	0.6882	0.6873	0.6872	0.3749	0.6877	0.6877

Table 29: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 5 EEG channels

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B32_SL50_O50_D0.2	0.7326	0.7303	0.7607	0.5047	0.7442	0.7442
	B32_SL50_O80_D0.5	0.7301	0.7244	0.7767	0.5185	0.7430	0.7430
	B32_SL50_O80_D0.2	0.7289	0.7265	0.7535	0.4916	0.7383	0.7383
	B64_SL50_O80_D0.2	0.7280	0.7262	0.7491	0.4858	0.7368	0.7368
	B128_SL50_O50_D0.2	0.7261	0.7230	0.7584	0.4965	0.7385	0.7385
LSTM	B32_SL50_O80_D0.2	0.6602	0.6596	0.6598	0.3203	0.6605	0.6605
	B32_SL50_O80_D0.5	0.6602	0.6596	0.6598	0.3203	0.6605	0.6605
	B64_SL50_O80_D0.2	0.6589	0.6573	0.6574	0.3146	0.6572	0.6572
	B64_SL50_O80_D0.5	0.6589	0.6573	0.6574	0.3146	0.6572	0.6572
	B256_SL50_O50_D0.2	0.6520	0.6510	0.6511	0.3029	0.6519	0.6519
CNN-LSTM	B64_SL50_O80_D0.2	0.7366	0.7360	0.7480	0.4909	0.7429	0.7429
	B128_SL50_O50_D0.2	0.7355	0.7333	0.7627	0.5093	0.7468	0.7468
	B32_SL50_O50_D0.2	0.7326	0.7319	0.7481	0.4892	0.7411	0.7411
	B64_SL50_O50_D0.2	0.7317	0.7307	0.7489	0.4896	0.7407	0.7407
	B64_SL50_O80_D0.5	0.7289	0.7246	0.7660	0.5058	0.7405	0.7405
Transformer	B64_SL50_O80_D0.2	0.7272	0.7252	0.7493	0.4854	0.7362	0.7362
	B64_SL50_O80_D0.5	0.7146	0.7112	0.7442	0.4690	0.7252	0.7252
	B64_SL50_O50_D0.5	0.7036	0.7008	0.7305	0.4454	0.7152	0.7152
	B64_SL50_O50_D0.2	0.6886	0.6886	0.6932	0.3861	0.6929	0.6929
	B256_SL50_O80_D0.2	0.6740	0.6740	0.6762	0.3527	0.6765	0.6765

Table 30: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 2 EEG channels

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B32_SL50_O80_D0.2	0.7370	0.7299	0.7952	0.5445	0.7511	0.7511
	B32_SL50_O50_D0.2	0.7355	0.7277	0.8036	0.5542	0.7529	0.7529
	B128_SL50_O80_D0.2	0.7337	0.7250	0.8021	0.5485	0.7490	0.7490
	B64_SL50_O50_D0.2	0.7326	0.7247	0.8009	0.5487	0.7501	0.7501
	B64_SL50_O80_D0.2	0.7309	0.7222	0.7977	0.5413	0.7460	0.7460
LSTM	B64_SL50_O50_D0.2	0.6295	0.6294	0.6320	0.2644	0.6324	0.6324
	B64_SL50_O50_D0.5	0.6295	0.6294	0.6320	0.2644	0.6324	0.6324
	B32_SL50_O50_D0.2	0.6285	0.6284	0.6342	0.2676	0.6334	0.6334
	B128_SL50_O50_D0.2	0.6266	0.6266	0.6311	0.2619	0.6308	0.6308
	B128_SL50_O50_D0.5	0.6266	0.6266	0.6311	0.2619	0.6308	0.6308
CNN-LSTM	B64_SL50_O50_D0.2	0.7308	0.7245	0.7865	0.5318	0.7468	0.7468
	B32_SL50_O80_D0.2	0.7293	0.7201	0.7986	0.5406	0.7447	0.7447
	B32_SL50_O50_D0.2	0.7233	0.7152	0.7882	0.5267	0.7406	0.7406
	B64_SL50_O80_D0.2	0.7232	0.7140	0.7897	0.5256	0.7384	0.7384
	B128_SL50_O80_D0.2	0.7224	0.7128	0.7910	0.5261	0.7378	0.7378
Transformer	B64_SL50_O50_D0.2	0.6585	0.6393	0.7524	0.4267	0.6803	0.6803
	B128_SL50_O80_D0.2	0.6520	0.6361	0.7198	0.3854	0.6689	0.6689
	B128_SL50_O50_D0.2	0.6454	0.6327	0.7036	0.3650	0.6636	0.6636
	B256_SL50_O50_D0.2	0.6407	0.6323	0.6813	0.3366	0.6562	0.6562
	B256_SL50_O50_D0.5	0.6257	0.5998	0.7255	0.3667	0.6491	0.6491

Table 31: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 1 EEG channel

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B32_SL50_O80_D0.2	0.7337	0.7236	0.8117	0.5582	0.7498	0.7498
	B256_SL50_O80_D0.2	0.7321	0.7220	0.8091	0.5539	0.7481	0.7481
	B64_SL50_O80_D0.2	0.7313	0.7209	0.8100	0.5539	0.7475	0.7475
	B32_SL50_O80_D0.5	0.7293	0.7183	0.8110	0.5529	0.7458	0.7458
	B64_SL50_O80_D0.5	0.7293	0.7182	0.8116	0.5535	0.7458	0.7458
LSTM	B64_SL50_O50_D0.2	0.7326	0.7288	0.7718	0.5174	0.7462	0.7462
	B64_SL50_O50_D0.5	0.7326	0.7288	0.7718	0.5174	0.7462	0.7462
	B128_SL50_O80_D0.2	0.7244	0.7180	0.7736	0.5100	0.7377	0.7377
	B128_SL50_O80_D0.5	0.7244	0.7180	0.7736	0.5100	0.7377	0.7377
	B32_SL50_O50_D0.5	0.7233	0.7220	0.7420	0.4747	0.7327	0.7327
CNN-LSTM	B64_SL50_O80_D0.2	0.7402	0.7315	0.8112	0.5640	0.7556	0.7556
	B128_SL50_O80_D0.2	0.7337	0.7240	0.8086	0.5550	0.7496	0.7496
	B64_SL50_O80_D0.5	0.7309	0.7205	0.8091	0.5527	0.7471	0.7471
	B64_SL50_O50_D0.2	0.7289	0.7195	0.8059	0.5502	0.7474	0.7474
	B256_SL50_O80_D0.2	0.7272	0.7163	0.8074	0.5473	0.7436	0.7436
Transformer	B128_SL50_O80_D0.2	0.6772	0.6592	0.7696	0.4594	0.6957	0.6957
	B256_SL50_O80_D0.2	0.6724	0.6532	0.7672	0.4519	0.6911	0.6911
	B256_SL50_O50_D0.2	0.6604	0.6410	0.7568	0.4327	0.6823	0.6823
	B128_SL50_O50_D0.2	0.6482	0.6254	0.7522	0.4155	0.6712	0.6712
	B128_SL50_O50_D0.5	0.6144	0.5791	0.7456	0.3716	0.6405	0.6405

Table 32: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: 1.6 km/h with ICA using 16 components

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B32_SL50_O80_D0.2	0.5919	0.5915	0.5975	0.1939	0.5964	0.5964
	B128_SL50_O50_D0.2	0.5880	0.5873	0.5968	0.1917	0.5948	0.5948
	B256_SL50_O50_D0.2	0.5863	0.5813	0.6071	0.2052	0.5983	0.5983
	B32_SL50_O50_D0.2	0.5846	0.5843	0.5910	0.1811	0.5901	0.5901
	B64_SL50_O50_D0.5	0.5812	0.5688	0.6174	0.2143	0.5978	0.5978
LSTM	B128_SL50_O80_D0.2	0.6085	0.6085	0.6101	0.2205	0.6104	0.6104
	B128_SL50_O80_D0.5	0.6085	0.6085	0.6101	0.2205	0.6104	0.6104
	B256_SL50_O50_D0.2	0.5949	0.5931	0.5930	0.1863	0.5933	0.5933
	B256_SL50_O50_D0.5	0.5949	0.5931	0.5930	0.1863	0.5933	0.5933
	B256_SL50_O80_D0.2	0.5948	0.5948	0.5976	0.1952	0.5976	0.5976
CNN-LSTM	B32_SL50_O80_D0.2	0.5970	0.5959	0.6061	0.2094	0.6033	0.6033
	B128_SL50_O50_D0.5	0.5957	0.5802	0.6442	0.2566	0.6142	0.6142
	B64_SL50_O50_D0.5	0.5932	0.5856	0.6211	0.2278	0.6071	0.6071
	B128_SL50_O50_D0.2	0.5923	0.5894	0.6083	0.2106	0.6024	0.6024
	B32_SL50_O50_D0.5	0.5915	0.5853	0.6156	0.2196	0.6044	0.6044
Transformer	B32_SL50_O80_D0.5	0.6637	0.6583	0.6631	0.3213	0.6582	0.6582
	B256_SL50_O80_D0.2	0.6304	0.6304	0.6329	0.2659	0.6330	0.6330
	B64_SL50_O50_D0.2	0.6171	0.6168	0.6182	0.2369	0.6188	0.6188
	B128_SL50_O80_D0.2	0.6078	0.6065	0.6064	0.2131	0.6066	0.6066
	B128_SL50_O50_D0.2	0.6060	0.6059	0.6113	0.2220	0.6107	0.6107

Table 33: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: 2.5 km/h with ICA using 16 components

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B32_SL50_O50_D0.2	0.4974	0.4563	0.5356	0.0555	0.5217	0.5217
	B128_SL50_O80_D0.5	0.4952	0.4274	0.5473	0.0622	0.5204	0.5204
	B128_SL50_O50_D0.2	0.4949	0.4424	0.5410	0.0594	0.5216	0.5216
	B32_SL50_O50_D0.5	0.4872	0.4000	0.5658	0.0718	0.5196	0.5196
	B128_SL50_O50_D0.5	0.4863	0.3967	0.5674	0.0716	0.5190	0.5190
LSTM	B32_SL50_O50_D0.5	0.5111	0.5042	0.5257	0.0489	0.5233	0.5233
	B32_SL50_O80_D0.2	0.4996	0.4816	0.5177	0.0318	0.5143	0.5143
	B32_SL50_O80_D0.5	0.4996	0.4816	0.5177	0.0318	0.5143	0.5143
	B64_SL50_O80_D0.2	0.4926	0.4789	0.5066	0.0123	0.5057	0.5057
	B64_SL50_O80_D0.5	0.4926	0.4789	0.5066	0.0123	0.5057	0.5057
CNN-LSTM	B256_SL50_O80_D0.2	0.5219	0.4721	0.5830	0.1219	0.5447	0.5447
	B64_SL50_O80_D0.2	0.5170	0.5031	0.5363	0.0666	0.5306	0.5306
	B256_SL50_O80_D0.5	0.5078	0.4218	0.6198	0.1311	0.5359	0.5359
	B128_SL50_O80_D0.2	0.5056	0.4866	0.5259	0.0462	0.5207	0.5207
	B32_SL50_O80_D0.2	0.5037	0.4965	0.5154	0.0293	0.5140	0.5140
Transformer	B32_SL50_O50_D0.2	0.5385	0.3500	0.2692	0	0.5000	0.5000
	B64_SL50_O50_D0.5	0.5385	0.3500	0.2692	0	0.5000	0.5000
	B32_SL50_O80_D0.2	0.5333	0.3478	0.2667	0	0.5000	0.5000
	B32_SL50_O80_D0.5	0.5333	0.3478	0.2667	0	0.5000	0.5000
	B64_SL50_O80_D0.5	0.5333	0.3478	0.2667	0	0.5000	0.5000

Table 34: Results from Hold-out method in slip-like perturbations protocol - first iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with ICA using 16 components

	Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	B128_SL50_O50_D0.2	0.5119	0.4383	0.5619	0.0804	0.5261	0.5261
	B64_SL50_O50_D0.2	0.5086	0.4590	0.5350	0.0537	0.5206	0.5206
	B256_SL50_O50_D0.2	0.5078	0.4242	0.5626	0.0753	0.5226	0.5226
	B32_SL50_O50_D0.2	0.5037	0.4505	0.5281	0.0424	0.5159	0.5159
	B256_SL50_O50_D0.5	0.5029	0.3909	0.5974	0.0871	0.5195	0.5195
LSTM	B32_SL50_O50_D0.5	0.5300	0.5079	0.5490	0.0871	0.5386	0.5386
	B32_SL50_O80_D0.2	0.5135	0.4880	0.5388	0.0670	0.5289	0.5289
	B32_SL50_O80_D0.5	0.5135	0.4880	0.5388	0.0670	0.5289	0.5289
	B256_SL50_O50_D0.2	0.5004	0.4724	0.5128	0.0222	0.5096	0.5096
	B256_SL50_O50_D0.5	0.5004	0.4724	0.5128	0.0222	0.5096	0.5096
CNN-LSTM	B128_SL50_O80_D0.2	0.5259	0.5034	0.5532	0.0932	0.5408	0.5408
	B128_SL50_O80_D0.5	0.5232	0.4822	0.5689	0.1080	0.5423	0.5423
	B256_SL50_O80_D0.2	0.5226	0.4784	0.5716	0.1099	0.5422	0.5422
	B32_SL50_O80_D0.2	0.5195	0.4957	0.5457	0.0796	0.5347	0.5347
	B32_SL50_O50_D0.2	0.5160	0.4493	0.5641	0.0872	0.5297	0.5297
Transformer	B32_SL50_O80_D0.2	0.5303	0.3465	0.2652	0	0.5000	0.5000
	B32_SL50_O80_D0.5	0.5303	0.3465	0.2652	0	0.5000	0.5000
	B256_SL50_O50_D0.2	0.5210	0.4821	0.5484	0.0785	0.5319	0.5319
	B32_SL50_O50_D0.2	0.5185	0.3415	0.2593	0	0.5000	0.5000
	B32_SL50_O50_D0.5	0.5185	0.3415	0.2593	0	0.5000	0.5000

Appendix B

Leave-one-out Results from Slip-like Perturbations Protocol - first iteration

The following tables are related to the results obtained in the first leave-one-out iteration, mentioned in Chapter 5.

Table 35: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: 1.6 km/h with CAR using 16 EEG channels

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6667	0.6136	0.6697	0.3213	0.6584	0.6584
	Std dev	0.0945	0.0616	0.0734	0.1213	0.0688	0.0688
LSTM	Mean	0.6281	0.5933	0.6242	0.2545	0.6364	0.6364
	Std dev	0.0463	0.0525	0.0524	0.0834	0.0476	0.0476
CNN-LSTM	Mean	0.6585	0.6067	0.6591	0.3049	0.6513	0.6513
	Std dev	0.0928	0.0669	0.0768	0.1289	0.0653	0.0653
Transformer	Mean	0.6223	0.5754	0.6188	0.2352	0.6223	0.6223
	Std dev	0.0698	0.0694	0.0693	0.1216	0.0648	0.0648

Table 36: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: 1.6 km/h with CAR using 5 EEG channels

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6646	0.5961	0.7061	0.3482	0.6542	0.6542
	Std dev	0.1060	0.0715	0.0630	0.0937	0.0572	0.0572
LSTM	Mean	0.6244	0.5888	0.6277	0.2575	0.6361	0.6361
	Std dev	0.0454	0.0640	0.0593	0.0933	0.0463	0.0463
CNN-LSTM	Mean	0.6603	0.5940	0.6936	0.3353	0.6529	0.6529
	Std dev	0.1038	0.0765	0.0709	0.1078	0.0622	0.0622
Transformer	Mean	0.6438	0.5838	0.6501	0.2783	0.6361	0.6361
	Std dev	0.0897	0.0623	0.0598	0.1011	0.0638	0.0638

Table 37: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: 1.6 km/h with CAR using 2 EEG channels

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6700	0.5945	0.7169	0.3502	0.6506	0.6506
	Std dev	0.1256	0.0993	0.0460	0.0980	0.0712	0.0712
LSTM	Mean	0.6457	0.5816	0.6671	0.2956	0.6394	0.6394
	Std dev	0.0963	0.0748	0.0736	0.1119	0.0661	0.0661
CNN-LSTM	Mean	0.6823	0.6189	0.7154	0.3733	0.6683	0.6683
	Std dev	0.1031	0.0720	0.0575	0.0871	0.0537	0.0537
Transformer	Mean	0.6548	0.5858	0.6712	0.3000	0.6386	0.6386
	Std dev	0.1121	0.0799	0.0730	0.1160	0.0686	0.0686

Table 38: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: 1.6 km/h with CAR using 1 EEG channel

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6834	0.6176	0.7175	0.3705	0.6627	0.6627
	Std dev	0.1134	0.0715	0.0462	0.0747	0.0499	0.0499
LSTM	Mean	0.6444	0.5708	0.6727	0.2922	0.6330	0.6330
	Std dev	0.1144	0.0830	0.0741	0.1149	0.0740	0.0740
CNN-LSTM	Mean	0.6841	0.6228	0.7120	0.3730	0.6700	0.6700
	Std dev	0.1034	0.0685	0.0643	0.0957	0.0548	0.0548
Transformer	Mean	0.6478	0.5653	0.6859	0.2989	0.6291	0.6291
	Std dev	0.1291	0.1002	0.0688	0.1148	0.0745	0.0745

Table 39: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: 2.5 km/h with CAR using 16 EEG channels

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6297	0.5813	0.6436	0.2592	0.6198	0.6198
	Std dev	0.0632	0.0641	0.0772	0.1201	0.0469	0.0469
LSTM	Mean	0.6036	0.5719	0.6020	0.2052	0.6055	0.6055
	Std dev	0.0544	0.0759	0.0655	0.1179	0.0529	0.0529
CNN-LSTM	Mean	0.6251	0.5738	0.6419	0.2535	0.6167	0.6167
	Std dev	0.0713	0.0831	0.0739	0.1228	0.0515	0.0515
Transformer	Mean	0.5478	0.3470	0.2739	0	0.5000	0.5000
	Std dev	0.1819	0.0657	0.0909	0	0	0

Table 40: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: 2.5 km/h with CAR using 5 EEG channels

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6577	0.6095	0.6713	0.3105	0.6446	0.6446
	Std dev	0.0531	0.0550	0.0808	0.1154	0.0396	0.0396
LSTM	Mean	0.6003	0.5666	0.5988	0.1942	0.5976	0.5976
	Std dev	0.0447	0.0620	0.0547	0.0930	0.0383	0.0383
CNN-LSTM	Mean	0.6595	0.6133	0.6606	0.2996	0.6427	0.6427
	Std dev	0.0648	0.0628	0.0798	0.1261	0.0488	0.0488
Transformer	Mean	0.5528	0.3505	0.3236	0.0074	0.5007	0.5007
	Std dev	0.1798	0.0642	0.1618	0.0234	0.0021	0.0021

Table 41: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: 2.5 km/h with CAR using 2 EEG channels

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.7016	0.6340	0.7252	0.3828	0.6671	0.6671
	Std dev	0.0841	0.0422	0.0610	0.0796	0.0311	0.0311
LSTM	Mean	0.6227	0.5906	0.6105	0.2221	0.6134	0.6134
	Std dev	0.0208	0.0527	0.0481	0.0800	0.0295	0.0295
CNN-LSTM	Mean	0.7077	0.6549	0.7068	0.3833	0.6798	0.6798
	Std dev	0.0665	0.0473	0.0755	0.1128	0.0420	0.0420
Transformer	Mean	0.5478	0.3470	0.2739	0	0.5000	0.5000
	Std dev	0.1819	0.0657	0.0909	0	0	0

Table 42: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: 2.5 km/h with CAR using 1 EEG channel

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6897	0.6303	0.7088	0.3656	0.6638	0.6638
	Std dev	0.0699	0.0531	0.0792	0.1094	0.0411	0.0411
LSTM	Mean	0.6909	0.6407	0.6995	0.3659	0.6705	0.6705
	Std dev	0.0491	0.0549	0.0841	0.1266	0.0457	0.0457
CNN-LSTM	Mean	0.7056	0.6506	0.7128	0.3866	0.6782	0.6782
	Std dev	0.0655	0.0489	0.0784	0.1144	0.0421	0.0421
Transformer	Mean	0.5478	0.3470	0.2739	0	0.5000	0.5000
	Std dev	0.1819	0.0657	0.0909	0	0	0

Table 43: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 16 EEG channels

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6340	0.5780	0.6457	0.2607	0.6205	0.6205
	Std dev	0.0744	0.0761	0.0737	0.1170	0.0483	0.0483
LSTM	Mean	0.6077	0.5749	0.6095	0.2163	0.6095	0.6095
	Std dev	0.0433	0.0616	0.0523	0.0883	0.0364	0.0364
CNN-LSTM	Mean	0.6343	0.5877	0.6386	0.2598	0.6247	0.6247
	Std dev	0.0668	0.0745	0.0772	0.1278	0.0529	0.0529
Transformer	Mean	0.5521	0.3489	0.2761	0	0.5000	0.5000
	Std dev	0.1802	0.0649	0.0901	0	0	0

Table 44: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 5 EEG channels

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6639	0.6116	0.6724	0.3127	0.6453	0.6453
	Std dev	0.0638	0.0544	0.0758	0.1113	0.0409	0.0409
LSTM	Mean	0.6026	0.5677	0.6004	0.1976	0.5994	0.5994
	Std dev	0.0496	0.0674	0.0584	0.1021	0.0435	0.0435
CNN-LSTM	Mean	0.6623	0.6154	0.6622	0.3036	0.6452	0.6452
	Std dev	0.0629	0.0553	0.0766	0.1175	0.0434	0.0434
Transformer	Mean	0.5521	0.3489	0.2761	0	0.5000	0.5000
	Std dev	0.1802	0.0649	0.0901	0	0	0

Table 45: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 2 EEG channels

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6947	0.6406	0.7013	0.3665	0.6693	0.6693
	Std dev	0.0667	0.0428	0.0731	0.1041	0.0369	0.0369
LSTM	Mean	0.6362	0.6039	0.6209	0.2451	0.6262	0.6262
	Std dev	0.0284	0.0557	0.0515	0.0869	0.0330	0.0330
CNN-LSTM	Mean	0.6839	0.6664	0.7064	0.4122	0.6320	0.6728
	Std dev	0.0594	0.0661	0.0645	0.0993	0.1485	0.0338
Transformer	Mean	0.5387	0.3754	0.3058	0.0430	0.5243	0.5243
	Std dev	0.1521	0.1207	0.1528	0.0909	0.0512	0.0512

Table 46: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 1 EEG channel

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.6888	0.6289	0.7091	0.3656	0.6627	0.6627
	Std dev	0.0724	0.0493	0.0770	0.1094	0.0414	0.0414
LSTM	Mean	0.6912	0.6404	0.6960	0.3610	0.6688	0.6688
	Std dev	0.0559	0.0483	0.0782	0.1151	0.0406	0.0406
CNN-LSTM	Mean	0.6962	0.6390	0.7070	0.3711	0.6687	0.6687
	Std dev	0.0727	0.0475	0.0736	0.1090	0.0423	0.0423
Transformer	Mean	0.5521	0.3489	0.2761	0	0.5000	0.5000
	Std dev	0.1802	0.0649	0.0901	0	0	0

Table 47: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: 1.6 km/h with ICA using 16 components

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.5954	0.5746	0.6115	0.2142	0.6055	0.6055
	Std dev	0.0443	0.0614	0.0550	0.0946	0.0431	0.0431
LSTM	Mean	0.5687	0.5587	0.5782	0.1604	0.5842	0.5842
	Std dev	0.0421	0.0586	0.0342	0.0638	0.0341	0.0341
CNN-LSTM	Mean	0.5813	0.5619	0.5989	0.1898	0.5928	0.5928
	Std dev	0.0400	0.0596	0.0466	0.0829	0.0374	0.0374
Transformer	Mean	0.5889	0.5870	0.5971	0.1911	0.5941	0.5941
	Std dev	0.0421	0.0426	0.0428	0.0850	0.0422	0.0422

Table 48: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: 2.5 km/h with ICA using 16 components

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.5533	0.5317	0.5809	0.1527	0.5741	0.5741
	Std dev	0.0618	0.0787	0.0502	0.0948	0.0462	0.0462
LSTM	Mean	0.5444	0.5305	0.5586	0.1190	0.5620	0.5620
	Std dev	0.0501	0.0624	0.0398	0.0779	0.0417	0.0417
CNN-LSTM	Mean	0.5444	0.5280	0.5676	0.1290	0.5629	0.5629
	Std dev	0.0590	0.0714	0.0539	0.0968	0.0449	0.0449
Transformer	Mean	0.4547	0.3474	0.2990	0.0540	0.5232	0.5232
	Std dev	0.1362	0.1429	0.1993	0.1147	0.0489	0.0489

Table 49: Results from Leave-one-out method in slip-like perturbations protocol - first iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with ICA using 16 components

Model		Accuracy	F1-Score	Precision	MCC	Recall	Specificity
CNN	Mean	0.5550	0.5313	0.5830	0.1553	0.5751	0.5751
	Std dev	0.0592	0.0762	0.0495	0.0903	0.0437	0.0437
LSTM	Mean	0.5389	0.5237	0.5528	0.1071	0.5557	0.5557
	Std dev	0.0432	0.0570	0.0316	0.0622	0.0337	0.0337
CNN-LSTM	Mean	0.5451	0.5235	0.5694	0.1323	0.5644	0.5644
	Std dev	0.0554	0.0723	0.0531	0.0921	0.0411	0.0411
Transformer	Mean	0.4306	0.2960	0.2153	0	0.5000	0.5000
	Std dev	0.1140	0.0703	0.0570	0	0	0

Appendix C

Hold-out Results from Slip-like Perturbations Protocol - second iteration

The following tables are related to the results obtained in the second hold-out iteration, mentioned in Chapter 5. Each table contains the 2 best executions obtained in the first hold-out iteration with CNN-LSTM, since it was the best DL model in the first hold-out iteration. Every execution is identified by the following syntax: Bx_SL50_Oy_Dz.

Table 50: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: 1.6 km/h with CAR using 16 EEG channels

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.7305	0.5604	0.5654	0.1934	0.6431	0.6431
B64_SL50_O50_D0.2	0.6983	0.5422	0.5585	0.1798	0.6381	0.6381

Table 51: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: 1.6 km/h with CAR using 5 EEG channels

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.6352	0.5077	0.5495	0.1597	0.6288	0.6288
B64_SL50_O50_D0.2	0.5873	0.4800	0.5433	0.1423	0.6169	0.6169

Table 52: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: 1.6 km/h with CAR using 2 EEG channels

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.8102	0.6252	0.6077	0.2827	0.6856	0.6856
B64_SL50_O50_D0.2	0.7729	0.6029	0.5935	0.2658	0.6888	0.6888

Table 53: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: 1.6 km/h with CAR using 1 EEG channel

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B64_SL50_O50_D0.2	0.8669	0.6759	0.6600	0.3566	0.6987	0.6987
B128_SL50_O50_D0.2	0.8623	0.6629	0.6487	0.3301	0.6832	0.6832

Table 54: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: 2.5 km/h with CAR using 16 EEG channels

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.5779	0.4585	0.5216	0.0706	0.5578	0.5578
B64_SL50_O50_D0.2	0.4845	0.4095	0.5181	0.0595	0.5491	0.5491

Table 55: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: 2.5 km/h with CAR using 5 EEG channels

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.7739	0.5622	0.5576	0.1533	0.6020	0.6020
B64_SL50_O50_D0.2	0.7491	0.5432	0.5450	0.1246	0.5864	0.5864

Table 56: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: 2.5 km/h with CAR using 2 EEG channels

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.8309	0.6104	0.5980	0.2290	0.6337	0.6337
B64_SL50_O50_D0.2	0.8293	0.6099	0.5973	0.2289	0.6346	0.6346

Table 57: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: 2.5 km/h with CAR using 1 EEG channel

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.8545	0.6169	0.6128	0.2342	0.6216	0.6216
B64_SL50_O50_D0.2	0.8524	0.6211	0.6145	0.2435	0.6295	0.6295

Table 58: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 16 EEG channels

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.6759	0.5352	0.5605	0.1906	0.6502	0.6502
B64_SL50_O50_D0.2	0.6408	0.5110	0.5505	0.1624	0.6306	0.6306

Table 59: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 5 EEG channels

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.6697	0.5302	0.5579	0.1833	0.6449	0.6449
B64_SL50_O50_D0.2	0.6594	0.5219	0.5536	0.1705	0.6356	0.6356

Table 60: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 2 EEG channels

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B64_SL50_O50_D0.2	0.8016	0.6301	0.6124	0.3067	0.7093	0.7093
B128_SL50_O50_D0.2	0.7999	0.6286	0.6113	0.3045	0.7084	0.7084

Table 61: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 1 EEG channel

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.8661	0.6698	0.6557	0.3433	0.6893	0.6893
B64_SL50_O50_D0.2	0.8532	0.6637	0.6441	0.3382	0.6984	0.6984

Table 62: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: 1.6 km/h with ICA using 16 components

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.5326	0.4501	0.5415	0.1371	0.6134	0.6134
B64_SL50_O50_D0.2	0.4806	0.4178	0.5370	0.1212	0.5992	0.5992

Table 63: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: 2.5 km/h with ICA using 16 components

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B64_SL50_O50_D0.2	0.4893	0.4134	0.5201	0.0664	0.5548	0.5548
B128_SL50_O50_D0.2	0.4282	0.3788	0.5220	0.0708	0.5570	0.5570

Table 64: Results from Hold-out method in slip-like perturbations protocol - second iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with ICA using 16 components

Run	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
B128_SL50_O50_D0.2	0.5292	0.4425	0.5323	0.1070	0.5885	0.5885
B64_SL50_O50_D0.2	0.5130	0.4332	0.5315	0.1042	0.5861	0.5861

Appendix D

Leave-one-out Results from Slip-like Perturbations Protocol - second iteration

The following tables are related to the results obtained in the second leave-one-out iteration, mentioned in Chapter 5. All executions were made with the CNN architecture along with the best hyper-parameter combination for each trial, since it provided the best results in the first LOO iteration.

Table 65: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: 1.6 km/h with CAR using 16 EEG channels

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S5	0.5875	0.5592	0.6545	0.2549	0.6051	0.6051
S6	0.5208	0.4411	0.6363	0.1579	0.5457	0.5457
S7	0.5800	0.5205	0.7492	0.3214	0.6036	0.6036
S8	0.7271	0.5694	0.5748	0.2241	0.6679	0.6679
S9	0.6878	0.5371	0.5569	0.1757	0.6357	0.6357
S10	0.4970	0.4334	0.5409	0.0560	0.5192	0.5192
Mean	0.6000	0.5101	0.6188	0.1983	0.5962	0.5962
Std dev	0.0909	0.0590	0.0780	0.0910	0.0553	0.0553

Table 66: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: 1.6 km/h with CAR using 5 EEG channels

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S5	0.6273	0.6005	0.7176	0.3557	0.6453	0.6453
S6	0.6053	0.5755	0.6893	0.3060	0.6237	0.6237
S7	0.5781	0.5175	0.7484	0.3180	0.6018	0.6018
S8	0.8565	0.6799	0.6569	0.3735	0.7223	0.7223
S9	0.8449	0.6515	0.6329	0.3147	0.6863	0.6863
S10	0.5076	0.4108	0.6410	0.1387	0.5341	0.5341
Mean	0.6699	0.5726	0.6810	0.3011	0.6356	0.6356
Std dev	0.1457	0.0977	0.0457	0.0838	0.0660	0.0660

Table 67: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: 1.6 km/h with CAR using 2 EEG channels

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S5	0.6165	0.5732	0.7363	0.3494	0.6291	0.6291
S6	0.6611	0.6459	0.7130	0.3802	0.6697	0.6697
S7	0.5783	0.5060	0.7618	0.3126	0.5933	0.5933
S8	0.8911	0.7360	0.7136	0.4782	0.7676	0.7676
S9	0.8669	0.6886	0.6675	0.3863	0.7227	0.7227
S10	0.5417	0.4640	0.6555	0.1875	0.5565	0.5565
Mean	0.6926	0.6023	0.7079	0.3490	0.6565	0.6565
Std dev	0.1499	0.1063	0.0404	0.0964	0.0795	0.0795

Table 68: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: 1.6 km/h with CAR using 1 EEG channel

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S5	0.6148	0.5761	0.7151	0.3304	0.6269	0.6269
S6	0.6769	0.6596	0.7437	0.4259	0.6861	0.6861
S7	0.5632	0.4810	0.7565	0.2841	0.5787	0.5787
S8	0.8894	0.7284	0.7086	0.4618	0.7555	0.7555
S9	0.8966	0.6917	0.7116	0.3864	0.6764	0.6764
S10	0.5444	0.4596	0.6919	0.2144	0.5599	0.5599
Mean	0.6975	0.5994	0.7212	0.3505	0.6473	0.6473
Std dev	0.1582	0.1122	0.0241	0.0924	0.0733	0.0733

Table 69: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: 2.5 km/h with CAR using 16 EEG channels

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S4	0.5565	0.5547	0.5553	0.1101	0.5549	0.5549
S5	0.5945	0.5842	0.6153	0.2160	0.6012	0.6012
S6	0.4897	0.3552	0.5759	0.0472	0.5073	0.5073
S7	0.5917	0.5300	0.7518	0.3267	0.6059	0.6059
S8	0.6688	0.5053	0.5322	0.0996	0.5770	0.5770
S9	0.7995	0.5772	0.5684	0.1713	0.6072	0.6072
S10	0.5759	0.5490	0.6184	0.2016	0.5858	0.5858
Mean	0.6109	0.5222	0.6025	0.1675	0.5770	0.5770
Std dev	0.0987	0.0784	0.0728	0.0924	0.0360	0.0360

Table 70: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: 2.5 km/h with CAR using 5 EEG channels

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S4	0.5870	0.5763	0.6072	0.2005	0.5938	0.5938
S5	0.6448	0.6319	0.6835	0.3347	0.6526	0.6526
S6	0.5383	0.4641	0.6352	0.1690	0.5528	0.5528
S7	0.5682	0.4894	0.7585	0.2939	0.5836	0.5836
S8	0.7390	0.5501	0.5534	0.1532	0.6098	0.6098
S9	0.8367	0.5523	0.5524	0.1047	0.5522	0.5522
S10	0.6407	0.6230	0.6919	0.3389	0.6497	0.6497
Mean	0.6507	0.5553	0.6403	0.2278	0.5992	0.5992
Std dev	0.1049	0.0627	0.0763	0.0941	0.0411	0.0411

Table 71: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: 2.5 km/h with CAR using 2 EEG channels

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S4	0.6231	0.5781	0.7600	0.3762	0.6361	0.6361
S5	0.6220	0.5977	0.6830	0.3109	0.6320	0.6320
S6	0.6988	0.6901	0.7385	0.4432	0.7059	0.7059
S7	0.5790	0.5063	0.7668	0.3168	0.5941	0.5941
S8	0.8533	0.6396	0.6278	0.2829	0.6565	0.6565
S9	0.8739	0.6300	0.6407	0.2616	0.6216	0.6216
S10	0.6667	0.6472	0.7371	0.4089	0.6763	0.6763
Mean	0.7024	0.6127	0.7077	0.3429	0.6461	0.6461
Std dev	0.1165	0.0591	0.0570	0.0676	0.0370	0.0370

Table 72: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: 2.5 km/h with CAR using 1 EEG channel

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S4	0.6481	0.6197	0.7635	0.4197	0.6671	0.6671
S5	0.6298	0.6122	0.6891	0.3312	0.6450	0.6450
S6	0.7057	0.6977	0.7556	0.4724	0.7182	0.7182
S7	0.6171	0.5730	0.7742	0.3905	0.6390	0.6390
S8	0.8196	0.5969	0.5857	0.2038	0.6212	0.6212
S9	0.8597	0.5856	0.5948	0.1732	0.5791	0.5791
S10	0.6826	0.6714	0.7396	0.4338	0.6963	0.6963
Mean	0.7089	0.6224	0.7004	0.3464	0.6523	0.6523
Std dev	0.0949	0.0459	0.0801	0.1165	0.0467	0.0467

Table 73: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 16 EEG channels

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S4	0.5463	0.5160	0.5790	0.1334	0.5563	0.5563
S5	0.6288	0.6183	0.6571	0.2922	0.6358	0.6358
S6	0.5646	0.5217	0.6283	0.1982	0.5765	0.5765
S7	0.6016	0.5490	0.7387	0.3315	0.6151	0.6151
S8	0.7936	0.6241	0.6083	0.3000	0.7078	0.7078
S9	0.7164	0.5563	0.5660	0.1998	0.6511	0.6511
S10	0.5000	0.4343	0.5275	0.0383	0.5133	0.5133
Mean	0.6216	0.5457	0.6150	0.2133	0.6080	0.6080
Std dev	0.1023	0.0651	0.0691	0.1042	0.0649	0.0649

Table 74: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 5 EEG channels

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S4	0.5491	0.5192	0.5827	0.1397	0.5591	0.5591
S5	0.6510	0.6322	0.7106	0.3675	0.6603	0.6603
S6	0.6321	0.6130	0.6836	0.3220	0.6412	0.6412
S7	0.5808	0.5123	0.7505	0.3094	0.5956	0.5956
S8	0.8398	0.6745	0.6478	0.3803	0.7446	0.7446
S9	0.8074	0.6331	0.6144	0.3078	0.7071	0.7071
S10	0.5250	0.4350	0.6338	0.1473	0.5405	0.5405
Mean	0.6550	0.5742	0.6605	0.2820	0.6355	0.6355
Std dev	0.1235	0.0862	0.0580	0.0986	0.0754	0.0754

Table 75: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 2 EEG channels

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S4	0.6528	0.6283	0.7336	0.3905	0.6632	0.6632
S5	0.6519	0.6252	0.7407	0.3957	0.6626	0.6626
S6	0.6733	0.6630	0.7120	0.3914	0.6806	0.6806
S7	0.6007	0.5411	0.7696	0.3521	0.6149	0.6149
S8	0.8826	0.7241	0.6991	0.4576	0.7629	0.7629
S9	0.8921	0.6998	0.7032	0.3998	0.6966	0.6966
S10	0.5537	0.4852	0.6647	0.2117	0.5680	0.5680
Mean	0.7010	0.6238	0.7176	0.3712	0.6641	0.6641
Std dev	0.1334	0.0851	0.0339	0.0769	0.0616	0.0616

Table 76: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with CAR using 1 EEG channel

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S4	0.6565	0.6284	0.7553	0.4137	0.6676	0.6676
S5	0.6584	0.6363	0.7343	0.3974	0.6685	0.6685
S6	0.6872	0.6771	0.7296	0.4229	0.6947	0.6947
S7	0.5998	0.5404	0.7657	0.3481	0.6140	0.6140
S8	0.8763	0.7147	0.6889	0.4411	0.7575	0.7575
S9	0.8967	0.6937	0.7121	0.3901	0.6793	0.6793
S10	0.5593	0.4842	0.7058	0.2471	0.5742	0.5742
Mean	0.7049	0.6250	0.7274	0.3800	0.6651	0.6651
Std dev	0.1312	0.0843	0.0273	0.0656	0.0585	0.0585

Table 77: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: 1.6 km/h with ICA using 16 components

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S5	0.5638	0.5430	0.5922	0.1634	0.5725	0.5725
S6	0.5621	0.5620	0.5621	0.1242	0.5621	0.5621
S7	0.5654	0.5642	0.5695	0.1377	0.5682	0.5682
S8	0.5185	0.5000	0.5321	0.0581	0.5264	0.5264
S9	0.4753	0.4125	0.5329	0.1076	0.5881	0.5881
S10	0.5366	0.5133	0.5586	0.1033	0.5455	0.5455
Mean	0.5370	0.5158	0.5579	0.1157	0.5605	0.5605
Std dev	0.0355	0.0568	0.0229	0.0357	0.0217	0.0217

Table 78: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: 2.5 km/h with ICA using 16 components

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S4	0.5605	0.5553	0.5696	0.1348	0.5653	0.5653
S5	0.5967	0.5891	0.6132	0.2155	0.6026	0.6026
S6	0.5350	0.5104	0.5575	0.1006	0.5440	0.5440
S7	0.4749	0.4259	0.4779	-0.0347	0.4864	0.4864
S8	0.5951	0.5718	0.6397	0.2417	0.6045	0.6045
S9	0.5070	0.4325	0.5358	0.1182	0.5975	0.5975
S10	0.5580	0.5565	0.5569	0.1135	0.5566	0.5566
Mean	0.5467	0.5202	0.5644	0.1271	0.5653	0.5653
Std dev	0.0448	0.0667	0.0524	0.0896	0.0423	0.0423

Table 79: Results from Leave-one-out method in slip-like perturbations protocol - second iteration. Dataset: combination of 1.6 km/h and 2.5 km/h with ICA using 16 components

Test Subject	Accuracy	F1-Score	Precision	MCC	Recall	Specificity
S4	0.4815	0.4750	0.4773	-0.0447	0.4780	0.4780
S5	0.5333	0.5109	0.5535	0.0948	0.5420	0.5420
S6	0.5802	0.5797	0.5833	0.1657	0.5824	0.5824
S7	0.5778	0.5712	0.5903	0.1733	0.5831	0.5831
S8	0.5136	0.4958	0.5256	0.0466	0.5212	0.5212
S9	0.4772	0.4105	0.5270	0.0886	0.5727	0.5727
S10	0.5103	0.5023	0.5171	0.0328	0.5157	0.5157
Mean	0.5248	0.5065	0.5392	0.0796	0.5422	0.5422
Std dev	0.0417	0.0576	0.0396	0.0766	0.0398	0.0398

Appendix E

ADLs Results - Hold-out

The following table is related to the results obtained with the hold-out method, mentioned in Chapter 6.

Table 80: Results obtained from ADLs protocol - Hold-out

Dataset	Model	Batch Size	Window Size	Regularisation	Epochs	Dropout	Accuracy Train	Accuracy Val	Accuracy Test
CAR	CNN	32	150	L1	28	0.2	0.3438	0.06552	0.09007
CAR	CNN	32	750	L1	24	0.2	0.2813	0.1429	0.2614
CAR	CNN	128	750	L1	27	0.2	0.2813	0.09472	0.1654
CAR	CNN-LSTM	32	750	L1	26	0.2	0.375	0.446	0.2679
CAR	CNN-LSTM	128	750	L1	26	0.2	0.2656	0.1762	0.1606
ICA	CNN	32	750	L1	24	0.2	0.5938	0.2147	0.08552
ICA	CNN	128	750	L1	34	0.2	0.5	0.4667	0.24444
ICA	CNN-LSTM	32	750	L1	35	0.2	0.4375	0.6102	0.4491
ICA	CNN-LSTM	128	750	L1	26	0.2	0.3984	0.1894	0.1468

Appendix F

ADLs Results - Leave-one-out

The following tables are related to the results obtained with the leave-one-out method using the CNN-LSTM model, mentioned in Chapter 6.

Table 81: Results obtained from ADLs protocol with CAR dataset - Leave-one-out

Subject	Acc_Test	Acc_Train	Acc_Validation
S1	0.1525	0.4063	0.3145
S2	0.2757	0.5625	0.3591
S3	0.1781	0.4375	0.2955
S4	0.0971	0.4063	0.228
S5	0.5258	0.25	0.3314
S6	0.3088	0.375	0.2611
Mean	0.256333	0.406267	0.298267
Std dev	0.153621	0.100778	0.047725

Table 82: Results obtained from ADLs protocol with ICA dataset - Leave-one-out

Subject	Acc_Test	Acc_Train	Acc_Validation
S1	0.0378	0.625	0.5094
S2	0.3671	0.5938	0.5588
S3	0.3148	0.5625	0.5742
S4	0.3667	0.6563	0.5953
S5	0.3611	0.5313	0.4607
S6	0.4578	0.5625	0.6346
Mean	0.31755	0.588567	0.5555
Std dev	0.14471	0.046002	0.062151