université
PARIS-SACLAY

Deep Learning for Predictive Simulation of the Gait and Post-Treatment Functional Benefits in Neurological Diseases

Apprentissage profond pour la simulation prédictive de la marche et les bénéfices fonctionnels post-traitement dans les maladies neurologiques

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n°580 Sciences et technologies de l'information et de la communication (STIC)
Spécialité de doctorat: Informatique mathématique

Graduate school: Sciences de l'ingénierie et des systèmes

Référent: Université d'Évry Val d'Essonne

Thèse préparée dans l'unité de recherche IBISC (Université Paris-Saclay, Univ Evry), sous la direction de **Vincent VIGNERON** (**Professeur des Universités**), le co-direction de **Sonia GARCIA-SALICETTI** (**Professeur des Universités**), et la co-encadrement de **Omar GALARRAGA** (**Ingénieur Recherche**).

Thèse soutenue à Paris-Saclay, le 08th **of October, 2024, par**

Adil KHAN

Composition du Jury

Membres du jury avec voix délivérative

Jérôme BOUDY Professeur (TELECOM SudParis)	**Président**
Zulfiqar ALI MEMON Professeur (FAST-National University Karachi)	**Rapporteur & Examinateur**
Dan ISTRATE ECC Grade A (Université de Technologie de Compiègne)	**Rapporteur & Examinateur**
Eric DESAILLY Directeur Recherche and Innovation (Foundation Ellen Poidatz)	**Examinateur**

Thèse de doctorat

NNT : 2024UPAST107

To my family, whose unwavering support and encouragement have been my constant source of strength throughout this journey.

To my parents, for their boundless love, sacrifices, and belief in my dreams. Your faith in me has been my guiding light.

To my partner, for your endless patience, understanding, and for being my rock during the most challenging times. Your love and support have been invaluable.

To my friends, for your encouragement, laughter, and for always being there to lift my spirits.

And to my mentors and colleagues, whose guidance and wisdom have been instrumental in shaping this work.

This thesis is dedicated to all of you.

Acknowledgements

This Ph.D. thesis is the fruit of a collaboration between the Higher Education Commission (HEC) of Pakistan and Campus France. HEC financially supports this project for "predicting post-treatment gait and functional benefits in neurological diseases" and directed locally by Professors Sonia Garcia-Salicetti (Télécom SudParis) and Vincent Vigneron (Univ Evry, université Paris-Saclay). Dr. Omar Galarraga is the clinical expert and co-supervisor of the project at UGECAM Coubert (Seine-et-Marne territory).

Completing this PhD thesis has been a long and challenging journey, and I could not have reached this milestone without the support and guidance of many individuals.

First and foremost, I would like to express my deepest gratitude to my advisors, Prof. Sonia Garcia-Salicetti, Prof. Vincent Vigneron and Dr. Omar Galarraga. Your expertise, patience, and constant encouragement have been invaluable throughout this process. Your insightful feedback and unwavering support have pushed me to grow both personally and academically.

I am also profoundly grateful to the members of my dissertation committee, Prof. Jérôme BOUDY, Prof. Zulfiqar ALI MEMON, Prof. Dan ISTRATE, and Eric DESAILLY. Your constructive criticism and thoughtful suggestions have greatly improved the quality of this work.

A special thanks to my family. To my parents, your unconditional love and belief in my abilities have been my motivation and strength. To my partner and son, Murk and Abdul Hadi, thank you for your patience, understanding, and for being my constant source of support and comfort.

I would like to acknowledge my colleagues and friends for moral support throughout the process. Your camaraderie and encouragement have made this journey more enjoyable and less daunting.

I am grateful to the staff and faculty at University of Paris-Saclay for providing a supportive and stimulating environment for research. Your assistance and resources have been crucial to the completion of this thesis.

Finally, I would like to thank all the participants and collaborators involved in my research. Without your contributions, this work would not have been possible.

This thesis is a reflection of the collective support, guidance, and inspiration I have received from all of you. Thank you.

Acronyms

The following acronyms are used in this manuscript:

AI Artificial intelligence

ML Machine learning

DL Deep learning

MLP Multi-layer perceptron

STL Single-task learning

MTL Multi-task learning

MOE Mixture of experts

NLP Natural language processing

CV Computer vision

GRU Gated recurrent unit

CP Cerebral palsy

CGA Clinical gait analysis

LSTM Long short-term memory

CNN Convolutional neural network

MAE Mean absolute error

MSE Mean-squared error

RMSE Root-mean-squared error

SE Standard error

BTX-A Botulinum toxin type A

TBI Traumatic brain injury

SCI Spinal cord injury

MS Multiple sclerosis

CMA Clinical movement analysis

QGA Quantified gait analysis

DNN Deep neural network

RNN Recurrent neural network

IMU Inertial measurement unit

HPA High-pass algorithm

Bi-LSTM Bi-directional LSTM

MTD Medical treatment data

IC Initial contact

TS Terminal swing

KNN k-nearest neighbor

CLSTM Convolutional long short-term memory

QAC Quality assessment criteria

EMG Electromyography

RQ Research question
PD Parkinson's disease
FMG Force-myography
TCN Temporal convolutional neural network
GCNM Graph convolutional network model
COP Center of pressure
GAN Generative adversarial network
FSR Force-sensitive resistor
UPDRS Unified parkinson disease rating scale
HC Healthy controls
ALS Amyotrophic lateral sclerosis
PD Parkinson's disease
HD Huntington's disease
NDD Neurodegenerative disease
GEI Gait energy image
TP True positive
TN True negative
FP False positive
FN False negative
FOG Freezing of gait
BMI Body-Mass index
AUC Area under curve
FO Foot off
FS Foot strike
GRF Ground reaction force
RQ Research question
PRISMA Preferred reporting items for systematic literature review and meta-analysis
SLR Systematic literature review

Contents

I Context and Data **9**

1 Introduction **11**
1.1 Context and Rationale . . . 11
1.2 Objectives . . . 12
1.3 Thesis Outline / Organization . . . 13
1.4 List of Publications . . . 13

2 State of the art **15**
2.1 Introduction . . . 16
2.1.1 Related Surveys . . . 17
2.1.2 Background Information . . . 18
2.1.2.1 Quantified Gait Phases: Related Information . . . 18
2.1.2.2 Deep Learning Algorithms . . . 19
2.2 Research Design . . . 21
2.2.1 Research Questions . . . 21
2.2.2 Research Objectives . . . 22
2.2.3 Search strategy to retrieve primary studies . . . 22
2.2.4 Article Screening and Selection Criteria . . . 22
2.2.5 Quality Assessment . . . 23
2.3 Systematic Mapping Study Results . . . 24
2.3.1 Classification Studies . . . 24
2.3.1.1 Detection of Gait phases and Gait events . . . 25
2.3.1.2 Classification of Healthy and Pathological gait . . . 26
2.3.1.3 Fall-risk Assessment . . . 28
2.3.1.4 Prediction of Freezing of Gait . . . 30
2.3.1.5 Other classification studies . . . 31
2.3.2 Regression Studies . . . 31
2.3.2.1 Future sub-sequence forecasting . . . 32
2.3.2.2 Sensor-to-sensor sequence estimation . . . 34
2.3.2.3 Condition/joint translation . . . 37
2.3.2.4 Clinical scores prediction . . . 39
2.3.3 Gait analysis using Multi-task learning . . . 40
2.3.4 Major Datasets Used in Gait Analysis . . . 42
2.3.4.1 OULP-Age “OU-ISIR Gait Database, Large Population Dataset with Age” . . . 42
2.3.4.2 Bath Natural Environment HAR Data Set . . . 43
2.3.4.3 3d gait database . . . 43
2.3.4.4 Walking gait dataset . . . 43
2.3.4.5 Multi-modal gait symmetry (MMGS) dataset . . . 43

2.3.4.6 Pathological gait dataset . . . 44
2.3.4.7 The Daphnet dataset . . . 44
2.4 Identified Gaps . . . 44
2.5 Discussion and Future Research Direction . . . 46
2.5.1 Gait Applications . . . 46
2.5.2 Multi-task learning . . . 46
2.5.3 Datasets . . . 46
2.5.4 Future Recommendations . . . 47
2.6 Conclusion . . . 47

3 Gait Analysis: from Aquisition to Data Analysis 49
3.1 Dataset Description . . . 49
3.2 Kinematic Data . . . 50
3.3 Medical Treatment Data . . . 53

II Post-treatment Gait Prediction using Deep Learning 59

4 Treatment Outcome Prediction Using Multi-Task Learning: Application to Botulinum Toxin in Gait Rehabilitation 61
4.1 Introduction . . . 62
4.2 Data . . . 62
4.3 Description of the models . . . 63
4.3.1 Long Short-Term Memory . . . 63
4.3.2 Bi-directional LSTM . . . 65
4.3.3 Experimental Setup . . . 67
4.4 Results . . . 68
4.5 Discussion and Conclusion . . . 70
4.5.1 Comparison to Previous Works . . . 71
4.5.2 Limitations . . . 72
4.5.3 Conclusion . . . 72

5 Phase-based Gait Prediction after Botulinum Toxin Treatment Using Deep Learning 75
5.1 Introduction . . . 76
5.2 Data . . . 76
5.3 Description of the Models . . . 77
5.4 Results . . . 78
5.4.1 Analysis on gait phases and whole cycle . . . 78
5.4.2 Visualizing predictions of the best model (MTD-GM) . . . 80
5.4.3 Comparison between cycle-based prediction, phased-based prediction, and related works . . . 80
5.5 Discussion and Conclusion . . . 81

6 Post-Treatment Gait Prediction after Botulinum Toxin Injections Using Deep Learning with an Attention Mechanism 89
6.1 Introduction . . . 89
6.2 Description of Models . . . 90
6.3 Experimental Setup and Results . . . 91
6.4 Conclusion . . . 96

7 Post-Treatment Gait and Treatment Prediction after Botulinum Toxin Treatment Using Multi-task Learning **101**
7.1 Introduction . . . 101
7.2 Description of Experiments . . . 102
7.2.1 Experiment 1: Presence/Absence of MTD as auxiliary task . . . 103
7.2.2 Experiment 2: Predicting MTD dosage as auxiliary task . . . 104
7.2.3 Experiment 3: Gait phase prediction as auxiliary task . . . 107
7.3 Discussion and Conclusion . . . 108

8 Conclusions and perspectives **113**
8.1 Conclusions . . . 113
8.2 Limitations . . . 115
8.3 Perspectives . . . 116

List of Figures

1.1 Example of the outcome of BTX-A treatment on gait 12

2.1 Phases in the gait cycle. [1] 19
2.2 Architecture of typical Convolutional Neural Network (CNN). [2] 20
2.3 Architecture of typical Recurrent Neural Network (RNN) [3] 20
2.4 Flowchart on the methodology of article selection. 23
2.5 Applications in the literature for classification studies. 24
2.6 Deep learning (DL) techniques in the literature used for classification studies. 25
2.7 Applications in the literature for regression studies. 32
2.8 DL techniques in the literature used for regression studies. 33

3.1 Gait cycles per patient. 51
3.2 Proportion of stance and swing phase of each patient. 51
3.3 Clinical gait analysis 52
3.4 Gait kinematics of healthy subject 53
3.5 Pre-treatment Clinical Gait Analysis (CGA) (Kinematics) 54
3.6 Post-treatment CGA (Kinematics) 55
3.7 Unilaterally affected patient (hemiplegic). 56
3.8 Bilaterally affected patient (paraplegic). 56
3.9 Bilaterally affected patient (tetraplegic). 56
3.10 Process of converting patient's gait cycles into normalized gait phases (Stance and Swing). 57

4.1 Flowchart of data pre-processing for experiment 1. 63
4.2 LSTM unit. 64
4.3 LSTM architectures (Model 1 and Model 2) 65
4.4 First Bi-LSTM architecture (Model 3) 66
4.5 MTL architectures with Bi-LSTM sub-models 67
4.6 Comparison of the post-treatment gait trajectories (CP Patient) 73
4.7 Comparison of the post-treatment gait trajectories (MS Patient) 74

5.1 Flowchart of data pre-processing for experiment 2. 76
5.2 Process of converting patient's gait cycles into normalised gait phases (Stance and Swing). 77
5.3 DL architectures with Bi-LSTM sub-models. 79
5.4 *Cont.* 84
5.4 Comparison of MTD-GM prediction and post-treatment kinematics 85
5.5 Few outliers. 85

6.1 Attention Mechanism 91
6.2 MTD-DM-Att model and MTD-GM-Att model 92
6.3 Post-treatment gait trajectories (Knee joint). 94

6.4 Post-treatment gait trajectories (Ankle joint). 95

7.1 MTL-based Model and MTL-based Attention-Model 104
7.2 MTL-based Attention-Model (Predicting MTD dosage) 105
7.3 MTL-based Model and MTL-based Att-Model (Gait-phase) 107

List of Tables

2.1 A summary of related surveys . . . 18
2.2 Research questions for a literature review . . . 21
2.3 Selected keywords for extracting related literature . . . 23
2.4 Detection of Gait Phases and Gait Events . . . 27
2.5 Classification of Healthy and Pathological Gait . . . 29
2.6 Prediction of fall-risk and Freezing of Gait (FOG) . . . 30
2.7 Remaining classification studies . . . 32
2.8 Future sub-sequence forecasting regression studies . . . 35
2.9 Sensor-to-sensor sequence estimation regression studies . . . 38
2.10 Condition/joint translation and remaining regression studies . . . 41
2.11 Gait analysis using Multi-task learning (MTL) . . . 42
2.12 Detail of Datasets . . . 45

3.1 Patient database description. . . . 50
3.2 Considered injected muscles and their frequencies in the database (43 patients). . . . 58

4.1 Hyper-parameter selection for Long Short-term Memory (LSTM)-like models. . . . 68
4.2 Performance of different models in predicting post-treatment gait trajectories . . . 69
4.3 Performance of different models in predicting post-treatment knee gait . . . 70
4.4 Performance of different models in predicting post-treatment ankle gait . . . 71
4.5 Performance comparison of the prediction methods . . . 72

5.1 Hyper-parameter selection for for both models. . . . 78
5.2 Performance of both models in predicting post-treatment gait trajectories of stance phase 80
5.3 Performance of both models in predicting post-treatment gait trajectories of swing phase 81
5.4 Performance of both models in predicting post-treatment gait trajectories of a complete cycle . . . 82
5.5 Performance of both models for both joints in predicting post-treatment gait trajectories 83
5.6 Performance comparison to related works. . . . 86

6.1 Performance of both models in predicting post-treatment gait trajectories . . . 93
6.2 Performance comparison of the study with other prediction studies . . . 97

7.1 Considered injected muscles and their frequencies in the database (43 patients). . . . 103
7.2 Performance of both Model 1 and Model 2 . . . 105
7.3 Performance of both Model 3 and Model 4 . . . 106
7.4 Performance of both Model 5 and Model 6 . . . 108

Part I

Context and Data

Chapter 1

Introduction

1.1 Context and Rationale

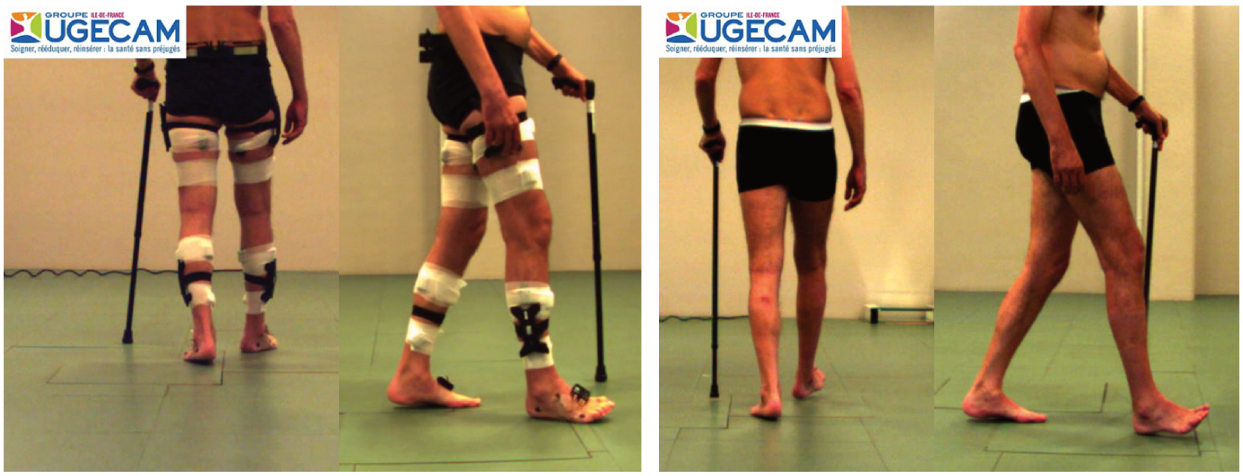
Walking is the most human common activity of daily living and, at the same time, a very complex one. It encompasses various aspects of the nervous system, musculoskeletal apparatus, and cardiorespiratory system. The way a person walks is significantly impacted by factors such as age, personality, mood but also various comorbidities. Gait and balance disorders become more prevalent with aging, often exacerbated by neurodegenerative diseases such as Parkinson's disease (PD).

In this thesis, we focus on neurological diseases that are not necessarily age-related, such as Multiple Sclerosis (MS) [4], traumatic brain injury (TBI), spinal cord injury (SCI), cerebral palsy (CP), and stroke. These conditions impact gait in various ways, leading to disorders such as fatigue, weakness, sensory loss, ataxia, and spasticity. Consequently, doctors often recommend rehabilitation therapy alongside pharmacologic treatments for patients with these impairments. Spasticity is a symptom of movement disorders, characterized by an increase in speed-dependent tonic stretch reflexes (muscle tone), causing overactive stretch reflexes and exaggerated tendon jerks. This condition is a component of upper motor neuron syndrome [5].

Their impacts on gait cover several gait disorders: fatigue, weakness, sensory loss, ataxia, and spasticity. As a result, doctors frequently advise patients with these impairments to receive rehabilitation treatment in addition to their ongoing pharmacologic care. A rise in tonic stretch reflexes (muscle tone) that depends on speed is a symptom of the movement disorder spasticity. The stretch reflexes become overactive, causing tendons to jerk more than usual. This is one part of the upper motor neuron syndrome [5]. Intramuscular injections of Botulinum Toxin type-A (BTX-A) are a common treatment for spasticity, shown to improve BTX-A both lower and upper limb functions [6], thereby enhancing movements such as walking [7] (see Figure 1.1).

In practice, decision-making is based on a patient's medical history, physical examination, and Clinical Movement Analysis (CMA). CMA consists of studying movement troubles and identifying their plausible causes based on the biomechanical interpretation of instrumental measures [8]. If certain quality criteria are fulfilled, CMA data are sufficiently reliable for clinical interpretation [9]. CMA techniques can be used to analyze lower limb movement (e.g., walking, climbing stairs, running, etc.). Several research studies have shown that CMA, especially CGA, helps a lot with the diagnosis and treatment of many neurological diseases, including CP [10], MS [11], and hemiparesis after a stroke [7].

Artificial Intelligence (AI) and Machine Learning (ML) techniques have become almost ubiquitous in our daily lives by supporting or guiding our decisions and providing recommendations. Therefore, it is not surprising that ML approaches are becoming increasingly popular in precision medicine and fulfill an increasing demand for new healthcare solutions, in particular, a better understanding of pathological processes. Among AI and ML methods, Deep Neural Network (DNN) [12] have already



(a) **(b)**

Figure 1.1: Example of the outcome of BTX-A treatment on gait (**a**) before treatment (**b**) after BTX-A treatment.

shown spectacular results in clinical decision-making aid [13]. DNNs require a significant amount of data to be properly trained. However, available experimental databases are often limited in size, which makes it impractical to construct DNNs for prediction models. Medical data is often different, complicated, incomplete, uncertain, multimodal, and multilevel. This makes it much harder to use and raises questions about the development of prediction models [14]. ML models need to handle different types of data about the patient, like pictures, time series, discrete clinical data, and so on. They also need to be able to connect this data to data about the treatment, which can be nominal, categorical (like the type of treatment) [15], or discrete (like doses). This requires that the model be taught a regression task between the data after and before the BTX-A treatment. Since these treatments are often a combination of several factors (e.g., several drug injections), it is necessary to be able to model their interactions.

Our contribution consists of proposing novel solutions for predicting the post-treatment gait trajectory of patients receiving BTX-A, as well as examining the potential interactions between various treatments. To address this regression task, we suggest a strategy to design DNNs-based and MTL-based models. MTL is particularly effective for modeling interactions between treatments, managing issues related to sparse data, and enhancing model robustness through knowledge sharing across different tasks [16]. This approach has been widely utilized in machine learning and the biomedical field to accommodate the diversity present in the data [16].

1.2 Objectives

In this above-described context, the main goal of this work is to utilize DL and MTL techniques to develop a system capable of predicting the post-treatment gait kinematics of adults with various neurological disorders. This prediction will be based on pre-treatment gait analysis and a proposed treatment. This system aims to provide a preliminary prediction of the probable treatment outcome. This will enhance the understanding of outcomes and serve as a valuable decision-making tool for medical professionals. The specific objects include:

- Study and investigate the current ML and DL techniques for CGA.
- Organizing a database of clinical gait data for adults with MS, SCI, TBI, CP, and stroke that have been treated, including pre-and post-treatment CGA and Medical Treatment Data (MTD).

- Determining the most important features among all the available data.
- Designing several DL models to determine a mathematical relationship, given the MTD, between the pre-and post-treatment CGA data.
- Exploiting MTL-based models to improve post-treatment gait prediction.
- Assessing comparatively performance of all the prediction methods considered.

1.3 Thesis Outline / Organization

This document is structured into two main sections. Part I focuses on the context, state-of-the-art developments, data description, and conditioning. Part II is dedicated to the experimental works conducted for predicting post-treatment gait using various deep learning (DL) and multitask learning (MTL) techniques. Below is a brief description of the content in each chapter

- **Chapter 2:** provides a comprehensive review of the literature on the most widely used DL and MTL algorithms in gait analysis. This chapter discusses various methods and applications relevant to gait analysis, evaluates the strengths and weaknesses of each study based on the latest research, and examines the most commonly used datasets for gait analysis.
- **Chapter 3:** Outlines the relevant information considered for the experiments, including details on medical treatments and kinematics. It also describes the processes involved in data preprocessing and preparation for the models.
- **Chapters 4, 5, 6, and 7:** Detail the experimental procedures conducted for post-treatment gait prediction. These chapters present the results of each experiment and provide comparisons with other related studies.
- **Chapter 8:** Summarizes the key findings and contributions of this research, serving as a general conclusion. This chapter also offers recommendations for future work and constructive feedback aimed at further enhancing the contributions already made.

1.4 List of Publications

The academic publications completed for this PhD thesis, including accepted and pending articles, are listed in this section.

Published articles

1. Khan, A.; Hazart, A.; Galarraga, O.; Garcia-Salicetti, S.; Vigneron, V. Treatment Outcome Prediction Using Multi-Task Learning: Application to Botulinum Toxin in Gait Rehabilitation. Sensors 2022, 22, 8452. https://doi.org/10.3390/s22218452
2. A. Khan, O. Galarraga, S. Garcia-Salicetti and V. Vigneron, "Deep Learning for Quantified Gait Analysis: A Systematic Literature Review," in IEEE Access, doi: 10.1109/ACCESS.2024.3434513.
3. Khan, A.; Galarraga, O.; Garcia-Salicetti, S.; Vigneron, V. Phase-Based Gait Prediction after Botulinum Toxin Treatment Using Deep Learning. Sensors 2024, 24, 5343. https://doi.org/10.3390/s24165343.

Accepted

1. Khan, A.; Galarraga, O.; Garcia-Salicetti, S.; Vigneron, V. Post-Treatment Gait Prediction after Botulinum Toxin Injections Using Deep Learning with an Attention Mechanism. Accepted for presentation and publication at The 10th International Conference on Machine Learning, Optimization, and Data Science, September, 2024

Submitted

1. Prediction of Gait Outcome and Treatment using Multitask Learning and Attention (submitted to Applied Sciences)

Chapter 2

State of the art

Contents

2.1 Introduction 16
2.1.1 Related Surveys 17
2.1.2 Background Information 18
2.2 Research Design 21
2.2.1 Research Questions 21
2.2.2 Research Objectives 22
2.2.3 Search strategy to retrieve primary studies 22
2.2.4 Article Screening and Selection Criteria 22
2.2.5 Quality Assessment 23
2.3 Systematic Mapping Study Results 24
2.3.1 Classification Studies 24
2.3.2 Regression Studies 31
2.3.3 Gait analysis using Multi-task learning 40
2.3.4 Major Datasets Used in Gait Analysis 42
2.4 Identified Gaps 44
2.5 Discussion and Future Research Direction 46
2.5.1 Gait Applications 46
2.5.2 Multi-task learning 46
2.5.3 Datasets 46
2.5.4 Future Recommendations 47
2.6 Conclusion 47

Summary

Over the past few years, there has been notable advancement in the field of Quantified Gait Analysis (QGA), thanks to machine learning techniques. QGA and gait prediction are areas where DL techniques are gaining popularity. There has been a significant amount of attention from the scientific community on the application of gait analysis in various fields. Based on our understanding, there is a noticeable absence of a comprehensive review and current understanding of gait analysis utilizing DL and MTL models. Therefore, this chapter provides a comprehensive assessment of the current application of DL algorithms for QGA. This chapter takes a systematic approach to explore this topic in depth. We conducted a thorough search of three databases, namely Web of Science, IEEEXplore, and Scopus, to

identify relevant papers published from 1989 to October 2023. A total of 55 papers were considered eligible and included in this review. Approximately 46% of the studies that were identified utilized classification models to categorize gait phases and locomotion modes. Additionally, a significant portion of the studies (45%) utilized regression models to estimate and predict various kinematic and kinetic parameters, including joint angles, trajectories, moments, and torques. Interestingly, a notable 9% of the studies employed the use of MTL techniques in the realm of DL for gait analysis. We have also provided information on the most commonly utilized datasets for QGA.

The content of this chapter is based on the following paper:

- A. Khan, O. Galarraga, S. Garcia-Salicetti and V. Vigneron, "Deep Learning for Quantified Gait Analysis: A Systematic Literature Review," in IEEE Access, doi: 10.1109/ACCESS.2024.3434513.

2.1 Introduction

Gait analysis is a field of study in human bio-mechanics, aiming at quantifying the elements influencing the functionality of gross motor functions in locomotion. Measuring or estimating a variety of parameters is essential in gait analysis. This includes spatiotemporal parameters, electromyography (EMG) activity, kinematic, and kinetic parameters, which are observed during walking or other locomotion activities [17]. Applications for gait analysis are numerous, ranging from athletics to medical research. Gait analysis is frequently used in sports to evaluate athlete performance, avoid injuries, and give a training schedule [18, 19]. Gait analysis is used in clinical settings to describe certain gait disorders, track the progress of recovery, and judge how well certain treatments work [20, 21, 22]. Gait analysis has further uses, such as predicting the likelihood that an elderly individual may fall [23, 24]. Gait analysis can assist in the treatment planning process for individuals suffering from a variety of conditions, including MS [25], Stroke, TBI, CP [26], and SCIs [27].

Moreover, with the help of QGA, one can design and develop numerous wearable robotic systems, such as exoskeletons and orthoses, to help people with gait-related issues. An electromechanical system called an exoskeleton is made up of actuators, sensors, and controllers that work together to provide torque to joints [28]. Another class of assistive and corrective technology is orthoses, which are occasionally used interchangeably with exoskeletons. There is a distinction between the two, though. Herr claims that, in contrast to exoskeletons, which improve human capacities in all circumstances, orthoses are meant to help those who already have diseases [29].

The gait analysis can be carried out in several ways. The majority of clinical settings use a combination of patient self-reported assessments, qualitative evaluations conducted by professionals, and observation [30, 31]. Clinical observations, such as those made by physicians or physical therapists, can yield quantitative information about gait characteristics, including cadence, speed, distance traveled, and total walking time.

In a lab setting, QGA is frequently carried out utilizing gold standard measuring techniques, such as force plate and motion capture devices together. While force plates offer dynamic elements like ground response forces and moments, motion capture allows for accurate tracking of the spatial information of human movements in three dimensions. It is now feasible to evaluate gait outside of the lab using wearable sensor systems because of recent significant advancements in sensor systems and computational techniques[32, 33].

QGA is essential for identifying postural instabilities, detecting abnormal gait patterns, and evaluating clinical interventions and rehabilitation plans. The clinicians anticipate using QGA for diagnosis and treatment decision-making.

In this context, AI methods would be useful for gait analysis. These methods can effectively handle complicated, temporal, high-dimensional data [34, 35] and extract pertinent features. They create models that automatically learn from accessible sources, generate precise predictions, and exhibit intelligent behavior [36, 37]. DL, a subset of ML, is widely utilized in various fields, including medical

diagnosis [38, 39], pattern recognition [40, 41], image processing [42, 43], classification [44, 45], prediction analysis [46], and monitoring [45]. DL techniques have been used in various applications in the field of QGA, for the diagnosis of gait disorders [47, 48], the prediction of early intervention for fall-related risks associated with a disability or aging [49, 50, 23], the determination of motor recovery tasks [51, 52], and the planning of therapeutic or rehabilitation interventions [39].

DL approaches are used to analyze data, and images, and recognize patterns by developing algorithms that assist doctors in promptly and properly diagnosing a particular disease or disorder. Additionally, these algorithms may continuously learn, which enhances the diagnostic outcomes. However, there are situations when doctors must concurrently take into account the patient's symptoms, treatment options, probable side effects, another condition with similar symptomatology, past medical history, and several other factors. In turn, DL technology offers a way to help doctors by analyzing a very large amount of data and ensuring a comprehensive grasp of patient health records. Even DL methods can foretell the start of disease by analyzing the vast amounts of data amassed over time from tracking a person's health. The development of DL approaches for disease treatment is evidenced by the arrival of surgical robots to undertake numerous difficult operations. As a result, DL models can speed up diagnosis, improve patient monitoring, and help clinicians choose the best course of action, but they need computational resources at the implementation site. Additionally, early detection can stop mobility loss [35] and can lower healthcare expenditures, which are a growing concern for developing nations [53].

However, DL models performed well when they were provided with a sufficient amount of data. But, as we know, in medical applications this is not true, due to the complexity and cost of acquiring large amounts of real data from patients. On the other hand MTL can be used in QGA to cope with the problem of few data. MTL can indeed handle sparse data issues and create a more reliable model by utilizing information from various tasks [16]. Furthermore, MTL has been widely used in ML and biomedical areas [16] to handle the variability of data. For all these reasons, we are also reviewing studies that used MTL for QGA.

2.1.1 Related Surveys

A few surveys have been published on gait analysis, as described in Table 2.1. Only two of the five review studies analyzed were systematic reviews. Prasanth et al. [32] systematically reviewed wearable sensors and methods for real-time gait analysis. They focused on the most widely utilized sensors and methods in clinical settings for pathological gaits. The limitations of this study are: (i) the authors didn't mention the period of works included in the paper, (ii) 19 studies out of 113 were using ML (and fewer of them DL), (iii) commonly used datasets for gait analysis are not discussed. Kolaghassi et al. [54] performed a systematic review of gait analysis and prediction for lower limb robotic systems using intelligent algorithms. The limitations of this study are: (i) 3 studies out of 41 used DL, (ii) commonly used datasets for gait analysis are not discussed. Cicirelli et al. [55] reviewed a highly consistent work tackling gait analysis-related challenges: sensors, characteristics, and processing approaches have all been analyzed. Their paper lists the most popular processing methods for classification and clustering as well as feature extraction and selection. Nevertheless, this paper has some limitations: (i) it mostly reviewed those studies that covered neurodegenerative diseases (NDDs), (ii) only 3 studies using DL for gait analysis are reported, (iii) commonly used datasets for gait analysis are not discussed. Hutabarat et al. [33] provides a comprehensive overview of the latest developments in wearable sensors for gait analysis. The review covers a range of topics, including the utilization of wearable gait analysis systems, the types of sensor systems and their corresponding attachment locations, as well as the algorithms employed for analysis. Most of the studies in this article use inertial measurement unit (IMU) sensors to collect data. Finally, Khera et al. [56] aims to provide readers with an overview of the main approaches utilizing ML techniques for gait analysis and rehabilitation. It is worth noticing that few studies in this review discuss papers using DL for gait analysis.

Table 2.1: A summary of related surveys

No	Ref	Year	Studies included	Type of survey	Databases	Discussed datasets	Studies included using DL	Studies included using MTL	No. of studies
1	[32]	2021	Not mentioned	systematic review	Scopus, WOS, Cochrane, and PubMed	No	Less than 5%	No	113
2	[54]	2021	1989 – May 2020	systematic review	IEEE Explore and Scopus	No	Around 7%	No	41
3	[55]	2021	2011-2020	review	IEEE Explore, ScienceDirect, Scopus, PubMed, and ACM	No	Less than 5%	No	Not mentioned
4	[33]	2021	2011 – 2020	review	IEEE Explore, PubMed, Scopus, and WOS	No	Less than 5%	No	76
5	[56]	2020	1980 - 2019	review	IEEE Explore, WOS, PubMed, Mendeley, Elsevier, ScienceDirect, Springer Link, Cochrane, Wiley Online Library	No	Around 10%	No	43

From these observations summarized in Tab. 2.1, we propose to conduct a systematic and comprehensive review of gait analysis using DL. Therefore, our study focuses on articles published between 1989 to October 2023.

The primary contributions of this chapter are outlined below:

1. A systematic map of 55 primary studies based on the Preferred Reporting Items for Systematic literature review and Meta-Analysis (PRISMA) framework;

2. An assessment of the literature on gait analysis across four dimensions: DL approaches, datasets, quality metrics, and the impact of MTL techniques;

3. An overview of the challenges, opportunities, and recommendations for future research in the field.

This systematic literature review (SLR) offers a comprehensive and up-to-date examination of gait analysis across five key research areas. The analysis is thorough and provides valuable insights into the major challenges facing the field, as well as potential avenues for future research. To the best of our knowledge, no existing SLR comprehensively addresses all aspects of gait analysis.

The remainder of the article is structured as follows: Section 2.2 summarizes the research design of this SLR. The results of the systematic mapping study, along with the most pertinent publications based on quality assessment criteria, are covered in Section 2.3.Section 2.4 lists the problems and research gaps. Section 2.5 provides recommendations and future research topics. Finally, Section 2.6 summarizes the SLR.

2.1.2 Background Information

2.1.2.1 Quantified Gait Phases: Related Information

Gait describes how we walk.Gait analysis is the study of walking patterns, applicable not only to humans but also to other beings [57]. Multiple parameters are observed and evaluated in QGA. These parameters often fall within the normal ranges of healthy gait, which vary according to anthropometric

characteristics such as age, height, and limb lengths [58]. Parameters in abnormal gait often diverge from these ranges. Each of the parameters provided in this section has been detected, predicted, and examined using different methodologies in QGA.

Gait Phases: Gait is a periodic activity, with each period referred to as the gait cycle. The gait cycle, illustrated in Figure 2.1, consists of two main phases: the stance, , where the lower limb is in ground contact, and a swing phase, where there is no contact.
There are four segments to the stance phase: (1) initial contact: which starts with a heel strike, also known as the loading response, (2) midstance: the moment the foot is on the ground due to a dorsiflexion moment (3) terminal stance: when the heel starts to lift off the ground; and (4) pre-swing: the fi ground point of contact before the swing phase begins. The swing phase is divided into three segments: (1) initial swing, (2) mid-swing, and (3) terminal swing. In total, there are seven segments in a gait cycle, as shown in Figure 2.1. Forward movement is achieved by alternating between the right and left feet. There are periods of single support, where only one leg is in contact with the ground, and periods of double support, where both legs are in contact with the ground [58].

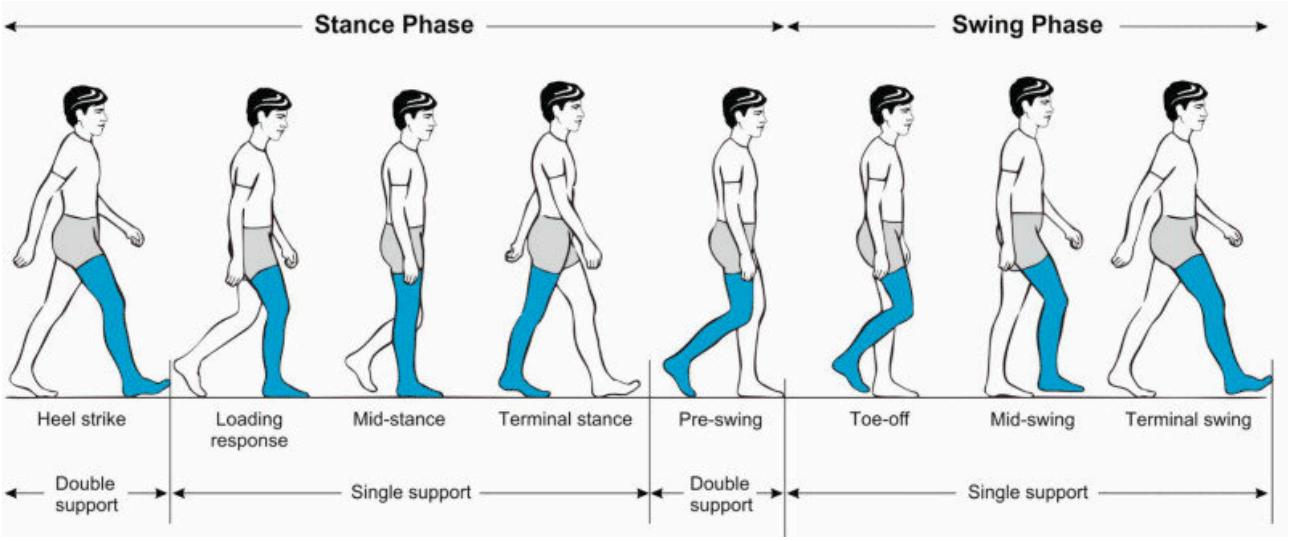


Figure 2.1: Phases in the gait cycle. [1]

Anatomical Planes in a Human: our body moves in three anatomical planes that can be used to explain exercise and other activities:

- The coronal (frontal) plane divides the body's anterior and posterior halves.
- The sagittal (longitudinal) plane divides the body's left and right halves.
- The body's top (superior) and lower (inferior) halves are divided by the transverse (axial) plane.

Joint Angle: Every gait cycle results in a periodic change in the hip, knee, ankle, and foot joint angles for the three planes. Most of the movement is visible in the sagittal plane [58]. Joint angles are thought to be a kinematic part of walking [17]. Their second-time derivatives, angular acceleration and velocity, are also often seen and recorded.

Moment (Gait Kinetic): Moments fall under the category of kinetic gait parameters. During the gait, muscles generate moments of force across joints.

2.1.2.2 Deep Learning Algorithms

Our focus in this study is on those studies that are using DL for gait analysis. In this section, we will describe DNNs that are commonly applied for gait analysis.

CNN: CNN is a type of DNN that typically takes 2D images as input. The CNN uses a kernel, smaller than the original image; which is swept across the image to perform a convolution operation between the kernel and a piece of the image. This process results in the construction of the network's next layer, known as the feature map. In Artificial Neural Networks (ANNs), feature maps are analogous to hidden layers. In contrast to hidden layers in artificial ANNs, where each node connects to every single input, a CNN maps a group of inputs, corresponding to the kernel size, to a single point on the feature map. This characteristic, known as sparsity of connections, is unique to CNNs. The sparsity of connections due to convolutions reduces the number of parameters in the model, thereby decreasing the memory storage space and computational power required. Another distinguishing feature of CNN algorithms is parameter sharing. The same set of weights is applied across the entire image to generate a single feature map. These weights are only modified when creating feature maps that extract different features. In addition to convolution layers, CNNs also include pooling layers. To generate the output, the final feature map is unrolled to a fully connected hidden layer after numerous alternating convolution and pooling processes [59] (shown in Fig 2.2).

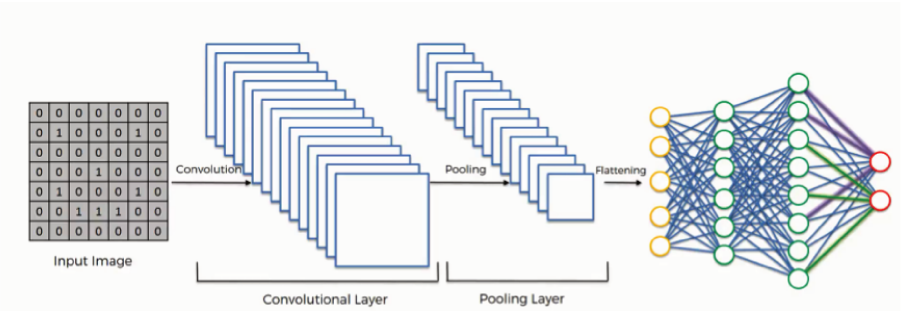


Figure 2.2: Architecture of typical CNN. [2]

RNNs: A RNN [60] is a type of ANN specifically designed to handle time series data or data that contains sequences. Ordinary feed-forward neural networks are intended for data points that are independent of each other. However, when dealing with sequential data where each data point depends on the preceding one, the neural network must be adapted to account for these dependencies. RNNs use the concept of "memory" to store the states or information of previous inputs, which helps in generating the next output in the sequence, as illustrated in Figure 2.3. A simple RNN that allows it to retain information across multiple time steps. This feedback loop can be unrolled over n time steps. By unfolding the network for k time steps, it can produce the output at time step $k+1$. The unfolded network resembles a feed-forward neural network but is capable of handling sequential dependencies.

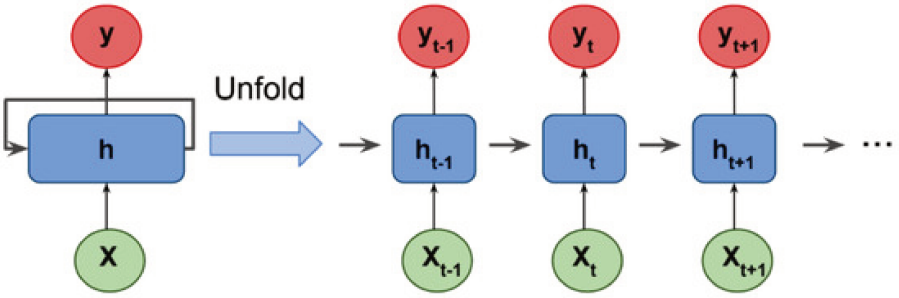


Figure 2.3: Architecture of typical RNN [3]

LSTM: The LSTM model [61] is a powerful recurrent neural system specifically designed to address the exploding and vanishing gradient problems that commonly occur when learning long-term dependencies, even with minimal time lags [62]. A standard LSTM unit consists of four main

components: a cell, an input gate, an output gate, and a forget gate. This forget gate was not originally part of the LSTM network, but was later introduced by Gers et al. [63] to enable the network to reset its state. The cell in an LSTM unit serves as a memory mechanism, capable of retaining values over arbitrary time intervals. The three gates (input, output, and forget) regulate the flow of information associated with the cell. This structure allows LSTM networks to effectively manage and process sequential data with long-term dependencies, making them particularly useful for tasks involving time series or sequential information [62].

Attention Mechanism: DL approaches, such as attention models, are employed to provide enhanced focus on specific components of input data [64]. In DL, the attention mechanism mimics the human brain's ability to concentrate on particular elements and recognize their importance. This mechanism is typically implemented in one of two ways within a network's architecture: self-attention which focuses on maintaining and quantifying interdependent relationships between input elements and general attention which concentrates on the relationships between input and output elements.

For instance, when observing an image, the human brain initially focuses on a specific aspect with high resolution while perceiving the surrounding areas with lower resolution. As the brain begins to comprehend the image, it dynamically adjusts its focal point to thoroughly understand all components. The attention mechanism in DL models operates similarly, allowing the network to prioritize certain parts of the input data dynamically. This approach enables the model to allocate computational resources more efficiently and effectively, leading to improved performance on various tasks, especially those involving complex or long-range dependencies in the data.

2.2 Research Design

In this study, we have applied systematic mapping as a research methodology for reviewing the literature [65]. This review is a 'systematic' way of exploring existing literature on gait analysis using DL. This SLR consists of four essential steps: planning and searching for primary studies, collecting studies, extracting data, and synthesizing data. The first step is to come up with research questions and goals. The criteria for choosing studies, selecting studies, coming up with keywords for research and search queries, and judging the quality of extracted studies are all part of the second step. In this step, we choose a group of keywords and boolean operators to extract only the most relevant papers from the literature. The data extraction step (third step) uses strategies for getting data from specific studies. The quality assessment of selected studies is the last step.

This methodical approach ensures a comprehensive and unbiased review of the literature, providing a clear overview of the current state of research in the field.

2.2.1 Research Questions

This literature review's primary goal is to investigate various DL techniques applied in gait analysis. Table 2.2 lists the research questions (RQs) to achieve the primary goal.

Table 2.2: Research questions for a literature review

RQ	Research Questions
RQ1	Which DL techniques have been used for QGA?
RQ2	What are the major databases used in QGA?
RQ3	What are the various metrics for evaluating the performance of different models used in QGA?
RQ4	How MTL can improve the performance of models in QGA?

2.2.2 Research Objectives

The research objectives of this literature review are given below:

- To investigate the existing traditional and advanced DL techniques and approaches for QGA.
- To explore the existing databases used for QGA.
- To investigate various evaluation metrics used to measure the performance of QGA.
- To investigate how MTL can help in better analysis of QGA.

2.2.3 Search strategy to retrieve primary studies

This part starts with choosing a set of keywords that reflect the topic of this literature review, along with Boolean operators. All authors participated in selecting keywords that helped us extract all related literature on gait analysis using DL. Table 2.3 shows the keywords we applied to databases to extract related studies. Query 1 was used to select all studies on gait analysis that used DL, and Query 2 was used to determine those studies that used MTL. In the end, we combine both queries using the OR operator. Three databases, IEEE, Web of Science, and SCOPUS have been searched for publications using the search query mentioned in Table 2.3. The search query was used to find relevant publications from the chosen databases published in English between 1989 and October 2023. It was applied to the article title, abstract, and keywords. We applied Query 1 and Query 2 separately; Query 1 identified 293 studies, and Query 2 identified 75 studies. The identical studies from different databases were removed; only distinctive copies were retained, reducing the number of papers to 203.

2.2.4 Article Screening and Selection Criteria

After duplicate records were eliminated, the remaining 203 studies were examined. The retrieved publications were screened based on their title, abstracts, and keywords. The authors decided on inclusion and exclusion criteria to find these studies. Articles were added or removed based on a majority vote on all inconsistencies. A final decision was also made in the event of a tie between all the authors. The screening of all articles using the title, abstract, and keyword-based screening method is shown in Figure 2.4. Additionally, only 55 publications—out of a total of 145—were chosen for primary studies; the remaining articles were disregarded.

We use the following inclusion criteria:

- The article must be published before October 2023.
- The article must be published in a journal, but concerning Query 2, we also selected conference papers due to limited studies that are using MTL techniques.
- The article exploits DL.
- The article must use a sensor(s) to collect gait data.
- The article must use only gait data.
- The purpose of the article must be QGA.

We use the following exclusion criteria:

- The article does not use DL.
- The article's domain is not the field of machine learning for QGA.
- The article does not use any other type of data except gait (like speech).

Table 2.3: Selected keywords for extracting related literature

Query 1	("Deep Learning" OR "deep learning") AND (Gait prediction)
Query 2	("Multi-task learning" OR "MTL") AND ("Gait")
Search Query	Query 1 OR Query 2

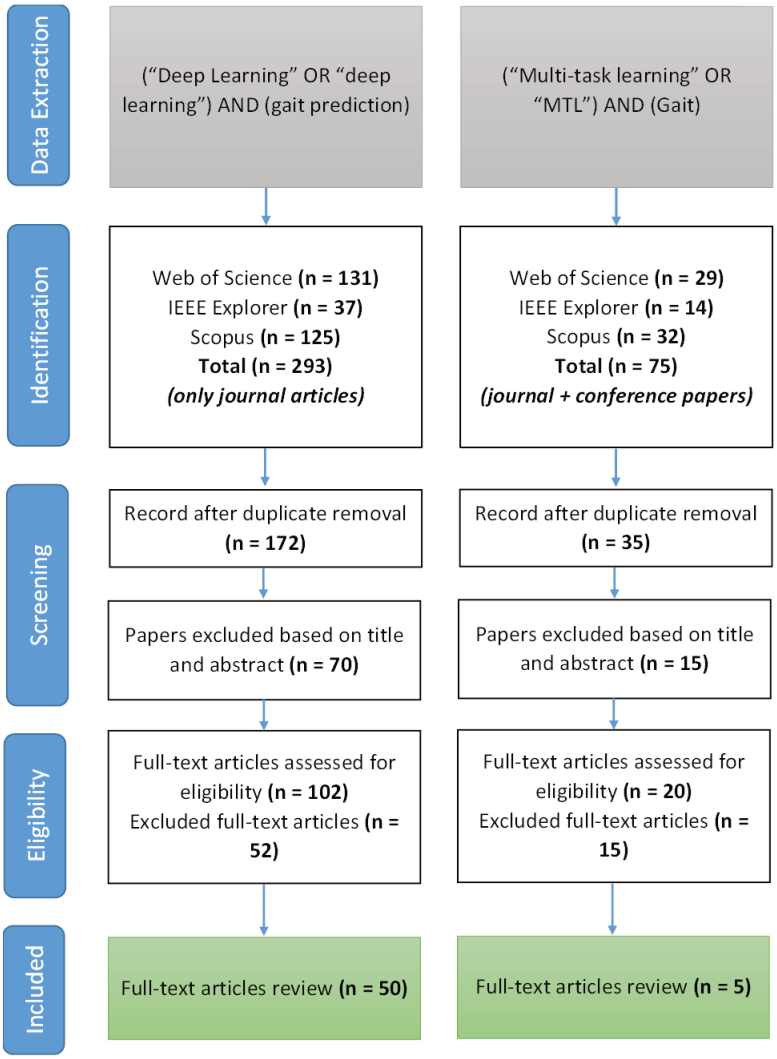


Figure 2.4: Flowchart on the methodology of article selection.

2.2.5 Quality Assessment

In this study, the Quality assessment criteria (QAC) were employed to evaluate the quality of 55 studies that were selected for analysis. The QAC served as a tool for assessing the rigor and validity of the research conducted in these studies. Through the use of these criteria, the researchers were able to determine the overall quality of the studies and draw conclusions based on the findings. The present study utilized the QAC to evaluate the suitability of a chosen primary investigation for meeting the objectives of our review. To assess the consistency of the selected primary studies, the collective authors posed a set of questions. Following are the questions to check the quality of the selected studies.

- Are the research objectives clearly stated?
- Is the proposed methodology well defined?

- Is DL model/architecture clearly defined?
- Is there enough information available for the dataset?
- Are evaluation metrics fully defined?
- Are the results properly interpreted and discussed?

The present study evaluated the results following a thorough quality assessment of each primary study.

2.3 Systematic Mapping Study Results

In this section, we discuss 55 papers that meet our inclusion/exclusion criteria. The performance of the models, the data used for training and validating the algorithms, the type of data used to make predictions, and quality metrics are examined. The results are discussed in five main groups: (1) classification studies, (2) regression studies (4) MTL studies, and (5) major datasets used in the literature for gait analysis.

2.3.1 Classification Studies

This section reviews the studies on the classification of gait phases, detection of gait events, recognition of gait anomalies, classification of normal and pathological gait, and fall-risk assessment (shown in Figure 2.5). The prevalence of LSTM-CNNs in recent literature demonstrates how widely used it is as an DL technique for classification studies. The study explores various techniques that have been employed in the field of DL, in addition to multi-layer perceptrons (MLPs). These techniques include ConvLSTMs, Graph CNNs, attention mechanisms, and transformers. These studies provide valuable insights into the performance of these models and their potential applications in the gait analysis field.

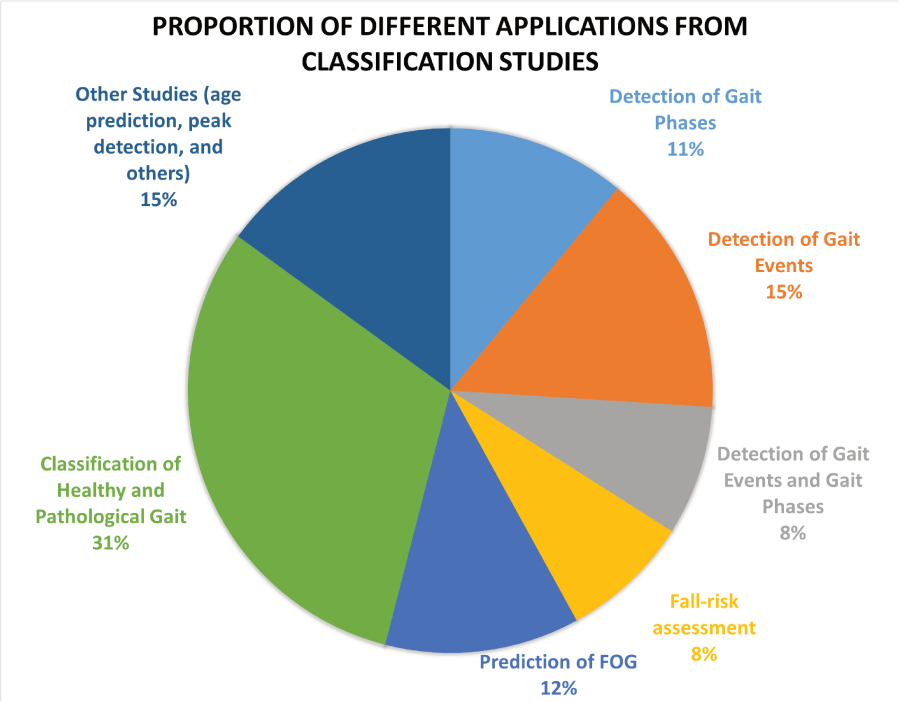


Figure 2.5: Applications in the literature for classification studies.

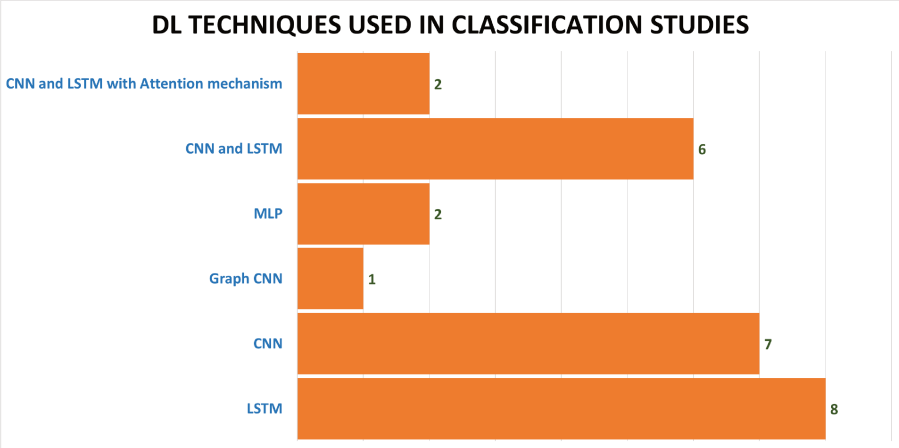


Figure 2.6: DL techniques in the literature used for classification studies.

2.3.1.1 Detection of Gait phases and Gait events

In the past, researchers have effectively utilized ML algorithms such as SVM [66], DT [67, 68], RF [66], and KNN [40] to detect gait phases or events[56]. IMUs and force-sensitive resistors (FSRs) were identified as the most suitable options for data collection for these tasks. The classification accuracy they achieved ranged from 83% to 100%. In recent years, DL has proven to be highly effective for this purpose. CNN and LSTM have been widely utilized. Data was collected using sEMG, IMU, ground reaction force (GRF), and FSR. Details are given in the following section.

The reviewed studies have demonstrated innovative methods for gait phase prediction and classification using diverse techniques. Morbidoni et al. [69] introduced a MLPs approach focusing on foot-floor-contact signal prediction and gait phase categorization, employing surface sEMG data from 23 healthy adults during level ground walking. Their utilization of MLPs with varying hidden layer configurations resulted in an impressive 94.94% accuracy, notably achieved by an MLP with 4 hidden layers after signal normalization. Su and Gutierrez-Farewik [70] employed LSTM networks to predict human gait trajectories and classify phases, achieving 95% accuracy in identifying gait phases, particularly excelling in detecting the swing phase. Additionally, a study by a different research group[71] utilized a graph convolutional network model (GCMM) to effectively group walking phases for lower limb exoskeleton control. This investigation, involving ten healthy male volunteers, exhibited remarkable performance in gait phase classification, reaching a maximum accuracy of 97.43%, outperforming established techniques like LSTM and DCNNs. These findings collectively indicate promising advancements in gait phase prediction, showcasing the potential of CNN, LSTM and GCMM techniques in improving accuracy and applicability across various terrains and control mechanisms, notably in scenarios involving exoskeletons.

In the realm of gait event detection, several studies present diverse approaches employing DL techniques for different pathologies. Filtjens et al.[72] focused on PD and Freezing of Gait (FOG), introducing a data-driven approach to enhance gait event detection accuracy crucial for assessing and managing gait disorders in PD. Their study showcased that a temporal convolutional neural network (TCN) outperformed RNNs like LSTMs, with TCN displaying F1-scores of 0.995 to 0.999 for various events in PD and FOG scenarios. Additionally, Lempereur et al.[47] proposed DeepEvent, utilizing bi-directional LSTMs (Bi-LSTMs) to automatically detect foot strike (FS) and foot off (FO) events in children with gait disorders, showcasing impressive area under the curve (AUC) values of around 0.995 to 0.997 for FS and FO events. Arshad et al.[73] explored gait event recognition in older individuals using single waist sensors and DL models, achieving up to 99.73% accuracy with CNN-BiGRU-Att models. Furthermore, a study by Dumphart et al.[74] introduced IntellEvent, a DL-based

algorithm exhibiting exceptional accuracy in detecting various gait events across diverse pathologies, surpassing state-of-the-art heuristic methods with detection rates above 99% for initial contact (IC) events and over 95% for FO events, achieved within milliseconds. Collectively, these studies show that DL methods can accurately find gait events in people with a variety of pathologies. They do this quickly and with great accuracy, which is important for accurate assessment and intervention plans.

Di et al.[75] aimed to devise a MLP architecture exclusively using knee-joint angle data from electrogoniometers to accurately identify gait phases and events. Conducted on 23 healthy adults without prior pathologies or orthopedic surgeries impacting leg mechanics, the study assessed average participant characteristics: mean height (173±10 cm), mass (63.3±12.4 kg), and age (23.8±1.9 years). By using MLP architecture, their method showed high accuracy (90.6±2.9%) in evaluating the timing of heel-strike and toe-off, working well even for people who were not in the initial training dataset. Meanwhile, Ling et al.[76] proposed a domain adaptive CNN (DACNN) model for gait phase and event recognition using surface EMG data. This model was designed to pre-train on a comfortable gait speed and adapt quickly to new gait speeds. Their study involved four healthy volunteers without muscle conditions, with an average age between 23 and 26 years. Using different CNN-based backbone models that had already been trained on comfortable walking speed data, the model was able to adapt to different walking speeds using architectures such as LeNet, AlexNet, and fusion-data-CNN (FDCNN). Their evaluation showcased promising accuracy in identifying gait phases at different speeds, underscoring its potential application across varied gait scenarios after minimal recalibration.

2.3.1.2 Classification of Healthy and Pathological gait

In the past, IMUs have shown a high success rate of around 95% in detecting abnormalities of gait, which accurately predict gait abnormalities with a high asymmetry score[56]. The ML models utilized in past studies included the hierarchical clustering algorithm [77], SVM [43], LDA [78], KNN [79], and Bayesian network [80]. Currently, CNNs and LSTMs have become popular choices for classifying healthy and abnormal gait patterns. Nowadays, data is collected using IMU, ECG, EMG, and GRF.

El Maachi et al. [81] proposed a new intelligent system for detecting PD, which utilizes the 1D-Convnet to analyze gait data. This study made use of the PhysioNet public database, consisting of 166 participants, 93 individuals with PD, and 73 control subjects. Eight sensors were positioned beneath each foot for every participant to measure vertical GRF. Their method also predicts PD severity. They employed the unified Parkinson disease rating scale (UPDRS), the most common PD assessment system [82]. The proposed algorithm achieved an accuracy of 98.7% for PD detection and 85.3% accuracy in predicting the severity of PD.

Alharthi et al. [83] proposed a 2D-DCNN model to classify PD and healthy subjects using GRF. They used the open-access benchmark dataset, PhysioNet [84]. Four convolutional layers, an average pooling layer, two fully connected layers, and 10 stacked layers comprise a 2D-DCNN model created for PD severity classification. Another network (parallel DCNN) produces the best classification accuracy, with mean performance and standard errors of 95.5% and 0.28%, respectively.

Sadeghzadehyazdi et al. [52] proposed an end-to-end CNN-LSTM model that captures spatiotemporal patterns for gait anomaly recognition using Kinect skeleton data. The sequence of normalized skeletons was used in the study, with a length of 50 frames per sequence. Three different datasets were used to evaluate the performance of the proposed model: the walking gait dataset, the MMGS dataset, and the pathological gait dataset. On the walking dataset, the CNN-LSTM model achieved 90.57% accuracy. On the MMGS dataset, it achieved 82.71%, and on the pathological gait dataset, it achieved 89.83% accuracy.

Zhu et al. [51] proposed a two-stream CNN for classifying healthy and pathological subjects. The authors claimed existing work mostly applied DL on individual joint features, and they don't consider inter-joint features due to the complexity of smaller-scale medical datasets. So the authors proposed an approach that explicitly takes individual and inter-joint features. One stream of CNN was used to

Table 2.4: Detection of Gait Phases and Gait Events

Research Article	Purpose	Model [Architecture]	Type of Data	Dataset	Performance Metric	Results
Morbidoni et al. [69]	Gait phase prediction	5 MLPs [1, 2, 3, 4, and 5 FCLs, respectively.]	sEMGS signals	23 Healthy adults	Accuracy	94.9% for seen subjects and 93.4% on unseen subjects
Su and Gutierrez-Farewik [70]	Gait phase prediction	LSTM-based [LUs and FCLs depends on the number of future steps to predict]	IMU	12 healthy subjects (thighs, shanks, feet, and pelvis)	Accuracy	95%
Di et al. [75]	Gait phases and event detection	MLP [3 FCLs with 512, 256, and 128 neurons]	electrogoniometer, foot-switches (for foot-floor-contact), and sEMG	23 healthy subjects	Mean Absolute Error (MAE) and Accuracy	MAE: 29.4±13.7ms for HS and 99.5±28.9ms for TO detection, Accuracy: 90.6%±2.9% for gait phase detection
Filtjens et al. [72]	Gait event detection	Temporal CNN	retro-reflective markers	15 patients with PD	F1-Score	For FOG-trials, F1 scores of 0.995 and 0.992 were obtained for IC and EC, respectively. For functional gait trials, F1 scores of 0.997 and 0.999 were obtained for IC and EC, respectively.
Lempereur et al. [47]	Gait event detection	DeepEvent [3 BLs (800 units)]	GRF	226 pathological children	Time and Confidence Interval (ms)	Foot Strike: 5.5 [0.9; 10.2] and Foot Off: 10.7 [5.4; 15.9]
Wu et al. [71]	Gait phase prediction	GCMM	goniometers and FSRss	10 Healthy subjects	Accuracy	97.34%
Ling et al. [76]	Gait phases and event detection	Domain Adaptive CNN	sEMG	4 healthy subjects	MAE and Accuracy	Accuracy: 58.13% for AlexNet, 81.56% for LeNet, and 81.53% for FDCNN. The average MAEs for gait event identification are 48, 85, and 66 ms, respectively.
Arshad et al. [73]	Gait event detection	CNN, RNN, CNN-RNN with and without attention mechanism	IMU	169 (94 healthy and 75 pathological) subjects	Accuracy	Highest accuracy were 99.73% for the CNN-BiGRU-Att model
Dumphart et al. [74]	Gait event detection	LSTM-based [3 BLs (200 units) + DrL + DL]	force plates	1272 (61 healthy and 1211 pathological) subjects	MAE	Detection rate for IC events is above 99% and for FO events is above 95%

Note: FCL = Fully Connected Layer, BL = Bi-LSTM Layer, DrL = Dropout Layer, and LU = LSTM units

learn the joint position, and another was responsible for learning the relative joint displacement. After that, they implemented a mid-layer fusion module to concatenate the results of both streams for better classification. They used a dataset of skeletons from 45 subjects, including four classes. 10 are healthy, 4 have joint problems, 18 have muscle weakness, and the rest 13 have neurological defects. The ages of the subjects range from 61 to 91 y.o.. Five of them were male, and 40 were female. 20 main joints were considered to form the subject's skeleton for input to the model. 5-fold cross-validation was used for the evaluation. Using their proposed methodology, they got an overall 95% accuracy, 92% precision, 92% recall, and 92% F1-measure.

Kaur et al. [85] introduced a data-driven methodology for stride classification in individuals with MS, PD, and healthy older adults (HOA). They involved 33 participants across different age groups

and conditions, performing two walking tasks on a sensor-equipped treadmill. The study utilized sixteen ML algorithms, including traditional supervised ML and ML models, to establish baseline performance. Multi-Scale RNN exhibited high accuracy in distinguishing between walking conditions but the 1D CNN model achieved the highest accuracy, scoring 79.3%.

J. Yu et al. [86] developed a real-time motion data-based AI system using RF and LSTM to predict geriatric stroke. The data were obtained from elderly Koreans walking with affixed wearable sensors. Shoulders and quadriceps were targeted for the placement of the sensors. After processing the data, they got 12 motion attributes, including angles and acceleration. The proposed system achieved high prediction accuracy: from 98.25% for the C4.5 DT model to 98.72% for RF, 96.60% for XGBoost, and 98.99% for LSTM.

Chen et al. [87] proposed a CNN model for recognizing sarcopenia disease comprising both hardware and software. The hardware is made up of multiple sensor modules (MSM), worn behind the ear and used to collect EMG and gait (EAG) data. Biomedical and motion sensor algorithms (Bodi algorithm) and leg health classification net (LCNet) are the parts of the software. 21 men and 34 women between 20 and 81 y.o. participated in this study. Their weights range between 38 and 112 kg, with between 12 and 29 kg of total muscle mass. Patients with serious damage to their legs or spines were turned away. We found that 19 successful subjects were at a high risk of sarcopenia. They used the Bodi method to figure out the gait features that feed an LCNet model, composed of a convolutional layer and two FC layers. LCNet has 94.41% accuracy, 91.58% precision, 95.81% specificity, and 91.58% sensitivity.

D. Thakur and S. Biswas [88] proposed an attention-based DL system utilizing smartphone accelerometer and gyroscope data to predict hemiplegic gait. They collected data from 28 participants, divided into normal and hemiplegic groups. The hemiplegic individuals, who had experienced a stroke, walked more than 15 steps on a flat surface. A novel Android smartphone application with tri-axial sensors collected data at a frequency of 50 Hz. A CNN-LSTM architecture automatically learns features from sensory data, leveraging the strengths of both CNN and LSTM for effective feature extraction and pattern recognition. The study recommended combining automatically learned and hand-engineered features from sensory data. An attention network adjusted the importance of the different features, achieving an impressive accuracy rate of 86%, a precision rate of 80%, a recall rate of 100%, and an F1-score of 88.89% on an unseen dataset, demonstrating the model's effectiveness.

2.3.1.3 Fall-risk Assessment

Tunca et al. [49] developed a Bi-LSTM model aimed at predicting fall risk in geriatric populations with neurological disorders. This study involved 76 subjects, with 37 categorized as high-fall-risk due to a history of falls within the year before data collection. They compared their model to other algorithms like MLP, RF, hidden Markov model (HMM), and SVM. The Bi-LSTM model utilized complete data sequences within a given window for sequence-to-label classification and achieved an accuracy of 89%, an AUC of 94.3% on the validation set, and 92.1% accuracy and 98.7% AUC on the test set.

Savadkoohi et al. [50] explored the recognition of human balance characteristics using One-One-One architecture to predict fall risk. Their study included 163 participants with diverse ages and health conditions. Assessing balance involved a protocol where participants maintained a stationary balance on a force plate for 60 seconds across various conditions. The One-One-One architecture comprises a 1D-convolutional layer, an LSTM layer, and a dense layer. The authors employed the Falls Efficacy Scale (FES) results to categorize falling into low, moderate, and high classes. The proposed model exhibited outstanding efficiency, showcasing precision, sensitivity, and accuracy rates of 100%, 100%, and 99.9% respectively.

Table 2.5: Classification of Healthy and Pathological Gait

Research Article	Purpose	Model [Architecture]	Data Type	Dataset	Performance Metric	Results
El Maachi et al. [81]	Detection of PD	1D-Convnet-based [two parts: the first part comprises 18 parallel 1D CNNs, while the second part is an FCN that operates on the concatenation of the 18 1D-CNNs]	vertical GRF	PhysioNet: imbalance dataset of 350 walks with 70% of PD walks (healthy subjects: 73 versus PD: 93)	Accuracy	98.70%
Alharthi et al. [83]	Classification of PD and healthy subjects	2D-DCNN [4 CLs + APL + 2 FCLs]	GRF	166 healthy and pathological subjects	Accuracy	95.5±0.28%
Sadeghzadehyazdi et al. [52]	Gait anomaly recognition	CNN-LSTM-based [3 CLs + MPL + FCL + BL]	skeleton images	Walking Gait dataset (9 participants), MMGS dataset (27 participants), and Pathological gait dataset (10 participants)	Accuracy	Accuracy on Walking Gait dataset: 90.75%, MMGS dataset: 83.64%, and Pathological gait dataset: 90.83%
Zhu et al. [51]	Classification of healthy and pathological subjects	Two-stream CNN [2 parallel CLs + fusion network (2 CLs + MPL + FCL)]	motion capture suit	45 subjects (healthy: 10, patient with joint problems: 4, muscle weakness: 18, patient with neurological defects: 13)	Accuracy, Precision, Recall, and F1-score	95% accuracy, 92% precision, 92% recall, and 92% F1-measure
Kaur et al. [85]	Classification of strides in individuals with MS, PD, and HOA	Nine traditional ML algorithms (LR, SVM, and RBF SVM kernels, DT, RF, AdaBoost, XGBoost, GBM, and MLP) and 7 DL algorithms (1D-CNN, ResNet, MSResNet, TCN, Vanilla RNN, LSTM, and Gated Recurrent Unit (GRU))	two digital cameras	33 subjects (healthy:14, patients with MS: 10, and patients with PD: 9)	AUC	1D CNN AUC: 79.3%, residual network: 78.1%
J. Yu et al. [86]	Prediction of geriatric stroke	C4.5 DT model, RF, XGBoost, and LSTM	motion, ECG, EMG, and foot sensor	Elderly Korean subjects	Accuracy	C4.5 DT model: 98.25%, RF: 98.72%, XGBoost: 96.60%, and LSTM: 98.99%
Chen et al. [87]	Recognition of sarcopenia disease	CNN-based [CL + 2 FCL]	IMU and EMG sensor	55 Pathological subjects	Accuracy, Precision, Specificity, and Sensitivity	94.41% accuracy, 91.58% precision, 95.81% specificity, and 91.58% sensitivity.
D. Thakur and S. Biswas [88]	Prediction of hemiplegia gait	CNN-LSTM with an attention mechanism	accelerometer and gyroscope signals	28 subjects (healthy and pathological)	Accuracy, Precision, Recall and F1-score	Accuracy: 86%, Precision of 80%, recall (sensitivity) 100% and F1-score of 88.89%

Note: BL = Bi-LSTM Layer, CL = Conv Layer, FCL = Fully Connected Layer, MPL = Max Pooling Layer, and APL = Average Pooling Layer

Table 2.6: Prediction of fall-risk and Freezing of Gait (FOG)

Research Article	Purpose	Model [Architecture]	Data Type	Dataset	performance Metric	Results
Tunca et al. [49]	Prediction of fall-risk assessment	LSTM-based [2 architectures: one with BL(10 units) + FCL(10 units), another with 2 BLs(100 and 20 units) + FCL(10 units)]	IMU, accelerometer, and gyroscope	76 pathological subjects	Accuracy	92.10%
Shalin et al. [89]	Prediction of FOG	LSTM-based [2 LLs(16 units)]	plantar pressure images	11 pathological subjects	Sensitivity and Specificity	Mean sensitivity: 82.1% (SD 6.2%) and mean specificity: 89.5% (SD 3.6%)
Savadkoohi et al. [50]	Prediction of fall-risk assessment	One-One-One Deep Neural Networks [Conv1D + LL + FCL]	force plates	163 pathological subjects	Accuracy, Precision, and Sensitivity	Accuracy: 99.9%, Precision: 100%, and Sensitivity: 100%
El-ziaat et al. [90]	Prediction of FOG	Deep Conv-LSTM [3 CLs + MPLs + 3 LLs]	sensors (IMU and accelerometer) data converted into spectrogram images	Daphnet (10 subjects) and Opportunity Dataset (12 subjects) - Healthy and pathological subjects	Accuracy	93.5%
Li et al. [91]	Prediction of FOG	Multimodal fusion strategy consisting Conv1D and LSTM	IMU and FSI	32 pathological subjects	Sensitivity, specificity, accuracy, AUC, EER, and F1 value	Sensitivity: 0.924, Specificity of: 0.983, Accuracy of: 96.3%, and F1 value of 0.943.

Note: LL = LSTM Layer, BL = Bi-LSTM Layer, CL = Conv Layer, FCL = Fully Connected Layer, and MPL = Max Pooling Layer

2.3.1.4 Prediction of Freezing of Gait

Several research projects have looked into using advanced DL methods, like LSTM models and hybrid architectures, to find and predict FOG in people who have been diagnosed with PD. Shalin et al. [89] conducted a study centered on LSTM utilization based on plantar pressure data. Their investigation included 11 male volunteers, focusing on real-time wearable applications and achieving impressive accuracy in FOG detection. The best FOG detection model achieved a mean sensitivity of 82.1%, a mean specificity of 89.5%, and 95% identification of freeze episodes. El-ziaat et al. [90] proposed a deep Conv-LSTM model with novel spectrogram image inputs derived from angular axes features. They evaluated their approach on benchmark datasets (Daphnet and Opportunity datasets, see Table 2.12) and achieved noteworthy accuracy, indicating potential advancements in FOG prediction accuracy. The best results from DL models were 97.6% for 2D CNN and 93.5% for hybrid Conv-LSTM. Li et al [91] adopted a multimodal fusion strategy, integrating force-sensitive insole (FSI) and IMU data from 32 individuals. Their feature-fusion-weighted model showcased excellent performance in predicting FOG duration and frequency, emphasizing robustness in monitoring FOG occurrences. The feature-fusion-weighted model yielded the best performance with a sensitivity of 92.4%, a specificity of 98.3%, an accuracy of 96.3%, and an F1 value of 0.943. In conclusion, these studies show that DL-based methods have a lot of potential for improving FOG detection and prediction. They also give us important information that can help doctors better manage and treat people with PD.

2.3.1.5 Other classification studies

The study [92] aimed to explore the effectiveness of a LSTM model using IMU sensor data to differentiate between age groups and walking surfaces. Participants walked on flat and uneven surfaces wearing athletic shoes while tethered to a safety harness in a motion-capture lab. Hu et al. classified the data into flat or uneven surfaces and young or elderly categories. This study included seventeen older (71.5±4.2 years, 165.7±9.3 cm, 67.6±12 kg) and eighteen young (27.0±4.7 years, 171.6±8.8 cm, 69.5±14.7 kg) adults. Two LSTM layers extracted information and the model predicted the most probable walking surface. Evaluation metrics, including accuracy, precision, recall, and F1-score, showed highly accurate and precise models, achieving 96.3% and 94.7% accuracy for type of surface and age prediction, respectively. With high AUC values of 0.97 and 0.96 for surface and age, these models demonstrate reliability for future research and applications.

Girka et al. [93] formulate the prediction of the appearance of the impact peak of GRF as a binary classification problem. They utilized kinematic data as predictors, specifically raw signals in the sagittal plane, which were collected using the Vicon motion capture system. 135 healthy individuals participated in the study, comprising both elite and recreational runners who underwent measurements. In total, the dataset was comprised of 1196 trials. This study aimed to investigate the anthropometric characteristics of the participants (body weight was 71.6±11.4 kg; height was 174.8±9.0 cm; leg length was 91.8±5.4 cm; Body-Mass Index (BMI) 23.3±2.3 kg/m^2). They proposed three layers CNN each subsequently followed by a spatial pooling layer. The accuracy and F-measure metrics were 81.09%±2.58% and 82.07%±2.31%, respectively.

Hernandez et al. [94] proposed a CNN architecture for detecting and predicting walking activity and gait period utilizing wearable sensors. Twelve healthy male individuals (from 24 to 34 y.o., between 1.74 and 1.79 meters tall, and weight between 77.6 and 85 kg) without abnormal gait were selected at the University of Leeds' Institute of Design, Robotics, and Optimization (iDRO). The authors gathered information from three IMU sensors mounted to the subjects' feet, shanks, and thighs. The IMU sensors possess nine degrees of freedom and can supply data from the accelerometer, gyroscope, and magnetometer. CNN was used for recognizing the walking activity and gait phases. Two layers CNN identified from these wearable sensors level ground walking, ramp ascent, and ramp descent and recognized walking activity and gait period with 100% and 98.63% accuracy, respectively.

Yen et al. [95] aimed to develop an LSTM model to accurately identify and categorize dynamically the data of the center of pressure (COP) and GRF to distinguish between straight walking and turns. Furthermore, the study seeks to uncover gait traits that have the potential to substitute EEG in forecasting walking directional intentions. Following that, the gait intention classification in the LSTM model would involve treating GRF and COP as characteristics. A total of ten healthy individuals (age of 24.3±1.79 years, height of 176±3.80 cm, and mass of 72±11.78 kg) were selected for the study. The sliding window achieved a maximum accuracy of 94.8% within a time frame of 0.7s.

2.3.2 Regression Studies

In this section, we review the studies that worked on the prediction of kinematic or kinetic sequences and the severity prediction of NDD (shown in Figure 2.7). In this part, a total of 24 studies were conducted to investigate the use of various neural network architectures in different applications. In the past, except DL, the kernel recursive least square method, principal component analysis, and best linear unbiased estimation were used for joint torque prediction [96], joint trajectory prediction [97, 98], generation of trajectories [99], and limb motion estimation [100]. Currently, LSTMs, CNNs, and a combination of both are the most used models. Other studies also utilized diverse DL models, including generative adversarial network (GAN), MLP with CNN, and LSTM with an attention mechanism (shown in Figure 2.7).

Table 2.7: Remaining classification studies

Research Article	Purpose	Model [Architecture]	Data Type	Dataset	performance Metric	Results
Hu et al. [92]	Detection of surface- and age-related differences in walking	LSTM-based [2 LLS + FCL]	triaxial accelerometer, gyroscope, and magnetometer data from an IMU	36 healthy subjects	Accuracy, precision, recall, and F1-score	Accuracy (96.3%, 94.7%), precision (96.4%, 95.2%), recall (96.3%, 94.7%), and f1-score (96.3%, 94.6%)
Girka et al. [93]	Prediction of the appearance of the impact peak of GRF	CNN-based [3 CLs + MPL]	GRF	135 healthy subjects	Accuracy	81.09%±2.58%
Hernandez et al. [94]	Detection of walking activity and gait period recognition	CNN-based [2 CLs + MPL]	IMU	12 healthy subjects	Accuracy	Activity recognition:100% and Gait period recognition:98.63%
Yen et al. [95]	Recognition of walking directional intention	LSTM-based	IMU	10 healthy subjects	Accuracy	94.79%

Note: FCL = Fully Connected Layer, CL = Conv Layer, LL = LSTM Layer, and MPL = Max Pooling Layer

We identify four categories of regression studies that will be detailed in the next sections: future sub-sequence forecasting based on time window input, sensor-to-sensor sequence estimation, condition/joint translation, and clinical scores prediction.

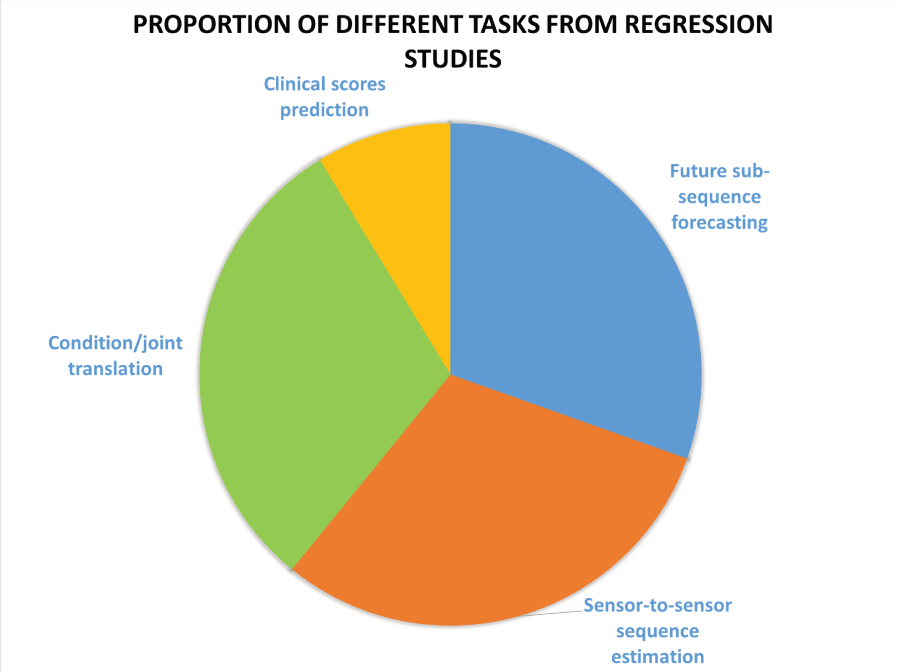


Figure 2.7: Applications in the literature for regression studies.

2.3.2.1 Future sub-sequence forecasting

This category includes studies that take biomechanical time series for a certain time window as input and output future values of the same signal. A typical example of this category of DL model considers $n_{\text{win,in}}$ values of a kinematic or kinetic sequence (i.e., knee flexion/extension during gait) from t_0 to t_{in} and predicts the next $n_{\text{win,out}}$ values of the same sequence from $t_{\text{in}} + T_s$ to $t_{\text{in}} + T_s + t_{\text{out}}$. $n_{\text{win,in}}$ and

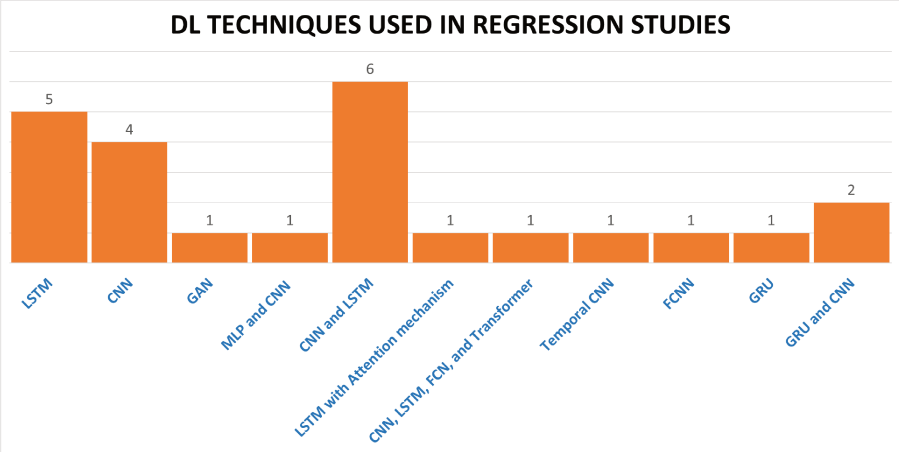


Figure 2.8: DL techniques in the literature used for regression studies.

$n_{\text{win,out}}$ are the number of time steps in the input and output windows, respectively; t_0 is the initial time of the signal; t_{in} and t_{out} are the duration of the input and output windows, respectively; and T_s is the sampling time.

Su and Gutierrez-Farewik [70] proposed an LSTM-based network to predict the angular velocity of the shank, foot, and thigh segments during gait. 12 healthy subjects (25–30 y.o.) were used in the study. The authors recorded 1500 to 1700 gait cycles per participant. For a forecast of 5 and 10-time steps (output window), the input window size was set to 10 and 30, respectively. They achieved correlation coefficient values greater than 0.98 in trajectory predictions.

Gholami et al. [101] proposed a shoe-mounted accelerometer for predicting sagittal plane lower extremity angles (hip, knee, and ankle) during running using a 1D-CNN. The study recorded motion and accelerometer data from 10 healthy male participants. The data recording protocol included 15 trials of running at varying speeds. Each participant was outfitted with 25 reflective markers. A 1D-CNN model was implemented, consisting of four convolutional layers and a max-pooling layer. In the context of inter-participant scenarios, the Root Mean Squared Error (RMSE) were lower than 6.5° and the average R^2 values for hip, knee, and ankle were 0.84, 0.93, and 0.73, respectively.

Sharma and Rombokas [102] proposed an LSTM-based to predict knee and ankle motion from the Xsens Awinda suit for locomotion data (consisting of 17 body-worn sensors) and an eye tracker for egocentric vision data. 23 healthy subjects were used in the study. For almost 12 hours, subjects were recorded as they moved through public classrooms, a large atrium, and stairs. The variety of the activities recorded made the prediction task more complex. Two parallel LSTMs were used, one for processing the optical flow and the other for processing joint angles. After that, both were combined using fusion layers to predict the trajectories of the knee and ankle on the sagittal plane.The authors showed that using vision enhances the accuracy of the predicted knee and ankle trajectories, particularly in crowded areas and when the visual environment offers information that does not simply appear in body movements. Overall, the predictions Normalized RMSE (NRMSE) for knee and ankle angles are 0.129±0.057 and 0.119±0.052, respectively.

Kolaghassi et al. [103] implemented CNN and LSTM models for forecasting gait trajectories of the ankle, knee, and hip up to 200ms in the future. They used the public data of children with neurological disorders collected by Gillette Children's Speciality Healthcare from 1994–2017. Their ages ranged from 4 to 19 y.o. 73% children have CP and the rest have other disorders. Only one lower limb was used for prediction. Input windows ranging from 50 to 1000ms were used to predict future trajectories (output windows ranging from 8 to 200ms). The results show that LSTMs does better than CNNs, and the difference gets more significant as the input and output window sizes get bigger. The most significant difference between the CNN and LSTM networks' MAEs was 0.91°.

Kolaghassi et al. [104] predicted the next 200-time steps based on the walking patterns of children in two groups: typically developing (TD) and patients with CP. The authors compared four DL models: FCN, CNN, LSTM, and Transformer. A dataset of hip, knee, and ankle flexion-extension angles was measured concurrently for both lower limbs. Training and validation were performed with TD data, whereas both TD and CP data were used for testing. According to the results, Transformer (1.17°) has the lowest MAEs for one-step-ahead TD gait prediction, whilst LSTM outperformed the other methods in long-term predictions for both CP and TD data (MAE of 13.41° and 9.36° for 200 time-steps ahead, respectively).

Karakish et al. [105] proposed an intelligent embedded micro-controller for foot angular velocity prediction using MLP and CNN. The input was the angular velocity and acceleration of the shank based on IMU data. Sliding windows were used for prediction, with 5 time-frames as input and 10 time-frames as output (about 200ms ahead). Five subjects from the human gait database (HUGaDB) were used in the study (those with non-corrupted gyroscope data). Two experiments were performed with and without current gait phase (CGP) information. The authors found that without CGP information, CNN obtained better results (R^2 of 0.944), but both achieved equivalent performance with CGP (R^2 of 0.956 and 0.952 for CNN and MLP respectively).

Kolaghassi et al. [106] investigated the efficiency of deep MLPs in accurately predicting gait angles of the hip, knee, and ankle in the sagittal plane at different speeds. The purpose of this research is to control exoskeletons. The online gait dataset provided by Camargo et al. [107] served as the basis for training the deep MLPs in our study. The dataset comprises gait data obtained from 22 physically fit individuals. Participants were recorded as they walked on a treadmill at 28 different speeds.The data was divided into three different speed ranges for training purposes: low, medium, and high. Each model had an identical architecture, including the number of layers and nodes per layer. Four models were trained using gait data at different speeds to predict gait data. According to the findings, when evaluated on the excluded speeds, the MAE found that the low- and high-speed models both showed a decline in performance ranging from roughly 43.7% to 90.7%. Conversely, when evaluated on the medium speeds that were not initially considered, the low high-speed model demonstrated a 2.8% enhancement in short-term predictions and a 9.8% improvement in long-term predictions.

Song et al. [108] suggested a way to use surface EMG and an LSTM network to predict joint angles online. Five healthy subjects took part in the data-collection experiment. The TrignoTM Wireless EMG System from Delsys was used in this investigation to collect sEMG signals. Simultaneous data collection was done on the plantar pressure signals, three joint angles (hip, knee, and ankle), and eight muscles of the right leg of five patients. For training the model, different inputs were used: sEMG (unimodal) and sEMG combined with plantar pressure (multimodal). The input sliding window was 20points (about 16.7ms) to predict four different predicted time intervals (50, 100, 150, and 200 ms). The range of the RMSE, MAE, and PCC obtained from the input of sEMG data were [1.63°, 3.20°], [1.27°, 2.36°], and[0.9747, 0.9935], respectively.

2.3.2.2 Sensor-to-sensor sequence estimation

This category includes studies that estimate kinematic or kinetic sequences from sensors based on regressions where the target output is computed with other types of sensors, usually to simplify computation and experimental settings (i.e., reduce the number of sensors). A typical example is a regression task between a kinematic signal computed with an optoelectronic system based on a biomechanical model and raw data from other sensors (i.e., IMU or accelerometers). In this case, the purpose is not to predict future values of the sequence but to estimate another sequence based on another one in the same time frame.

Senanayake et al. [109] aimed to utilize GAN to accurately transform IMU data into ankle joint angles that are comparable to those obtained through motion-capture-based inverse kinematics. The

Table 2.8: Future sub-sequence forecasting regression studies

Research Article	Purpose	Model [Architecture]	Data Type	Dataset	performance Metric	Results
Su and Gutierrez-Farewik [70]	Prediction of the angular velocity of the shank, foot, and thigh segments (next 5 and 10 time-steps)	LSTM-based [LUs and FCLs depends on the number of future steps to predict]	IMU	12 healthy subjects (thighs, shanks, feet, and pelvis)	RMSE and correlation coefficient	RMSE: 2.4±0.55 and correlation value > 0.98
Gholamie al. [101]	Forecasting gait trajectories of the ankle, knee, and hip up to 300ms	CNN-based [4 CLs + MPL]	IMU	10 healthy subjects	RMSE	< 3.5 and 6.5 in intra- and inter-participant scenarios
Sharma and Rombokas [102]	Prediction of the knee and ankle motion (next 1 gait cycle)	LSTM-based [2 parallel LLs + FL]	egocentric vision data using eye tracker and Xsens Awinda suit, consisting of 17 body-worn sensors	23 healthy subjects	RMSE and PCC	RMSE of knee and ankle: 0.129±0.057 and 0.119±0.052. PCC of knee and ankle: 0.799±0.205 and 0.637±0.264
Kolaghassi et al. [103]	Forecasting gait trajectories of the ankle, knee, and hip up to 200ms	CNN-based [4 CLs] and LSTM-based [4 LLs(128 units)]	VICON (motion capture system)	Gillette Children's Speciality Healthcare dataset	MAE	0.095–2.531 degrees for the LSTM and 0.129–2.840 degrees for the CNN
Kolaghassi et al. [104]	Forecasting gait trajectories of the ankle, knee, and hip up to 200ms	LSTM-based [2 LLs(100 units) + FCL], FCN [FCLs + sigmoid layer], CNN [4 CLs + 2 MPLs + FCL] and Transformer	IMU	10 typically developing children (TD) and 11 CP patients	Mean Squared Error (MSE) and MAE	FCN and Transformer with $MAEs$ for one-step-ahead predictions are between 1.17°-1.63°.
Karakish et al. [105]	Angular velocity prediction up to 200ms	MLP-bases and CNN-based	IMU	Five healthy subjects Human Gait Database (shank and foot)	$Bias$, MAE, RMSE, R^2, and PCC	R^2 of 0.956 and 0.952 for CNN and MLP respectively
Kolaghassi et al. [106]	Prediction of gait angles of the hip, knee, and ankle in the sagittal plane at different speeds	FCNN [5 FCLs(100 neurons) + FCL]	OpenSim's inverse kinematics tool (motion capture data)	22 healthy Subjects	MAE and MSE	please see the section for details
Song et al. [108]	Prediction of joints angles at four different time intervals (50, 100, 150, and 200 ms)	LSTM-based	sEMG	5 healthy subjects	RMSE, MAE, and PCC	[1.63°,3.20°], [1.27°, 2.36°], and [0.9747, 0.9935]

Note: FCL = Fully Connected Layer, CL = Conv Layer, LL = LSTM Layer, and MPL = Max Pooling Layer

study recruited nine participants with no prior history of lower limb pathology. The authors trained a GAN using gait data collected at two different walking speeds. The objective was to predict ankle kinematics solely from IMU data for a third walking speed. The results showed that the RMSE difference for ankle dorsiflexion was 3.8°, inversion was 2.1°, and axial rotation was 3.5°.

Rapp et al. [110] proposed a CNN and LSTM for estimating joint kinematics (hip, knee, and ankle) using IMUs. The authors used marker-based motion capture data. After removing data with marker-tracking errors, 420 subjects of walking trials and 580 subjects of running trials were chosen for further analysis. Separate predictive models were developed for each joint. To predict lower extremity kinematics on left-out subjects with a mean RMSE of less than 1.27°±0.38° in flexion/extension, less than 2.52°±0.98° in adduction/abduction, and less than 3.34°±1.02° in internal/external rotation.

Hernandez et al. [111] proposed a DeepConvLSTM (combining convolutional layers and LSTM modules) to estimate marker-based kinematics from IMU data during walking and running at different speeds. A total of 27 healthy participants were recruited. For each DOF, the average r, MAE, and mean error (ME) ranged from 0.67 (0.23) to 0.99 (0.01), 2.2 (0.9) to 5.1 (2.7), and -0.29 (2.06) to 0.85 (5.58), respectively, when all speed circumstances were taken into account. A significant connection (r: 0.7-0.9) is seen for the hip left rotation and ankle right/left inversion, while all other DOFs showed a significant association (r: 0.9). Moderate correlation (r: 0.4–0.7) is seen for the right hip rotation and lumbar extension.

Hossain et al. [112] presented a DeepBBWAE-Net (ensemble technique using base learners) for mapping joint angles using a reduced number of IMU sensors. Their study optimizes the sensor setup to achieve accurate joint angle estimation. Ten healthy individuals participated in the study. Two IMU sensors were utilized and placed on the participants' shoes. Multiple trials were conducted for each participant in different conditions: treadmill, overground, stair, and slope. They utilized the sagittal plane angles of the hip, knee, and ankle for both lower limbs. The authors implement five networks combining CNNs and GRU-based RNN as base learners for their framework. DeepBBWAE-Net achieved RMSE values of 3.77°, 4.62°, and 3.22° degrees for hip, knee, and ankle joints, resp., during treadmill walking and RMSE values of 4.32°, 4.28°, and 3.09° for the hip, knee, and ankle joints, resp., during overground walking.

Tan et al. [113] utilized 2 or 4 IMUs sensors for capturing data from 17 individuals with knee osteoarthritis to forecast knee flexion/extension via stacked Bi-LSTM. A total of 28 retroreflective markers were affixed to the pelvis and lower limbs of the participant. The performance of a single-leg model was compared to that of a double-leg model in terms of prediction error. The results indicated that the single-leg model exhibited significantly lower prediction error than the double-leg model. The RMSE was found to range from 7.04 (2.6) to 11.78 (6.04), while the MAE ranged from 5.99 (2.34) to 10.37 (5.44). R^2 score was also calculated and ranged from 0.85 to 0.99.

Heeb et al. [114] proposed ML and DL models to measure the power of the ankle joint using force-myography (FMG) sensors. Nine young, healthy men took part in this study and gave information that was used in it. Before the experiment, each person was given an FMG strap with eight FSR devices that they wore around their waist. About 2 inches above the ankle, the FMG strap was wrapped around the participant's right leg. During the experiment, each person walked on the force-plate-equipped treadmill for one minute at each of the five different speeds (0.4, 0.7, 1.0, 1.3, and 1.6 m/s). The authors implemented LSTM, 1-D CNN, and Cat Boost Regressor (CBR) to achieve the purpose of the study. Results showed that the LSTM model using time-domain features got a correlation coefficient of over $R = 0.91 \pm 0.07$, and the CNN model using raw features got $R = 0.89 \pm 0.13$.

Hossain et al. [115] introduced a new DL model called Kinetics-FM-DLR-Ensemble-Net. The model is designed to predict the moments of the hip, knee, and ankle joints, as well as the 3-D GRFs, using data from three IMU sensors placed on the thigh, shank, and foot. The model was tested under various walking conditions commonly encountered in daily life, including treadmills, level ground, stairs, and ramps. The primary component of the Kinetics-FM-DLR-Ensemble-Net model is

the Kinetics-FM-DLR-Net model. They utilized bagging methodologies by employing Kinetics-FM-DLR-Net to construct Kinetics-FM-DLR-Ensemble-Net. Kinetics-FM-DLR-Net is mostly made up of two Kinetics-FM-Nets. Each model was trained using two different loss functions, and then they were combined using a new method called Double Loss Regression (DLR). Kinetics-FM-Net is constructed by combining Kinetics-Net with a fusion module (FM) component. Kinetics-Net is constructed by incorporating various DL layers such as GRU, Conv1D, Conv2D, and fully connected dense layers. This study utilized two openly accessible datasets, Dataset A [107] and Dataset B [116]. Dataset A has data from twenty participants and it was employed to train a model. Dataset B has seventeen participants. The proposed model achieved RMSE of 4.32±0.81 and 7.43±0.71 for Dataset A and Dataset B, respectively.

2.3.2.3 Condition/joint translation

This category is similar to the previous one, but instead of estimating from different types of sensors for the same segment or joint movement, they estimate from another joint data (i.e., estimating knee kinematics from foot kinematics) or another condition (i.e., estimating running kinematics from gait kinematics) using the same type of sensors.

He et al. [117] proposed an LSTM-based model to investigate the relationship between upper and lower limb movements. They predicted the hip and knee movements on one side of the lower limb using the shoulder and elbow movements on the other side of the upper limb. The authors recruited 10 healthy participants. A unique Kinect-Treadmill data collection platform was created for walking at various velocities (3.0, 3.5, 4.0, and 4.5km/h). The average RMSE of the prediction of the lower limb for all velocities was around 2.4°±0.55°.

Chow et al. [118] compared accelerometer and gyroscope in forecasting level-ground running kinematics (hip and knee flexion/extension) from treadmill running kinematics using a CNN model. The training data for treadmill running kinematics were measured using a single IMU on the anteromedial side of the right tibia and level-ground running kinematics were measured by four IMUs placed on the lower extremities. Ten healthy recreational runners were included in the study.The results showed that the R^2 scores ranged from 0.85 to 0.96 for intraparticipant comparisons and 0.7 to 0.92 for interparticipant comparisons. The RMSE values of running kinematics ranged from 3.2° to 7.1° and from 6.3° to 6.8° in intraparticipant and interparticipant tests, respectively.

Sharifi Renani et al. [119] proposed a Bi-LSTM-based model to predict the three-dimensional hip and knee kinematics during gait using either experimentally measured IMU data, synthetically generated IMU data, or a combination of both (pelvis, left thigh, left shank, and left foot). The study used 30 participants. Two separate neural network models were developed to predict joint angles from the related IMU data, one for the knee and one for hip kinematics. Both architectures consist of a Bi-LSTM layer followed by two fully connected layers. The authors concluded that using synthetic data for training and prediction improves hip and knee results compared to measured data from subjects (RMSE of the hip: 2.3° and knee: 2.9°). When trained using synthetic data and measurements, models got better results than when using only IMU synthetic data (RMSE of the hip: 1.9° and knee: 1.7°).

Dey & Schilling [120] introduced an online or offline foot angle trajectory prediction network (FATP-N), which is based on temporal convolution. They predict the sagittal foot angular locations during natural walking (different speeds, cadences, and not necessarily in a straight line) based on the shank angular position measured with a single wearable IMU motion tracker sensor. They also evaluated their method with various baseline and cutting-edge data-driven methods (LSTM, GRU, LR, and GPR) for predicting gait trajectories. Shank and foot angular positions were measured on seven healthy people. Wireless motion tracker sensors (MTw Awinda, Xsens) were used to collect data at 100 Hz from the subject's shank and foot center of mass. A correlation coefficient of 0.98±0.01, an R^2 score of 0.95±0.01, and an NRMSE of 4.7±0.9% were recorded by the offline-trained FATP-N. Compared to offline forecasts, the performance of the online predictions was marginally worse.

Table 2.9: Sensor-to-sensor sequence estimation regression studies

Research Article	Purpose	Model [Architecture]	Data Type	Dataset	performance Metric	Results
Senanayake et al. [109]	Prediction of ankle joint angles from IMU data	GAN	IMU	9 healthy subjects	RMS	3.8°, 2.1° and 3.5° for dorsiflexion, inversion, and axial rotation, respectively.
Rapp et al. [110]	Prediction of hip, knee, and ankle joint angles from IMU data	Conv1D and LSTM-based	IMU	420 healthy subjects	RMSE	RMSE $<$ 1.27 (0.38) in flexion/extension, $<$ 2.52 (0.98) in adduction/abduction, and $<$ 3.34 (1.02) in internal/external rotation
Hernandez et al. [111]	Estimation of marker-based kinematics from IMU data	DeepConvLSTM [2 CLs + 2 LLs]	IMU	27 healthy subjects	MAE and R	MAE for the DOFs ranged from 2.2(0.9)° to 5.1(2.7)° with an average of 3.6(2.1)°.
Hossain et al. [112]	Prediction of hip, knee, and ankle joint angles from IMU data	CNN-GRU	IMU	10 healthy subjects	RMSE	6.93-29.0
Tan et al. [113]	Prediction of knee flexion/extension from IMU data	Bi-LSTM-based [2 BLs + DrL + FCL]	IMU	17 pathological subjects	RMSE , MAE, and Pearson's R	RMSE, MAE, and R ranged from 7.04(2.6) to 11.78(6.04), 5.99(2.34) to 10.37(5.44), and 0.85 to 0.99, respectively.
Heeb et al. [114]	Prediction of ankle joint power using FMG	LSTM-based [LL(1024 neurons) + 4 FCLs(256, 128, 64, and 1 neurons)], and CNN-based [3 CLs + 2 DrLs + 1D AP + 2 FCLs(128 and 1 neurons)]	FMG sensor	9 healthy subjects	R, RMSE , and MAE	LSTM: R = 0.91±0.07 and CNN: R = 0.89±0.13
Hossain et al. [115]	Prediction of joint moments and 3D GRFs from IMU data	Kinetics-FM-DLR-Ensemble-Net consists of GRU, Conv1D, Conv2D, and FCLs	IMU	Dataset A (20 healthy subjects) and Dataset B(17 healthy subjects)	NRMSE and PCC	4.32 and 0.929 for Dataset A. 7.43 and 0.886 for Dataset B

Note: FCL = Fully Connected Layer, CL = Conv Layer, LL = LSTM Layer, and MPL = Max Pooling Layer

Ding et al. [121] proposed the LSTM-based model with an attention mechanism for motion prediction of a subject's limb using the motion of complementary limbs. Twenty-one subjects' data were used in that study for experiments. A total of six modes of walking were measured: ground-level walking, stopped, ramp descent, ramp ascent, stair descent, stair ascent, and transition between them. Five sensors, consisting of a gyroscope and an accelerometer, are placed on the subject's body to extract the data. The authors implemented four different models: two without an attention mechanism and two with an attention mechanism. Two of them were using LSTM, and two of them were using Bi-LSTM. The study used two joints, the left ankle, and the left hip, for motion predictions (linear acceleration). Results showed that adding attention layers made 57% of the model parameters unnecessary. At the same time, the prediction error was lower than with the LSTM model without the attention mechanism. The NRMSE for the attention model was 9.06% for ankle acceleration and 7.64% for hip acceleration. The MAE was 1.43m/sec^2 and 3.20m/sec^2 for the ankle and hip, respectively.

Lee and Asbeck [122] presented a foot placement prediction method based on GRU. This method involves analyzing data from three IMU sensors, which are attached to the pelvis and feet, sequentially. Two DL models were developed, with the first model focused on estimating the progression of gait and the second model dedicated to predicting the next foot placement. Both models used GRU. The IMU data and the ground truth foot placement data are gathered using an XSens MVN Link suit. A total of ten healthy participants (8 male and 2 female) took part in the experiment. During the study, participants were instructed to walk naturally on a treadmill. The pre-trained, base, and fine-tuned models achieved mean distance errors of 6.99cm, 3.32cm, and 3.22cm, respectively.

Bajpai et al. [123] implemented an ANN called "foot2hip" that can accurately record gait kinematics over an extended period. Foot2hip utilizes foot kinematics and kinetics during walking to accurately predict the angles of the ankle, knee, and hip joints in the sagittal plane. The CNN-LSTM-DNN-based model (foot2hip) was composed of a series of layers, including convolution, max-pooling, LSTM, and dense layers. A locally created insole and an outsole were utilized to analyze the dynamics and movements of the foot, respectively. Data was gathered from a group of seven healthy male participants. The knee joint had an RMSE of 3.04±0.20 and a correlation coefficient of 0.97±0.01. The hip joint had an RMSE of 1.7±0.09 and a correlation coefficient of 0.95±0.01. Lastly, the ankle joint had an RMSE of 1.32±0.08 and a correlation coefficient of 0.91±0.02.

2.3.2.4 Clinical scores prediction

In this category, we regroup studies that predict continuous clinical variables (i.e. pathology severity scores) from gait data.

Berke Erdaş et al. [124] introduced a method for developing a disease severity grading system for NDDs, including amyotrophic lateral sclerosis (ALS), huntington's disease (HD), and PD. The proposed approach utilizes 1D and 2D CNN and gait data represented by a Quick Response code to achieve high levels of effectiveness and reliability. The PhysioNet database was utilized to evaluate the efficiency of the developed algorithms. The disease severity grades of Parkinson's patients from the HY scale, a worldwide scale, and the TFC degree of Huntington patients from the UHDRS scale, which measures symptom severity, were estimated. Since ALS patients have no disease degree information, the number of months since diagnosis has been estimated. They computed the R values for ALS, HD, and PD and found them to be (0.61, 0.39, and 0.7) and (0.88, 0.83, and 0.79), respectively for 1D and 2D CNN. The study also computed the R^2 value for three subsets, which resulted in values of (0.37, 0.15, and 0.49) and (0.79, 0.69, and 0.62), respectively.

Eguchi et al. [125] utilized gait videos to predict the score on the PD rating scale. They obtained 737 consecutive gait videos of 74 patients with PD and their corresponding neurologist-rated UPDRS scores retrospectively. A CNN model was used to predict the total UPDRS part III score and four subscores related to axial symptoms, bradykinesia, rigidity, and tremor. For the entire UPDRS part III score and the subscores of axial symptoms, bradykinesia, rigidity, and tremor in the test dataset,

the R^2 values between the model-predicted and neurologist-rated values were 0.59, 0.77, 0.56, 0.46, and 0.0, respectively. They reported that the performance was relatively low for patients with significant symptoms.

2.3.3 Gait analysis using Multi-task learning

Most of the studies using MTL are devoted to gait-based identification and authentication, not for clinical gait analysis. We could only find 5 studies, that come under the inclusion criteria of this review.

A study by Nait Aicha et al. [126] compared how well CNN, LSTM and a combined ConvLSTM used raw accelerometer data to predict falls in older adults. By incorporating auxiliary tasks like gender and age, the models showed improved performance. The dataset consisted of 296 subjects, 191 of whom had experienced at least one fall in the past six months. Their experiments involved various data divisions at both subject and sample levels, incorporating MTL to evaluate fall risk (primary task), and subject identification, age, gender, weight, and height prediction (auxiliary tasks). The ConvLSTM architecture generally outperformed CNN and LSTM, with the AUC of 0.75.

Yu et al. [127] came up with a Deep Multi-source Multi-task Learning (DMML) approach that gives a framework for assessing fall risk and PD severity based on accelerometer and gyroscope data. The goal of MTL is to improve performance of every single task by simultaneously evaluating the fall risk and PD severity. The authors used timed up-and-go (TUG) test data to evaluate the model. Five IMU sensors were put on 22 PD people aged 65 or over to track how they moved during the test. The training had a total of three stages: in stage 1, multi-layer CNN extract features that are specific to a source. In stage 2, these networks pull out general features that are common to all sources. Finally, the last layer is for MTL, where each task learns its features from the generic features learned in stage 2. Their model greatly outperforms the benchmark approaches, achieving an F-measure of 0.940 for assessing fall risks and an RMSE of 0.060 for measuring PD severities.

Zhang et al. [128] proposed a multi-task CNN for age estimation using gait. They used gender information as another task to improve age estimation. This study used a Large Population Dataset with Age (OULP-Age) dataset consisting of gait energy image (GEI). 28,923 subjects were used for training, 3000 for validation, and 31,923 for testing the model. The residual architecture of ConvNet was adopted for the model, following fully connected layers of each task. They got the 5.47 MAE for age estimation and 96.26% CCR.

Aoki et al. [129] implemented multi-task RNN (MRNN) to classify physically fatigued and non-fatigued gait cycles. According to the authors, there are significant intra-class variations in the gait cycle due to differences in the stance phase (which foot is swinging or supporting). One branch of MRNN is responsible for detecting the supporting foot as an auxiliary task, and the other is for the main task of fatigue classification. Eight healthy subjects were used for the experiments. The input for the model was the gait cycle, represented as 3-D coordinates of body joints. The model contains four independent RNN (IndRNN) layers, two common for both tasks, and two task-specific layers. The overall AUC in leave-one-subject-out validation was 0.86 ± 0.019, and in leave-one-day-out was 0.915.

Khan et al. [39] proposed a framework for patients with neurological diseases to optimize treatment outcomes. A regression strategy was used to predict how patients' joints would move after treatment using a multi-task architecture with LSTM and Bi-LSTM models. Medical treatment data (MTD) was added for context modeling, and a gating mechanism was used for model treatment interaction. A total of 38 patients were used in the study with different neurological diseases, e.g., CP, MS, SCI, stroke, or TBI. Data was collected using an optoelectronic Codamotion system. A total of 23 patients were bilaterally affected, and the rest of them were unilaterally affected. Only the sagittal plane of knee and ankle trajectories was used for the experiments. Seven models were used in the study; three were without MTL, and four used MTL in their architecture. In all models, there were five layers of LSTMs or Bi-LSTMs in a serial architecture (without MTL) or a parallel one (with MTL). Leave-one-

Table 2.10: Condition/joint translation and remaining regression studies

Research Article	Purpose	Model [Architecture]	Data Type	Dataset	performance Metric	Results
He et al. [117]	Prediction of lower limb using upper limbs	LSTM-based	Kinect sensor	10 healthy subjects	RMSE	2.4±0.55
Chow et al. [118]	Prediction of level-ground running kinematics (hip and knee flexion/extension values) by treadmill running kinematics	CNN-based [4 CLs + MPL]	IMU data + accelerometer and gyroscope	10 healthy subjects	R^2 and Normalized RMSE	Intraparticipant and interparticipant R^2 values: 0.85 to 0.96 and 0.7 to 0.92, NRMSE values: 3.6% to 10.8% and from 7.4% to 10.8%, respectively.
Sharifi Renani et al. [119]	Prediction of hip and knee kinematics using measured and synthetically generated IMU data	BiLSTM separate for Hip and Knee [BL + 2 FCLs]	IMU	30 healthy subjects	RMSE	Hip: 1.9 degrees and Knee: 1.7 degrees
Berke Erdaş et al. [124]	Severity prediction of NDD	1D CNN-based [1 CL + DrLs + FCLs] and 2D CNN-based [2 CLs + MPL + FCLs]	GRF data converted into QR images	PhysioNet (46 Healthy and pathological subjects)	R, R^2, MAE, Median Absolute Error ($MedAE$), MSE, and RMSE	R and R^2 values for ALS & Control: 0.88 and 0.79, HD & Control: 0.83 and 0.69, and PD & Control: 0.79 and 0.62.
Dey & Schilling. [120]	Prediction of foot angular locations during natural walking based on shank angular position	FATP-N (TCN) [3 temporal CLs(kernel size: 16, 32, and 64, respectively) + FCL]	Wireless motion tracker sensors (MTw Awinda, Xsens)	7 healthy subjects	R^2, NRMSE, and correlation coefficient	R^2 score: 0.95±0.01, NRMSE: 4.7±0.9%, and a correlation coefficient of 0.98±0.01.
Ding et al. [121]	Prediction of the motion of subject's leg using complementary limbs	LSTM with an attention mechanism	IMU	Bath Narual Environment HAR dataset (21 healthy subjects)	Mean Absolute Percentage Error ($MAPE$), MAE and NRMSE	Hip – Ankle (MAE (m/sec2): 1.43 – 3.20, NRMSE: 7.64% – 9.06%, and $MAPE$: 13.37% – 21.99%
Lee & Asbeck [122]	Foot placement prediction	GRU (single-layered)	IMU	10 healthy Subjects	RMSE and mean distance error (MDE)	please see the section for details
Bajpai & Joshi [123]	Prediction of the ankle, knee, and hip joints in the sagittal plan using foot kinetics and kimenatics	CNN-LSTM-DNN based model (foot2hip) [CL + MPL + LL + FCLs]	goniometer, pressure insole, and outsole with six Vertical-Cavity Surface-Emitting Laser (VCSEL) sensors	7 healthy subjects (kinetics and kinematics of the foot)	RMSE and PCC	3.04±0.20 and 0.97±0.01 for knee joint. 1.7±0.09 and 0.95±0.01 for the hip joint. 1.32±0.08 and 0.91±0.02 for ankle joint.
Eguchi et al. [125]	PD disease rating scale score prediction	CNN-based	gait video	74 patients	R^2	the R^2 values for the total UPDRS part III score was 0.59.

Note: LU = LSTM unit, BL = Bi-LSTM Layer, CL = Conv Layer, LL = Lstm Layer, FCL = Fully Connected Layer, MPL = Max Pooling Layer, BNL = Batch Normalization Layer, DrL = Dropout Layer, and FL = Fusion Layer

Table 2.11: Gait analysis using MTL

Research Article	Purpose	Model [Architecture]	Data Type	Dataset	Performance Metric	Results
Nait Aicha et al. [126]	Prediction of falls in older adults	CNN, LSTM, and ConvLSTM [5 CLs + LL + DL]	triaxial accelerometer	Fall risk assessment in older adults (FARAO) 296	AUC (10-fold cross-validation)	0.75
Yu et al. [127]	Prediction of the fall risk and severity of conditions of PD	Multi-source Multi-task Learning (DMML) - CNN [4 parallel layers (2 CLs + 2 non-linear layers + 2 MPLs) + DL(1000 neurons) + 2 parallel DLs (500 neurons)]	video	timed up and go (TUG) dataset	For fall risk, precision, recall, and F-Measure. For PD Severity RMSE (10-fold cross-validation)	Precision: 0.925, Recall: 0.958, FMeasure: 0.940 and RMSE: 0.060
Zhang et al. [128]	Age estimation using gait.	Multi-task CNN [CL + 2 parallel FCLs for each task]	Gait Energy Images	OULP-Age, 63,846 subjects	MAE and Correct Classification Rate (CCR)	MAE: 5.47 and CCR: 98.10%
Aoki et al. [129]	Classification of physically fatigued and non-fatigued gait cycles	Multi-task RNN	kinect sensor	8 healthy subjects	AUC, Sensitivity, and Specificity (Leave-one-subject-out Cross-validation)	AUC: 0.860±0.019 Sensitivity: 0.763±0.045 Specificity: 0.812±0.035
Khan et al. [39]	Optimize treatment outcomes (predicting post-gait trajectories of the knee and ankle).	MTL using LSTM and Bi-LSTM	optoelectronic Codamotion system	38 pathological patients with different diseases	RMSE and R^2	lowest RMSE of 5.60° for the knee and 3.77° for the ankle

Note: DL = Dense Layer, CL = Conv Layer, LL = Lstm Layer, FCL = Fully Connected Layer, and MPL = Max Pooling Layer

out cross-validation was used to assess the models' performance. Results showed that MTL models outperformed other models. The authors achieved the best results (lowest RMSE) of 5.60° for the knee in TBI patients and 3.77° for the ankle in CP patients. In conclusion, the MTL models had the best RMSE, ranging from 5.24° to 6.24°.

2.3.4 Major Datasets Used in Gait Analysis

The present section provides an overview of the primary publicly-available datasets utilized for DL applied to QGA. The majority of gait datasets available for research purposes consist of healthy subjects. Few databases contain both information on healthy and pathological gait. Most pathological gait datasets are not publicly available, usually for clinical data protection issues.
Two types of data are primarily available: (1) Vision-based: Images of skeletons, gait energy, silhouettes, and other objects. (2) Wearable sensor-based: IMU data, accelerometer, gyroscope, magnetometer, temperature, and other sensors are included in the sensor-based dataset. It is worth noting that only one dataset offers optoelectronic marker-based data (Rueangsirarak et al. [130]).

2.3.4.1 OULP-Age "OU-ISIR Gait Database, Large Population Dataset with Age"

OULP-Age [131] is a large population dataset of human gait for gait-based age estimation. It is an extended version of the OULP dataset. The databases consist of 63,846 subjects, with a balanced gender ratio of 32,753 females and 31,093 males. The ages of the participants ranged from 2 to 90 y.o.. The database is a collection of GEIs of subjects. The total length of video contained in this dataset is

44.3 hours. Each subject had a 2.5-s walking image sequence taken, with 640×480 pixel images and a 30 frames per second rate. More subjects belong to the age group 6–10, and fewer subjects belong to the age group 86–90.

2.3.4.2 Bath Natural Environment HAR Data Set

Bath Natural HAR dataset was published by Sherratt in 2020 [132]. The dataset consists of a total of 22 participants. All of them were healthy subjects. Five nine-axis IMU sensors consisting of a gyroscope, accelerometer, and magnetometer are placed onto the body: two on the legs at the ankles, two on the torso at the hips, and one on the chest. The three axes' linear acceleration, angular velocity, and magnetic field were measured with each IMU sensor. The vector sum of the three linear accelerations of the triaxial accelerometer was used to calculate the resulting acceleration. The sampling rate was 100 Hz, and 1.25 million samples were collected, corresponding to 209 minutes of data and 10,440 strides.

2.3.4.3 3d gait database

The dataset, published by Rueangsirarak et al. [130] in 2018, was collected from 45 senior Thai volunteers residing in Chiang Mai retirement villages and nursing institutions. The study protocol was approved by the Faculty of Associated Medical Sciences Ethics Council at Chiang Mai University. Three medical professionals from the university's College of Associated Medical Sciences evaluated the participants and categorized their gait conditions into four groups: healthy, muscle weakness, joint issues, and neurological defects. The participants, aged from 61 to 91, included 5 males and 40 females. The breakdown of gait conditions was as follows: 18 with muscle weakness, 13 with neurological defects; 4 with joint problems, 10 healthy individuals. Data collection involved a motion capture suit with reflective markers attached to the participants' bodies [133], producing temporal 3D marker placements. Following [134], the walking protocol consisted of four attempts along a 10-meter corridor at the subjects' typical gait speed, with two-minute rest periods between attempts: the first and last trials were designated for practice and cooling down, respectively, only the second and third trials were included in the final dataset. This comprehensive dataset provides valuable information for studying gait patterns in older adults with various musculoskeletal and neurological conditions, contributing to the development of automated diagnostic tools and interventions for age-related mobility issues.

2.3.4.4 Walking gait dataset

The walking gait dataset was designed by Nguyen and Meunier [135] in 2018. The dataset has nine categories of healthy and altered gait data. They instructed the test subjects to wear soles with 5, 10, and 15 cm padding, once on the right foot and once on the left. The individuals once wore a 4-kg weight ankle on each ankle to collect various anomalies. Each participant also walks with a normal gait. There are nine subjects in total. Each participant walks through eight abnormal and one normal sequence, each with 1200 frames. The dataset was represented in silhouette, skeleton, and point clouds. There were a total of 291,600 frames in the dataset.

2.3.4.5 Multi-modal gait symmetry (MMGS) dataset

The MMGS dataset was designed by Khokhlova et al. [136] in 2019. They collected the data using a single Kinect v.2 sensor placed before the test individual. Each individual approached the camera. The experiment involved 27 subjects, with 19 males and 8 females participating. The participants' average age is 30±7 y.o. The gait classes include normal walking, limping gait imitated by each participant wearing a 7-cm padding sole, and gait with knee injury-related issues, including a prosthesis worn

or after-fracture recovery gait. The latter category is replicated by instructing participants to walk without bending their right knee.

2.3.4.6 Pathological gait dataset

This dataset was designed by Jun et al. [137] in 2020. A system consisting of six calibrated Kinect sensors was used to capture the data for this set. To collect reliable data from various directions, the sensors have been calibrated to ensure they continuously acquire the same XYZ coordinate. A total of ten healthy participants acted out five distinct types of pathological gait, including antalgic, stiff-legged, lurching, steppage, and Trendelenburg gait. Every subject went through each gait class twenty times. This dataset includes six classes of normal and pseudo-pathological gait, with 120 sequences for each subject. These classes are in addition to the normal class. This dataset has the most significant number of frames overall, as well as the largest number of frames for each class, compared to the other datasets for skeleton-based gait anomaly recognition (SGAR) that are publicly available.

2.3.4.7 The Daphnet dataset

The Daphnet dataset was designed by Bachlin et al. [138] in 2009. Participants were people with idiopathic PD who had a history of FOG and could walk without help. Patients were not allowed to participate if they had serious problems with their eyesight or hearing, dementia, or signs of other neurological or orthopedic diseases. The local Human Subjects Review Committee permitted the study, which was done per the ethical rules of the Declaration of Helsinki. Ten people with PD took part in this study. The average age was 66.5±4.8 years. The average time of their disease was 13.7±9.67. Participants were asked to walk in a straight line (back and forth), do random walks, and do walking activities such as getting something to drink and entering the room. Two sessions were recorded for said activities. 500 minutes of data were recorded, equivalent to 8 hours and 20 minutes. 237 FOG events are identified in the dataset.

2.4 Identified Gaps

This section aims at identifying and discussing the significant gaps in gait analysis research. Specifically, we focus on a few aspects that require further investigation and development.

- Upon conducting a thorough review of the existing literature on gait analysis datasets, it has been determined that there is a limited number of datasets available on pathological gait (see Table 2.12). Specifically, only one dataset has been identified that pertains to pathological gait, and it includes only 10 subjects. The existing literature has extensively explored normal gait patterns over the past few years. However, there is a dearth of research studies that focus on pathological gait. Furthermore, the few studies that have been conducted in this area have predominantly utilized subjects as their primary research participants. The absence of publicly available datasets is a notable limitation in this area of research.

- The current body of literature on gait trajectory prediction is characterized by a lack of studies that utilize pathological gait. The majority of research studies have exploited normal gait for prediction purposes.

- The present study also aims to contribute to the existing literature on gait analysis by considering the MTL approach. Specifically, we seek to build upon the limited body of research in this area, as only five studies have been identified in the literature that have employed an MTL methodology

Table 2.12: Detail of Datasets

S#N	Name	Type	Size	Healthy/ Pathological	Pathology	Detail
1	OULP-Age	GEIs	63,846 subjects	Healthy		This dataset is used for age estimation.
2	Bath Natural Environment HAR Dataset	IMU	22 subjects	Healthy		This dataset is used to predict the motion of the leg.
3	3D Gait Database	Marker-based	45 subjects	Healthy+ Pathological	muscle weakness, joint issues, and neurological defects	This dataset is used to classify healthy and pathological subjects.
4	Walking gait dataset	pointcloud, skeleton, and silhouette	9 subjects	Healthy+ Abnormal	The individuals once wore a 4-kg weight ankle on each ankle to collect various anomalies.	This dataset is used to classify healthy and pathological subjects.
5	MMGS dataset	skeleton and silhouette	27 subjects	Healthy+ Pathological	wearing a 7-cm padding sole, and gait with knee injury-related issues, including a prosthesis worn or after-fracture recovery gait.	This dataset is used to classify healthy and pathological subjects.
6	Pathological gait dataset	skeleton	10 subjects	Healthy+ Pathological	five distinct types of pathological gait, including antalgic, stiff-legged, lurching, steppage, and Trendelenburg gait.	This dataset is used to classify healthy and pathological subjects.
7	The Daphnet dataset	accelerometer	10 subjects	Pathological	idiopathic PD	This dataset is used to classify patients with FOG problem.

for gait analysis. In recent years, the use of MTL has gained widespread success in various fields, including the field under consideration (see Table 2.11), as it offers a promising solution to address diverse problems.

- The present SLR has found that the availability of a public dataset containing information on pre-treatment gait and post-treatment gait is currently lacking. Most datasets on gait analysis were exploited for research studies on action recognition (out of the scope of this SLR). There are few datasets containing pathological gait data, and these are mainly employed to automatically distinguish between healthy and pathological gait patterns.
- Most of the studies used a very limited number of subjects, usually around ten. It makes it difficult to generalize methods used in the studies due to the huge variability in human gait.
- The present literature review has identified a single study [39] that focuses on the prediction of treatment outcomes through the utilization of gait analysis using DL. This study stands as the sole example of research in this area.

2.5 Discussion and Future Research Direction

The present study is an SLR that aims to provide an overview of DL algorithms and specifically MTL techniques that have been utilized for gait analysis. This paper delves into the topic of gait analysis and explores the major datasets that have been utilized in the existing literature. The present section presents a comprehensive discussion of our research findings and identifies several research directions on gait analysis.

2.5.1 Gait Applications

In all gait applications encountered in this SLR (see Pie charts 2.5 and 2.7), namely as phase detection, gait events detection, prediction of FOG, classification of healthy and pathological gait, future subsequence forecasting, and sensor-to-sensor sequence estimation, we note the increasing importance and strong potential of DL for prediction purposes, in particular in healthy gait, and classification tasks as well.

Specifically, CNN and LSTM networks have been identified as the major DL algorithms.

2.5.2 Multi-task learning

MTL has seen extensive use across a variety of scientific domains. Other gait-based applications, such as identification, authentication, and emotion recognition, which are out of the scope of this SLR, have also made use of this technology in the past.

All the reported studies using MTL for QGA concern classification tasks. The addition of MTL would be of great interest for regression purposes since auxiliary tasks usually improve the performance of the model for the principal tasks [129]. The MTL framework has two advantages: (a) it follows the divide and conquer principle and makes triggering the model easier; (b) it helps with model explainability. The price to pay is that the model is greedy, and the primary task's identification must be carried out very carefully.

2.5.3 Datasets

The majority of datasets that are publicly accessible are utilized for the classification of healthy and pathological gait. The availability of datasets for specific applications, such as age estimation and treatment outcome prediction, is limited.

2.5.4 Future Recommendations

We note that a significant amount of work on gait analysis utilizing DL has been done in recent years. After conducting an SLR, the following are some suggestions for the years to come:

- The utilization of MTL for gait analysis has been the subject of limited research studies. The utilization of the MTL approach has been proposed as a potential solution to address the challenge of gait analysis with increased complexity. This methodology has shown promise in its ability to effectively handle multiple tasks simultaneously in gait analysis, in particular because of the numerous variables to consider. By leveraging MTL, it may be possible to improve the accuracy and efficiency of gait prediction, ultimately leading to better outcomes for patients.

- The majority of gait analysis studies have utilized normal gait data for their analyses. However, there is a dire need for studies that examine pathological gait to better understand and address related issues.

- The scarcity of datasets about pathological gait is a significant constraint for research in the field of Clinical Gait Analysis. The production of a comprehensive public dataset covering pathological gait data is a challenge for researchers in this field. Such dataset would serve as a valuable resource for further research and analysis of gait-related disorders.

- The present study aims to address a gap in the existing literature by examining treatment outcome prediction. The ultimate purpose of CGA is to improve the patient's gait during rehabilitation. The treatment is a complex, personalized, data-driven decision process in personalized medicine. This decision relies on the experience of medical staff. In this context, AI can be an added value for decision aid.

  Specifically, it is noteworthy that only one reported study here has been conducted on this topic to date [39]. Therefore, the current investigation seeks to expand upon this limited body of research and contribute to a more comprehensive understanding of treatment outcome prediction. The current state of research in this area reveals a significant gap that requires further investigation.

- No publicly accessible dataset containing pre-treatment and post-treatment gait data has been identified. This lack of available data poses a significant challenge for researchers seeking to investigate the effects of treatment on gait patterns. Further exploration and potential collaboration with relevant institutions may be necessary to address this issue. The production of this particular dataset is also a crucial area of focus for researchers. Closer collaborations between clinicians, computer scientists, and engineers would help in this progress.
  Also, the use of synthetic pathological data using generative models could be a hint to overcome this limitation.

2.6 Conclusion

In the present work, we performed a comprehensive SLR and provided a broad overview of applications, DL methods, and datasets used in QGA. The findings of this review are divided into two major parts: classification and regression studies. Most of the classification studies focus on the classification of healthy and pathological gait, as well as gait event detection. On the other hand, the regression studies span future sub-sequence forecasting, sensor-to-sensor sequence estimation, condition/joint translations, and clinical scores prediction. CNN and LSTM have been mostly used as DL methods in these studies.

After literature screening, it has been depicted that more than half (around 55%) of current studies are mainly focusing on healthy gait. Most of the studies considering pathological gait use proprietary

datasets with a relatively small number of patients and a lack of external validation. Due to ethical constraints, most of these datasets are not publicly accessible. An effort for patient data sharing is to be made. Researchers can also work in the direction of synthetic gait data by using generative models.

In terms of applications, treatment outcome prediction applications, which are considered the "holy grail" in rehabilitation, are rare. In addition, whereas many QGA applications are based on classical Machine Learning algorithms [139, 56] considerably fewer AI-based QGA applications utilize DL approaches. Given the predictive potential of DL-based approaches, an effort should also be made towards studying treatment outcome prediction using DL. In this context, MTL approaches should also play a role in improving the performance of gait studies.

Chapter 3

Gait Analysis: from Aquisition to Data Analysis

Contents

3.1 Dataset Description 49
3.2 Kinematic Data 50
3.3 Medical Treatment Data 53

Summary

This chapter provides a comprehensive overview of the data used in this project and is divided into three main sections. The first section offers a detailed description of the data, including the acquisition process, protocols followed, the number of patients involved, and additional relevant information. The second section focuses on kinematic data, discussing its significance in this project and the preprocessing steps undertaken. The third section delves into the MTD given the pathological conditions of the patients involved, treatment-related information was incorporated into our models to enhance results. Detailed descriptions are provided in this section. While our research follows a specific direction, not all acquired laboratory data were utilized. This chapter includes all pertinent details.

3.1 Dataset Description

In this retrospective study, data was collected in the Movement Analysis Laboratory of the Rehabilitation Centre of UGECAM Coubert in France. In this laboratory, every single patient was an adult who was experiencing a variety of gait-related difficulties. Patients with central nervous system problems, such as CP, SCI, MS, stroke, or TBI, are included in this database. All of these patients had previously received treatment for spasticity using BTX-A injections and are followed-up at UGECAM Coubert. Patients who had an incomplete rehabilitation or who experienced complications throughout their recovery were excluded from the study. For patients who use various walking aids, trials with the same walking aid were selected for both the preoperative and postoperative phases of the study.

Concerning this study, every piece of information considered came from patients who took part in therapeutic activities. In total, $N_{\text{pat}} = 38$ patients at the start of the study were subjected to CGA both before and after receiving treatment for spasticity with botulinum toxin. The institution's research ethics committee approved utilizing these data after reviewing them. Although they were informed about the research, the patients did not object to the utilization of their data for research purposes.

The database is continuously expanding due to the new patients being brought into the UGECAM Laboratory regularly. For this reason the number of patients in some trials conducted during the three-year thesis research varied. The first experiments were conducted with 38 patients, as stated in Chapter 4. Later, the database was extended to 43 patients, corresponding to the most recent version of the database.

More precisely, in the first set of experiment (Experiment protocol 1 on 38 patients), a total of $N_{\text{limbs}} = 61$ lower limbs were treated, with $N_{\text{uni}} = 15$ patients (39.47%) being affected unilaterally (the right lower limb was impacted in 6 of them, and the left lower leg for the other 9). On the other hand, $N_{\text{bil}} = 23$ patients (60.53%) were affected bilaterally. The data sets include the CGA prior to treatment, specifics regarding medical treatment, and the CGA of patients following treatment. At the time of pre-treatment CGA, the average age of the patients was 46.67 years old, respectively. At the time of injection, the average age of the patients was 46.76 years old, and at the time of post-treatment CGA, the average age was 46.93 years old. The dataset contains individuals ranging in age from 21 to 75 years old. The comparison of the pre-treatment CGA to the post-treatment CGA took place roughly three months after the treatment. All patient information for the experiments is included in Table 3.1.

Table 3.1: Patient database description.

Total Patients	43
Age (Mean ± SD)	46.67 ± 13.43
Males/Females	26/17
Unilaterally/Bilaterally affected	19/24
Cerebral Palsy	4
Stroke	11
Multiple Sclerosis	12
Traumatic Brain Injury	3
Spinal Cord Injury	13

The number of gait cycles per patient when the database was extended to 43 patients can be seen in Figure 3.1. The minimum number of gait cycles for any patient was 12, and the maximum was 72. The average number of gait cycles per patient was 29. In the first experiment, we processed the full gait cycle, and in other experiments (Chapter 5), gait cycles were divided into two phases: the stance phase and the swing phase. Usually, a healthy person has 60% of the stance phase and 40% proportion of the swing phase. However, this work deals with pathological patients having different diseases, so we have different proportions for each phase. It can be seen in Figure 3.2.

3.2 Kinematic Data

Records of all CGA examinations carried out since 2016 can be found in the Movement Analysis Laboratory of the Rehabilitation Centre at the UGECAM. The acquisition of joint kinematics is commonly accomplished through the utilization of optoelectronic systems [8] or IMU systems [140]. In this work, kinematic data have been acquired using a Codamotion system that consists of four CX1 cameras at a frame rate of one hundred frames per second (fps) (see Figure 3.3). The participants walked straight, with or without technical aids (i.e., cane, rollator, tripod, etc.), through a 10-meter-long laboratory room. Patients walked back and forth throughout the gait hallway (trials). Each patient's multiple trials are recorded depending on the patient's capability. These data have been saved in a Matlab file.

Kinematic signals are not quite regular due to the variability of biomedical data [141]. This is the

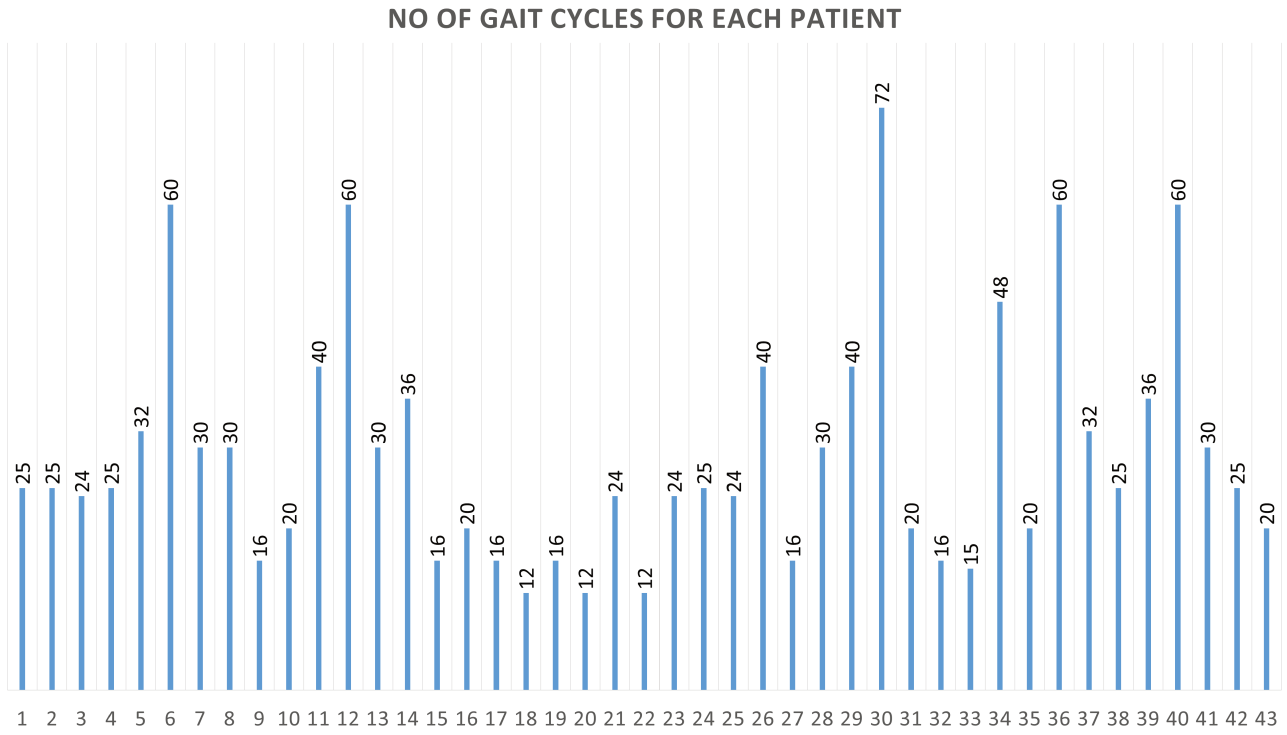


Figure 3.1: Gait cycles per patient. The X-axis represents patients, and the y-axis represents the total number of gait cycles.

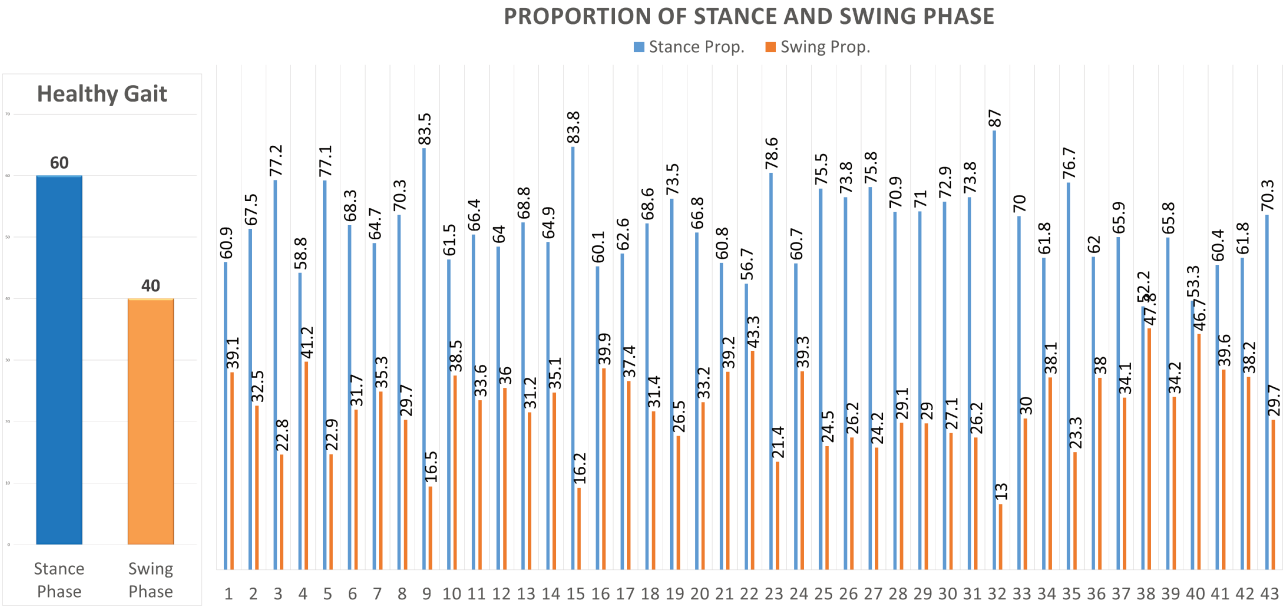


Figure 3.2: Proportion of stance and swing phase healthy gait (left side) and of each patient in our dataset (right side). The X-axis represents patients, and the Y-axis represents the proportion of each phase.

case even if walking is a cyclic process, not exactly periodic. Even when the same person walks, two separate gait cycles vary slightly in time and amplitude. This is true even if the individual walking is the same. A computation of an average gait cycle is performed for each lower limb in order to provide a unique representation. For this objective, kinematic signals are segmented into gait cycles, where a lower limb cycle is defined from two consecutive initial contacts. According to the findings of this

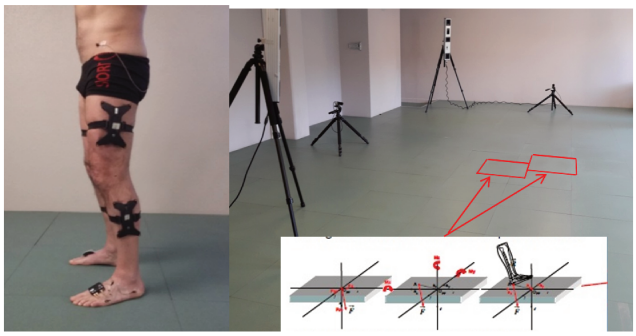


Figure 3.3: Clinical gait analysis. Different types of sensors are used to collect kinematic and kinetic data.

study [142], foot contacts were automatically identified by the utilization of the high pass algorithm (HPA). Following the localization of minima of the vertical component of the heel marker, HPA applies a high pass filter to the horizontal displacement of the forefoot marker and the heel marker in order to ascertain both initial contacts and toe-offs, which is the beginning of the swing phase.

The gait cycle is divided into two main phases: the stance and swing phases. In a healthy individual, the stance phase makes up around 60% of the gait cycle and can be further divided into initial contact (heel strike), loading response, mid-stance, terminal stance, and pre-swing. Both feet firmly touch the ground at the start and end of the stance phase. These double support periods make up around 10-12% of the gait cycle. The swing phase takes approximately 40% of the gait cycle. It is divided into three parts: the initial swing (toe-off), the mid-swing (tibia vertical), and the terminal swing, which concludes with the heel making contact with the ground. Figure 3.4 shows the average gait cycle of a healthy subject with five joints: knee, ankle, foot, pelvis, and hip on all three planes.

Our dataset includes kinematic signals from 5 joints: the pelvis, hip, knee, ankle, and foot. Each joint has three signals, resulting in a total of 15 signals. These three signals illustrate the trajectory of each joint on different planes: sagittal, frontal, and transverse (see Figures 3.5 and 3.6). For this study, we focused solely on the knee and ankle on the sagittal plane. This decision was made due to the fact that the majority of treatments are typically carried out concerning these particular joints.

Subjects in our dataset were affected on one side (unilaterally) or both sides (bilaterally). Patients are classified according to the limbs that are affected by a disease, and the categories are hemiplegia (the right or left limb is affected), paraplegia (both lower limbs are affected), and tetraplegia (both limbs are affected). Figure 3.7 shows kinematic signals of the knee and ankle of unilaterally affected patients (hemiplegia). Figures 3.8 and 3.9 show the kinematic signals of bilaterally affected patients (paraplegia and tetraplegia).

Following that, each gait cycle of selected joints was resampled to increments of 2% of the cycle, which resulted in 51 points to each cycle, as described in [143]. The interpolation technique was used for resampling processes [144]. The patient's data includes several gait cycles during the pre-treatment CGA and post-treatment CGA. Various trials were recorded for each individual. Multiple cycles of pre-treatment CGA were extracted during the trial.

Figure 3.10 shows the process for converting trials into normalized gait phases. The sagittal plane (flexion/extension) trajectories of the knee and ankle during a patient's whole trial, including numerous cycles, are depicted in Figures 3.10a and b, respectively. The cycles retrieved from the entire knee and ankle trials are shown in Figures 3.10c and d, respectively. Figures 3.10e and f show one cycle, and Figures 3.10(g-j) represent the segmentation of the gait cycle into stance phase and swing phases of knee and ankle joints.

We collected and saved the cycles of all patients. We analyzed a person's right and left cycles as separate samples in our data. We performed the identical procedure for the data on post-treatment CGA. Every pre-treatment cycle was linked to a desired post-treatment cycle.

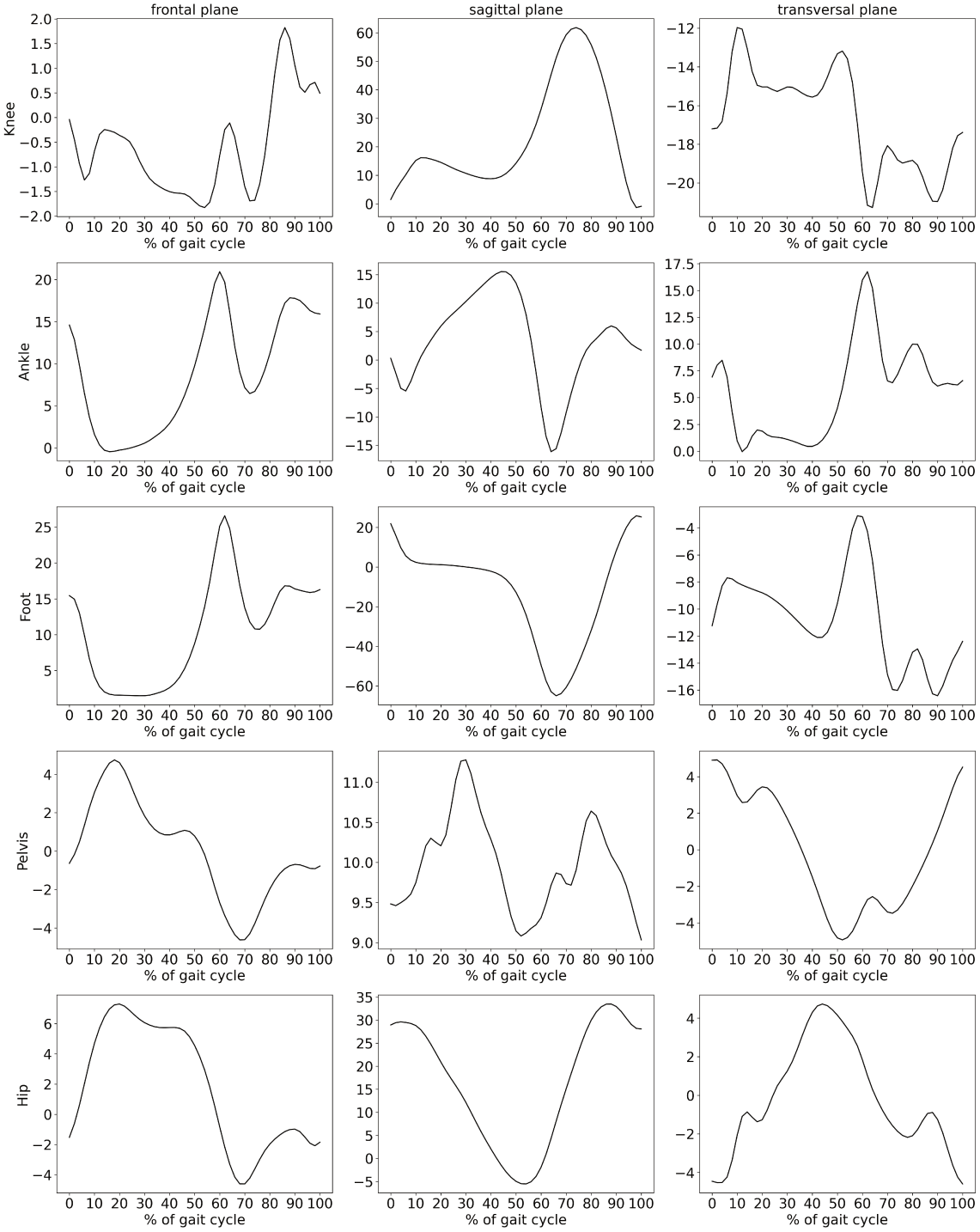
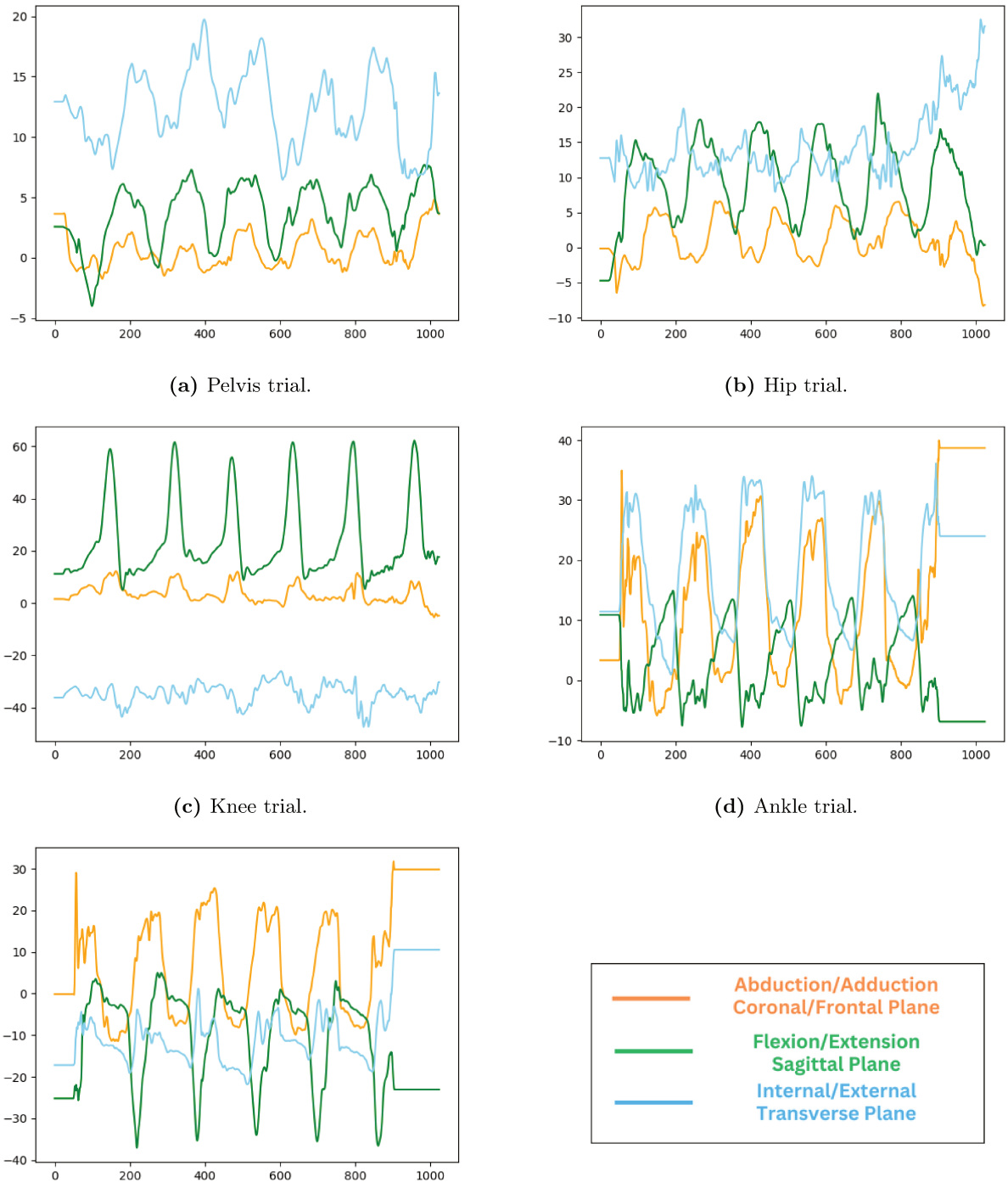


Figure 3.4: Gait kinematics of healthy subject (knee, ankle, foot, pelvis, and hip joints).

3.3 Medical Treatment Data

In this study, a total of 38 muscles have been considered for BTX-A injection. Each side has nineteen muscles (right and left lower limbs). Following is the list of muscles: Rectus Femoris, Vastus Lateralis, Vastus Medialis, Gracilis, Adductor Brevis, Adductor Longus, Adductor Magnus, Semitendinosus, Semimenbranosus, Biceps Femoris, Tibialis Anterior, Extensor Hallucis Longus, Tibialis Posterior,



(a) Pelvis trial.

(b) Hip trial.

(c) Knee trial.

(d) Ankle trial.

(e) Foot trial.

(f) Legends

Figure 3.5: Pre-treatment CGA (Kinematics) of the pelvis, hip, knee, ankle, and foot (random patient).

Gastrocnemius Lateral, Gastrocnemius Medial, Soleus, Peroneus Longus, Flexor Digitorum Longus, and Flexhallucis Longus. Three different molecules have been used for injections (Dysport, Xeomin, and Botox).

The most common method of treating spasticity is BTX-A through intramuscular injection. It has been demonstrated that BTX-A facilitates increases in the function of both the lower and upper



(a) Pelvis trial.

(b) Hip trial.

(c) Knee trial.

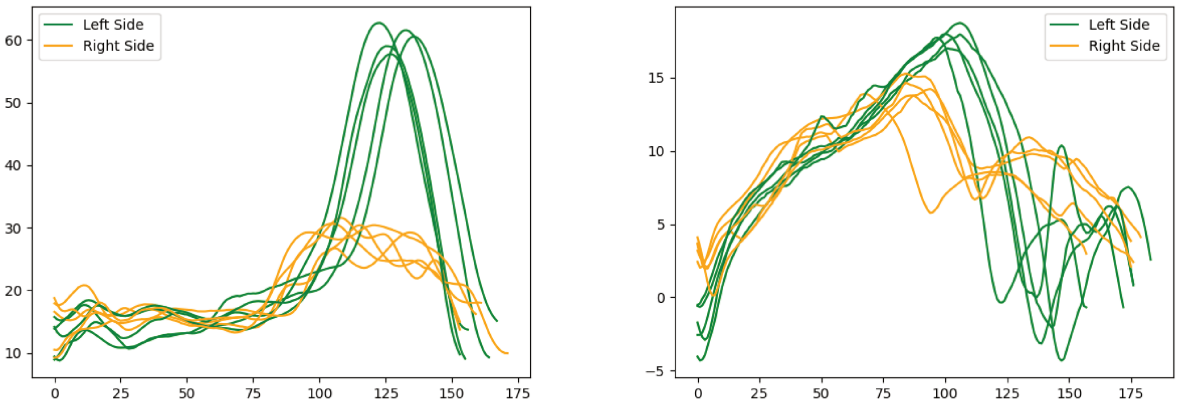
(d) Ankle trial.

(e) Foot trial.

(f) Legends

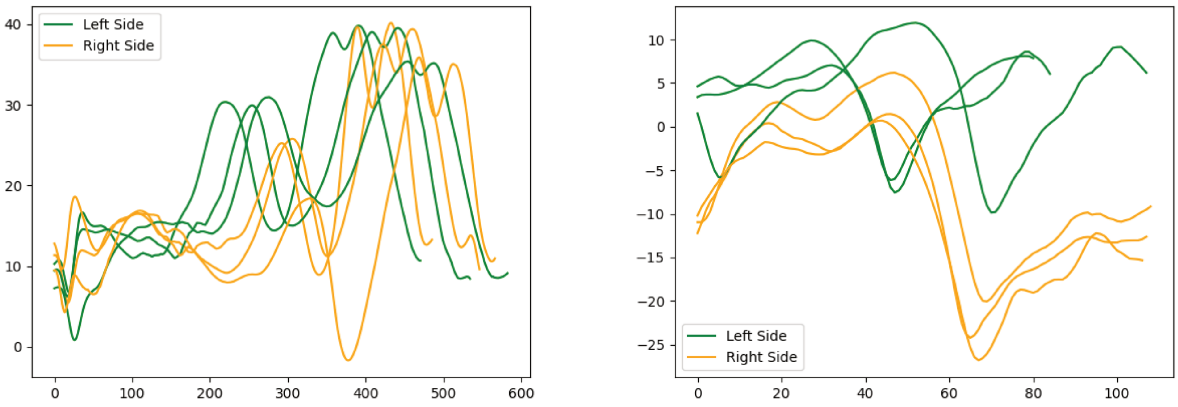
Figure 3.6: Post-treatment CGA (Kinematics) of the pelvis, hip, knee, ankle, and foot (random patient).

limbs [6], which in turn leads to enhancements in movement abilities such as walking [7]. There is a possibility that the minimum and maximum doses of BTX-A will change depending on the considered muscle [145]. Moreover, the total dose of BTX-A, which is the sum of doses for all of the muscles that are being treated, should not exceed the quantity that is advised for the patient and the muscles that are being considered (i.e., the upper limbs and the lower limbs separately).



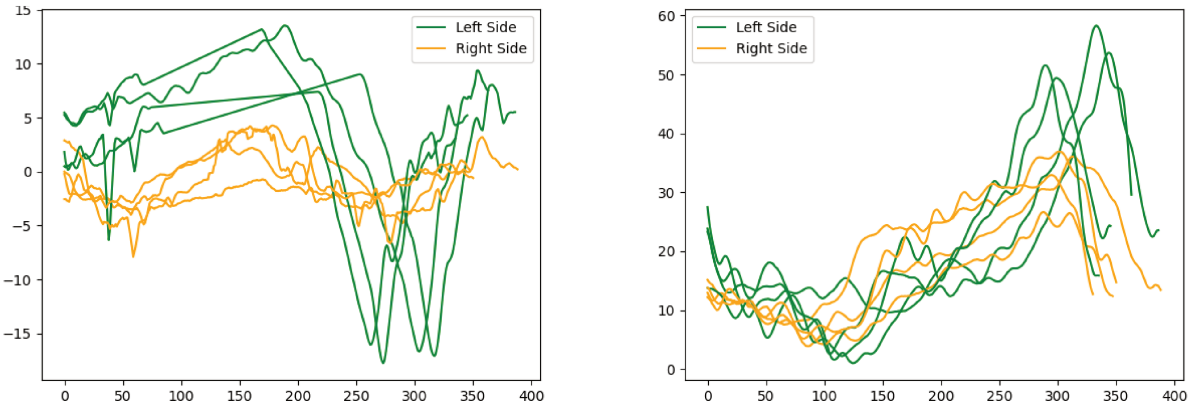
(a) Knee joint. **(b)** Ankle joint.

Figure 3.7: Knee and ankle kinematics of unilaterally affected patient (hemiplegic).



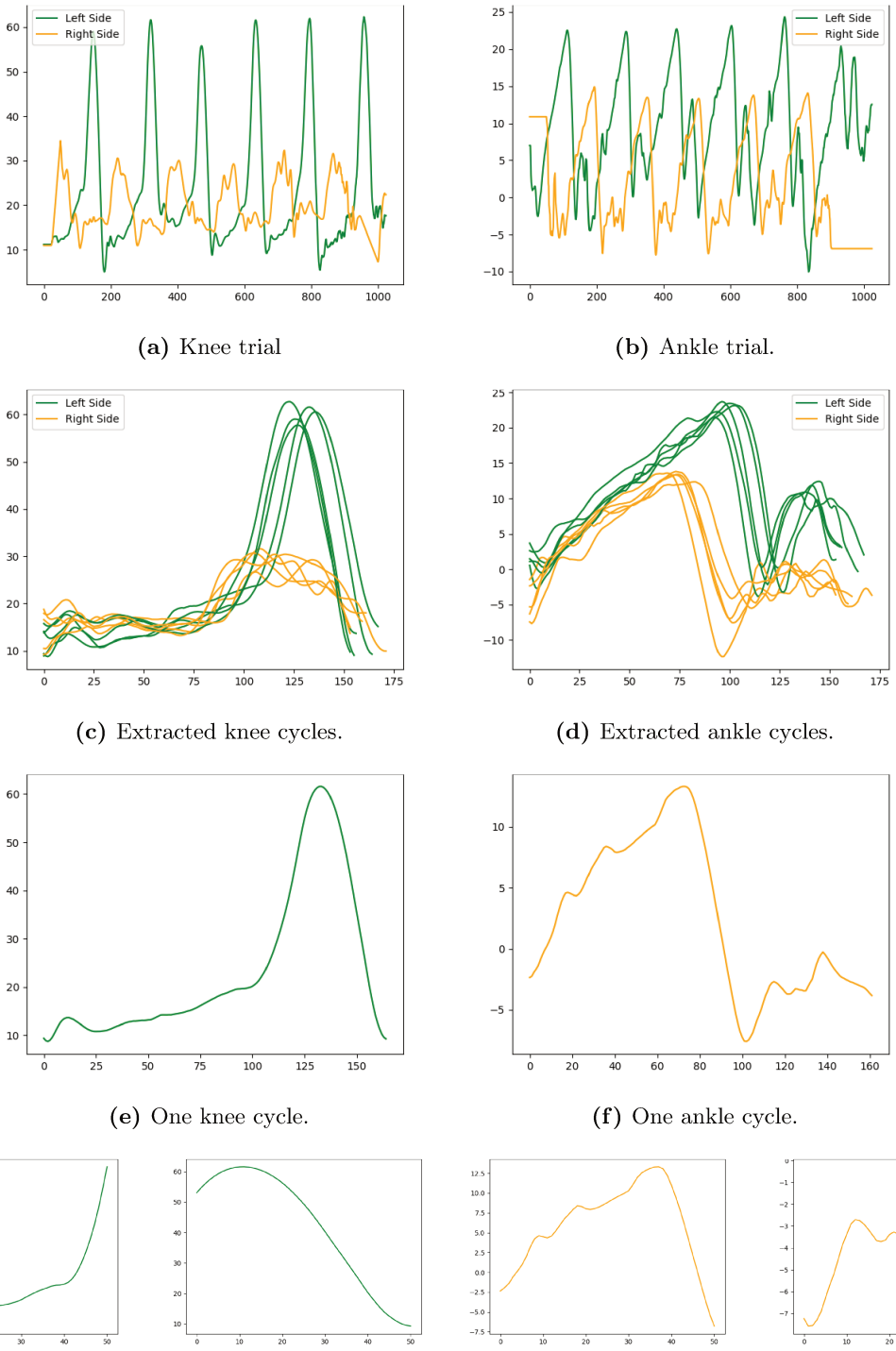
(a) Knee joint. **(b)** Ankle joint.

Figure 3.8: Knee and ankle kinematics of bilaterally affected patient (paraplegic).



(a) Knee joint. **(b)** Ankle joint.

Figure 3.9: Knee and ankle kinematics of bilaterally affected patient (tetraplegic).



(a) Knee trial

(b) Ankle trial.

(c) Extracted knee cycles.

(d) Extracted ankle cycles.

(e) One knee cycle.

(f) One ankle cycle.

(g) Stance phase of knee cycle.

(h) Swing phase of knee cycle.

(i) Stance phase of ankle cycle.

(j) Swing phase ankle cycle.

Figure 3.10: Process of converting patient's gait cycles into normalized gait phases (Stance and Swing).

BTX-A is a medicinal medicine that is very expensive, and the number of people who are using it has been growing over the past few years [146, 147]. Even though its effect on muscular function is thought to be reversible, the treatment with BTX-A is associated with hazards (that is, outcomes that are not desirable), and administration sessions ought to be spread out by at least three months.

Because of these factors, optimizing BTX-A treatment by selecting the appropriate muscles to be treated and the dose distribution is a complex undertaking that is of considerable importance and requires a comprehensive analysis of the patient's condition.

For this study, we examined the effects of injections into four different muscles: soleus, gastrocnemius (medial and lateral), semitendinosus, and rectus femoris. In addition, we established a separate category called "other muscles" to encompass all the remaining muscles that were included in our study (refer to Table 3.2). There were 24 different combinations of BTX-A injections of these four muscles.

A treatment binary code vector:

$$\mathbf{s}^j = (s_1^j, \ldots, s_c^j)^T, s_i^j \in \{0,1\}, i = 1 \ldots c (c = 5 \text{ as shown in Table 3.2})$$

was attributed to each lower limb i, with $s_i^j = 1$ if muscle i was injected in limb j, 0 otherwise, and

$$\mathbf{d}^j = (d_1^j, \ldots, d_5^j)^T, d_i^j \in \{0,1\}$$

is a binary vector for the disease of patient's limb j. There are five diseases: CP, MS, TBI, SCI, and stroke. ^{T} is the transpose operator.

Table 3.2: Considered injected muscles and their frequencies in the database (43 patients).

Muscle Number	**Muscle/Category**	**Injections limbs**	
		Number	**Proportion**
1	Soleus	53	28.3%
2	Gastrocnemius (Medialis and/or Lateralis)	51	27.2 %
3	Rectus Femoris	22	11.7%
4	Semitendinosus	14	7.4%
5	Other Muscle	47	25.1 %

Part II

Post-treatment Gait Prediction using Deep Learning

Chapter 4

Treatment Outcome Prediction Using Multi-Task Learning: Application to Botulinum Toxin in Gait Rehabilitation

Contents

4.1 Introduction 62
4.2 Data 62
4.3 Description of the models 63
4.3.1 Long Short-Term Memory 63
4.3.2 Bi-directional LSTM 65
4.3.3 Experimental Setup 67
4.4 Results 68
4.5 Discussion and Conclusion 70
4.5.1 Comparison to Previous Works 71
4.5.2 Limitations 72
4.5.3 Conclusion 72

Summary

This section presents an innovative predictive framework that utilizes MTL to enhance the accuracy of treatment outcome predictions for patients undergoing rehabilitation for spasticity—a common motor disorder often treated with Botulinum toxin type A (BTX-A). The main focus of the research in this section is on using a multi-task architecture made up of LSTM models, which are great for dealing with the fact that gait data in clinical settings is sequential and time-dependent. One of the key innovations of this study is the incorporation of MTD into the LSTM models, which includes detailed information about the types of treatments administered, specifically the muscles injected and the dosages used.

By integrating MTD, the study demonstrates that MTL models can significantly reduce prediction errors, as measured by the RMSE, for both knee and ankle kinematics. This was particularly notable in patients with TBI and CP, where the models achieved the lowest RMSE values, highlighting the potential of MTL to provide more accurate and clinically relevant predictions. The study's findings suggest that MTL, when paired with detailed treatment data, can outperform traditional approaches in predicting post-treatment trajectories, thus supporting better-informed clinical decisions and potentially improving rehabilitation strategies.

The content of this chapter is based on the following paper [39]:

- Khan, A.; Hazart, A.; Galarraga, O.; Garcia-Salicetti, S.; Vigneron, V. Treatment Outcome Prediction Using Multi-Task Learning: Application to Botulinum Toxin in Gait Rehabilitation. Sensors 2022, 22, 8452. https://doi.org/10.3390/s22218452

4.1 Introduction

Common motor impairments caused by neurological diseases like MS, TBI, SCI, and CP include fatigue, weakness, sensory loss, ataxia, and spasticity. Physicians often recommend rehabilitation as a supplement to pharmacologic treatment for individuals with impairments. Due to hyper-excitability, people with spasticity, a motor disorder related to upper motor neuron syndrome, have stronger tonic stretch reflexes and more pronounced tendon jerks. BTX-A intramuscular injections are a common spasticity treatment. Although BTX-A is a costly pharmaceutical product, its consumption has risen recently. Although its effect on muscle function is considered reversible, BTX-A treatment presents risks (i.e., undesirable effects), and injection sessions should be spaced at least 3 months apart. To optimize BTX-A treatment, it is crucial to carefully study the patient's condition and select the appropriate muscles and dose distribution.

AI and ML techniques have become commonplace in our daily lives, guiding decisions and offering recommendations. ML approaches are gaining popularity in precision medicine to meet the demand for new healthcare solutions, particularly in understanding pathological processes. Several CGA studies have used DL to predict gait trajectories, primarily for healthy gaits. For exoskeleton design, Su et al. [70] used LSTM to predict gait trajectories and phases (loading response, mid-stance, terminal stance, pre-swing, and swing). Zhu et al. [148] utilized an attention-based CNN-LSTM to predict knee and ankle joint trajectories in 60 milliseconds using lower and upper limb data. Zaroug et al. [149] developed an LSTM auto-encoder for predicting linear acceleration and angular velocity trajectories. Hernandez et al. [111] developed DeepConvLSTM, a hybrid network combining LSTM and CNN, to estimate kinematic trajectories. Jia et al. [150] developed a DNN for trajectory prediction using LSTM units and a feature fusion layer. Liu et al. [151] developed an LSTM-based model to predict two-time steps in the future using kinematic data from 35 subjects. Recent research by Kolaghassi et al. [103] examined the gait trajectories of children with neurological disorders. In this study, the complex issue involves the effects of multiple treatments (BTX-A) on gait trajectories.

Our contribution in this chapter consists of proposing a new solution to predict the BTX-A post-treatment gait trajectory of the patient and possibly the interaction between different treatments. This solution is an MTL architecture, alleviating the previously mentioned drawbacks: dataset size (number of patients), sample size (number of features), and feature diversity. To our knowledge, this is the first time MTL has been used for post-treatment gait trajectory prediction. This architecture comprises a collection of LSTM-shaped sub-models arranged in parallel or series. Each sub-model is used for one treatment, and each treatment corresponds to an injected muscle. These muscles are attached to the left and right knees and ankles. This MTL model learns to map pre-treatment gait sequences to post-treatment sequences. A gating mechanism with different architectures is proposed to control the influence of the treatments on the final prediction.

4.2 Data

In this section, we use the dataset already discussed in Section 3. The database comprises $N_{pat} = 38$ patients that underwent CGA before and after spasticity treatment with botulinum toxin. The details of the patients are listed in Table 3.1. In this work, we considered injections into four muscles: soleus, gastrocnemius (medial and lateral), semitendinosus, and rectus femoris. We also defined a fifth category called "other muscles", which groups all the other treated muscles (see Table 3.2).

The patient's data consists of multiple gait cycles at the time of pre-treatment CGA and post-treatment CGA. Different trials were recorded for each patient. In one trial, there were multiple cycles of pre-treatment CGA. Gait data were extracted in the following order (see Figure 4.1): (1) load Matlab files patient by patient, (2) extract joint angles data, (3) convert trials into cycles, and (4) normalize gait cycles. We extracted all the cycles of all patients and stored them. We separated a person's right and left cycles since we considered them different samples in the data. We performed the same procedure for post-treatment CGA data. Each pre-treatment cycle was associated with a target post-treatment cycle. Note that the number of cycles per patient varies from one patient to another.

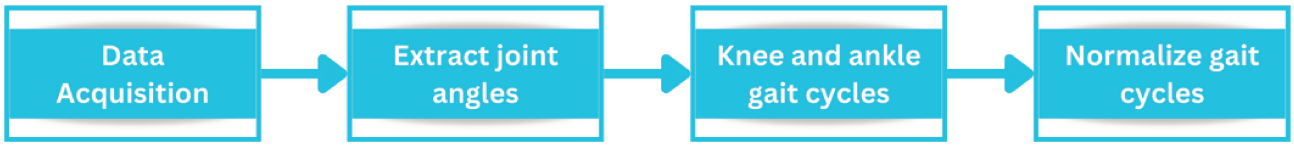


Figure 4.1: Flowchart of data pre-processing for experiment 1. We acquired trajectories of patient trials from the database and extracted all cycles from trials for knee and ankle joints. Finally, we applied normalization to the gait cycles.

There are 5 joints (pelvis, hip, knee, ankle, and foot) and three signals per joint in our dataset, leading to 15 signals. These three signals represent the projections of the trajectory of each joint, respectively, on the sagittal, frontal, and transverse planes. This study only considered the knee and ankle on the sagittal plane because most treatments were performed around these joints.

In this study, we applied different DL models to predict post-treatment CGA. As we know, DL models work on fixed-length sequences, so gait cycles have been normalized by 51 points each. By considering both the pre- and post-treatment cycles of each patient, a total of $n = 1,622$ gait strides were analyzed with 38 patients. We process the full gait cycle as input to the model. For any patient's limb j, the input vector is an angular time series $\mathbf{x}^j = (x_1^j, \ldots, x_m^j)^T \in [-180, +180]^m$, and the target vector is $\mathbf{t}^j = (t_1^j, \ldots, t_m^j)^T$, with $m = 51 \times 2 = 102$. Let $\mathcal{D} = \{\mathbf{x}^j, \mathbf{t}^j, \mathbf{d}^j, \mathbf{s}^j\}_{j=1}^n$ be the input–target training set.

In the end, our dataset contains $n = 1,622$ samples and 210 features; the first feature represents the ID (patient identifier); the second to 103rd are the features of the pre-treatment CGA; then, $c = 5$ features describe the presence or absence of botulinum toxin injection according to the muscle categories; finally, the last 102 features concern the post-treatment CGA of a patient.

An input matrix $\mathbf{x}$ and a target output matrix $\mathbf{y}$ were constructed using the parameters of n training samples, f features (the sagittal plane of the ankle and knee), l_{in} input size, and l_{out} output size. Pre- and post-treatment data were centered and reduced by the standard deviation. The goal was to construct a model with $g(.)$ that maps $\hat{\mathbf{y}} = g(\mathbf{x})$, where $\hat{\mathbf{y}}$ is a value that is very close to the actual value of $\mathbf{y}$.

4.3 Description of the models

4.3.1 Long Short-Term Memory

When training, early recurrent networks had difficulty remembering information for longer periods, such as several thousand-time steps. Hochreiter et al. [61] introduced a particular memory cell capable of retaining information for long periods of time. The LSTM can read and write to its memory. More importantly, this memory never goes through an activation function. This effectively combats the trailing gradient problem [152] and makes forming this pattern very stable. The original LSTM works with a series of input signals $\mathbf{x}_t$. It has a so-called hidden state $\mathbf{h}_t$ and a cell state $\mathbf{c}_t$ of the same size as $\mathbf{x}_t$. The cell state $\mathbf{c}_t$ is the model's memory. The hidden state $\mathbf{h}_t$ is the model's prediction of $\mathbf{x}_t$.

The LSTM equations are defined by the following set of matrix equations:

$$
\begin{aligned}
A &= \mathbf{h}_t \|^1 \mathbf{x}_t && (4.1)\\
ldf_t &= \sigma(W_f A + b_f) && (4.2)\\
\mathbf{i}_t &= \sigma(W_i A + b_i) && (4.3)\\
\mathbf{o}_t &= \sigma(W_O A + b_O) && (4.4)\\
\mathbf{d}_t &= \tanh(W_d A + b_d) && (4.5)\\
\mathbf{c}_{t+1} &= ldf_t \circ \mathbf{c}_t + \mathbf{i}_t \circ \mathbf{d}_t && (4.6)\\
\mathbf{h}_{t+1} &= \mathbf{o}_t \circ \tanh(\mathbf{c}_{t+1}) && (4.7)
\end{aligned}
$$

where $\|^1$ is the concatenation operator, $\circ$ is the Hadamard product, σ is the logistic function, W are weight matrices, and b are biases. The basic idea is that the model takes the input $\mathbf{x}_t$ and the previous prediction of the current input $\mathbf{h}_t$, updates its internal memory $\mathbf{c}_t$ to $\mathbf{c}_{t+1}$, and then makes, a new prediction $\mathbf{h}_{t+1}$ based on $\mathbf{c}_{t+1}$, $\mathbf{h}_t$, and $\mathbf{x}_t$.

The original LSTM could have multiple parallel memory cells $\mathbf{c}_t$, but in practice, mostly only one memory cell is used; the description of the LSTM was limited to one $\mathbf{c}_t$. All the gate functions (Equations (4.2)–(4.4)) are fully connected layers, $y = f(Wx + b)$ with a sigmoid activation function. The data flow in the LSTM is illustrated in Figure 4.2).

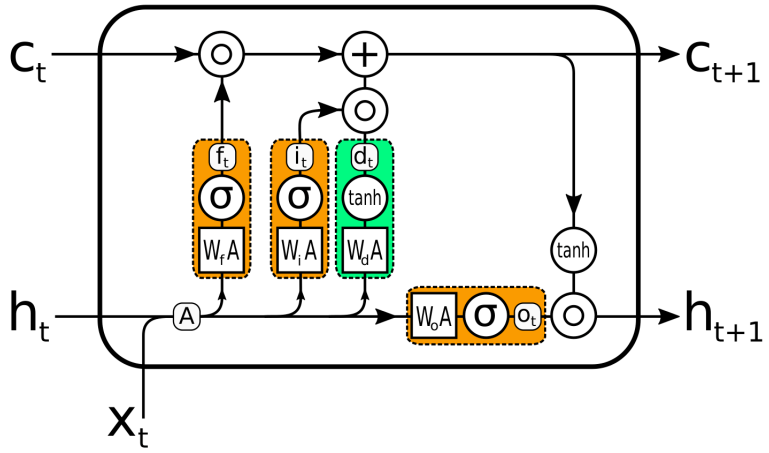


Figure 4.2: LSTM unit. The gates, which decide which part of the information to pass on, are orange. Green is the update to the memory cell.

Furthermore, the role of $\mathbf{h}_t$ is not strictly fixed to predict $\mathbf{x}_t$. It can be any prediction series connected to the input series $\mathbf{x}_t$. For example, if $\mathbf{x}_t$ was the number of people who entered (or left) a building in the last hour, then $\mathbf{h}_t$ could be the current number of people inside the building (with appropriate scaling, so it fits the output range $[-1, 1]$).

For this study, we used several variants of LSTM. The five categories of treatments are reported in Table 3.2: BTX-A injection of the first four muscles and the fifth category of injections in all other muscles. An LSTM layer represents each treatment. According to the DL architecture used, hidden states represent the presence or absence of treatments by BTX-A in the five muscles.

While the LSTM is well suited to prediction tasks of time series, knowledge about future events is sometimes necessary for the correct prediction. Therefore, the term "future" is relative to t and means the following data points. Of course, the next/future data points must already be known to be included in the prediction. Salehinejad et al. [153] identified two strategies to integrate the knowledge of future events into an LSTM model: bi-directional recurrent neural network (RNN) [154] and delayed input, the second approach consisting of delaying the signal by a delay τ:

- **Model 1 (control):** LSTM was used with pre-treatment CGA data and post-treatment CGA data. Treatments were not considered in this experiment. The model was implemented using five layers of LSTM units, with 51 units per layer, one unit for each point of a cycle. Note that each unit receives a pair of inputs for the knee and ankle, respectively. The final layer is fed into a dense layer of 102 neurons (2×51 values), which is then reshaped to obtain the desired output, as shown in Figure 4.3a. In this model, we initialized the values of the cell and hidden states to 0.
- **Model 2:** A total of five treatments and pre- and post-treatment CGA data were included in this model, displayed in Figure 4.3b. In this architecture, the values were initialized according to the medical treatment. Suppose one patient had muscle 1 and muscle 3 injected (Table 3.2). In that case, each layer's components of the hidden state vectors in LSTM layer 1 and LSTM layer 3 are initialized to 1, and the other layers' hidden state vectors is initialized as 0. In this model, we also initialized the cell state as 0.

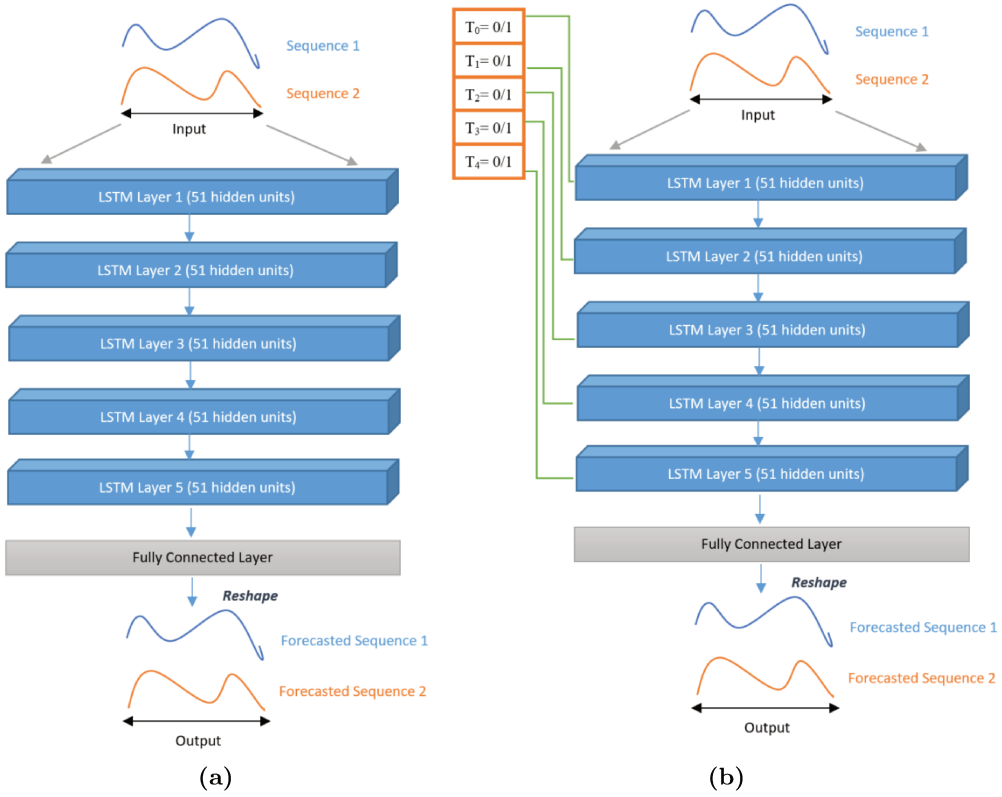


Figure 4.3: LSTM architectures (Model 1 and Model 2) proposed in this work: (**a**) without MTD; (**b**) with MTD.

4.3.2 Bi-directional LSTM

The entire signal must be known for this approach. Two LSTM models were trained in parallel, one on the input series (forward) and the other on the reverse input series (backward), starting with the last input and then the next-to-last, and so on. Thus, for each t, there were two hidden states: $h_{1,t}$ and $h_{2,t}$ among the two available models. $h_{1,t}$ only contains information about the past, and $h_{2,t}$ only contains information about the future. Together, they have information about the whole signal, and the final prediction $f(h_{1,t}, h_{2,t})$ was made using the two hidden states. This method has the disadvantage that two models must be trained; therefore, the number of parameters and the training time are doubled.

We studied the bi-directional LSTM (Bi-LSTM) architecture and considered two experiments, namely with and without MTD, as previously presented for LSTM:

- **Model 3:** A Bi-LSTM, as depicted in Figure 4.4. As shown in Figure4.4, the model has mainly the same structure as the previous Model 1 (same number of layers and units in each layer). The final layer's hidden state is fed into a fully connected layer. As in Model 1, we initialized the values of the cell state and hidden state of each layer to 0.

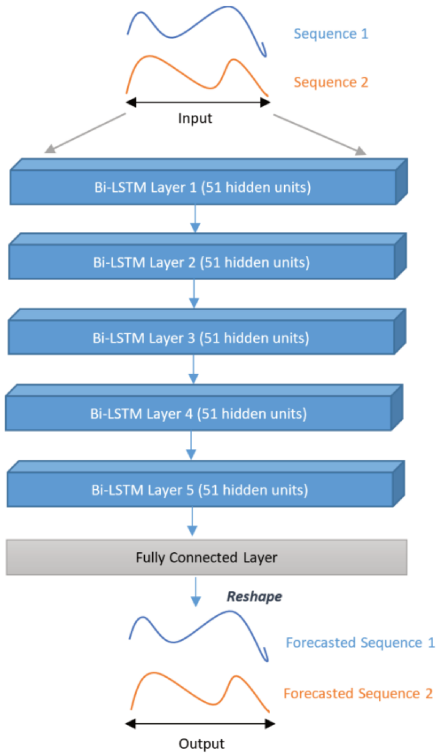


Figure 4.4: First Bi-LSTM architecture (Model 3) proposed in this work without considering MTD.

- **Model 4:** This model takes into account MTD in a multi-task architecture of Bi-LSTM models. Indeed, five Bi-LSTM models work in parallel while incorporating MTD as in Model 2. Each Bi-LSTM has 51 units, each receiving an input pair for the knee and ankle, respectively. Input X is fed to the five Bi-LSTM sub-models, and the cell state of all such sub-models was initialized to 0. Furthermore, the hidden states of all sub-models were initialized according to the presence or absence of MTD (as discussed in Model 2). This architecture has two fully connected layers: the first layer concatenates the outputs of all the sub-models; the second maps the output of the first layer to 102 neurons as per the desired output, as shown in Figure 4.5a.

- **Model 5:** This model is also a multi-task architecture of Bi-LSTM sub-models, as in Model 4. However, in this case, MTD is considered differently, with a gating mechanism. Instead of passing MTD as a hidden state of each Bi-LSTM sub-model, we incorporated them at the end of such sub-models by multiplying each sub-model's output by its corresponding binary value of MTD. In other words, if there is any treatment, it will be used further in the model; otherwise, it will be discarded (multiplying with 0), as illustrated in Figure 4.5b. By performing this experiment, we wanted to assess the impact of this gating mechanism compared to MTD internal processing by each sub-model, as done in Model 4.

- **Model 6 and 7:** In both models, we replaced the first fully connected layer (FC Layer 01) (see Figure 4.5a,b) with a convolutional layer (see Figure 4.5c,d), with kernel size (5,2) and

stride (3,2). As there are five Bi-LSTM sub-models and each has an output of size 2×102, we concatenated such outputs and reshaped them into a matrix of size (10 × 102), then given as the input to the convolutional layer. Finally, the convolutional layer's output is fed into a fully connected layer of size 102. Model 6 incorporates MTD, as in Model 4, through the internal states of the sub-models. Model 7 uses the same gating mechanism as in Model 5.

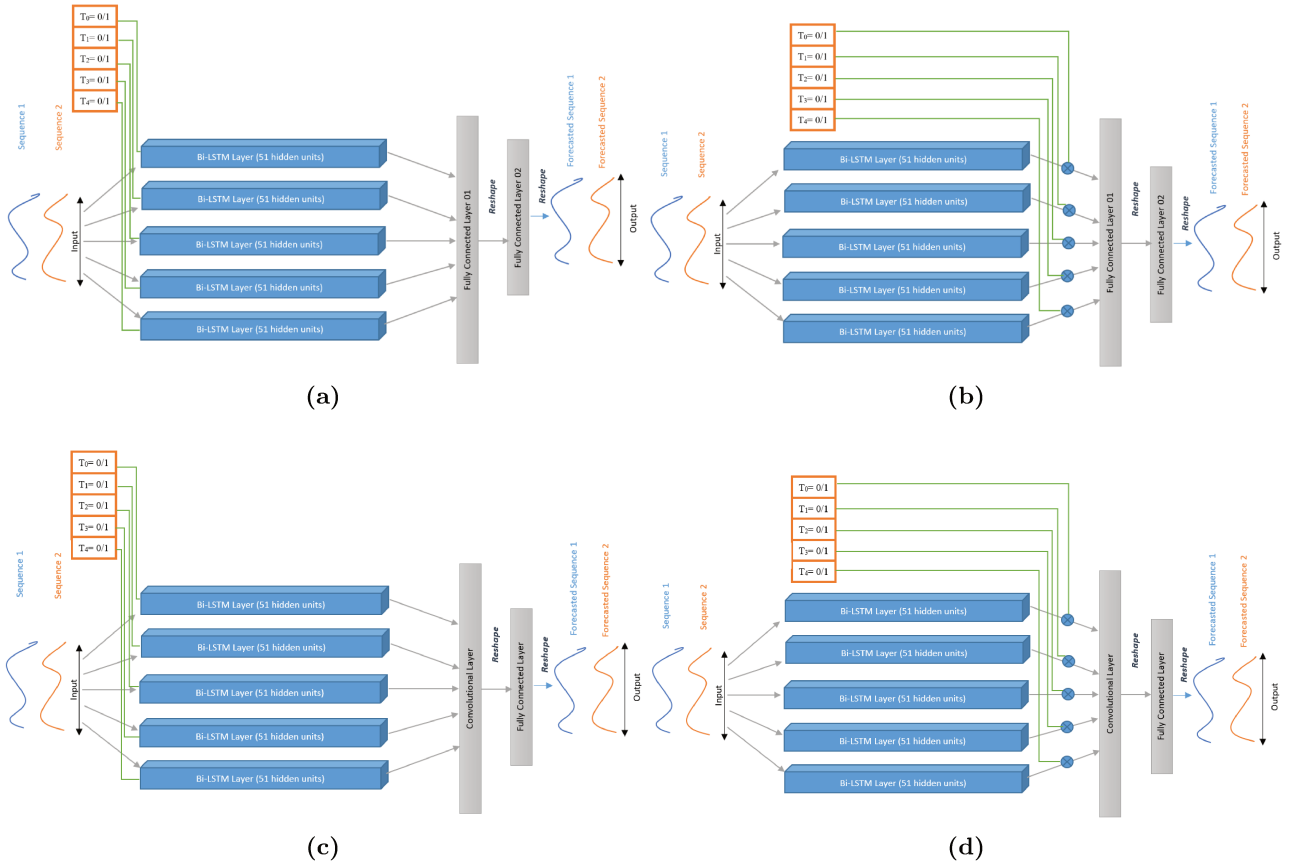


Figure 4.5: Multi-task learning architectures with Bi-LSTM sub-models; (**a**) Model 4: processing MTD internally in each sub-model; (**b**) Model 5: incorporating MTD through a gating mechanism; (**c**) Model 6: processing MTD internally in each sub-model using the Conv layer; (**d**) Model 7: incorporating MTD through a gating mechanism using the Conv layer

4.3.3 Experimental Setup

CGA data consists of 1622 combination pre-treatment and post-treatment gait cycles of 38 patients. Leave-one-out cross-validation was used to assess the models' performance. For each iteration, we used 37 patients to train the model and one to test. Ultimately, we took the RMSE of all tested patients for each model. The mini-batches were used throughout the training process, and the size of each batch was 16. We chose the RMSE as the loss function for optimizing the deep learning models and used the ADAM optimizer for learning. We tried different learning rates and selected the best possible values. We report in Table 4.1 all details concerning the models' hyper-parameters.

We calculated the RMSE to see how closely the predicted trajectories of the knee and ankle, $\hat{Y}$, matched the actual trajectories of the knee and ankle, Y. The following equation for the RMSE can be derived if we assume that n represents the number of testing samples, f represents the number of features, and l_{out} represents the output size.

Table 4.1: Hyper-parameter selection for LSTM, Bi-LSTM, and other architecture variants. The MTD column is used for medical treatment data (included/not included).

Model No.	Model Type	MTD	LSTM Layers (Units)	Conv. Layer	FC Layers (Units)	Learning Rate
Model 1	LSTM (serial)	No	5 layers (51)	None	1(102)	0.005
Model 2	LSTM (serial)	Yes	5 layers (51)	None	1(102)	0.005
Model 3	Bi-LSTM (serial)	No	5 layers (51)	None	1(102)	0.005
Model 4	MTL, 5 Bi-LSTMs	Yes	1 layer per sub-model (51)	None	2 (1020 & 102)	0.005
Model 5	MTL, 5 gated Bi-LSTMs	Yes	1 layer per sub-model (51)	None	2 (1020 & 102)	0.005
Model 6	MTL, 5 Bi-LSTMs + Conv Layer	Yes	1 layer per sub-model (51)	Kernel (5,2), stride (3,2)	1(102)	0.005
Model 7	MTL, 5 gated Bi-LSTM + Conv Layer	Yes	1 layer per sub-model (51)	Kernel (5,2), stride (3,2)	1(102)	0.001

$$RMSE = \sqrt{\frac{1}{nfl_{out}} \sum_{i=1}^{n} \sum_{j=1}^{f} \sum_{k=1}^{l_{out}} (y_{i,j,k} - \hat{y}_{i,j,k})^2} \tag{4.8}$$

We also calculated the standard error (SE) to measure the variation of the RMSE with respect to each disease and the R^2 score to check how well the data fit the regression model. SE is calculated using

$$SE = \frac{s}{\sqrt{n}}. \tag{4.9}$$

where s is the standard deviation of prediction with respect to a particular disease and n is the total number of patients having a particular disease. The coefficient of determination (R^2) as follows:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2}, \tag{4.10}$$

can be interpreted as the proportion of the variance in the dependent variable that is predictable from the independent variables (worst value $-\infty$, best value= $+1$), opposite the MSE, which magnifies the error if the model outputs a very bad prediction (worst value $+\infty$, best value= 0). We compared and evaluated the models' performance using these measures.

4.4 Results

We evaluated Models 1 to 7 on our dataset with the above-mentioned metrics and displayed the results in Table 4.2. The lowest average RMSE values and the highest R^2 scores are displayed in bold; they correspond to the best prediction model according to the diseases reported in Table 4.2. From Table 4.2, we noticed that Model 4 outperformed other models in predicting post-treatment gait trajectories for patients with MS and TBI. Furthermore, Model 6 performed better for SCI patients than all other architectures. Model 7 outperformed other models of patients having stroke and CP. We noticed that, in all cases, the MTL architectures achieved better performance globally on both knee and ankle signals.

Tables 4.3 and 4.4) report the performance scores of the prediction of gait trajectories for knee and ankle, respectively. In Table 4.3, the best prediction for the knee angle was obtained for TBI patients by Model 5 with an average RMSE of 5.60°and $R^2 = 0.72$. Furthermore, for all diseases, the MTL

architectures outperformed the others. Model 6 gave the best prediction for MS and SCI regarding RMSE and Model 4 gave the R^2 score for the same pathologies. On the other hand, for stroke patients, Model 7 had the best average RMSE, and Model 6 had the best coefficient of determination. We notice in Table 4.4 that the best RMSE for the ankle was 3.77°, which is lower than that obtained for the knee (5.60°). However, even though the RMSE was usually lower (thus better) for the ankle, the R^2 scores were also lower (thus worse). In particular, all the R^2 of the ankle angle were negative for stroke.

Table 4.2: Performance of different models in predicting post-treatment gait trajectories with respect to different diseases.

Model No.	Model Type	SCI	MS	Stroke	CP	TBI
		No. of Patients				
		11	**12**	**9**	**3**	**3**
		No. of Cycles				
		474	**530**	**322**	**148**	**148**
		RMSE Mean° ± Standard Error				
		R^2 Score				
Model 1	**LSTM (serial)**	6.82 ± 0.79	6.89 ± 1.18	8.11 ± 1.04	7.66 ± 1.09	5.87 ± 0.84
		0.53	0.54	0.53	0.61	0.63
Model 2	**LSTM (serial)**	6.71 ± 0.5	6.77 ± 0.97	8.03 ± 1.07	7.23 ± 1.0	7.63 ± 0.94
		0.46	0.52	0.65	0.58	0.64
Model 3	**Bi-LSTM (serial)**	6.9 ± 1.17	6.38 ± 1.17	7.06 ± 1.04	7.2 ± 1.07	7.78 ± 1.04
		0.60	0.65	0.65	0.68	0.51
Model 4	**MTL, 5 Bi-LSTMs**	6.26 ± 1.12	**5.8 ± 1.01**	6.99 ± 1.04	6.57 ± 1.24	**5.24 ± 0.77**
		0.66	**0.66**	0.66	0.67	**0.67**
Model 5	**MTL, 5 gated Bi-LSTMs**	6.67 ± 1.26	6.11 ± 1.43	7.73 ± 1.04	6.22 ± 1.71	6.07 ± 0.94
		0.59	0.55	**0.68**	0.64	0.56
Model 6	**MTL, 5 Bi-LSTMs + Conv Layer**	**5.75 ± 1.12**	6.08 ± 1.11	7.16 ± 1.12	6.2 ± 1.22	6.58 ± 0.57
		0.61	0.61	0.65	0.69	0.60
Model 7	**MTL, 5 gated Bi-LSTMs + Conv Layer**	6.31 ± 1.21	7.59 ± 1.46	**6.24 ± 1.0**	**6.00 ± 1.75**	7.02 ± 0.66
		0.52	0.52	0.60	**0.60**	0.42

Bold entries denote the lowest average RMSE and maximum R^2 over all limbs having a given disease.

From a different perspective, Figures 4.6 and 4.7 illustrate the pre-treatment, real post-treatment, predicted post-treatment of the patient, and standard course trajectories (healthy subject) of two patients. The Y-axis represents the ankle dorsiflexion or knee flexion, and the X-axis represents the gait cycle of a patient.

Figure 4.6 compares the prediction of different models on the knee and ankle joints of a patient diagnosed with CP. These figures differentiate the prediction between the MTL models and others. Figures 4.6a–c illustrate the predictions on the knee angles made by Model 1, Model 2, and Model 3, which are not MTL models. Figure 4.6d shows the corresponding prediction of Model 7, an MTL model. The predictions of post-treatment gait from Model 7 were better than others. In other words, it was closer than that patient's expected post-treatment gait trajectory (average of all his/her target gait cycles in the training set). On the other hand, Figure 4.6e-h compares the prediction of the ankle joint of the same patient. Figure 4.6g,h illustrates the prediction of Model 1 and Model 3, respectively. Figure 4.6g,h show the predictions of Model 4 and Model 7, respectively, which are MTL models. We noticed that the predicted post-treatment trajectory in Figure 4.6g was better than the first two models, which were serial, and we see in Figure 4.6h the significant improvement of the prediction at the end of the gait cycle, between 80% and 100%, compared to Figure 4.6g. On this patient, the MTL models also performed better on the ankle joint.

Figure 4.7 compares the trajectories of the knee and ankle joints of another patient diagnosed with MS. Figures 4.7a,b represent the predictions of the knee angles made by Model 1 and Model 2, which

Table 4.3: Performance of different models in predicting post-treatment knee gait with respect to different diseases.

Model No.	Model Type	SCI	MS	Stroke	CP	TBI
		No. of Patients				
		11	12	9	3	3
		No. of Cycles				
		474	530	322	148	148
		RMSE Mean° ± Standard Error R^2 Score				
Model 1	**LSTM (serial)**	7.73 ± 0.09	8.05 ± 0.11	8.62 ± 0.21	10.16 ± 0.13	6.66 ± 0.09
		0.67	0.65	−0.08	0.70	0.59
Model 2	**LSTM (serial)**	7.58 ± 0.08	8.26 ± 0.09	7.85 ± 0.17	8.56 ± 0.12	8.05 ± 0.27
		0.72	0.62	0.04	0.70	0.55
Model 3	**Bi-LSTM (serial)**	8.11 ± 0.13	7.41 ± 0.12	7.77 ± 0.21	7.42 ± 0.11	7.89 ± 0.28
		0.67	0.73	0.16	0.78	0.41
Model 4	**MTL, 5 Bi-LSTMs**	7.51 ± 0.08	7.23 ± 0.14	7.14 ± 0.018	**6.75 ± 0.11**	5.81 ± 0.13
		0.76	**0.77**	0.26	**0.80**	0.61
Model 5	**MTL, 5 gated Bi-LSTMs**	7.62 ± 0.10	7.23 ± 0.11	8.02 ± 0.25	7.00 ± 13	**5.60 ± 0.06**
		0.71	0.64	0.45	0.77	**0.72**
Model 6	**MTL, 5 Bi-LSTMs + Conv Layer**	**6.94 ± 0.09**	**6.78 ± 0.14**	7.19 ± 0.25	8.63 ± 0.18	8.24 ± 0.17
		0.69	0.70	**0.48**	0.79	0.62
Model 7	**MTL, 5 gated Bi-LSTMs + Conv Layer**	8.14 ± 0.12	8.52 ± 0.13	**6.21 ± 0.14**	7.82 ± 0.14	5.94 ± 0.07
		0.59	0.66	0.08	0.78	0.34

Bold entries denote the lowest average RMSE and maximum R^2 over all limbs having a given disease.

are not MTL models. Figures 4.7c,d represent the prediction of the knee angles made by Model 4 and Model 6, respectively, which are MTL models. We can see that MTL models had better predictions than the first two. The predicted post-treatment trajectories were closer to the real post-treatment trajectories. The last four Figures 4.7e–h compare the trajectories of the ankle joint. Figures 4.7e-g represent the predictions of Model 1, Model 2, and Model 3. Although Model 3 is not an MTL model, its predictions were much better than the first two serial models. However, the MTL model (Model 5) prediction in Figure 4.7h was better than all other models for this particular patient. In general, as shown by Tables 4.2-4.4, MTL performed better for almost every patient.

4.5 Discussion and Conclusion

In this study, we used MTL to design an LSTM model and its variants to predict the post-treatment trajectory of adults with an abnormal gait. To the best of our knowledge, this specific prediction task, which exhibits greater inter- and intra-subject variability compared to the courses of normal adults, has not been addressed before in the literature using MTL.

To forecast the trajectories of the knee and the ankle in the sagittal plane, we used LSTM-based models. LSTM was chosen because it has been successfully applied to sequential data, and it can capture long-term dependencies through its learning [59]. To better evaluate the performance of MTL on a given problem, we also implemented serial models using LSTM.

The RMSE was used to compare the results of both sorts of models. The RMSE of the MTL models was lower for all types of patients (different pathologies). The MTL models also gave the highest R^2, better explaining the total variance of the target than the serial models. The MTL models performed better than the serial models in our problem of multiple treatment combinations. MTL architectures allow the introduction of medical treatment metadata into the model. Instead of performing a simple post-pre regression task, our results imply that introducing the treatment information (i.e., muscles

Table 4.4: Performance of different models in predicting post-treatment ankle gait with respect to different diseases.

Model No.	Model Type	SCI	MS	Stroke	CP	TBI
				No. of Patients		
		11	**12**	**9**	**3**	**3**
				No. of Cycles		
		474	**530**	**322**	**148**	**148**
				RMSE Mean° ± Standard Error		
				R^2 Score		
Model 1	**LSTM (serial)**	5.91 ± 0.08	5.73 ± 0.08	7.61 ± 0.16	5.16 ± 0.14	5.09 ± 0.13
		0.52	0.50	**-3.75**	0.49	0.08
Model 2	**LSTM (serial)**	5.85 ± 0.07	5.29 ± 0.07	8.21 ± 0.21	5.89 ± 0.10	7.22 ± 0.36
		0.19	0.37	−4.48	0.37	0.15
Model 3	**Bi-LSTM (serial)**	5.69 ± 0.007	5.35 ± 0.07	6.34 ± 0.15	6.99 ± 0.09	5.66 ± 0.15
		0.40	0.68	−3.94	0.44	0.34
Model 4	**MTL, 5 Bi-LSTMs**	5.01 ± 0.08	**4.38 ± 0.08**	6.85 ± 0.19	6.4 ± 0.12	**4.68 ± 0.13**
		0.54	**0.69**	−4.13	0.37	**0.54**
Model 5	**MTL, 5 gated Bi-LSTMs**	4.56 ± 0.06	5.39 ± 0.10	7.14 ± 0.22	**3.77 ± 0.05**	4.93 ± 0.10
		0.44	0.47	−4.20	**0.50**	0.37
Model 6	**MTL, 5 Bi-LSTMs + Conv Layer**	5.72 ± 0.06	5.00 ± 0.06	7.44 ± 0.32	5.45 ± 0.14	6.54 ± 0.36
		0.46	0.63	−5.14	0.47	−0.03
Model 7	**MTL, 5 gated Bi-LSTMs + Conv Layer**	**4.49 ± 0.06**	6.66 ± 0.19	**6.26 ± 0.26**	4.17 ± 0.09	10.63 ± 0.26
		0.25	0.46	−4.85	0.23	−0.83

Bold entries denote the lowest average RMSE and maximum R^2 over all limbs having a given disease.

treated by BTX-A) improves performance.

Overall, the best prediction was obtained for TBI using the Bi-LSTM with MTL (Model 4) architecture. Table 4.2 shows only a 5.24°average difference in actual and predicted trajectories and $R^2 = 0.73$. The best maximum average RMSE error between actual and predicted trajectories was 6.24°for stroke patients, using the MTL architecture with gated Bi-LSTM and a convolutional layer (Model 7). For the knee and ankle, the best results were 6.75°($R^2 = 0.80$) and 3.77°($R^2 = 0.5$), respectively, for CP patients. The RMSE was usually higher for the knee than the ankle, despite having higher coefficients of determination. This suggests that the models could explain the variance of the knee angle better, but the amplitudes in the knee were higher than in the ankle. Moreover, no proposed model could adequately explain the variance of the ankle angle for patients with strokes (only negative R^2 scores).

4.5.1 Comparison to Previous Works

Since this is the first time that the whole kinematic signals for the knee and ankle on the sagittal plane were predicted for botulinum toxin treatment, it is difficult to compare our performance to other works. Nevertheless, we can compare our methods for the predictions of peak knee and ankle on sagittal planes reported by Roche. [7] for rectus femoris botulinum toxin injection of patients with stroke (Table 4.5). In this case, the R^2 score of the proposed method for stroke was better for peak knee flexion but worse for peak ankle dorsiflexion. Since the compared models were not trained and tested with the same databases, this comparison must be taken with caution.

We also compared our performances to the predictions of the whole postoperative kinematic curves for patients with CP. Even though the proposed methods were not tested on the same databases, these performances were better than the postoperative predictions for CP reported by Galarraga et al. [15], Niiler et al. [155], and Niiler [156], as shown in Table 4.5.

Table 4.5: Performance comparison of the prediction methods. LinReg and MLinReg correspond to linear regression and multiple LinReg, respectively, in [7, 15]. PCA stands for principal component analysis. NN99 and NN01 correspond to feedforward neural networks, respectively, in [155, 156].

Model	Knee Flexion	Ankle DorsiFlexion
	R^2 score	
Model 5 Stroke †	0.62	−4.69
Model 6 Stroke †	0.49	−4.24
LinReg Stroke † [7]	0.24	0.43
	Mean RMSE (°)	
Model 4 CP	6.8	6.4
Model 5 CP	7.0	3.8
PCA+MLinReg CP [15]	9.0	7.5
NN99 CP [155]	9.7	6.7
NN01 CP [156]	9.2	not reported

† Only peak flexion and not over the whole time series.

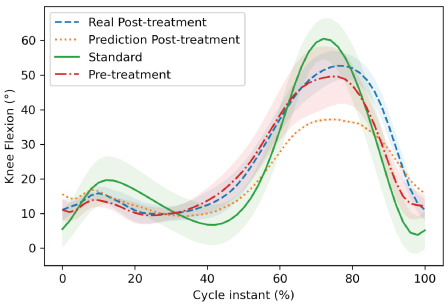
4.5.2 Limitations

Besides the lack of external validation, the main limitation of the proposed models is the relatively small size of the database. DL models usually need large amounts of data to be properly trained. Unfortunately, this is rarely the case in biomedical databases. Another limitation of the model is that it does not consider other aspects of the patient, such as psychological factors, age, stress, and social environment, which play a major role in the rehabilitation and, thus, in the treatment outcome.

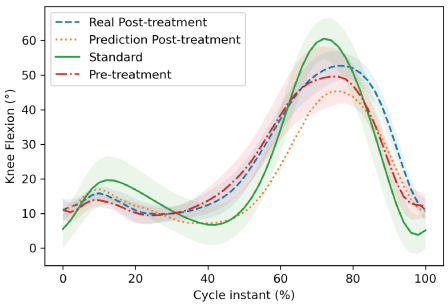
4.5.3 Conclusion

It was concluded from the results that the number of patients and type of disease did not directly affect the model's performance. More precisely, we can say that inter- and intra-subject variability affected the model's performance more than the number of patients (samples) and type of disease. Table 3.1 gives a detailed description of the number of patients with each disease. The minimum number of patients was 3 with CP and TBI diseases, while the maximum number of patients was 12 with MS disease. We noticed that the RMSE of CP and TBI patients were 6.00°and 5.24°, respectively. On the other hand, the RMSE of MS patients was 5.8°. This showed that having four times more patients for a given disease than others did not significantly affect the RMSE value.

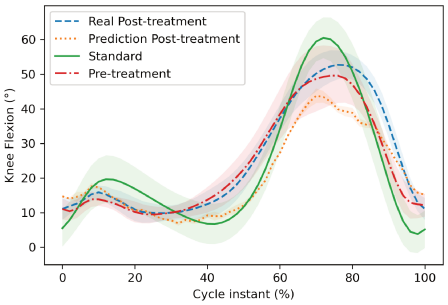
Finally, Bi-LSTM combined with MTL was highly effective at increasing the total quantity of information accessible to the model, enhancing the context provided to the algorithm. Future work will focus on MTL models with Bi-LSTM networks to exploit more precise information about treatments, such as the dose information, to further enhance the context given to the model.



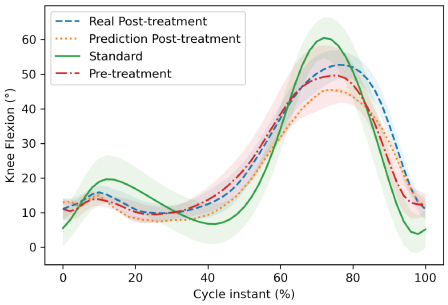
(a) Model 1: LSTM (serial)



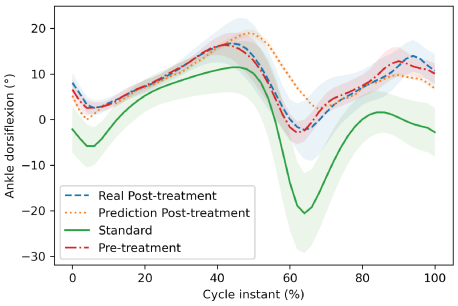
(b) Model 2: LSTM with MTD (serial)



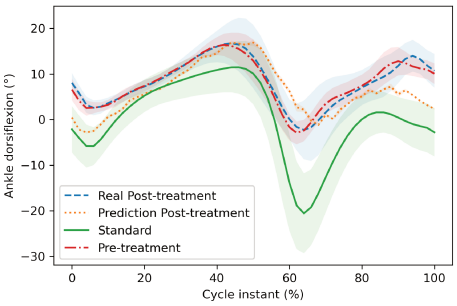
(c) Model 3: Bi-LSTM (serial)



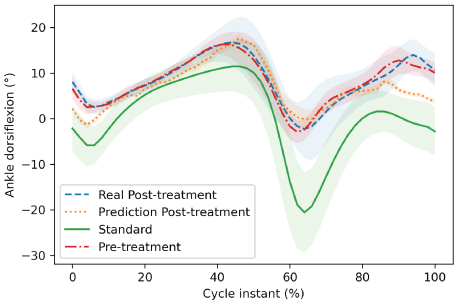
(d) Model 7: MTL, 5 gated Bi-LSTMs + Conv Layer



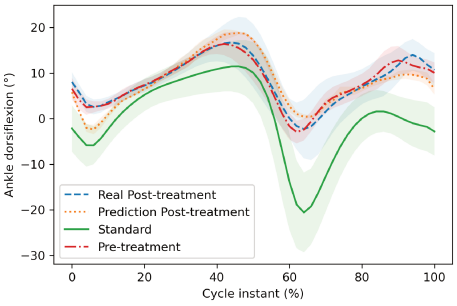
(e) Model 1: LSTM (serial)



(f) Model 2: Bi-LSTM (serial)

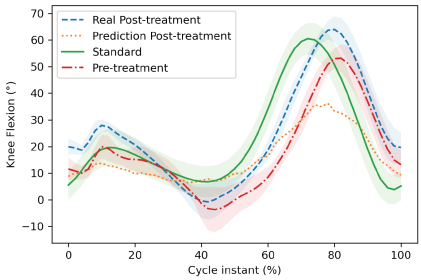


(g) Model 4: MTL, 5 Bi-LSTMs

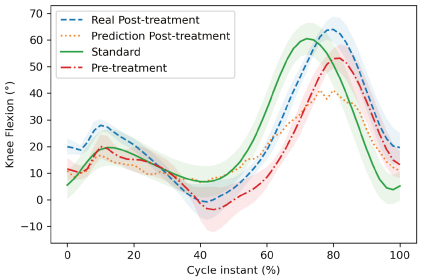


(h) Model 7: MTL, 5 gated Bi-LSTMs + Conv Layer

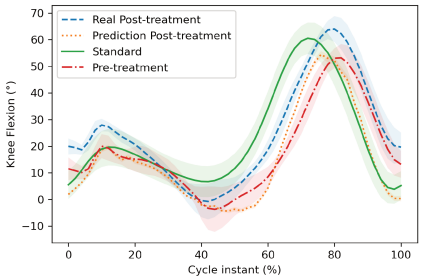
Figure 4.6: Comparison of the post-treatment gait trajectory of the knee and ankle joints in a patient diagnosed with CP.



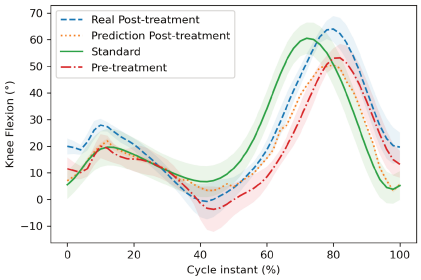
(**a**) Model 1: LSTM (Serial)



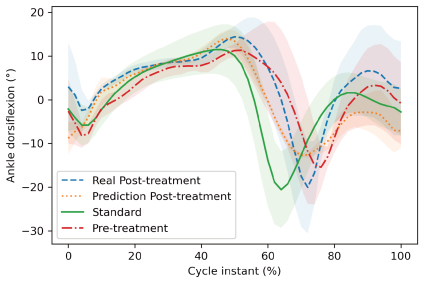
(**b**) Model 2: LSTM with MTD (Serial)



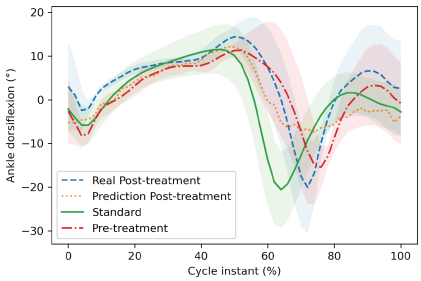
(**c**) Model 4: MTL, 5 Bi-LSTMs



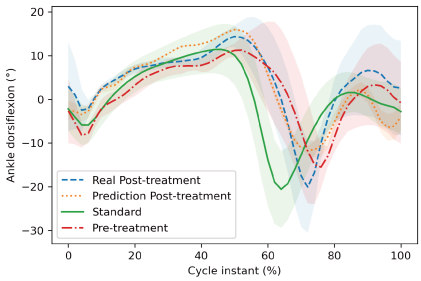
(**d**) Model 6: MTL, 5 Bi-LSTMs + Conv Layer



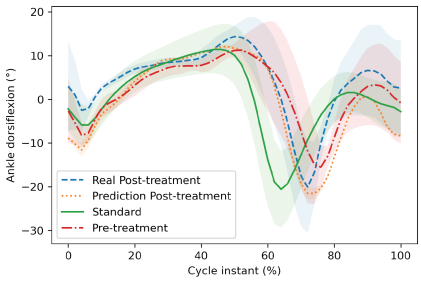
(**e**) Model 1: LSTM (Serial)



(**f**) Model 2: LSTM with MTD (Serial)



(**g**) Model 3: Bi-LSTM (Serial)



(**h**) Model 5: MTL, 5 Gated Bi-LSTMs

Figure 4.7: Comparison of the post-treatment gait trajectory of the knee and ankle joint in a patient diagnosed with MS.

Chapter 5

Phase-based Gait Prediction after Botulinum Toxin Treatment Using Deep Learning

Contents

5.1 Introduction 76
5.2 Data 76
5.3 Description of the Models 77
5.4 Results 78
5.4.1 Analysis on gait phases and whole cycle 78
5.4.2 Visualizing predictions of the best model (MTD-GM) 80
5.4.3 Comparison between cycle-based prediction, phased-based prediction, and related works 80
5.5 Discussion and Conclusion 81

Summary

In this section, based on the good results obtained in Chapter 4, we study splitting the learning process between stance and swing phases rather than processing the full gait cycle. The two best models from Chapter 4 were used to predict the stance and swing phases of the post-treatment gait cycle. Indeed, as we work on pathological gait, signals show more variance than in normal gait. This is especially the case in the swing phase. We observe that the stance phase lasts longer for patients than for healthy subjects given that the swing phase is more challenging for patients. This fact increases variance in the swing phase compared to the stance phase for pathological cycles. For this reason, in this chapter, we devote an MTL model to each phase separately. This proposal increases the capability of the global model to predict the complete cycle after treatment accurately. We perform this study based on the segmentation done on the pre-CGA data, then exploited on post-CGA. In other words, we combine both phases of each cycle using the proportion of stance and swing phases found pre-CGA. In this section, five more patients are added to the database ($N_{pat} = 43$).

The content of this chapter is based on the following paper:

- Khan, A.; Galarraga, O.; Garcia-Salicetti, S.; Vigneron, V. Phase-Based Gait Prediction after Botulinum Toxin Treatment Using Deep Learning. Sensors 2024, 24, 5343. https://doi.org/10.3390/s24165343.

5.1 Introduction

In this chapter, we split the learning process into stance and swing phases for to adapt to the parallel architectures proposed in [39]. Indeed, as we work on pathological gait, signals show more variance than in normal gait [157]. We note that the duration of the stance phase is often longer for patients than for healthy subjects, for example, in post-stroke Hemiparesis [158]. It is a fact that stance and swing phases obey different biomechanical constraints; therefore, we propose in this section to train separate models on each phase to enhance post-treatment CGA prediction. This strategy increases the global model's capability to predict the post-treatment full gait cycle accurately. The following section shows a significant improvement in prediction quality on pathological gait cycles.

5.2 Data

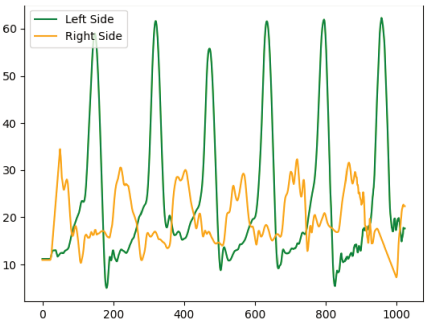
The $N_{\text{pat}} = 43$ patients who participated in the experiments performed in this chapter. Twenty-six of them were male, and seventeen were female. Their ages ranged from 21 to 75 years old. The time lag between pre- and post-treatment CGA was between 3 and 6 weeks. The lower limbs of individuals had undergone treatment, with $N_{\text{uni}} = 19$ patients (44.18%) were affected unilaterally (10 right limbs and 9 left limbs), and $N_{\text{bil}} = 24$ patients (55.82%) bilaterally. Four categories of injected muscles were considered: soleus, gastrocnemius (medial and lateral), semitendinosus, and rectus femoris. A fifth category called "other muscle(s)" groups all the other treated muscles [see Table 3.2].

The participants were recorded walking straight, with or without technical aids (i.e., cane, rollator, tripod, etc.), through a 10-meter-long laboratory room. Four CX1 motions tracked the coordinates of anatomical markers the patients wore in 3-D—capture units. Patients walked back and forth throughout the gait hallway (trials). Each patient's multiple trials are recorded depending on the patient's capability. 3-D gait kinematics were computed following the recommendations of the International Society of Biomechanics, based on the marker coordinate data. Each trial was divided into cycles, which were then divided into stance phases from initial contact to toe-off and swing phases from toe-off to subsequent initial contact [see Figure 5.2]. We found gait events, like first contacts and toe-offs, from force platform data and by using the HPA algorithm to extract them automatically [142]. A human expert validated and modified all gait events as necessary. The process of extracting cycles from trials to normalized phases is shown in Figure 5.1.

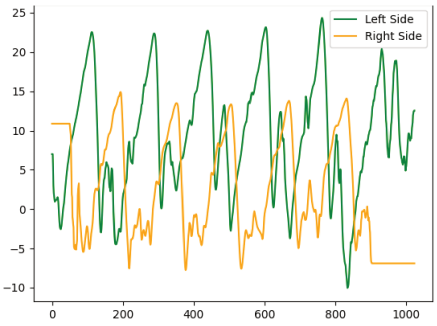


Figure 5.1: Flowchart of data pre-processing for experiment 2. We acquired trajectories of patient trials from the database and extracted all cycles from trials for knee and ankle joints. After that, we selected each cycle and segmented them into stance and swing phases. At last, we applied normalization to the phases.

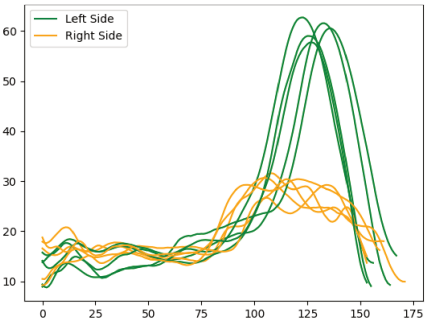
We consider a person's right and left cycles as different samples. Each pre-treatment cycle phase is associated with a target post-treatment cycle phase of the same patient and phase, leading to $n = 2518$ samples. Note that the number of cycles per patient varies from one patient to another. The data are centered and reduced by the standard deviation for normalization purposes. The goal is to use $g(\mathbf{x})$ to make a model that maps $\hat{\mathbf{y}} = g(\mathbf{x})$, where $\hat{y}$ is an estimation of $\mathbf{y}$. $\mathbf{x}$ is the input vector (an angular time series of pre-treatment gait kinematics) and $\mathbf{y}$ is the target vector (an angular time series of post-treatment gait kinematics) with size $m = 51 \times 2 = 102$.



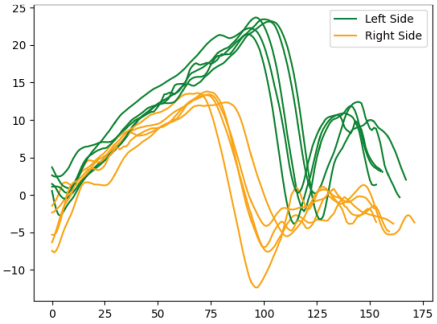
(a) Ankle trial



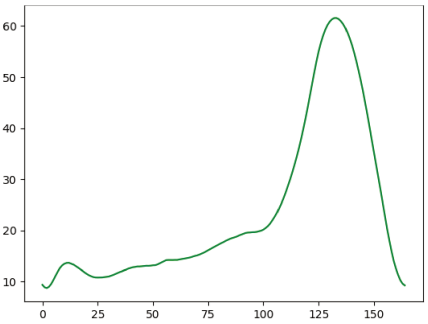
(b) Knee trial.



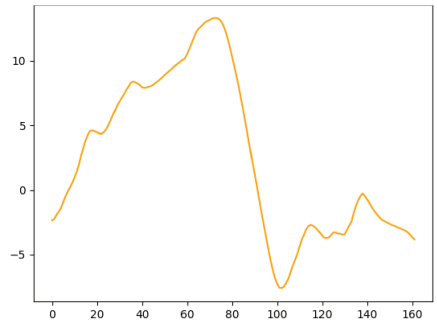
(c) Extracted knee cycles.



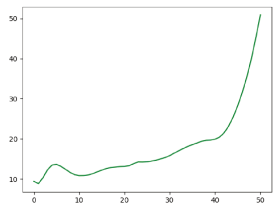
(d) Extracted ankle cycles.



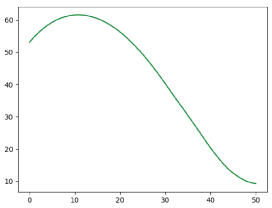
(e) One knee cycle.



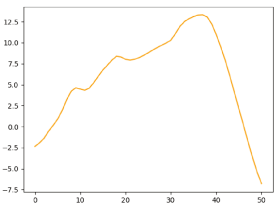
(f) One ankle cycle.



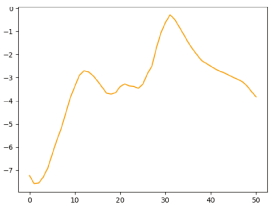
(g) Stance phase of knee cycle.



(h) Swing phase of knee cycle.



(i) Stance phase of ankle cycle.



(j) Swing phase ankle cycle.

Figure 5.2: Process of converting patient's gait cycles into normalised gait phases (Stance and Swing).

5.3 Description of the Models

In this section, the best models from our previous chapter 4 were used based on a bidirectional LSTM (Bi-LSTM), a variant of LSTM. In both models [see Figure 5.3], we use five parallel layers of Bi-LSTM.

Each layer is responsible for treatment according to the five categories of injected muscles, as reported in Table 3.2. Each Bi-LSTM layer has 51 units. Note that each unit gets a pair of inputs for the knee and ankle, respectively. We train separate models for stance and swing phases. For comparison with our previous experiment, we combined both predictions using the proportion of stance and swing phases of the pre-treatment gait cycle. Thereby, we retrieve the prediction for a full cycle of the knee and ankle joints.

MTD-Driven Model (MTD-DM) : The input vectors $\mathbf{x}$ and $\mathbf{s}$ are sent to the five Bi-LSTM sub-models in this architecture. The pre-treatment knee and ankle kinematic signals are represented by vector $\mathbf{x}$. The Medical Treatment Data (MTD) are represented by vector $\mathbf{s}$. We initialize the cell states of the LSTM as 0. The MTD (vector $\mathbf{s}$) is handled by the treatment supervisor and was used to set the values of the hidden states h. For example, if a patient had injections in muscles 1 and 3 (Table 7.1), then the states $\mathbf{h}_{1,t}$, $\mathbf{h}_{2,t}$ in Bi-LSTM, layers 1 and 3 are set to 1, and the other layers' (2, 4, and 5) hidden states are set to 0. The results of the five Bi-LSTM sub-models are concatenated to form a single tensor 'output,' a one-dimensional vector, and reshaped. The reshaped output is then processed through two subsequent fully connected layers (FC1 and FC2) to predict the post-treatment kinematics (see Figure 5.3a).

MTD-Gated Model (MTD-GM) : This architecture uses a gating mechanism to handle MTD. Instead of passing the MTD as a hidden state of each sub-model, the treatment supervisor is exploited downstream, multiplying each sub-model's output by the associated binary value of the MTD. Figure 5.3b shows that if there is a treatment, it will be used in the model, but if there is none, it will be neglected (multiplied by 0). The results of the remaining Bi-LSTM sub-models are concatenated and reshaped. The reshaped output is then processed through two subsequent fully connected layers (FC1 and FC2) to predict the post-treatment kinematics (see Figure 5.3b).

Experimental setup

Leave-one-out cross-validation is used to assess model performance. For each iteration, $N_{train} = N_{\text{pat}} - 1$ patients were used to train the model and one to test it. Mini-batches with $bs = 16$ samples were used during the training process. The ADAM optimizer [159] controlled the MSE loss function for training DL models. RMSE [160], Standard Error (SE) [161], and coefficient of determination (R^2) [162] are used to evaluate the performance of the proposed models.

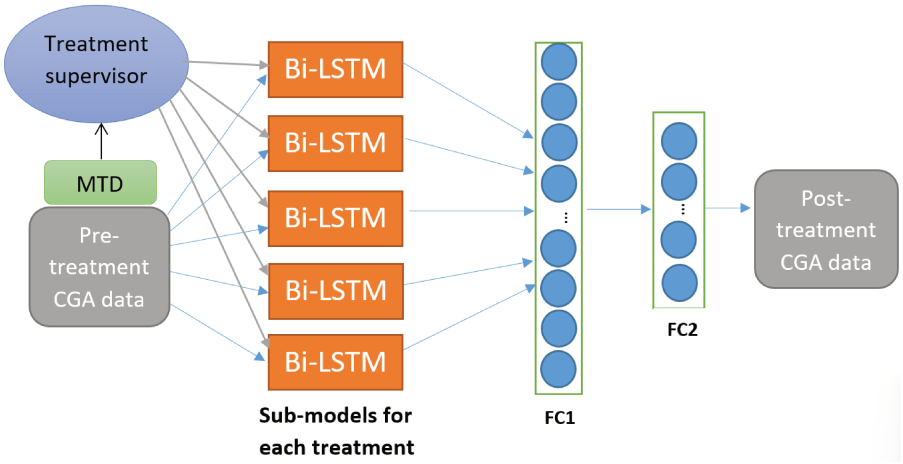
Table 5.1: Hyper-parameter selection for for both models.

Model No.	Model Type	LSTM Layers (Units)	FC Layers (Units)	Learning Rate
Model 1 (MTD-DM)	**MTL, 5 Bi-LSTMs**	1 layer per sub-model (51)	2 (1020 & 102)	0.005
Model 2 (MTD-GM)	**MTL, 5 gated Bi-LSTMs**	1 layer per sub-model (51)	2 (1020 & 102)	0.005

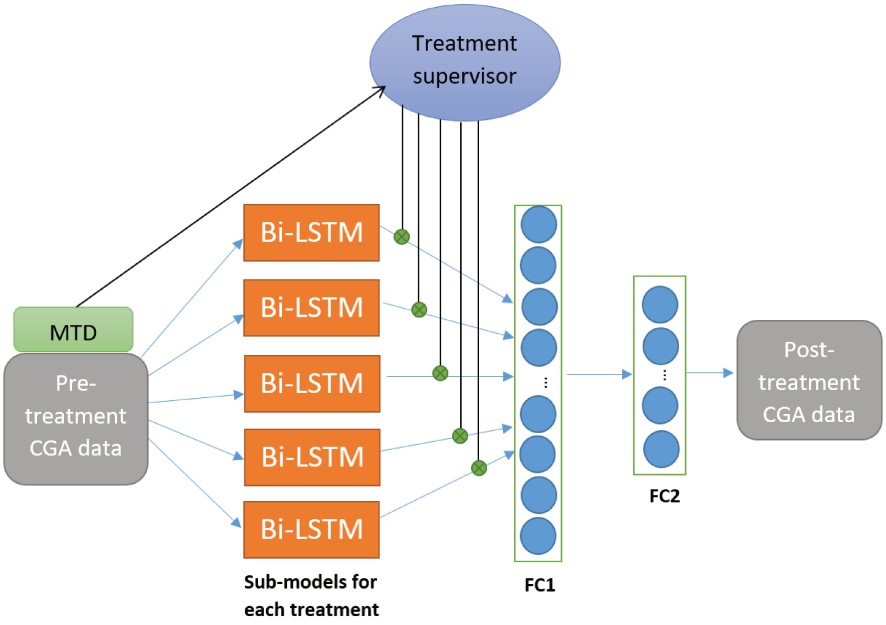
5.4 Results

5.4.1 Analysis on gait phases and whole cycle

43 participants are included in this study, of whom 38 were already included in our previous experiment (Chapter 4), and 5 new ones are added. We evaluate the two DL models on the new dataset using the above-mentioned evaluation metrics. Tables 5.2, 5.3, 5.4, 5.5 report the performance of both models in the prediction of post-treatment gait kinematics of the stance phase, swing phase, and complete gait cycle with respect to the disease: the bold entries in the following tables represent the best predictions (lowest $RMSE$ and highest R^2). We notice the performances of both models are equivalent in the stance phase for the knee and ankle joints (average $RMSE$ between 1.79°and 3.17°). On the contrary,



(a) MTD-DM



(b) MTD-GM

Figure 5.3: Deep Learning architectures with Bi-LSTM sub-models. (a) Medical Treatment Data-Driven Model (MTD-DM), and (b) Medical Treatment Data Gated Model (MTD-GM).

for the swing phase, MTD-GM is much better than MTD-DM, decreasing the RMSE from 3.08°and 7.01°(knee) for MTD-DM to 2.43°and 3.89°(ankle) for MTD-GM. With the MTD-GM, in the swing phase, the R^2 score is better for the knee than for the ankle. It is the opposite for the ankle for the stance phase.

As reported in Table 5.4, we combine stance and swing phase predictions to obtain the prediction for each complete cycle. We note that the MTD-GM outperforms MTD-DM, showing much better R^2 scores for the knee (higher than 0.8) and ankle (higher than 0.7). Table 5.5 reports the overall prediction results on both joints for MTD-DM and MTD-GM. We observe that our results improve the R^2 scores, which become higher than 0.9 in all cases. The average $RMSE$ for the stance phase of the knee, stance phase of the ankle, swing phase of the knee, swing phase of the ankle, complete knee, and complete ankle are 2.61°, 1.98°, 3.84°, 2.43°, 2.99°, and 2.21°, respectively.

Table 5.2: Performance of both models in predicting post-treatment gait trajectories of stance phase with respect to different diseases.

Model	Model Type	Body Joint	Spinal Cord Injury (SCI)	Multiple Sclerosis (MS)	Stroke	Cerebral Palsy (CP)	Traumatic Brain Injury (TBI)
			No. of Patients				
			13	**12**	**11**	**4**	**3**
			No. of Cycles				
			400	**351**	**260**	**138**	**74**
			RMSE Mean (°) ± Standard Error				
			R^2 Score				
MTD-DM	**MTL, 5 Bi-LSTMs**	Knee	2.6 ± 0.76 0.54	2.48 ± 0.73 0.79	**2.65 ± 0.78** **0.64**	**2.62 ± 1.08** **0.62**	3.17 ± 1.23 0.18
		Ankle	2.09 ± 0.79 0.87	1.83 ± 0.64 0.93	2.02 ± 0.5 0.67	1.99 ± 0.8 0.85	2.44 ± 0.78 0.84
MTD-GM	**MTL, 5 gated Bi-LSTMs**	Knee	**2.59 ± 0.84** **0.52**	**2.48 ± 0.7** **0.80**	2.7 ± 0.78 0.62	2.65 ± 1.08 0.46	**2.89 ± 1.03** **0.32**
		Ankle	**2.09 ± 0.75** **0.85**	**1.79 ± 0.61** **0.93**	**2.01 ± 0.49** **0.68**	**1.96 ± 0.82** **0.85**	**2.25 ± 0.75** **0.86**

5.4.2 Visualizing predictions of the best model (MTD-GM)

Figures 5.4 represent the average predictions on all cycles of five patients, each suffering from different diseases. The ankle dorsiflexion or knee flexion is shown on the Y-axis, and the patient's gait cycle is shown on the X-axis. In all these graphs, we observe that each patient's average predicted gait trajectory is very close to the average actual trajectory. Hence, it shows that our model performs well in predicting the post-treatment gait trajectory of disabled patients, as reported in Table 5.4.

However, in a few patients, post-treatment CGA kinematics are very difficult to predict. We report in Figure 5.5 that the post-treatment cycles for these patients show more variability between cycles than before treatment. This is not the case for most of the other patients. We investigate this phenomenon in the following. Figure 5.5a shows the trajectories of the right limb of a patient with SCI. The knee kinematics of this patient show more inter-cycle variability than most patients. The MTD-GM model predicted this patient's post-treatment kinematics with an average $RMSE$ of 6.06°and R^2 of -0.16 for the knee joint. This is worse than the average $RMSE$ of 3.05°. In Figure 5.5b, we notice that the prediction error for the stance phase is higher than for the swing phase. Figures 5.5c show the trajectories of a patient with CP. Both pre- and post-treatment gait kinematics of this patient show high variance between cycles. The prediction of this patient shows an average $RMSE$ of 3.02°and an R^2 score of 0.27 for the knee joint. It is difficult for our DL models to predict the post-treatment gait trajectories, which show so much variance.

5.4.3 Comparison between cycle-based prediction, phased-based prediction, and related works

It is difficult to compare the performance to other research because this is the first time that the entire knee and ankle kinematic signals on the sagittal plane have been anticipated for BTX-A treatment, except for our own work, which was reported in the last chapter. In this chapter, we worked on kinematic prediction per phase (stance and swing) of gait cycles, while in Chapter 4, complete cycles

Table 5.3: Performance of both models in predicting post-treatment gait trajectories of swing phase with respect to different diseases.

Model	Model Type	Body Joint	Spinal Cord Injury (SCI)	Multiple Sclerosis (MS)	Stroke	Cerebral Palsy (CP)	Traumatic Brain Injury (TBI)
			No. of Patients				
			13	**12**	**11**	**4**	**3**
			No. of Cycles				
			400	**351**	**260**	**138**	**74**
			RMSE Mean (°) ± Standard Error **R^2 Score**				
MTD-DM	**MTL, 5 Bi-LSTMs**	Knee	6.23 ± 2.05 0.31	7.01 ± 2.2 0.59	7.0 ± 1.73 -1.56	6.72 ± 2.29 -3.86	6.8 ± 1.17 0.26
		Ankle	5.15 ± 2.05 -3.35	5.15 ± 2.2 -1.18	4.36 ± 1.73 -5.16	5.3 ± 2.29 -4.36	3.08 ± 1.17 -1.34
MTD-GM	**MTL, 5 gated Bi-LSTMs**	Knee	**3.52 ± 1.2** **0.72**	**3.89 ± 1.24** **0.84**	**2.85 ± 1.04** **0.52**	**2.88 ± 1.49** **0.63**	**4.86 ± 1.28** **0.57**
		Ankle	**2.62 ± 1.2** **0.06**	**2.29 ± 1.24** **0.53**	**2.63 ± 1.04** **-1.45**	**1.71 ± 1.49** **0.15**	**2.43 ± 1.28** **-0.25**

were considered. The results of the complete cycles of this chapter and the last chapter are compared in Table 5.6. It is important to note that five additional patients are considered in this chapter. Thus, the datasets are similar but not identical. For MS and TBI diseases, we had the same patients in both studies, but the difference in prediction performance is evident. We can see that the R^2 score for stroke, TBI, and CP was negative in the last chapter, and in our new approach, there is an improvement in achieving more than 0.7 in the R^2 score.

Furthermore, we compare our results with those described by [7] for forecasts of peak knee and ankle on sagittal plane for rectus femoris BTX-A injection in patients with stroke [see Table 5.6]. In this case, our proposed method (MTD-GM) for stroke had a higher R^2 score for knee flexion and ankle dorsiflexion. However, since the models were not trained and tested with the same databases, this comparison should be taken into account. In addition, we also compare how well our model predicted the whole post-treatment kinematics for patients with CP. Different databases were used to test the proposed methods. Still, the results were better than what Galarraga et al. (2017) [163], Niiler et al. (1999) [155], and Niiler (2002) [164] said would happen with CP after surgery, as shown in Table 5.6. MTD-GM predicted knee flexion and ankle dorsiflexion with an average RMSE of 2.9°and 1.92°, respectively.

5.5 Discussion and Conclusion

In this chapter, we used two DL models using Bi-LSTM modules to predict the effects of BTX-A injections on gait kinematics. As far as we know, except for our previous experiments in Chapter 4, no other study in the literature has used DL to look at this specific prediction task. The prediction task we assess in this chapter is more challenging than predicting normal gait due to the high variability existing between and within subjects in pathological gait. DL architectures make it possible to add information about medical treatments to the model. Previous studies have shown that adding information about the treatment—specifically, which muscles received BTX-A treatment—improves performance instead of performing a straightforward post-to-pre regression task.

Table 5.4: Performance of both models in predicting post-treatment gait trajectories of a complete cycle with respect to different diseases.

Model	Model Type	Body Joint	Spinal Cord Injury (SCI)	Multiple Sclerosis (MS)	Stroke	Cerebral Palsy (CP)	Traumatic Brain Injury (TBI)
			No. of Patients				
			13	**12**	**11**	**4**	**3**
			No. of Cycles				
			400	**351**	**260**	**138**	**74**
			RMSE Mean (^{o}) ± Standard Error **R^2 Score**				
MTD-DM	**MTL, 5 Bi-LSTMs**	Knee	4.18 ± 0.97 0.70	4.58 ± 1.11 0.85	4.58 ± 1.09 0.69	4.82 ± 1.23 -0.17	4.26 ± 0.77 0.72
		Ankle	3.61 ± 0.99 0.55	3.35 ± 0.8 0.81	2.92 ± 0.85 0.47	3.6 ± 1.02 0.44	2.86 ± 0.73 0.75
MTD-GM	**MTL, 5 gated Bi-LSTMs**	Knee	**3.05 ± 0.7 0.80**	**3.08 ± 0.69 0.93**	**2.9 ± 0.79 0.88**	**2.88 ± 0.99 0.80**	**3.28 ± 0.78 0.80**
		Ankle	**2.37 ± 0.8 0.83**	**2.02 ± 0.54 0.94**	**2.3 ± 0.44 0.70**	**1.92 ± 0.72 0.86**	**2.44 ± 0.49 0.82**

In this chapter, we predict knee and ankle kinematics in the sagittal plane, but we perform this prediction per gait phase separately and only by exploiting the best DL architectures proposed in Chapter 4. These best architectures are based on Bi-LSTM sub-models. The best DL model incorporates the presence or absence of treatments thanks to a gating mechanism (MTD-GM model). This model predicts post-treatment gait kinematics with a lower $RMSE$ and higher R^2 score than the MTD-DM model and for all the considered pathologies. MTD-DM introduces treatments as the initialization of hidden units of each Bi-LSTM sub-model. In MTD-GM instead, we perform random initialization of hidden units and apply the gating mechanism to the outputs of Bi-LSTM sub-models, each devoted to a unique muscle treatment. In our previous experiment, this model (gated model) was not the best in all situations. This improvement is due to the approach proposed in this work: splitting the learning process of each DL architecture between the stance and swing phases. This proposal allows for coping with the high variance present in pre-treatment and post-treatment gait, both between individuals and within each individual. We notice the effectiveness of our approach by exploiting the advantages of DL architectures for each gait phase separately. We note that the gated model reaches a relative improvement of around 55% on both joints. We have thus proven the increased capability of the global model (after combining the predictions of both phases) in accurately predicting the complete gait cycles after treatment.

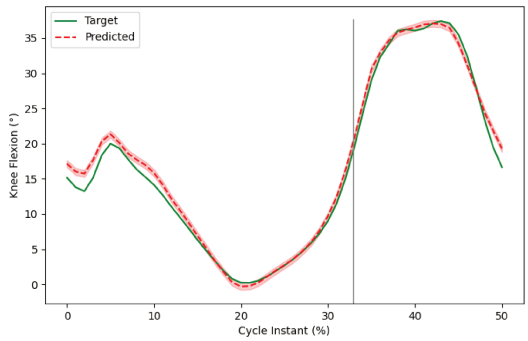
If we examine results per phase, for the stance phase, both models' performances (MTD-DM and MTD-GM) are equivalent. For the swing phase, the difference between both models is important. The gated model (MTD-GM) is much better: it often divides the $RMSE$ by a factor of 2. The R^2 score is generally good for the knee (higher than 0.7), but close to 0 for the ankle. Moreover, no proposed model was able to adequately explain the variance of the swing phase of the ankle joint (note that MTD-DM performed badly, as shown with negative R^2 scores).

We conclude that using treatment data as an initialization of hidden variables in each Bi-LSTM sub-model is not enough to cope with the strong variance encountered in the data, especially in the swing phase. Indeed, for this phase we remark that a more drastic treatment supervisor on the outputs of the Bi-LSTM sub-models is required. In this work, we do not address the segmentation between

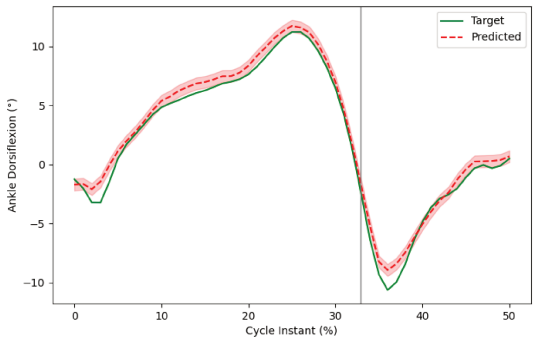
Table 5.5: Performance of both models for both joints in predicting post-treatment gait trajectories with respect to different diseases.

Model	Model Type	Body Joint	Spinal Cord Injury (SCI)	Multiple Sclerosis (MS)	Stroke	Cerebral Palsy (CP)	Traumatic Brain Injury (TBI)
			No. of Patients				
			13	**12**	**11**	**4**	**3**
			No. of Cycles				
			400	**351**	**260**	**138**	**74**
			RMSE Mean (°) ± Standard Error R^2 **Score**				
MTD-DM	MTL, 5 Bi-LSTMs	Knee and Ankle	4.02 ± 0.76 0.87	4.12 ± 0.79 0.91	3.91 ± 0.79 0.89	4.33 ± 1.02 0.77	3.76 ± 0.61 0.86
MTD-GM	MTL, 5 gated Bi-LSTMs		**2.78 ± 0.61 0.94**	**2.65 ± 0.53 0.96**	**2.65 ± 0.57 0.95**	**2.48 ± 0.75 0.94**	**2.97 ± 0.57 0.90**

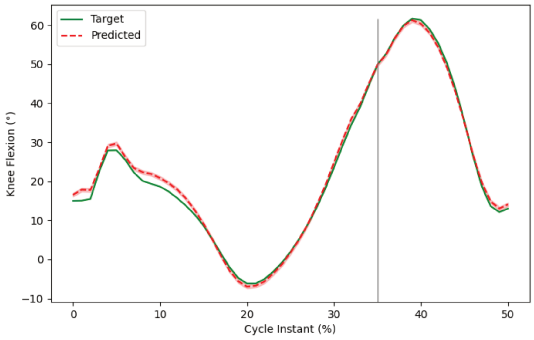
phases of gait cycles and focus on the prediction quality of our models, by exploiting the segmentation of pre-treatment gait cycles on post-treatment gait cycles.



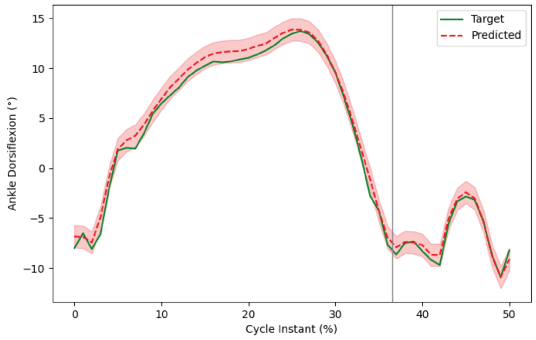
(a) Knee joint of patient with CP.



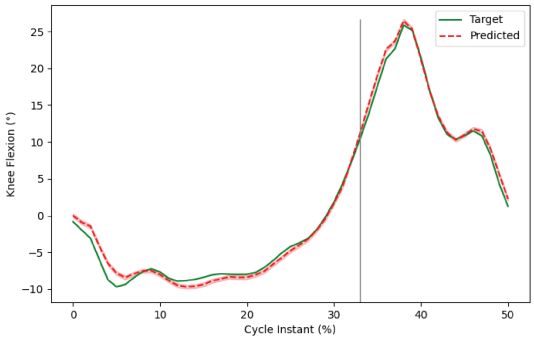
(b) Ankle joint of patient with CP.



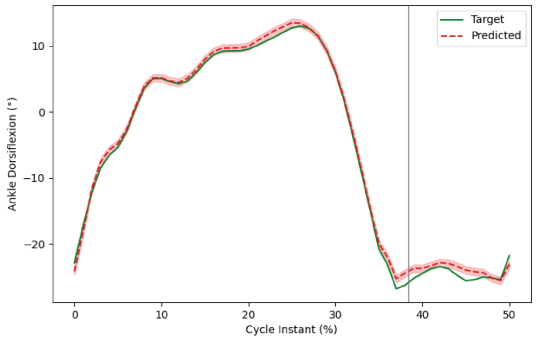
(c) Knee joint of patient with MS.



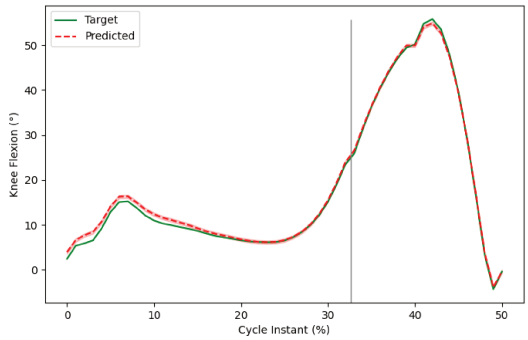
(d) Ankle joint of patient with MS.



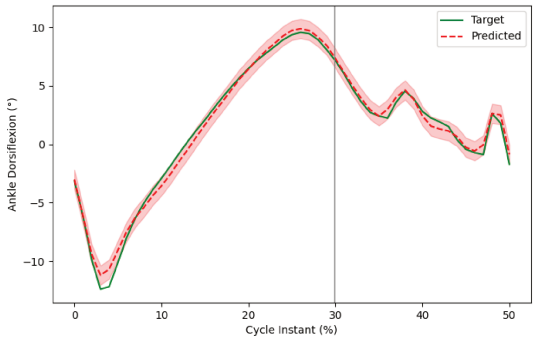
(e) Knee joint of patient with Stroke.



(f) Ankle joint of patient with Stroke.

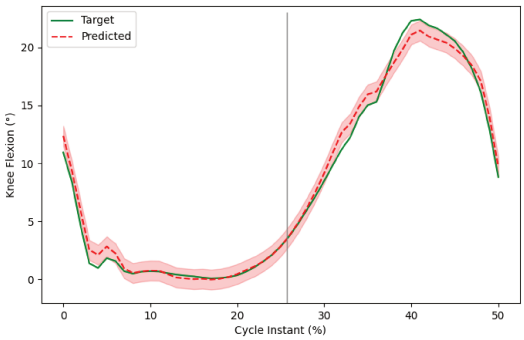


(g) Knee joint of patient with SCI.

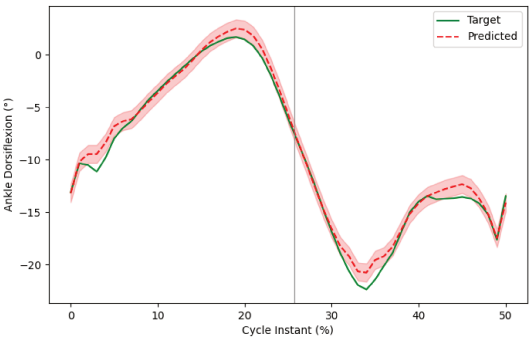


(h) Ankle joint of patient with SCI.

Figure 5.4: *Cont.*

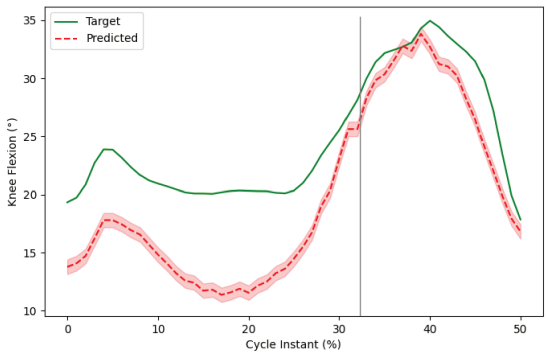


(i) Knee joint of patient with TBI.

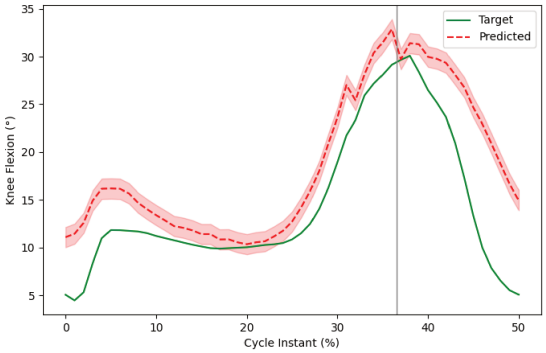


(j) Ankle joint of patient with TBI.

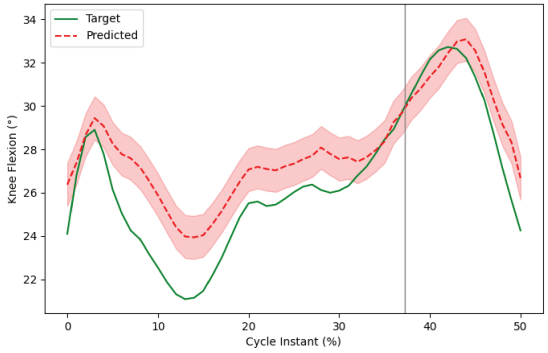
Figure 5.4: Comparison of MTD-GM prediction and post-treatment kinematics for a patient per disease.



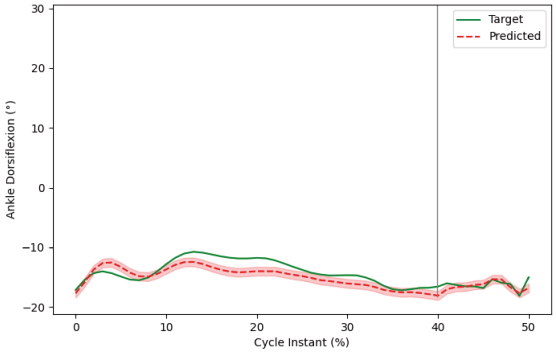
(a) Knee joint of patient with SCI.



(b) Knee joint of patient with TBI.



(c) Knee joint of patient with CP.



(d) Ankle joint of patient with Stroke.

Figure 5.5: Few outliers.

Table 5.6: Performance comparison to related works.

Model	Model Type	Body Joint	Spinal Cord Injury (SCI)	Multiple Sclerosis (MS)	Stroke	Cerebral Palsy (CP)	Traumatic Brain Injury (TBI)
			RMSE Mean (°) ± Standard Error R^2 Score				
Model 1 4	**MTL, 5 Bi-LSTMs**	**Knee**	7.51 ± 1.67 0.63	7.23 ± 1.69 0.63	7.14 ± 1.08 0.01	6.75 ± 1.73 0.7	5.81 ± 1.33 0.54
		Ankle	5.01 ± 0.99 0.38	4.38 ± 1.15 0.55	6.85 ± 1.61 -3.58	6.4 ± 1.19 -0.01	4.68 ± 0.82 0.4
Model 2 4	**MTL, 5 gated Bi-LSTMs**	**Knee**	7.62 ± 1.89 0.53	7.23 ± 2.01 0.48	8.02 ± 1.09 0.21	7.00 ± 2.42 0.68	5.60 ± 1.4 0.65
		Ankle	4.56 ± 1.05 0.27	5.39 ± 1.37 0.28	7.14 ± 1.44 -3.65	3.77 ± 1.41 -0.16	4.93 ± 1.0 -0.5
MTD-DM	**MTL, 5 Bi-LSTMs**	**Knee**	4.18 ± 0.97 0.70	4.58 ± 1.11 0.85	4.58 ± 1.09 0.69	4.82 ± 1.23 -0.17	4.26 ± 0.77 0.72
		Ankle	3.61 ± 0.99 0.55	3.35 ± 0.8 0.81	2.92 ± 0.85 0.47	3.6 ± 1.02 0.44	2.86 ± 0.73 0.75
MTD-GM	**MTL, 5 gated Bi-LSTMs**	**Knee**	**3.05 ± 0.7 0.80**	**3.08 ± 0.69 0.93**	**2.9 ± 0.79 0.88**	**2.88 ± 0.99 0.80**	**3.28 ± 0.78 0.80**
		Ankle	**2.37 ± 0.8 0.83**	**2.02 ± 0.54 0.94**	**2.3 ± 0.44 0.70**	**1.92 ± 0.72 0.86**	**2.44 ± 0.49 0.82**
LigReg Stroke* [7] (R^2 **Score**)	**Linear Regression**	**Knee**	-	-	0.24	-	-
		Ankle	-	-	0.43	-	-
MLinReg CP [15] ($RMSE(°)$)	**Multiple Linear Regression**	**Knee**	-	-	-	9.0	-
		Ankle	-	-	-	7.5	-
NN99 CP [155] ($RMSE(°)$)	**NN**	**Knee**	-	-	-	9.7	-
		Ankle	-	-	-	6.7	-
NN01 CP [156] ($RMSE(°)$)	**NN**	**Knee**	-	-	-	9.2	-
		Ankle	-	-	-	-	-

Chapter 6

Post-Treatment Gait Prediction after Botulinum Toxin Injections Using Deep Learning with an Attention Mechanism

Contents

6.1 Introduction **89**
6.2 Description of Models **90**
6.3 Experimental Setup and Results **91**
6.4 Conclusion **96**

Summary

In this chapter, we exploited attention mechanisms to enhance the learning process for post-treatment gait predictions on the full gait cycle. In chapter 4, we processed the full gait cycle, but the results were not so good compared to the prediction results in 5, when we split the gait cycle into two phases: the stance and swing phases. But as we have seen in chapter 5, after processing both phases separately, the combination of both phases of the gait cycle was done using the segmentation done on pre-CGA data. So, in this chapter, an attention mechanism has been integrated into DL models to process the full gait cycle to solve this problem. In this chapter, forty-three patients were used for experiments, the same as last chapter 5. The difference is only that full cycles are processed in this chapter rather than splitting the cycle into two phases. The two best models from the previous two chapters underwent modification (addition of an attention mechanism).

The content of this chapter is based on the following paper:

- Post-Treatment Gait Prediction after Botulinum Toxin Injections Using Deep Learning with an Attention Mechanism (accepted for presentation and publication at The 10th International Conference on Machine Learning, Optimization, and Data Science)

6.1 Introduction

The incorporation of an attention mechanism is one of the enhancements that have been made to the two best architectures discovered in Chapters 4 and 5. Concentrating on essential time steps makes this mechanism superior to an LSTM layer when it comes to the processing of long-time sequences. The attention mechanisms assist neural networks in processing information by enabling them to focus

on particular aspects of complex input. Particularly noteworthy is that Arshad et al. [73] improved gait event recognition in older individuals by utilizing Bi-LSTM and an attention mechanism with an accuracy of 99.73%. Similarly, D. Thakur and S. Biswas [88] utilized attention-based deep learning to improve hemiplegic gait prediction by using smartphone accelerometers and gyroscope data. Ding et al. [121] also used LSTM with attention to increase the accuracy of limb motion prediction for subjects. This methodology demonstrates significant advancements in predicting abnormal gait cycles and significantly improves the model's capacity to predict the entire gait cycle after treatment.

RNNs, specifically LSTM [62] and GRU [165] neural networks, are widely recognized as the most advanced methods for sequence modeling and transduction problems. These methods excel in language modeling and machine translation [166, 167, 60]. Self-attention, also known as intra-attention, is an attention mechanism that connects different positions within a single sequence to calculate a representation of the sequence. Self-attention has proven effective in reading comprehension, abstractive summarization, textual entailment, and learning sentence representation tasks [168, 169, 170, 171].

In this thesis, when dealing with pathological gait sequences, the attention method is particularly interesting to enhance important time steps in the signal, especially when it is complex and also combines MTD coming from different models (different muscles). Furthermore, when all information is combined, this leads to long-time sequences. The attention mechanism works better in dealing with long-time sequences than just using an LSTM layer because it gives more weight to essential time steps [172]. They let the network focus on different parts of the complex input, one at a time [172]. The goal is to break up big jobs into smaller, more manageable chunks, done one after the other. This method is like how our minds deal with new problems: they are broken down into smaller jobs that can be done more efficiently. Focusing models on specific tasks can help them improve and pay more attention to important information.

In the context of NLP, one way to describe an attention function is by mapping a query (a group of questions packed into a matrix, Q) and a set of key-value (K, V) pairs to a weighted output. In this mapping, all the elements involved—the "query," "keys," "values," and output—are represented as vectors in an embedded space [172]. Self-attention constructs in adding up the weighted values, where the weight measures how close Q and K match. The input includes queries and keys of dimension d_k, along with values of dimension d_v. We calculate the dot products of the query with all keys, divide each by $\sqrt{d_k}$, and then use a softmax function to determine the weights on the values (see Figure 6.1). We calculate the output matrix as follows [172]:

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T/\sqrt{d_k})V \tag{6.1}$$

In this thesis, we exploit a self-attention mechanism to process the pre-treatment gait kinematics (the knee and ankle joint signals) and focus on pertinent weights (attention scores) to important parts of signals for a given gait angular kinematics sequence (α_1, α_1, ..., α_n). We construct the matrix QK^T as follows: [$<\alpha_i,\alpha_j>$] and normalize it by its dimension and apply on each element a Softmax function. We obtain this way the weights (w_1, w_2, ..., w_n) in the [0,1] interval. Therefore, the final weighted sequence is $<w_1\alpha_1$, $w_2\alpha_2$, ..., $w_n\alpha_n>$ which is an enhanced version of signal. In the following, I explain how it is exploited in the models' architecture.

6.2 Description of Models

Five parallel layers of Bi-LSTM are used in both models (see Figure 6.2). As stated in Table 3.2, each layer is in charge of a particular treatment. The five categories of injected muscles correspond to the five treatments. Bi-LSTM layers each contained 51 units. Remember that every unit receives two inputs for the ankle and knee, respectively.

MTD-Driven Model with an Attention Mechanism (MTD-DM-Att)

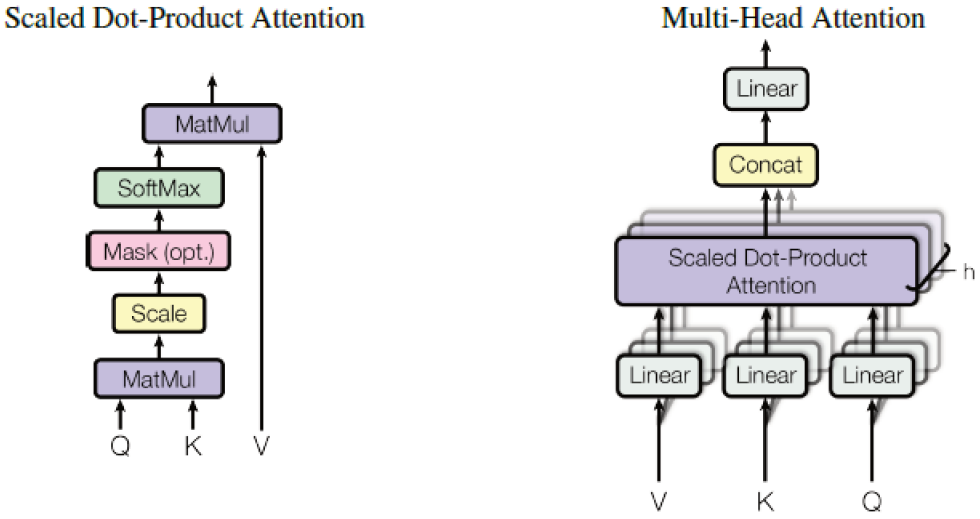


Figure 6.1: (left) Scaled Dot-Product Attention. (right) Multi-Head Attention consists of several attention layers running in parallel.

In this architecture, the input vector $\mathbf{x}$ is directed towards the five Bi-LSTM sub-models. The cell states of the LSTM are initialized at 0. The values of the hidden states $\mathbf{h}$ were determined using the Medical Treatment Data (MTD) vector $\mathbf{s}$. As an illustration, when a patient receives injections in muscles 1 and 3 (as shown in Table 3.2), the hidden states $\mathbf{h}_{1,t}$, $\mathbf{h}_{2,t}$ in Bi-LSTM layers 1 and 3 are assigned a value of 1, while the hidden states in layers 2, 4, and 5 are set to 0. This architecture incorporates an attention mechanism. The results of the five Bi-LSTM sub-models are concatenated to form a single tensor 'output,' that is, one-dimensional vector and reshaped. The reshaped output is then processed through an attention layer to enhance the accuracy of predictions. The attention mechanism enables the model to selectively concentrate on specific sequence segments while making predictions, potentially capturing complex relationships in a larger context. Next, the output of the attention layer is fed into two fully connected layers, FC1 and FC2, to produce the final output, as shown in Figure 6.2a.

MTD-Gated Model with an Attention Mechanism (MTD-GM-Att) This architecture utilizes a gating mechanism to manage MTD effectively. We utilized the MTD differently by incorporating it downstream. We multiplied the output of each sub-model by the corresponding binary value of the MTD. Figure 6.2b demonstrates that the model includes any existing treatment while disregarding those that do not (multiplied by 0). It utilizes an attention mechanism similar to MTD-DM-Att followed by two fully connected layers, FC1 and FC2, to produce the final output, as shown in Figure 6.2b.

6.3 Experimental Setup and Results

Leave-one-out cross-validation is commonly employed to evaluate the performance of a model. During each iteration, the model was trained using $N_{train} = N_{pat}$ - 1 patients, while one patient was reserved for testing purposes. Throughout the training process, mini-batches consisting of 16 samples were utilized. The ADAM optimizer-controlled mean square error (MSE) loss function was the basis for training DL models [159]. The performance of the proposed models is evaluated using RMSE [160], standard error (SE) [161], and the coefficient of determination R^2 [162].

This study included 43 participants. The previous assessment measures from previous chapters

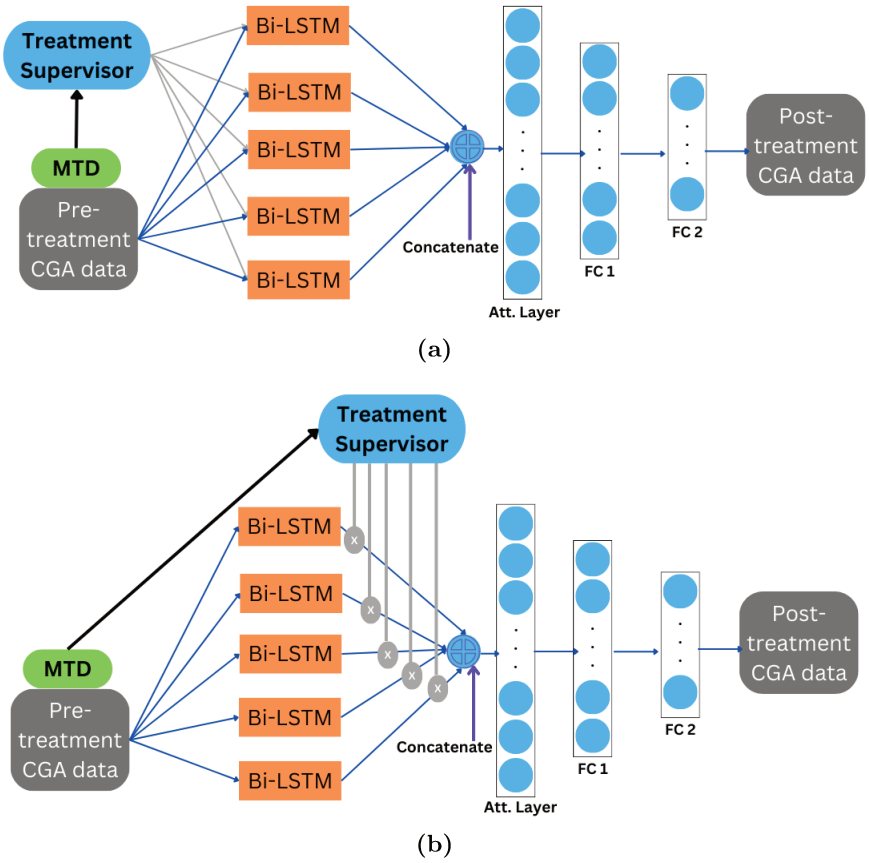


Figure 6.2: (a) MTD-DM architecture with an attention mechanism: processing of MTD in each sub-model, and (b) MTD-GM architecture with an attention mechanism: incorporating MTD through a gating mechanism

were used to evaluate the two models on the new dataset. Table 6.1 shows how well the MTD-DM-Att and MTD-GM-Att models predicted knee, ankle, and both joints post-treatment gait trajectories. The bold items in the tables below represent the most accurate predictions, with the lowest RMSE and highest R^2 values.

Both models performed similarly with the same RMSE. When predicting TBI patients for the knee joint, MTD-GM-Att outperforms MTD-DM-Att with an RMSE of 2.45 and a mean R^2 score of 0.93. The ankle joint prediction of MTD-DM-Att and MTD-GM-Att for CP patients was comparable, with the best RMSE of 1.97 and 1.87, respectively.

Table 6.1 also shows the overall prediction results for MTD-DM-Att and MTD-GM-Att joints. Using both joints has led to good performance in both models, with a mean R^2 score of 0.94 and 0.95 across all patients. As seen in Table 6.1, each disease has a variable number of cycles (samples). SCI has the most patients, and MTD-DM-Att predicted knee, ankle, and both joints better than MTD-GM-Att. MS has 12 patients, and MTD-GM-Att performed somewhat better than MTD-DM-Att.

From a different point of view, the following graphs (Figures 6.3 and 6.4) show the paths (mean of actual post-treatment CGA and mean of predicted post-treatment CGA) of five patients with different conditions for both models. The Y-axis shows ankle dorsiflexion or knee flexion, and the X-axis shows the patient's gait cycle. Figures 6.3 and 6.4 show the average forecasts for five patients with different conditions on all cycles. All graphs show that each patient's average predicted gait trajectory matches their average actual trajectory. As reported in the preceding sections, our models predict disabled

patients' post-treatment gait trajectory effectively.

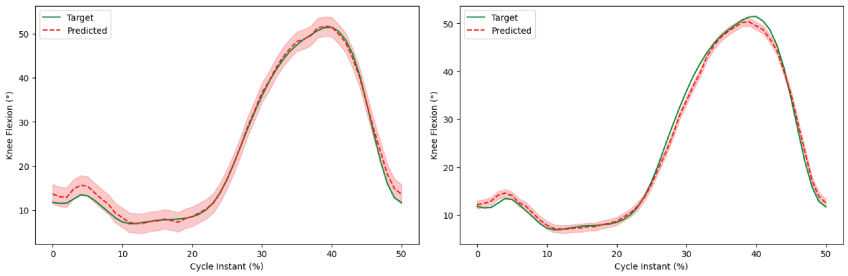
The first column in both figures shows MTD-DM-Att model predictions. The last column provides MTD-GM-Att model predictions for the same patient. The plots show that both models predict slightly differently. Both models have similar mean predictions but different SDs. Unless two ankle joint predictions are similar, MTD-DM-Att models predict knee and ankle joints with more deviation from the mean. The MTD-GM-Att model performed better.

Comparison to previous works

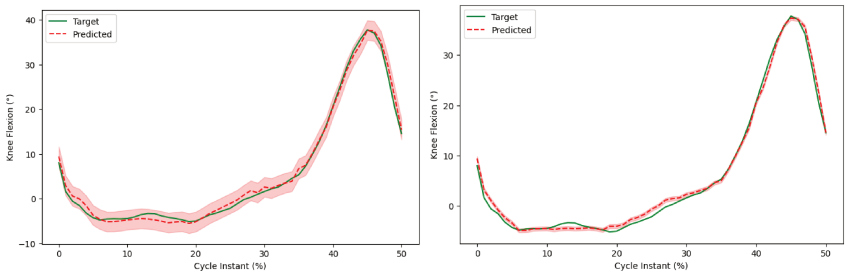
Except for our previous study [39], it is difficult to compare our results to other studies because this is the first time that the whole knee and ankle kinematic signals on the sagittal plane have been anticipated for botulinum toxin treatment. So, in Table 6.2, we compare the outcomes of this investigation to the previous study. The results show that adding an attention mechanism to models improves the results significantly. We had the same patient in both MS and TBI disease tests, but the difference in predictions is clear.

Table 6.1: Performance of both models in predicting post-treatment gait trajectories concerning different diseases.

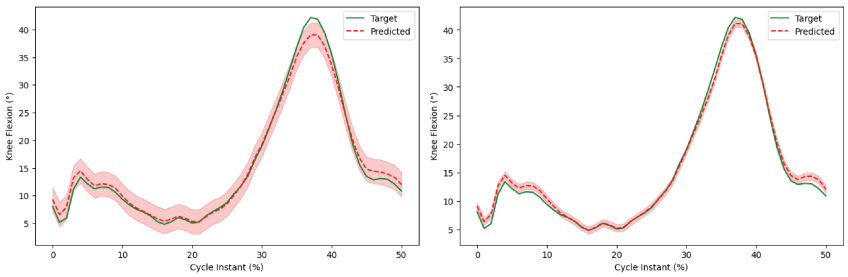
Model	**Body Joint**	**SCI**	**MS**	**Stroke**	**CP**	**TBI**	**All Patients**
		No. of Patients					
		13	**12**	**11**	**4**	**3**	**43**
		No. of Cycles					
		400	**351**	**260**	**138**	**74**	**1223**
		RMSE Mean (°) ± Standard Error					
		R^2 Score					
MTD-DM with an Attention Mechanism	Knee	**2.96 ± 0.69**	3.45 ± 0.8	**2.89 ± 0.83**	2.62 ± 0.93	3.15 ± 0.81	3.06 ± 0.79
		0.85	0.91	**0.84**	0.86	0.87	0.87
	Ankle	**2.15 ± 0.65**	2.21 ± 0.52	**2.27 ± 0.49**	1.97 ± 0.97	**2.2 ± 0.56**	**2.18 ± 0.59**
		0.87	0.92	**0.66**	0.84	**0.83**	**0.83**
	Knee & Ankle	**2.63 ± 0.57**	2.94 ± 0.58	**2.65 ± 0.59**	2.38 ± 0.78	2.75 ± 0.6	2.71 ± 0.6
		0.95	0.95	**0.94**	0.95	0.91	0.94
MTD-GM with an Attention Mechanism	Knee	3.02 ± 0.64	**3.43 ± 0.77**	2.93 ± 0.8	**2.57 ± 0.45**	**2.45 ± 0.83**	**3.03 ± 0.76**
		0.82	**0.91**	0.84	**0.88**	**0.93**	**0.87**
	Ankle	2.17 ± 0.6	**2.2 ± 0.49**	2.29 ± 0.53	**1.87 ± 0.91**	2.28 ± 0.66	**2.18 ± 0.56**
		0.86	**0.92**	0.64	**0.87**	0.82	**0.83**
	Knee & Ankle	2.68 ± 0.52	**2.92 ± 0.56**	2.68 ± 0.6	**2.82 ± 0.78**	**2.18 ± 0.56**	**2.69 ± 0.58**
		0.94	**0.95**	0.94	**0.96**	**0.94**	**0.95**



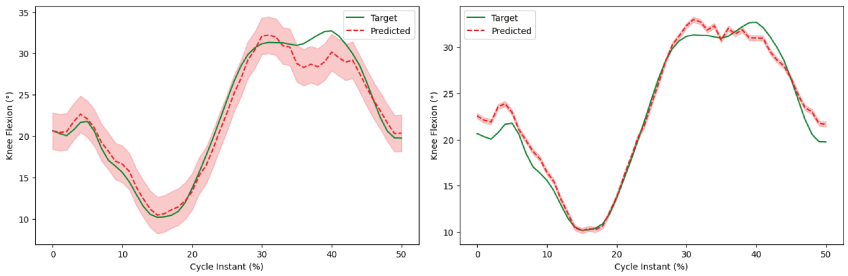
(a) CP Patient **(b)** CP Patient



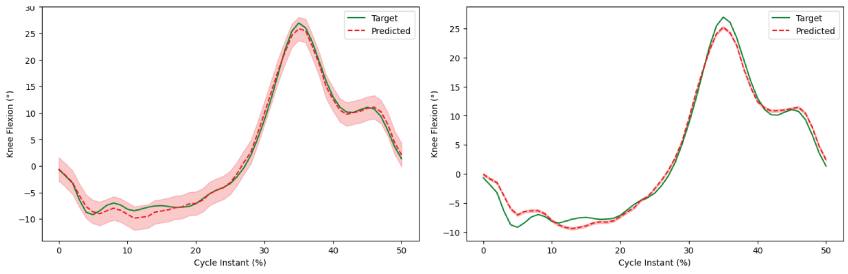
(c) MS Patient **(d)** MS Patient



(e) TBI Patient **(f)** TBI Patient

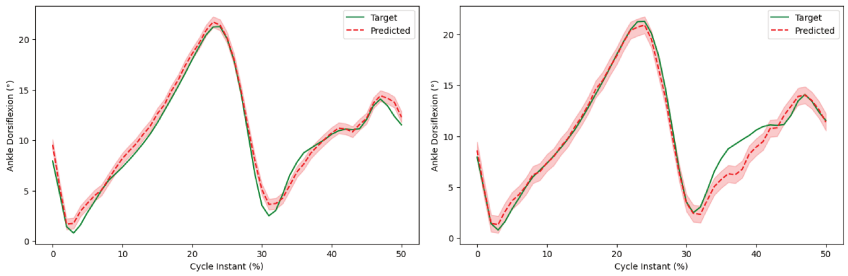


(g) SCI Patient **(h)** SCI Patient

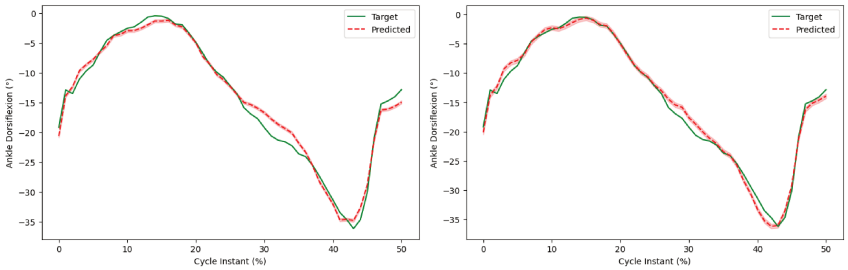


(i) Stroke Patient **(j)** Stoke Patient

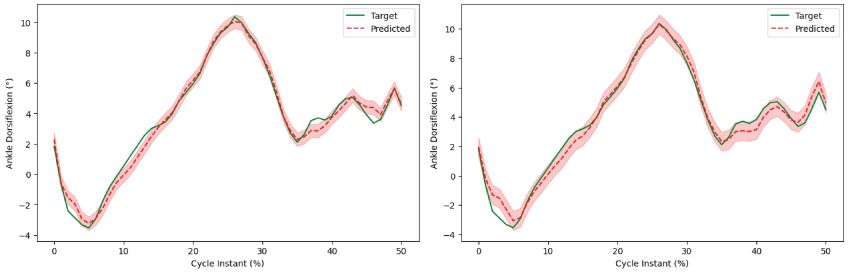
Figure 6.3: Few examples of post-treatment gait trajectory of the knee joint for one patient of each disease predicted by MTD-DM-Att and MTD-GM-Att models.



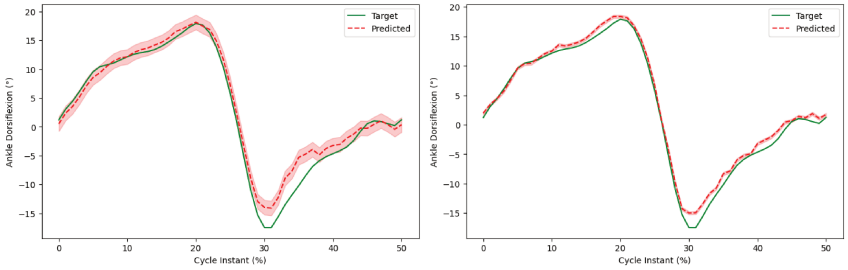
(a) CP Patient **(b)** CP Patient



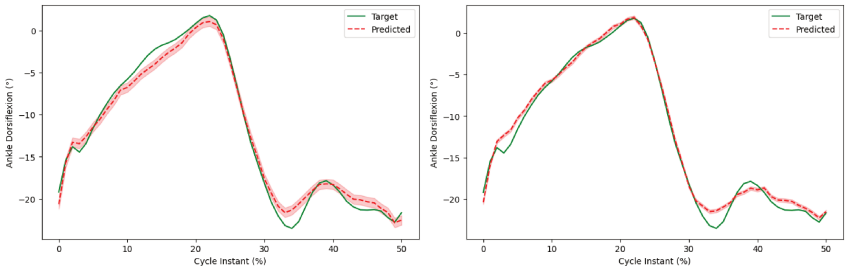
(c) MS Patient **(d)** MS Patient



(e) TBI Patient **(f)** TBI Patient



(g) SCI Patient **(h)** SCI Patient



(i) Stroke Patient **(j)** Stoke Patient

Figure 6.4: Few examples of post-treatment gait trajectory of the ankle joint for one patient of each disease predicted by MTD-DM-Att and MTD-GM-Att models.

Last time we used MTD-GM to predict MS patients, we had an RMSE of 7.23°and an R^2 of 0.48. In this investigation, the RMSE was 3.43°, and the R^2 was 0.91. Using MTD-GM, we obtained an RMSE of 5.60°and an R^2 of 0.65 for TBI. However, this time, we got the best RMSE of 2.45°and R^2 of 0.93. We can observe that the R^2 score for stroke, TBI, and CP patients was negative in the previous study for ankle joints, but this time, we have made significant progress and achieved more than a 0.8 R^2 score.

Additionally, we can contrast our results with those from [7] for predicting peak knee and ankle on sagittal planes for rectus femoris botulinum toxin injections in stroke patients (Table 6.2). In this situation, the proposed approach for stroke (MTD-GM-Att) showed a higher R^2 score for both peak knee flexion and peak ankle dorsiflexion. Because the models being compared were not trained and tested on the same databases, this comparison should be approached cautiously.

Furthermore, we investigated how well our model predicted entire postoperative kinematic trajectories for people with CP. Even though the proposed approaches were not tested on the same datasets, their results were significantly better than those of Galarraga et al. [173], Niiler et al. [155], and Niiler [156], as shown in Table 6.2. MTD-GM-ATT model had an average RMSE of 2.57°for knee flexion and 1.87°for ankle dorsiflexion.

6.4 Conclusion

In this chapter, we used two models using Bi-LSTM modules and attention mechanisms to predict the effects of BTX-A injections on gait trajectories. As far as we know, except for our previous work [39], no other study on the subjects in pathological gait. This study introduced novel MTL architectures incorporating attention mechanisms to predict post-BTX-A treatment gait trajectories in patients with musculoskeletal and neurological disorders. By leveraging CGA data and treatment specifics, the proposed models, MTD-DM-Att and MTD-GM-Att, demonstrated significant enhancements in post-treatment forecasting knee and ankle joint movements. The results showcased the efficacy of these models across various pathologies, with both demonstrating comparable predictive performance but excelling in specific patient cohorts. Comparison with previous studies underscored the superiority of our proposed approaches in predicting postoperative kinematic trajectories, showcasing substantial improvements in accuracy, particularly in predicting abnormal gait cycles post-BTX-A treatment. The findings of this study pave the way for more precise and personalized treatment strategies for individuals with musculoskeletal and neurological impairments, offering a promising avenue for enhancing rehabilitation outcomes.

Table 6.2: Performance comparison of the study with other prediction studies. LinReg and MLinReg correspond to linear regression and multiple LinReg, respectively, in [7, 173]. NN99 and NN01 correspond to feedforward neural networks, respectively, in [155, 156].

Model	**Model Type**	**Body Joint**	**SCI**	**MS**	**Stroke**	**CP**	**TBI**
			RMSE Mean (°) ± Standard Error				
			R^2 **Score**				
Model 1 [39]	MTL, 5 Bi-LSTMs	Knee	7.51 ± 1.67 0.63	7.23 ± 1.69 0.63	7.14 ± 1.08 0.01	6.75 ± 1.73 0.7	5.81 ± 1.33 0.54
		Ankle	5.01 ± 0.99 0.38	4.38 ± 1.15 0.55	6.85 ± 1.61 -3.58	6.4 ± 1.19 -0.01	4.68 ± 0.82 0.4
Model 2 [39]	MTL, 5 gated Bi-LSTMs	Knee	7.62 ± 1.89 0.53	7.23 ± 2.01 0.48	8.02 ± 1.09 0.21	7.00 ± 2.42 0.68	5.60 ± 1.4 0.65
		Ankle	4.56 ± 1.05 0.27	5.39 ± 1.37 0.28	7.14 ± 1.44 -3.65	3.77 ± 1.41 -0.16	4.93 ± 1.0 -0.5
MTD-DM-Att Model	MTL, 5 Bi-LSTMs	Knee	**2.96 ± 0.69** **0.85**	3.45 ± 0.8 0.91	**2.89 ± 0.83** **0.84**	2.62 ± 0.93 0.86	3.15 ± 0.81 0.87
		Ankle	**2.15 ± 0.65** **0.87**	2.21 ± 0.52 0.92	**2.27 ± 0.49** **0.66**	1.97 ± 0.97 0.84	**2.2 ± 0.56** **0.83**
MTD-GM-Att Model	MTL, 5 gated Bi-LSTMs	Knee	3.02 ± 0.64 0.82	**3.43 ± 0.77** **0.91**	2.93 ± 0.8 0.84	**2.57 ± 0.45** **0.88**	**2.45 ± 0.83** **0.93**
		Ankle	2.17 ± 0.6 0.86	**2.2 ± 0.49** **0.92**	2.29 ± 0.53 0.64	**1.87 ± 0.91** **0.87**	2.28 ± 0.66 0.82
LigReg Stroke* [7] (R^2 Score)	Linear Regression	Knee	-	-	0.24	-	-
		Ankle	-	-	0.43	-	-
MLinReg CP [173] ($RMSE(°)$)	Multiple Linear Regression	Knee	-	-	-	9.0	-
		Ankle	-	-	-	7.5	-
NN99 CP [155] ($RMSE(°)$)	NN	Knee	-	-	-	9.7	-
		Ankle	-	-	-	6.7	-
NN01 CP [156] ($RMSE(°)$)	NN	Knee	-	-	-	9.2	-
		Ankle	-	-	-	-	-

Chapter 7

Post-Treatment Gait and Treatment Prediction after Botulinum Toxin Treatment Using Multi-task Learning

Contents

7.1 Introduction 101
7.2 Description of Experiments 102
7.2.1 Experiment 1: Presence/Absence of MTD as auxiliary task 103
7.2.2 Experiment 2: Predicting MTD dosage as auxiliary task 104
7.2.3 Experiment 3: Gait phase prediction as auxiliary task 107
7.3 Discussion and Conclusion 108

Summary

In previous chapters, we tried to increase the performance of post-treatment gait prediction using different models and techniques. In this chapter, we are exploiting the MTL technique to see how it performs in predicting post-gait treatment trajectories. We are exploring the prediction of two tasks instead of just one task (prediction of post-treatment gait trajectories). Three different types of experiments are performed in this chapter. The main task of predicting post-gait trajectories is the same, but auxiliary tasks are changed. In experiment 1, post-treatment gait and treatment output (presence or absence of treatment) are predicted. In experiment 2, the prediction of BTX-A treatment dosage with the main task is performed. In the last experiment, the auxiliary task predicted the gait phase (stance or swing). The best models from previous chapters are modified to perform two tasks simultaneously (MTL). The input of the models is pre-treatment gait (knee and ankle joint), and the outputs are post-treatment gait and treatment information or gait phase. In this chapter, forty-three patients were used for experiments, the same as in the last chapter 6.

The content of this chapter is based on the following paper:

- Post-Treatment Gait and Treatment Prediction after Botulinum Toxin Treatment Using Multi-task Learning to be published

7.1 Introduction

Multi-task learning (MTL) aims to enhance generalization by utilizing domain-specific information found in the training signals of related tasks [174]. In the era of deep learning, MTL refers to creating

networks that can learn shared representations from multiple supervisory signals. This is in contrast to the single-task scenario, where each individual task is handled separately by its own network. Multi-task networks offer a range of advantages. Firstly, the amount of memory needed is significantly reduced due to the layer sharing inherent in the system. Additionally, by avoiding the need to repeatedly calculate the features in the shared layers for each task, they are able to achieve faster inference speeds. One key aspect to consider is the potential for enhanced performance when the tasks involved share complementary information or act as a regularizer for each other. MTL has been widely used across all applications of ML, including computer vision [175], speech recognition [176], natural language processing [177], drug discrovery [178], and biomedical areas [16].

In quantified gait analysis (QGA), we face the problem of few data. MTL can indeed handle sparse data issues and create a more reliable model by utilizing information from various tasks [16]. Nait Aicha et al. [126] used MTL and compared how well CNN, LSTM and a combined ConvLSTM used raw accelerometer data to predict falls in older adults. By incorporating auxiliary tasks like gender and age, the models showed improved performance. Yu et al. [127] came up with a Deep Multi-source Multi-task Learning (DMML) approach that gives a framework for assessing fall risk and PD severity based on accelerometer and gyroscope data. The goal of MTL is to improve the performance of every single task by simultaneously evaluating the fall risk and PD severity. Zhang et al. [128] proposed a multi-task CNN for age estimation using gait. They used gender information as another task to improve age estimation. Aoki et al. [129] implemented multi-task RNN (MRNN) to classify physically fatigued and non-fatigued gait cycles. They achieved improved results compared to state-of-art.

Within the field of DL, MTL is commonly achieved through either hard or soft parameter sharing of hidden layers. Hard parameter sharing is the predominant method for MTL in neural networks and has been used since [179]. Typically, it is implemented by sharing the hidden layers across all tasks while maintaining multiple task-specific output layers. Implementing hard parameter sharing significantly mitigates the likelihood of overfitting. Indeed, in [180] the authors demonstrated that the possibility of overfitting the shared parameters is considerably lower by a factor of N (where N represents the number of tasks) compared to overfitting the task-specific parameters, specifically the output layers. From an intuitive standpoint, it is logical that as we increase the number of tasks we learn simultaneously, our model needs to discover a representation encompassing all tasks. Consequently, our likelihood of overfitting on the original task decreases. Conversely, in soft parameter sharing, each task is equipped with a model with distinct parameters. The parameters of the model are regularised to promote similarity among them. In this chapter, we have used a hard parameter-sharing approach to predict, on one hand, post-gait and MTD, and on the other hand, post-gait and gait phase.

7.2 Description of Experiments

In all experiments, there are two outputs for models: one is primary tasks (post-treatment gait trajectories prediction), and the other is a prediction of MTD (presence/absence or full dosage) or prediction of gait phase (stance or swing). In this chapter, we have used $N_{pat} = 43$ patients for experiments. In the first experiment, only the presence or absence of treatment in different muscles is predicted along with post-treatment gait. The second experiment predicted the dosage of BTX-A treatment per muscle with the main task. The last experiment predicted the gait phase (stance of swing) along with the main task. In our dataset, three types of molecules were used for treatment: Dysport, Xeomin, and Botox. These molecules differ in units; one unit of Botox is equivalent to 2.5 units of Xeomin/Dysport. We rescale the botox molecule and normalize values of different molecules on the same scale from 0 to 1.

A hard parameter-sharing method is used in all models. Five parallel layers of Bi-LSTM are shared in all models. Table 7.1 states that each layer controls a particular treatment. The five categories of injected muscles correspond to the five treatments. Bi-LSTM layers each contained 51 units. Remem-

ber that every unit receives two inputs for the ankle and knee, respectively.

Table 7.1: Considered injected muscles and their frequencies in the database (43 patients).

Muscle Number	Muscle/Category	Injections limbs	
		Number	Proportion
1	Soleus	53	28.3%
2	Gastrocnemius (Medialis and/or Lateralis)	51	27.2 %
3	Rectus Femoris	22	11.7%
4	Semitendinosus	14	7.4%
5	Other Muscle	47	25.1 %

7.2.1 Experiment 1: Presence/Absence of MTD as auxiliary task

Following two models (show Figure 7.1) are responsible for the prediction of the presence/absence of treatment with post-treatment gait trajectories.

Model 1. MTL-based Model (Presence/Absence of MTD): In this architecture, the input vector **x** is directed towards the five Bi-LSTM sub-models. The cell states and hidden states of the LSTM are initialized at 0. The results of the five Bi-LSTM sub-models are concatenated and then processed through a dense layer. Next, the output of the dense layer is fed into task-specific layers for each task. There is one task-specific dense layer before output. Task 1 (the primary task referred to as FC2 in 7.1a) is a regression task responsible for predicting post-treatment gait trajectories. Task 2 (auxiliary task) is a binary classification task with five units (referred to as FC3 in 7.1a) responsible for checking the presence/absence of treatment in muscles (see Table 7.1).

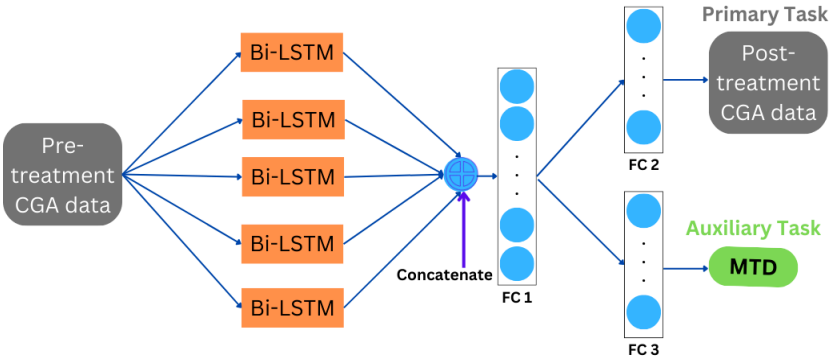
Model 2. MTL-based Attention-Model (Presence/Absence of MTD): The first part of this architecture is the same as the last architecture, having five Bi-LSTM sub-models. The results of the five Bi-LSTM sub-models are concatenated and then processed through an attention layer to enhance the accuracy of the predictions. Next, the output of an attention layer is fed into a dense layer (FC1 in Figure 7.1b). Finally, the output of the dense layer is fed into task-specific layers (FC2 and FC3) for each task. This architecture is also responsible for the same tasks as the previous model.

Leave-one-out cross-validation is employed to evaluate the performance of all models. During each iteration, the model was trained using $N_{train} = N_{pat}$ - 1 patients, while one patient was reserved for testing purposes. Throughout the training process, mini-batches consisting of 16 samples were utilized. The ADAM optimizer, MSE (Task 1) loss function and binary cross-entropy (Task 2) were the basis for training DL models [159]. The performance of the proposed models is evaluated using accuracy (Task 2), RMSE [160], standard error (SE) [161], and the coefficient of determination R^2 (Task 1) [162].

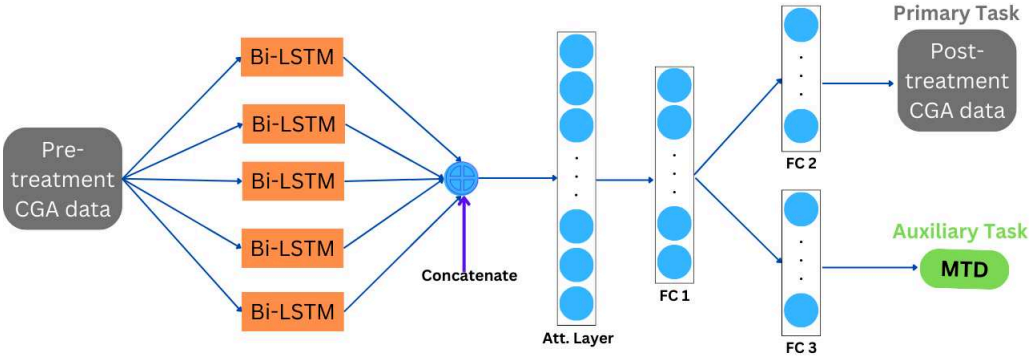
Results

Table 7.2 shows how well Models 1 and 2 predicted knee, ankle, and both joints post-treatment gait trajectories. These models also predicted the presence/absence of MTD. The bold items in the tables below represent the most accurate predictions, with the lowest RMSE, highest R^2 values, and best accuracies.

In predicting of presence/absence of MTD, both models (Model 1 and 2) performed similarly with the same accuracy (%), except for TBI patients. When predicting TBI patients, Model 2 outperforms



(a)



(b)

Figure 7.1: MTL models with two outputs (prediction of treatment and prediction of post-treatment gait trajectories) (a) MTL-based Model, and (b) MTL-based Attention-Model (Att-Model)

Model 1 with an accuracy 100.0. Regarding the prediction of post-treatment gait trajectories, Model 2, with an attention mechanism, performed better than Model 1 in terms of knee, ankle, and both joints. The average RMSE and R^2 scores on all patients for both joints using Model 2 are 2.73 $\pm$ 0.6 and 0.95, respectively. However, R^2 scores for both models are similar in most cases, except for TBI patients. It is noted that there is only a major difference in the results of TBI patients. Otherwise, there are very minor differences in the results of both models (Models 1 and 2). Furthermore, the ankle joint is better predicted for both models than the knee joint with RMSE of 2.31 $\pm$ 0.59 and 2.22 $\pm$ 0.57 for Model 1 and 2, respectively. Additionally, we can see that the accuracy of stroke patients in predicting the presence/absence of MTD is 100.0 for both models. On the other hand, the accuracy of CP patients is 75.0 for both models.

7.2.2 Experiment 2: Predicting MTD dosage as auxiliary task

The following two models are responsible for dosage value prediction of treatment with post-treatment gait trajectories.

Model 3. MTL-based Attention-Model (Predicting MTD Dosage): The architecture of this model is the same as Model 2. The difference is only in the output of Task 2. Task 2 is a regression task with four units responsible for checking the treatment dosage in the first four muscles (see Table 7.1). We don't consider the last category because it combines different muscles.

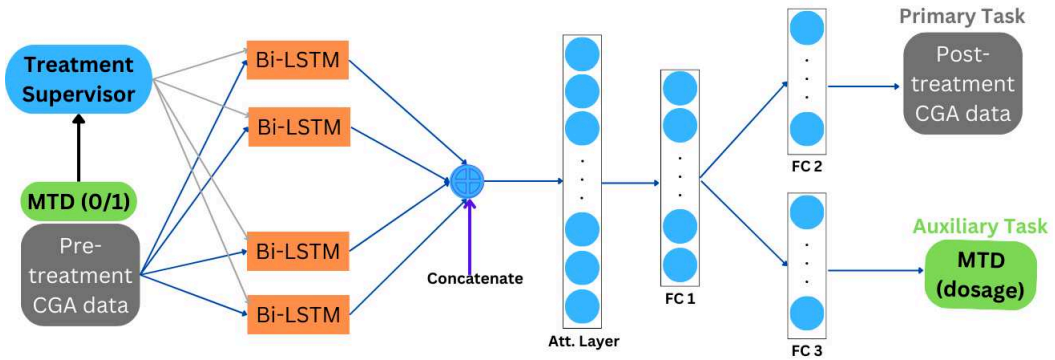
Model 4. MTL-based MTD-driven Attention-Model (Predicting MTD Dosage): The architecture of this model is the same as Model 2. The one difference is the output of Task 2, as

Table 7.2: Performance of both Model 1 and Model 2 in predicting the presence/absence of MTD and post-treatment gait trajectories (**Experiment 1 results**).

Model	Body Joint	SCI	MS	Stroke	CP	TBI	All Patients
				No. of Patients			
		13	**12**	**11**	**4**	**3**	**43**
				No. of Cycles			
		400	**351**	**260**	**138**	**74**	**1223**
				RMSE Mean (°) ± Standard Error			
				R^2 Score			
MTL-based Model (presence/ absence of MTD)	Knee	3.08 ± 0.68 0.85	3.5 ± 0.86 0.90	2.96 ± 0.79 0.84	2.81 ± 0.98 0.83	4.14 ± 0.77 0.70	3.22 ± 0.79 0.85
	Ankle	2.34 ± 0.66 0.84	2.3 ± 0.54 0.91	2.3 ± 0.5 0.64	1.96 ± 0.98 0.86	2.62 ± 0.67 0.75	2.31 ± 0.59 0.81
	Knee & Ankle	2.78 ± 0.59 0.94	3.01 ± 0.61 **0.95**	2.7 ± 0.57 **0.94**	2.46 ± 0.8 **0.95**	3.5 ± 0.6 0.83	2.84 ± 0.61 0.94
	Accuracy	**95.38**	**80.0**	**100.0**	**75.0**	86.67	**89.77**
MTL-based Att-Model (presence/ absence of MTD)	Knee	2.93 ± 0.7 0.87	3.48 ± 0.79 0.91	2.89 ± 0.85 0.84	2.66 ± 0.99 0.85	3.24 ± 0.8 0.85	3.07 ± 0.8 0.87
	Ankle	2.19 ± 0.63 0.86	2.28 ± 0.52 0.92	2.27 ± 0.49 0.65	1.96 ± 0.92 0.86	2.2 ± 0.59 0.83	2.22 ± 0.57 0.83
	Knee & Ankle	**2.63 ± 0.56** **0.95**	**2.98 ± 0.56** **0.95**	**2.65 ± 0.62** **0.94**	**2.38 ± 0.78** **0.95**	**2.8 ± 0.59** **0.91**	**2.73 ± 0.6** **0.95**
	Accuracy	**95.38**	**80.0**	**100.0**	**75.0**	**100.0**	**90.7**

Bold entries denote the best Accuracy (%), lowest average RMSE, and maximum R^2 over all limbs having a given disease.

reported in Model 3. Another difference is that MTD is processed as input using hidden layers of Bi-LSTM, and full dosage is predicted (see Figure 7.2). The first four categories of Table 7.1 are used. If there is any treatment, we feed 1 into a hidden layer of the respective submodel; otherwise, we feed 0.



(a)

Figure 7.2: MTL model with two outputs (prediction of MTD dosage and prediction of post-treatment gait trajectories): MTL-based MTD-driven Attention-Model

Remaining experiment protocol is the same for this experiment as well. But as Task 2 (prediction of BTX-A dosage) is a regression task here so, instead of binary cross-entropy, MSE is used as a loss

function for Task 1.

Results

Table 7.3 reports the results of Models 3 and 4 in the prediction of dosage of MTD and post-treatment gait trajectories. The bold items in the tables below represent the most accurate predictions, with the lowest RMSE, MAE, and highest R^2 values. It is interesting to see while predicting dosage of MTD with Task 2, Model 3 (without adding treatment information in hidden layers of Bi-LSTM submodels) performed better than Model 4 (with added information). The best results for the knee joint and ankle joint for all patients by Model 3 are (RMSE: 3.02 ± 0.78 and 2.19 ± 0.58 R^2 score: 0.88 and 0.83, respectively). Model 3 predicted all patients with RMSE of 2.68 ± 0.59 on both joints and R^2 score of 0.95. Results with prediction of MTD dosage are little better than prediction of presence/absence of MTD with post-treatment gait trajectories.

For MTD dosage prediction, Model 3 outperformed Model 4. Model 3 predicted a dosage of MTD for all patients with an average RMSE or MAE of 0.01. On the other hand, Model 4 was predicted to have an average RMSE or MAE of 0.02.

Table 7.3: Performance of both Model 3 and Model 4 in predicting dosage of MTD and post-treatment gait trajectories (**Experiment 2 results**).

Model	**Body Joint**	**SCI**	**MS**	**Stroke**	**CP**	**TBI**	**All Patients**
		No. of Patients					
		13	**12**	**11**	**4**	**3**	**43**
		No. of Cycles					
		400	**351**	**260**	**138**	**74**	**1223**
		RMSE Mean (°) ± Standard Error R^2 **Score**					
MTL-based Att-Model (Predicting MTD dosage)	Knee	2.87 ± 0.63 0.89	3.39 ± 0.8 0.91	2.89 ± 0.8 0.85	2.56 ± 0.95 0.88	3.23 ± 0.88 0.85	3.02 ± 0.78 0.88
	Ankle	2.12 ± 0.62 0.87	2.22 ± 0.52 0.92	2.31 ± 0.49 0.64	1.91 ± 0.94 0.87	2.37 ± 0.69 0.80	2.19 ± 0.58 0.83
	Knee & Ankle	**2.56 ± 0.53** **0.95**	**2.9 ± 0.58** **0.96**	**2.66 ± 0.59** **0.94**	**2.31 ± 0.75** **0.96**	**2.87 ± 0.68** **0.91**	**2.68 ± 0.59** **0.95**
	RMSE	**0.01**	**0.01**	**0.01**	**0.02**	**0.02**	**0.01**
	MAE	**0.01**	**0.01**	**0.01**	**0.01**	**0.02**	**0.01**
MTL-based MTD-driven Att-Model (Predicting MTD dosage)	Knee	2.86 ± 0.61 0.89	3.43 ± 0.78 0.91	2.87 ± 0.80 0.85	2.62 ± 0.91 0.88	3.58 ± 0.72 0.81	3.05 ± 0.75 0.88
	Ankle	2.32 ± 0.62 0.84	2.24 ± 0.55 0.92	2.33 ± 0.53 0.64	1.85 ± 1.02 0.87	2.55 ± 0.64 0.77	2.28 ± 0.59 0.82
	Knee & Ankle	2.65 ± 0.53 0.88	2.94 ± 0.58 **0.96**	2.66 ± 0.61 **0.94**	2.31 ± 0.82 **0.96**	3.13 ± 0.57 0.88	2.74 ± 0.59 0.94
	RMSE	0.02	0.02	**0.01**	**0.02**	0.03	0.02
	MAE	**0.01**	0.02	**0.01**	0.02	0.02	0.02

Bold entries denote the lowest average RMSE, MAE, and maximum R^2 over all limbs having a given disease.

7.2.3 Experiment 3: Gait phase prediction as auxiliary task

Experiment 3 differs from the previous experiment. Because in both those experiments, we were trying to predict MTD with post-treatment gait trajectories. However, in this experiment, we are trying to predict the gait phase (stance and swing) that may help us increase performance for the main task. Normal/healthy person has 60% of the stance phase and 40% of the swing phase in one gait cycle. However, we are working on patient data in this project, which is not the case here. If you can refer to Figure 3.2, it is visible almost all patients have different proportions of each phase. It will increase some of the performance of our main task if we can predict the gait phase correctly.

Model 5. MTL-based Model (gait-phase): The architecture of this model (see Figure 7.3a) is the same as Model 1. The difference is only in the output of Task 2. Here, Task 2 has 51 units and is responsible for predicting the gait phase at each time point. Our gait cycles have been normalized to 51 points; this has already been discussed in the chapter on Data.

Model 6. MTL-based Att-Model (gait-phase): The architecture of this model (see Figure 7.3b) is the same as Model 2. The difference is only in the output of Task 2, as reported in Model 5.

The experimental protocol for this experiment is also the same as in Experiment 1 because, in both tasks, Task 2 is a classification task.

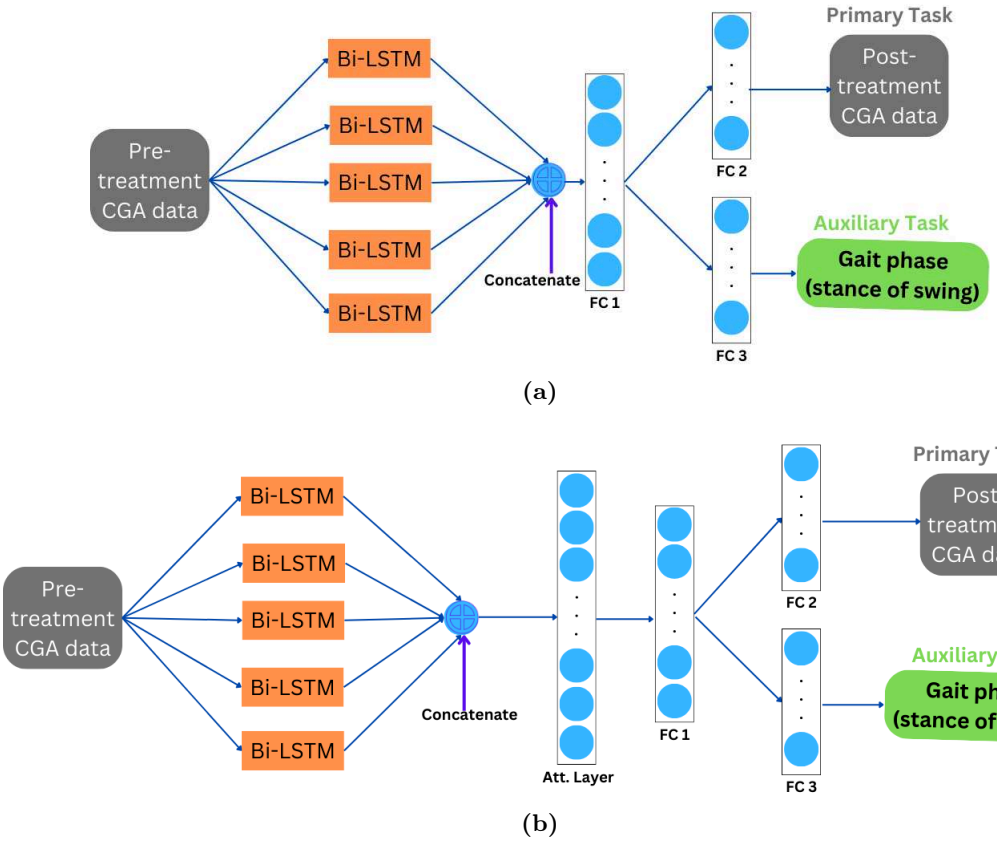


Figure 7.3: MTL models with two outputs (prediction of gait-phase and prediction of post-treatment gait trajectories) (a) MTL-based Model, and (b) MTL-based Att-Model

Results

Table 7.4 reports the results of Models 5 and 6's results in predicting gait-phase and post-treatment gait trajectories. The bold items in the tables below represent the most accurate predictions, with the lowest RMSE, highest R^2 values, and best accuracy. As we have seen in Experiment 1, Model 2 (with

an attention mechanism) outperformed Model 1 (without an attention mechanism); the same scenario repeats here. Model 6 (with an attention mechanism) performed better overall as well. The best results for the knee joint and ankle joint for all patients by Model 6 are (RMSE: 3.09 ± 0.84 and 2.23 ± 0.6 R^2 score: 0.87 and 0.83, respectively). Model 6 predicted all patients with RMSE of 2.74 ± 0.62 on both joints and R^2 score of 0.99. Interestingly, we can note the significant difference in TBI patients for both models, as we have seen in previous experiments. Model 5 predicted TBI patients with an RMSE of 3.48 ± 0.6 and R^2 score of 0.84. On the other hand, Model 6 predicted an RMSE of 2.89 ± 0.62 and R^2 score of 0.90.

For gait-phase prediction, both models performed similarly, with only a minor difference (0.14%).

Table 7.4: Performance of both Model 5 and Model 6 in predicting gait phase and post-treatment gait trajectories (**Experiment 3 results**).

Model	**Body Joint**	**SCI**	**MS**	**Stroke**	**CP**	**TBI**	**All Patients**
				No. of Patients			
		13	**12**	**11**	**4**	**3**	**43**
				No. of Cycles			
		400	**351**	**260**	**138**	**74**	**1223**
				RMSE Mean (°) ± Standard Error			
				R^2 Score			
MTL-based Model (gait-phase)	Knee	3.17 ± 0.69 0.85	3.58 ± 0.88 0.90	2.96 ± 0.81 0.84	2.93 ± 1.03 0.79	4.08 ± 0.81 0.71	3.28 ± 0.82 0.85
	Ankle	2.34 ± 0.69 0.84	2.36 ± 0.55 0.91	2.33 ± 0.51 0.63	2.06 ± 0.96 0.84	2.64 ± 0.66 0.75	2.34 ± 0.6 0.81
	Knee & Ankle	2.83 ± 0.59 **0.94**	3.08 ± 0.62 0.94	2.71 ± 0.58 0.94	2.58 ± 0.81 0.94	3.48 ± 0.6 0.84	2.89 ± 0.62 **0.94**
	Accuracy	81.6	85.13	**90.98**	83.82	**99.35**	86.64
MTL-based Att-Model (gait-phase)	Knee	2.87 ± 0.78 0.87	3.46 ± 0.86 0.91	2.93 ± 0.85 0.84	2.77 ± 0.88 0.84	3.48 ± 0.85 0.82	3.09 ± 0.84 0.87
	Ankle	2.16 ± 0.7 0.87	2.3 ± 0.53 0.91	2.32 ± 0.5 0.64	1.94 ± 0.097 0.86	2.29 ± 0.61 0.82	2.23 ± 0.6 0.83
	Knee & Ankle	**2.58 ± 0.62** 0.88	**2.99 ± 0.62** **0.95**	**2.7 ± 0.59** **0.95**	**2.44 ± 0.76** **0.95**	**2.98 ± 0.62** **0.90**	**2.74 ± 0.62** **0.94**
	Accuracy	**81.75**	**85.3**	**90.98**	**84.8**	98.69	**86.78**

Bold entries denote the lowest average RMSE, maximum R^2, and best accuracy over all limbs having a given disease.

7.3 Discussion and Conclusion

In this chapter, we explored the MTL to improve prediction of post-treatment gait trajectories. Three experiments were performed, and in each experiment, two models were designed to see the performance. Each model has two outputs, one is our primary task (prediction of post-treatment gait trajectories), and the other tasks are auxiliary (prediction of presence/absence of MTD or dosage of MTD or gait-phase prediction). After analyzing results in Tables 7.2, 7.3, and 7.4, it can be noted that there is a minor difference in the final results of our primary task. Models 2 and 3 get the best R^2 scores of 0.95 when predicting both joints (the knee and ankle). Model 3 also outperformed other models in terms of RMSE (2.68 ± 0.59).

If we compare the results of this chapter with previous chapters, then we can see the best results

we got in Chapter 6 by MTD-GM-Att model were a little bit better (RMSE: 2.69 ± 0.58 and R^2 score: 0.95).

However, in this chapter, models perform other predictions in addition to pos-treatment gait predictions, and yet results remain comparable to the best of Chapter 6. Moreover, the dosage of MTD is very important to predict. In Model 4, the input was pre-treatment gait trajectories and the presence/absence of MTD. The outputs were the dosage of MTD and post-treatment gait trajectories. Models 5 and 6 were responsible for predicting the proportion of the stance and swing phase along with the primary task. As we know, we don't have a regular proportion of stance and swing phases among patients. So, it can also be helpful for clinicians to see if, after treatment, there is any improvement in the proportion of both phases.

Chapter 8

Conclusions and perspectives

8.1 Conclusions

Over the past few years, there has been notable advancement in the field of QGA, thanks to ML techniques. QGA and gait prediction are areas where DL techniques gained popularity, as discussed in detail in Chapter 2. Although there has been a significant amount of attention from the scientific community to the application of DL techniques for healthy gait prediction, there is a noticeable absence of treatment outcome prediction for pathological gait. We have proposed different DL models for predicting how BTX-A treatment affects a patient's gait in different diseases, such as MS, CP, stroke, TBI, and SCI, using supervised machine learning. In this framework, pre-treatment gait and treatment information were used to predict post-treatment gait kinematics. We explored a variety of DL techniques to predict post-treatment gait, including LSTM, Bi-LSTM, ensemble learning, attention mechanisms, and MTL. We used regression analyses to look at the relationship between kinematic features following treatment and kinematics before treatment using clinician-provided treatment information. Treatment information consists of different injected muscles with BTX-A injections from both sides of the limbs.

To the best of our knowledge, the utilized database is the only database consisting of pre-treatment CGA, post-treatment CGA, and MTD. Indeed, most studies considering pathological gait use proprietary datasets with relatively few patients and lack external validation.

Chapter 2 gives background on QGA and talks about the latest DL strategies for QGA. It also presents the main datasets used in QGA and their strengths and weaknesses. The findings of Chapter 2 were divided into two major parts: classification and regression studies. Most classification studies focused on classifying healthy and pathological gait and on gait event detection. On the other hand, regression studies span future sub-sequence forecasting, sensor-to-sensor sequence estimation, condition/joint translations, and clinical score prediction. CNN and LSTM have been mostly used as DL methods in these studies.

Chapter 3 presents in detail the data collection protocol and dataset preparation.

Chapter 4 has addressed two different approaches to predicting the post-treatment gait of knee and ankle joints (serial and parallel processing). We developed seven different DL models, consisting of LSTM and BiLSTM. Parallel models outperformed serial models. MTD was fed into models in two different ways: (1) within the Bi-LSTM network's hidden layers and (2) through a gating mechanism. These architectures aim to model interactions among various treatment combinations when multiple muscles are injected simultaneously. Our findings after this chapter were: (a) the number of patients and type of disease did not directly affect the model's performance; (b) inter- and intra-subject variability affected the model's performance more than the number of patients (samples) and type of disease; and (c) parallel Bi-LSTM models were highly effective at increasing the total quantity of information accessible to the model and enhancing the context provided to the algorithm. Furthermore, the main

limitation of the proposed models was the relatively small size of the database. DL models usually need large amounts of data to be trained appropriately. Unfortunately, this is rarely the case in biomedical databases.

In Chapter 5, we proposed splitting the learning process into each phase of the cycle rather than the complete cycle to increase the size of the dataset. The models proposed in Chapter 4 were working on a complete gait cycle, but in Chapter 5, the full cycle was segmented into two phases (stance and swing phase). This technique increased the number of samples in the database. The two best models from Chapter 4 were chosen to process both gait cycle phases. Each model processed both phases separately. The results were combined to compare with the old results of Chapter 4. The findings of Chapter 5 were: (a) after the segmentation of the gait cycle into phases, the prediction of post-treatment gait kinematics (knee and ankle joint) was improved with a lower RMSE and higher R^2 score compared to the previous chapter and that for all the considered pathologies (relative improvement of around 55% on both joints); and (b) treatment data as initialization of hidden variables in each Bi-LSTM sub-model is not enough to cope with the strong variance encountered in the data, especially in the swing phase. Moreover, the limitation of Chapter 5 was that we do not address the segmentation between phases of gait cycles and focus on the prediction quality of our models by exploiting the segmentation of pre-treatment gait cycles on post-treatment gait cycles.

Chapter 6 has dealt with the limitations of the previous chapter. The chapter exploited the attention mechanism to cope with high-variance problems in the swing phase and segmentation process. We introduced an architecture incorporating attention mechanisms to predict post-BTX-A treatment gait trajectories in patients with musculoskeletal and neurological disorders. By leveraging CGA data and treatment specifics, the proposed models, MTD-DM-Att and MTD-GM-Att, demonstrated significant enhancements in post-treatment forecasting of knee and ankle joint movements. The results of the chapter showcased the efficacy of these models across various pathologies, with both demonstrating comparable predictive performance but excelling in specific patient cohorts, thanks to the introduction of the self-attention mechanism. Compared to previous approaches, our proposed methods in Chapter 6 were much better at predicting kinematic trajectories after treatment, showing significant accuracy improvements, mainly when predicting abnormal gait cycles after BTX-A treatment. Although tested on a relatively small sample, the chapter's findings encouraged more precise and personalized treatment strategies for individuals with musculoskeletal and neurological impairments, offering a promising avenue for enhancing rehabilitation outcomes. The model proposed in chapter 6 MTD-GM-Attention model (MTD-GM-Att) outperformed all other proposed models in this thesis, with the best RMSE of 2.69 ± 0.58 and R^2 score of 0.95. This model consists of five BiLSTM sub-models and has an attention mechanism. MTD information in that model was processed using a gating mechanism; typically, LSTM has been used to process long-term sequences. However, the attention mechanism concentrates on essential time steps, making this mechanism superior to an LSTM. This model outperformed other models due to using both BiLSTM and attention mechanisms.

MTL has been exploited in Chapter 7. The previous chapters only focused on prediction and improvement in performance of post-treatment gait using pre-treatment gait and MTD information. But in Chapter 7, we explored the prediction of two tasks instead of just one task (prediction of post-treatment gait trajectories). Three different types of experiments, with two models each, were performed in the chapter. The main task of predicting post-gait trajectories was the same, but auxiliary tasks were changed. Each model had two outputs; one was our primary task (prediction of post-treatment gait trajectories), and the other tasks were (prediction of presence/absence of MTD or dosage or MTD or gait-phase prediction). If we compare the results of this chapter with those of previous chapters, we can see there was very little improvement in the prediction of post-treatment gait. However, in this chapter, models perform other predictions besides post-treatment gait predictions (RMSE: 2.68 ± 0.58 and R^2: 0.95), yet results remain comparable to the best of Chapter 6 (RMSE: 2.69 ± 0.58 and R^2: 0.95). Moreover, the dosage of MTD is very important to predict. We got very

good prediction accuracy in predicting dosage of MTD.

The principal contributions of this study could be resumed with the following points:

- Development of a database with pre-treatment gait, post-treatment gait kinematics, and MTD information
- It gives a quantified prediction (estimated post-treatment gait kinematics)
- It simultaneously considers several treatments of BTX-A injections (and their combinations)
- It considers multiple gait patterns in different diseases (MS, CP, SCI, TBI, stroke)
- It can predict multiple tasks using MTL (main task with auxiliary task).

This study holds significant scientific and societal importance, as it explores the medical application of BTX-A treatment for adults with neurological gait difficulties. Firstly, the predictive models provide clinicians and gait analysts with kinematic curves, which they can use to assess the most probable treatment outcome.

Indeed, the prediction of gait outcomes based on treatment can assist clinicians in making informed decisions regarding various treatment combinations. For instance, when doctors evaluate treatments, they can test the model with and without the treatments in question and make a final decision by comparing the probable outcomes. The preview of the anticipated post-treatment gait will also assist the medical team in their treatment discussions, as there may be differing professional opinions on the expected outcome among team members (i.e. two different doctors within the same team).

Furthermore, when evaluating the effectiveness of a treatment, the prediction provides an estimate of the patient's potential for functional improvement, which typically varies. This could serve as a source of inspiration for the patient to pursue a specific treatment. Enhancing psychological motivation can increase patient engagement in recovery and rehabilitation, resulting in a more favorable outcome. If the optimal treatment is identified, we can estimate the maximum potential for improvement in the patient.

Globally, this research serves as a foundation for predicting post-treatment gait, although its effectiveness is limited by the complex nature of the problem and the limited size of the database. Even though there is still a long way to go, this study is a significant step towards improving the evaluation of BTX-A treatment.

8.2 Limitations

A major drawback of this study is the limited number of samples concerning the numerous parameters that need to be estimated and the small number of patients who underwent pre-treatment and post-treatment CGA compared to the multitude of kinematic variables and treatments being considered. In this particular context, we have used leave-one-person out cross-validation to cope with the limited size of the dataset.

Additionally, in the future the proposed models do not consider other commonly used CGA data for treatment assessment, such as kinetics or electromyography, among others. However, the absence of such data can pose a significant challenge as it may lead to a major missing data problem due to their unavailability. For instance, when a patient relies on a technical walking aid such as canes or a K-Walker, it becomes challenging to calculate kinetic data accurately.

However, without external validation, it is best to restrict the use of the proposed methods to patients and treatments that closely resemble those in the database being considered. No other databases with MTD were available to test our models.

Given that the methods used are purely statistical, there are no mechanical limitations when predicting gait. However, this lack of constraints can sometimes lead to physically impossible solutions.

Nevertheless, considering the inherent limitations of the training data, it is possible to mitigate the potential drawbacks of these solutions.

8.3 Perspectives

The approaches developed in this work serve as a foundation for future advancements in predicting post-treatment gait analysis. This section offers recommendations to enhance and broaden the methodologies presented and overcome their primary limitations.

- Since it is a highly complex mathematical problem, it is crucial in this era of big data to expand the databases by including pre-treatment and post-treatment CGA data, along with MTD. One of the primary challenges in building and managing these databases is the absence of post-treatment CGA. Clinicians typically rely on the CGA to determine the appropriate treatment rather than using it to assess the treatment's effectiveness. Furthermore, we strongly encourage laboratories worldwide that have addressed this research question to collaborate and exchange their data. Nevertheless, sharing data raises concerns regarding the legal aspects surrounding medical and personal information, even if the data is anonymized. Establishing agreements regarding the CGA and treatment data, such as the kinematic model, data representation, and preprocessing methods, is also necessary.

- In this work, we only consider two joints (knee and ankle) and one plane (sagittal) due to limited treatment information. One can enhance this work in the future by processing all the joints and all three planes of the lower limb and seeing how it performs using all joints.

- One can generate synthetic data of pre-treatment, post-treatment CGA, and MTD to increase the size of the database. Our proposed methods can be validated using synthetic data or other databases.

- These predicted post-treatment gait kinematic curves can be utilized to animate a 3-D avatar, demonstrating the potential gait to individuals unfamiliar with analyzing kinematic signals. This can help the clinician and the patient better understand the possible results. Without a thorough preview of the most probable outcome, there is a higher chance that the patient may misinterpret the qualitative explanation provided by the medical team. This is undoubtedly a significant issue in today's world.

Bibliography

[1] W. Pirker and R. Katzenschlager, "Gait disorders in adults and the elderly: A clinical guide," *Wiener Klinische Wochenschrift*, vol. 129, no. 3-4, pp. 81–95, 2017.

[2] M. Badole, "Create cnn model and optimize using keras tuner - deep learning," Jun 2021.

[3] D. Utebayeva, L. Ilipbayeva, and E. T. Matson, "Practical study of recurrent neural networks for efficient real-time drone sound detection: A review," *Drones*, vol. 7, no. 1, p. 26, 2022.

[4] J. McLoughlin, C. Barr, M. Crotty, S. Lord, and D. Sturnieks, "Association of postural sway with disability status and cerebellar dysfunction in people with multiple sclerosis: a preliminary study," *International journal of MS care*, vol. 17, no. 3, pp. 146–151, 2015.

[5] L. Blumhardt, *Multiple Sclerosis Dictionary*. Taylor & Francis, 2004.

[6] L.-C. Sun, R. Chen, C. Fu, Y. Chen, Q. Wu, R. Chen, X. Lin, and S. Luo, "Efficacy and Safety of Botulinum Toxin Type A for Limb Spasticity after Stroke: A Meta-Analysis of Randomized Controlled Trials," *BioMed Research International*, vol. 2019, pp. 1–17, Apr. 2019.

[7] N. Roche, J. Boudarham, A. Hardy, C. Bonnyaud, and D. Bensmail, "Use of gait parameters to predict the effectiveness of botulinum toxin injection in the spastic rectus femoris muscle of stroke patients with stiff knee gait," *European Journal of Physical and Rehabilitation Medicine*, vol. 51, pp. 361–370, Aug. 2015.

[8] R. W. Baker, *Measuring Walking: A Handbook of Clinical Gait Analysis*. London: MacKeith Press, 1 ed., 2013.

[9] J. L. McGinley, R. Baker, R. Wolfe, and M. E. Morris, "The reliability of three-dimensional kinematic gait measurements: A systematic review," *Gait & Posture*, vol. 29, pp. 360–369, Apr. 2009.

[10] D. Moon and A. Esquenazi, "Instrumented Gait Analysis: A Tool in the Treatment of Spastic Gait Dysfunction," *JBJS Reviews*, vol. 4, p. 1, Jun. 2016.

[11] B. Zörner, L. Filli, K. Reuter, S. Kapitza, L. Lörincz, T. Sutter, D. Weller, M. Farkas, C. S. Easthope, A. Czaplinski, M. Weller, and M. Linnebank, "Prolonged-release fampridine in multiple sclerosis: Improved ambulation effected by changes in walking pattern," *Multiple Sclerosis*, vol. 22, pp. 1463–1475, Oct. 2016.

[12] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[13] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine*, vol. 25, pp. 44–56, Jan. 2019.

[14] S. Barnes, S. Saria, and S. Levin, "An Evolutionary Computation Approach for Optimizing Multilevel Data to Predict Patient Outcomes," Mar. 2018.

[15] O. A. Galarraga C., V. Vigneron, B. Dorizzi, N. Khouri, and E. Desailly, "Predicting postoperative gait in cerebral palsy," *Gait & Posture*, vol. 52, pp. 45–51, 2017.

[16] Y. Zhang and Q. Yang, "A survey on multi-task learning," *IEEE Transactions on Knowledge and Data Engineering*, 2021.

[17] C. Prakash, R. Kumar, and N. Mittal, "Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges," *Artificial Intelligence Review*, vol. 49, pp. 1–40, 2018.

[18] Y.-S. Lee, C.-S. Ho, Y. Shih, S.-Y. Chang, F. J. Róbert, and T.-Y. Shiang, "Assessment of walking, running, and jumping movement features by using the inertial measurement unit," *Gait & posture*, vol. 41, no. 4, pp. 877–881, 2015.

[19] Q. Mei, J. Fernandez, W. Fu, N. Feng, and Y. Gu, "A comparative biomechanical analysis of habitually unshod and shod runners based on a foot morphological difference," *Human movement science*, vol. 42, pp. 38–53, 2015.

[20] B. H. Dobkin, X. Xu, M. Batalin, S. Thomas, and W. Kaiser, "Reliability and validity of bilateral ankle accelerometer algorithms for activity recognition and walking speed after stroke," *Stroke*, vol. 42, no. 8, pp. 2246–2250, 2011.

[21] C. Punin, B. Barzallo, R. Clotet, A. Bermeo, M. Bravo, J. P. Bermeo, and C. Llumiguano, "A non-invasive medical device for parkinson's patients with episodes of freezing of gait," *Sensors*, vol. 19, no. 3, p. 737, 2019.

[22] S. Del Din, A. Godfrey, and L. Rochester, "Validation of an accelerometer to quantify a comprehensive battery of gait characteristics in healthy older adults and parkinson's disease: toward clinical and at home use," *IEEE journal of biomedical and health informatics*, vol. 20, no. 3, pp. 838–847, 2015.

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

[24] T. Virmani, H. Gupta, J. Shah, and L. Larson-Prior, "Objective measures of gait and balance in healthy non-falling adults as a function of age," *Gait & posture*, vol. 65, pp. 100–105, 2018.

[25] F. Di Russo, M. Berchicci, R. L. Perri, F. R. Ripani, and M. Ripani, "A passive exoskeleton can push your life up: application on multiple sclerosis patients," *PloS one*, vol. 8, no. 10, p. e77348, 2013.

[26] Z. F. Lerner, D. L. Damiano, and T. C. Bulea, "A lower-extremity exoskeleton improves knee extension in children with crouch gait from cerebral palsy," *Science translational medicine*, vol. 9, no. 404, p. eaam9145, 2017.

[27] A. S. Khan, D. C. Livingstone, C. L. Hurd, J. Duchcherer, J. E. Misiaszek, M. A. Gorassini, P. J. Manns, and J. F. Yang, "Retraining walking over ground in a powered exoskeleton after spinal cord injury: a prospective cohort study to examine functional gains and neuroplasticity," *Journal of neuroengineering and rehabilitation*, vol. 16, pp. 1–17, 2019.

[28] V. Kumar, Y. V. Hote, and S. Jain, "Review of exoskeleton: history, design and control," in *2019 3rd international conference on recent developments in control, automation & power engineering (RDCAPE)*, pp. 677–682, IEEE, 2019.

[29] H. Herr, "Exoskeletons and orthoses: classification, design challenges and future directions," *Journal of neuroengineering and rehabilitation*, vol. 6, pp. 1–9, 2009.

[30] M. B. Nebel, E. L. Sims, F. J. Keefe, V. B. Kraus, F. Guilak, D. S. Caldwell, J. J. Pells, R. Queen, and D. Schmitt, "The relationship of self-reported pain and functional impairment to gait mechanics in overweight and obese persons with knee osteoarthritis," *Archives of Physical Medicine and Rehabilitation*, vol. 90, no. 11, pp. 1874–1879, 2009.

[31] M. G. Benedetti, V. Agostini, M. Knaflitz, V. Gasparroni, M. Boschi, and R. Piperno, "Self-reported gait unsteadiness in mildly impaired neurological patients: an objective assessment through statistical gait analysis," *Journal of neuroengineering and rehabilitation*, vol. 9, pp. 1–7, 2012.

[32] H. Prasanth, M. Caban, U. Keller, G. Courtine, A. Ijspeert, H. Vallery, and J. Von Zitzewitz, "Wearable sensor-based real-time gait detection: A systematic review," *Sensors*, vol. 21, no. 8, p. 2727, 2021.

[33] Y. Hutabarat, D. Owaki, and M. Hayashibe, "Recent advances in quantitative gait analysis using wearable sensors: a review," *IEEE Sensors Journal*, vol. 21, no. 23, pp. 26470–26487, 2021.

[34] D. T. Lai, R. K. Begg, and M. Palaniswami, "Computational intelligence in gait research: a perspective on current applications and future challenges," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 687–702, 2009.

[35] E. Dolatabadi, B. Taati, and A. Mihailidis, "An automated classification of pathological gait using unobtrusive sensing technology," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 12, pp. 2336–2346, 2017.

[36] J. Paulo, P. Peixoto, and P. Amorim, "Trajectory-based gait pattern shift detection for assistive robotics applications," *Intelligent Service Robotics*, vol. 12, pp. 255–264, 2019.

[37] P. Kutilek and J. Hozman, "Prediction of lower extremities movement using characteristics of angle-angle diagrams and artificial intelligence," in *2011 E-Health and Bioengineering Conference (EHB)*, pp. 1–4, IEEE, 2011.

[38] R. Begg and J. Kamruzzaman, "Neural networks for detection and classification of walking pattern changes due to ageing," *Australasian Physics & Engineering Sciences in Medicine*, vol. 29, pp. 188–195, 2006.

[39] A. Khan, A. Hazart, O. Galarraga, S. Garcia-Salicetti, and V. Vigneron, "Treatment outcome prediction using multi-task learning: Application to botulinum toxin in gait rehabilitation," *Sensors*, vol. 22, no. 21, p. 8452, 2022.

[40] A. d. M. e. Souza and M. R. Stemmer, "Extraction and classification of human body parameters for gait analysis," *Journal of Control, Automation and Electrical Systems*, vol. 29, pp. 586–604, 2018.

[41] H.-m. Shim and S. Lee, "Multi-channel electromyography pattern classification using deep belief networks for enhanced user experience," *Journal of Central South University*, vol. 22, no. 5, pp. 1801–1808, 2015.

[42] W. Wei, C. McElroy, and S. Dey, "Human action understanding and movement error identification for the treatment of patients with parkinson's disease," in *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 180–190, IEEE, 2018.

[43] D. Leightley, J. S. McPhee, and M. H. Yap, "Automated analysis and quantification of human mobility using a depth sensor," *IEEE journal of biomedical and health informatics*, vol. 21, no. 4, pp. 939–948, 2016.

[44] I. McLoughlin, H. Zhang, Z. Xie, Y. Song, and W. Xiao, "Robust sound event classification using deep neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 3, pp. 540–552, 2015.

[45] W. Zeng, S. A. Ismail, and E. Pappas, "Classification of gait patterns in patients with unilateral anterior cruciate ligament deficiency based on phase space reconstruction, euclidean distance and neural networks," *Soft Computing*, vol. 24, no. 3, pp. 1851–1868, 2020.

[46] B. Xiong, N. Zeng, H. Li, Y. Yang, Y. Li, M. Huang, W. Shi, M. Du, and Y. Zhang, "Intelligent prediction of human lower extremity joint moment: an artificial neural network approach," *IEEE Access*, vol. 7, pp. 29973–29980, 2019.

[47] M. Lempereur, F. Rousseau, O. Rémy-Néris, C. Pons, L. Houx, G. Quellec, and S. Brochard, "A new deep learning-based method for the detection of gait events in children with gait disorders: Proof-of-concept and concurrent validity," *Journal of biomechanics*, vol. 98, p. 109490, 2020.

[48] A. Foroutannia, M.-R. Akbarzadeh-T, and A. Akbarzadeh, "A deep learning strategy for emg-based joint position prediction in hip exoskeleton assistive robots," *Biomedical Signal Processing and Control*, vol. 75, p. 103557, 2022.

[49] C. Tunca, G. Salur, and C. Ersoy, "Deep learning for fall risk assessment with inertial sensors: Utilizing domain knowledge in spatio-temporal gait parameters," *IEEE journal of biomedical and health informatics*, vol. 24, no. 7, pp. 1994–2005, 2019.

[50] M. Savadkoohi, T. Oladunni, and L. A. Thompson, "Deep neural networks for human's fall-risk prediction using force-plate time series signal," *Expert systems with applications*, vol. 182, p. 115220, 2021.

[51] M. Zhu, Q. Men, E. S. Ho, H. Leung, and H. P. Shum, "A two-stream convolutional network for musculoskeletal and neurological disorders prediction," *Journal of Medical Systems*, vol. 46, no. 11, p. 76, 2022.

[52] N. Sadeghzadehyazdi, T. Batabyal, and S. T. Acton, "Modeling spatiotemporal patterns of gait anomaly with a cnn-lstm deep neural network," *Expert Systems with Applications*, vol. 185, p. 115582, 2021.

[53] G. Guo, K. Guffey, W. Chen, and P. Pergami, "Classification of normal and pathological gait in young children based on foot pressure data," *Neuroinformatics*, vol. 15, pp. 13–24, 2017.

[54] R. Kolaghassi, M. K. Al-Hares, and K. Sirlantzis, "Systematic review of intelligent algorithms in gait analysis and prediction for lower limb robotic systems," *IEEE Access*, vol. 9, pp. 113788–113812, 2021.

[55] G. Cicirelli, D. Impedovo, V. Dentamaro, R. Marani, G. Pirlo, and T. R. D'Orazio, "Human gait analysis in neurodegenerative diseases: a review," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 229–242, 2021.

[56] P. Khera and N. Kumar, "Role of machine learning in gait analysis: a review," *Journal of Medical Engineering & Technology*, vol. 44, no. 8, pp. 441–467, 2020.

[57] A. Aristotle, *On the gait of animals*. Virginia Tech, 2001.

[58] D. Levine, J. Richards, and M. W. Whittle, *Whittle's gait analysis*. Elsevier health sciences, 2012.

[59] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[60] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," *Advances in neural information processing systems*, vol. 27, 2014.

[61] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[62] S. Hochreiter and J. Schmidhuber, "Lstm can solve hard long time lag problems," *Advances in neural information processing systems*, vol. 9, 1996.

[63] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," *Neural computation*, vol. 12, no. 10, pp. 2451–2471, 2000.

[64] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, 2021.

[65] B. Takkouche and G. Norman, "Prisma statement," *Epidemiology*, vol. 22, no. 1, p. 128, 2011.

[66] S. Jiang, Y. Pang, D. Wang, Y. Yang, Z. Yang, Y. Yang, and T.-L. Ren, "Gait recognition based on graphene porous network structure pressure sensors for rehabilitation therapy," in *2018 IEEE International Conference on Electron Devices and Solid State Circuits (EDSSC)*, pp. 1–2, IEEE, 2018.

[67] J. D. Farah, N. Baddour, and E. D. Lemaire, "Design, development, and evaluation of a local sensor-based gait phase recognition system using a logistic model decision tree for orthosis-control," *Journal of neuroengineering and rehabilitation*, vol. 16, no. 1, pp. 1–11, 2019.

[68] A. Thongsook, T. Nunthawarasilp, P. Kraypet, J. Lim, and N. Ruangpayoongsak, "C4. 5 decision tree against neural network on gait phase recognition for lower limp exoskeleton," in *2019 First International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP)*, pp. 69–72, IEEE, 2019.

[69] C. Morbidoni, A. Cucchiarelli, S. Fioretti, and F. Di Nardo, "A deep learning approach to emg-based classification of gait phases during level ground walking," *Electronics*, vol. 8, no. 8, p. 894, 2019.

[70] B. Su and E. M. Gutierrez-Farewik, "Gait trajectory and gait phase prediction based on an lstm network," *Sensors*, vol. 20, no. 24, p. 7127, 2020.

[71] X. Wu, Y. Yuan, X. Zhang, C. Wang, T. Xu, and D. Tao, "Gait phase classification for a lower limb exoskeleton system based on a graph convolutional network model," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 5, pp. 4999–5008, 2021.

[72] B. Filtjens, A. Nieuwboer, N. D'cruz, J. Spildooren, P. Slaets, and B. Vanrumste, "A data-driven approach for detecting gait events during turning in people with parkinson's disease and freezing of gait," *Gait & Posture*, vol. 80, pp. 130–136, 2020.

[73] M. Z. Arshad, A. Jamsrandorj, J. Kim, and K.-R. Mun, "Gait events prediction using hybrid cnn-rnn-based deep learning models through a single waist-worn wearable sensor," *Sensors*, vol. 22, no. 21, p. 8226, 2022.

[74] B. Dumphart, D. Slijepcevic, M. Zeppelzauer, A. Kranzl, F. Unglaube, A. Baca, and B. Horsak, "Robust deep learning-based gait event detection across various pathologies," *Plos one*, vol. 18, no. 8, p. e0288555, 2023.

[75] F. Di Nardo, C. Morbidoni, A. Cucchiarelli, and S. Fioretti, "Recognition of gait phases with a single knee electrogoniometer: A deep learning approach," *Electronics*, vol. 9, no. 2, p. 355, 2020.

[76] Z.-Q. Ling, J.-C. Chen, G.-Z. Cao, Y.-P. Zhang, L.-L. Li, W.-X. Xu, and S.-B. Cao, "A domain adaptive convolutional neural network for semg-based gait phase recognition against to speed changes," *IEEE Sensors Journal*, vol. 23, no. 3, pp. 2565–2576, 2022.

[77] M. Devanne, H. Wannous, M. Daoudi, S. Berretti, A. Del Bimbo, and P. Pala, "Learning shape variations of motion trajectories for gait analysis," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, pp. 895–900, IEEE, 2016.

[78] S. S. Nair, R. M. French, D. Laroche, and E. Thomas, "The application of machine learning algorithms to the analysis of electromyographic patterns from arthritic patients," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 2, pp. 174–184, 2009.

[79] Y. Zhang and Y. Ma, "Application of supervised machine learning algorithms in the classification of sagittal gait patterns of cerebral palsy children with spastic diplegia," *Computers in biology and medicine*, vol. 106, pp. 33–39, 2019.

[80] L. Van Gestel, T. De Laet, E. Di Lello, H. Bruyninckx, G. Molenaers, A. Van Campenhout, E. Aertbeliën, M. Schwartz, H. Wambacq, P. De Cock, *et al.*, "Probabilistic gait classification in children with cerebral palsy: A bayesian approach," *Research in Developmental Disabilities*, vol. 32, no. 6, pp. 2542–2552, 2011.

[81] I. El Maachi, G.-A. Bilodeau, and W. Bouachir, "Deep 1d-convnet for accurate parkinson disease detection and severity prediction from gait," *Expert Systems with Applications*, vol. 143, p. 113075, 2020.

[82] C. Ramaker, J. Marinus, A. M. Stiggelbout, and B. J. Van Hilten, "Systematic evaluation of rating scales for impairment and disability in parkinson's disease," *Movement disorders: official journal of the Movement Disorder Society*, vol. 17, no. 5, pp. 867–876, 2002.

[83] A. S. Alharthi, A. J. Casson, and K. B. Ozanyan, "Gait spatiotemporal signal analysis for parkinson's disease detection and severity rating," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1838–1848, 2020.

[84] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[85] R. Kaur, R. W. Motl, R. Sowers, and M. E. Hernandez, "A vision-based framework for predicting multiple sclerosis and parkinson's disease gait dysfunctions-a deep learning approach," *IEEE Journal of Biomedical and Health Informatics*, 2022.

[86] J. Yu, S. Park, C. M. B. Ho, S.-H. Kwon, K.-H. Cho, and Y. S. Lee, "Ai-based stroke prediction system using body motion biosignals during walking," *The Journal of Supercomputing*, pp. 1–23, 2022.

[87] I. Chen, P.-Y. Yeh, T.-C. Chang, Y.-C. Hsieh, and C.-L. Chin, "Sarcopenia recognition system combined with electromyography and gait obtained by the multiple sensor module and deep learning algorithm," *Sensors and Materials*, vol. 34, no. 6, pp. 2403–2425, 2022.

[88] D. Thakur and S. Biswas, "Attention-based deep learning framework for hemiplegic gait prediction with smartphone sensors," *IEEE Sensors Journal*, vol. 22, no. 12, pp. 11979–11988, 2022.

[89] G. Shalin, S. Pardoel, E. D. Lemaire, J. Nantel, and J. Kofman, "Prediction and detection of freezing of gait in parkinson's disease from plantar pressure data using long short-term memory neural-networks," *Journal of neuroengineering and rehabilitation*, vol. 18, no. 1, pp. 1–15, 2021.

[90] H. El-ziaat, N. El-Bendary, and R. Moawad, "A hybrid deep learning approach for freezing of gait prediction in patients with parkinson's disease," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 4, 2022.

[91] B. Li, Y. Li, Y. Sun, X. Yang, X. Zhou, and Z. Yao, "A monitoring method of freezing of gait based on multimodal fusion," *Biomedical Signal Processing and Control*, vol. 82, p. 104589, 2023.

[92] B. Hu, P. Dixon, J. Jacobs, J. Dennerlein, and J. Schiffman, "Machine learning algorithms based on signals from a single wearable inertial sensor can detect surface-and age-related differences in walking," *Journal of biomechanics*, vol. 71, pp. 37–42, 2018.

[93] A. Girka, J.-P. Kulmala, and S. Äyrämö, "Deep learning approach for prediction of impact peak appearance at ground reaction force signal of running activity," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 23, no. 14, pp. 1052–1059, 2020.

[94] U. M. Hernandez, M. Awad, and A. Dehghani-Sanij, "Learning architecture for the recognition of walking and prediction of gait period using wearable sensors," *Neurocomputing*, vol. 470, pp. 1–10, 2022.

[95] Y.-L. Yen, S.-K. Ye, J. N. Liang, and Y.-J. Lee, "Recognition of walking directional intention employed ground reaction forces and center of pressure during gait initiation," *Gait & Posture*, vol. 106, pp. 23–27, 2023.

[96] Y. Ma, X. Wu, C. Wang, Z. Yi, and G. Liang, "Gait phase classification and assist torque prediction for a lower limb exoskeleton system using kernel recursive least-squares method," *Sensors*, vol. 19, no. 24, p. 5449, 2019.

[97] K. Tanghe, F. De Groote, D. Lefeber, J. De Schutter, and E. Aertbeliën, "Gait trajectory and event prediction from state estimation for exoskeletons during gait," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 1, pp. 211–220, 2019.

[98] A. M. Boudali, P. J. Sinclair, and I. R. Manchester, "Predicting transitioning walking gaits: Hip and knee joint trajectories from the motion of walking canes," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 9, pp. 1791–1800, 2019.

[99] H. Vallery, E. H. Van Asseldonk, M. Buss, and H. Van Der Kooij, "Reference trajectory generation for rehabilitation robots: complementary limb motion estimation," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 17, no. 1, pp. 23–30, 2008.

[100] M. Hassan, H. Kadone, T. Ueno, Y. Hada, Y. Sankai, and K. Suzuki, "Feasibility of synergy-based exoskeleton robot control in hemiplegia," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 6, pp. 1233–1242, 2018.

[101] M. Gholami, C. Napier, and C. Menon, "Estimating lower extremity running gait kinematics with a single accelerometer: A deep learning approach," *Sensors*, vol. 20, no. 10, p. 2939, 2020.

[102] A. Sharma and E. Rombokas, "Improving imu-based prediction of lower limb kinematics in natural environments using egocentric optical flow," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 699–708, 2022.

[103] R. Kolaghassi, M. K. Al-Hares, G. Marcelli, and K. Sirlantzis, "Performance of deep learning models in forecasting gait trajectories of children with neurological disorders," *Sensors*, vol. 22, no. 8, p. 2969, 2022.

[104] R. Kolaghassi, G. Marcelli, and K. Sirlantzis, "Deep learning models for stable gait prediction applied to exoskeleton reference trajectories for children with cerebral palsy," *IEEE Access*, vol. 11, pp. 31962–31976, 2023.

[105] M. Karakish, M. A. Fouz, and A. ELsawaf, "Gait trajectory prediction on an embedded microcontroller using deep learning," *Sensors*, vol. 22, no. 21, p. 8441, 2022.

[106] R. Kolaghassi, G. Marcelli, and K. Sirlantzis, "Effect of gait speed on trajectory prediction using deep learning models for exoskeleton applications," *Sensors*, vol. 23, no. 12, p. 5687, 2023.

[107] J. Camargo, A. Ramanathan, W. Flanagan, and A. Young, "A comprehensive, open-source dataset of lower limb biomechanics in multiple conditions of stairs, ramps, and level-ground ambulation and transitions," *Journal of Biomechanics*, vol. 119, p. 110320, 2021.

[108] Q. Song, X. Ma, and Y. Liu, "Continuous online prediction of lower limb joints angles based on semg signals by deep learning approach," *Computers in Biology and Medicine*, p. 107124, 2023.

[109] D. Senanayake, S. Halgamuge, and D. C. Ackland, "Real-time conversion of inertial measurement unit data to ankle joint angles using deep neural networks," *Journal of Biomechanics*, vol. 125, p. 110552, 2021.

[110] E. Rapp, S. Shin, W. Thomsen, R. Ferber, and E. Halilaj, "Estimation of kinematics from inertial measurement units using a combined deep learning and optimization framework," *Journal of Biomechanics*, vol. 116, p. 110229, 2021.

[111] V. Hernandez, D. Dadkhah, V. Babakeshizadeh, and D. Kulić, "Lower body kinematics estimation from wearable sensors for walking and running: A deep learning approach," *Gait & Posture*, vol. 83, pp. 185–193, 2021.

[112] M. S. B. Hossain, J. Dranetz, H. Choi, and Z. Guo, "Deepbbwae-net: A cnn-rnn based deep superlearner for estimating lower extremity sagittal plane joint kinematics using shoe-mounted imu sensors in daily living," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 8, pp. 3906–3917, 2022.

[113] J.-S. Tan, S. Tippaya, T. Binnie, P. Davey, K. Napier, J. Caneiro, P. Kent, A. Smith, P. O'Sullivan, and A. Campbell, "Predicting knee joint kinematics from wearable sensor data in people with knee osteoarthritis and clinical considerations for future machine learning models," *Sensors*, vol. 22, no. 2, p. 446, 2022.

[114] O. Heeb, A. Barua, C. Menon, and X. Jiang, "Building effective machine learning models for ankle joint power estimation during walking using fmg sensors," *Frontiers in neurorobotics*, vol. 16, 2022.

[115] M. S. B. Hossain, Z. Guo, and H. Choi, "Estimation of lower extremity joint moments and 3d ground reaction forces using imu sensors in multiple walking conditions: A deep learning approach," *IEEE Journal of Biomedical and Health Informatics*, 2023.

[116] T. Tan, D. Wang, P. B. Shull, and E. Halilaj, "Imu and smartphone camera fusion for knee adduction and knee flexion moment estimation during walking," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 2, pp. 1445–1455, 2022.

[117] J. He, Z. Guo, Z. Shao, J. Zhao, and G. Dan, "An lstm-based prediction method for lower limb intention perception by integrative analysis of kinect visual signal," *Journal of healthcare engineering*, vol. 2020, 2020.

[118] D. H. K. Chow, L. Tremblay, C. Y. Lam, A. W. Y. Yeung, W. H. W. Cheng, and P. T. W. Tse, "Comparison between accelerometer and gyroscope in predicting level-ground running kinematics by treadmill running kinematics using a single wearable sensor," *Sensors*, vol. 21, no. 14, p. 4633, 2021.

[119] M. Sharifi Renani, A. M. Eustace, C. A. Myers, and C. W. Clary, "The use of synthetic imu signals in the training of deep learning models significantly improves the accuracy of joint kinematic predictions," *Sensors*, vol. 21, no. 17, p. 5876, 2021.

[120] S. Dey and A. F. Schilling, "A function approximator model for robust online foot angle trajectory prediction using a single imu sensor: Implication for controlling active prosthetic feet," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 2, pp. 1467–1475, 2022.

[121] G. Ding, A. Plummer, and I. Georgilas, "Deep learning with an attention mechanism for continuous biomechanical motion estimation across varied activities," *Frontiers in Bioengineering and Biotechnology*, vol. 10, 2022.

[122] S.-W. Lee and A. Asbeck, "A deep learning-based approach for foot placement prediction," *IEEE Robotics and Automation Letters*, 2023.

[123] R. Bajpai and D. Joshi, "Foot2hip: A deep neural network model for predicting lower limb kinematics from foot measurements," *IEEE/ASME Transactions on Mechatronics*, 2023.

[124] Ç. Berke Erdaş, E. Sümer, and S. Kibaroğlu, "Cnn-based severity prediction of neurodegenerative diseases using gait data," *Digital Health*, vol. 8, p. 20552076221075147, 2022.

[125] K. Eguchi, I. Takigawa, S. Shirai, I. Takahashi-Iwata, M. Matsushima, T. Kano, H. Yaguchi, and I. Yabe, "Gait video-based prediction of unified parkinson's disease rating scale score: a retrospective study," *BMC neurology*, vol. 23, no. 1, p. 358, 2023.

[126] A. Nait Aicha, G. Englebienne, K. S. Van Schooten, M. Pijnappels, and B. Kröse, "Deep learning to predict falls in older adults based on daily-life trunk accelerometry," *Sensors*, vol. 18, no. 5, p. 1654, 2018.

[127] S. Yu, H. Chen, R. Brown, and S. Sherman, "Motion sensor-based assessment on fall risk and parkinson's disease severity: A deep multi-source multi-task learning (dmml) approach," in *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 174–179, IEEE, 2018.

[128] S. Zhang, Y. Wang, and A. Li, "Gait-based age estimation with deep convolutional neural network," in *2019 International Conference on Biometrics (ICB)*, pp. 1–8, IEEE, 2019.

[129] K. Aoki, H. Nishikawa, Y. Makihara, D. Muramatsu, N. Takemura, and Y. Yagi, "Physical fatigue detection from gait cycles via a multi-task recurrent neural network," *IEEE Access*, vol. 9, pp. 127565–127575, 2021.

[130] W. Rueangsirarak, J. Zhang, N. Aslam, E. S. Ho, and H. P. Shum, "Automatic musculoskeletal and neurological disorder diagnosis with relative joint displacement from human gait," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 12, pp. 2387–2396, 2018.

[131] C. Xu, Y. Makihara, G. Ogi, X. Li, Y. Yagi, and J. Lu, "The ou-isir gait database comprising the large population dataset with age and performance evaluation of age estimation," *IPSJ Transactions on Computer Vision and Applications*, vol. 9, no. 1, pp. 1–14, 2017.

[132] F. Sherratt, "Bath natural environment har data set," Dec 2020.

[133] T. D. Collins, S. N. Ghoussayni, D. J. Ewins, and J. A. Kent, "A six degrees-of-freedom marker set for gait analysis: repeatability and comparison with a modified helen hayes set," *Gait & posture*, vol. 30, no. 2, pp. 173–180, 2009.

[134] F. Han, B. Reily, W. Hoff, and H. Zhang, "Space-time representation of people based on 3d skeletal data: A review," *Computer Vision and Image Understanding*, vol. 158, pp. 85–105, 2017.

[135] T.-N. Nguyen and J. Meunier, "Walking gait dataset: point clouds, skeletons and silhouettes," *DIRO, University of Montreal, Tech. Rep*, vol. 1379, 2018.

[136] M. Khokhlova, C. Migniot, A. Morozov, O. Sushkova, and A. Dipanda, "Normal and pathological gait classification lstm model," *Artificial intelligence in medicine*, vol. 94, pp. 54–66, 2019.

[137] K. Jun, Y. Lee, S. Lee, D.-W. Lee, and M. S. Kim, "Pathological gait classification using kinect v2 and gated recurrent neural networks," *IEEE Access*, vol. 8, pp. 139881–139891, 2020.

[138] M. Bachlin, M. Plotnik, D. Roggen, I. Maidan, J. M. Hausdorff, N. Giladi, and G. Troster, "Wearable assistant for parkinson's disease patients with the freezing of gait symptom," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 436–446, 2009.

[139] F. Kohnehshahri, A. Merlo, D. Mazzoli, M. Bò, and R. Stagni, "Machine learning applied to gait analysis data in cerebral palsy and stroke: A systematic review," *Gait & Posture*, vol. 111, 04 2024.

[140] S. Cardarelli, A. Mengarelli, A. Tigrini, A. Strazza, F. D. Nardo, S. Fioretti, and F. Verdini, "Single imu displacement and orientation estimation of human center of mass: A magnetometer-free approach," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, pp. 5629–5639, 2020.

[141] S. Boudaoud, *Analyse de la variabilité de forme des signaux : Application aux signaux électrophysiologiques*. PhD thesis, Université de Nice-Sophia Antipolis, 2006.

[142] E. Desailly, Y. Daniel, P. Sardain, and P. Lacouture, "Foot contact event detection using kinematic data in cerebral palsy children and normal adults gait," *Gait & posture*, vol. 29, no. 1, pp. 76–80, 2009.

[143] M. H. Schwartz and A. Rozumalski, "The gait deviation index: a new comprehensive index of gait pathology," *Gait & posture*, vol. 28, no. 3, pp. 351–357, 2008.

[144] C. De Boor, "Calculation of the smoothing spline with weighted roughness measure," *Mathematical Models and Methods in Applied Sciences*, vol. 11, no. 01, pp. 33–41, 2001.

[145] "Notice patient - DYSPORT 500 UNITES SPEYWOOD, poudre pour solution injectable - Base de données publique des médicaments."

[146] H. Tribouillard, *Elaboration d'une methodologie de validation des indications hors autorisation de mise sur le marche de la toxine botulique de type A au centre hospitalo-universitaire de Lille : Exemple du bavage chez l'adulte*. PhD thesis, Faculte de Pharmacie, Universite de Lille, 2018.

[147] D. Battagli, *Utilisation des toxines botuliques aux Hospices Civils de Lyon : Bilan 2015 des indications*. PhD thesis, Faculte de Pharmacie, Universite Claude Bernard - Lyon 1, 2017.

[148] C. Zhu, Q. Liu, W. Meng, Q. Ai, and S. Q. Xie, "An attention-based cnn-lstm model with limb synergy for joint angles prediction," in *2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pp. 747–752, IEEE, 2021.

[149] A. Zaroug, D. T. Lai, K. Mudie, and R. Begg, "Lower limb kinematics trajectory prediction using long short-term memory neural networks," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 362, 2020.

[150] L. Jia, Q. Ai, W. Meng, Q. Liu, and S. Q. Xie, "Individualized gait trajectory prediction based on fusion lstm networks for robotic rehabilitation training," in *2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pp. 988–993, IEEE, 2021.

[151] D.-X. Liu, X. Wu, C. Wang, and C. Chen, "Gait trajectory prediction for lower-limb exoskeleton based on deep spatial-temporal model (dstm)," in *2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM)*, pp. 564–569, IEEE, 2017.

[152] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp. 249–256, JMLR Workshop and Conference Proceedings, 2010.

[153] H. Salehinejad, J. Baarbe, S. Sankar, J. Barfett, E. Colak, and S. Valaee, "Recent advances in recurrent neural networks," *ArXiv*, vol. abs/1801.01078, 2018.

[154] M. Schuster and K. Paliwal, "Bidirectional recurrent neural networks," *Signal Processing, IEEE Transactions on*, vol. 45, pp. 2673 – 2681, 12 1997.

[155] T. A. Niiler, J. G. Richards, F. Miller, J. Q. Sun, and P. Castagno, "Reliability of predictions of post-operative gait in rectus transfer patients using FFT neural networks," in *Proceedings of the Gait & Clinical Movement Analysis Society Conference*, 1999.

[156] T. A. Niiler, *Efficacy of predictions of post-operative gait in rectus transfer patients using neural networks*. PhD thesis, University of Delaware, 2001.

[157] S. Klejman, J. Andrysek, A. Dupuis, and V. Wright, "Test-Retest Reliability of Discrete Gait Parameters in Children With Cerebral Palsy," *Archives of Physical Medicine and Rehabilitation*, vol. 91, pp. 781–787, May 2010.

[158] O. Galarraga, A. Le Saux, and F. Hareb, "O 083-a new method for computing gait deviation scores in hemiparesis," *Gait & Posture*, vol. 65, pp. 170–171, 2018.

[159] A. KingaD, "A methodforstochasticoptimization," *Anon. InternationalConferenceon Learning Representations. SanDego: ICLR*, vol. 7, 2015.

[160] C. M. Bishop, *Pattern Recognition And Machine Learning*. New York: Springer-Verlag, 2nd ed., 2006.

[161] D. G. Altman and J. M. Bland, "Standard deviations and standard errors," *Bmj*, vol. 331, no. 7521, p. 903, 2005.

[162] G. James, D. Witten, T. Hastie, R. Tibshirani, *et al.*, *An introduction to statistical learning*, vol. 112. Springer, 2013.

[163] O. C. Galarraga, V. Vigneron, N. Khouri, B. Dorizzi, and E. Desailly, "Predictive simulation of surgery effect on cerebral palsy gait," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 20, pp. S85–S86, Oct. 2017.

[164] T. A. Niiler, J. G. Richards, and F. Miller, "Concurrent surgeries are a factor in predicting success of rectus transfer outcomes," *Gait & Posture*, vol. 26, pp. 76–81, June 2007.

[165] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[166] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.

[167] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[168] J. Cheng, L. Dong, and M. Lapata, "Long short-term memory-networks for machine reading," *arXiv preprint arXiv:1601.06733*, 2016.

[169] Ł. Kaiser and S. Bengio, "Can active memory replace attention?," *Advances in Neural Information Processing Systems*, vol. 29, 2016.

[170] A. P. Parikh, O. Täckström, D. Das, and J. Uszkoreit, "A decomposable attention model for natural language inference," *arXiv preprint arXiv:1606.01933*, 2016.

[171] R. Paulus, C. Xiong, and R. Socher, "A deep reinforced model for abstractive summarization," *arXiv preprint arXiv:1705.04304*, 2017.

[172] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[173] O. A. Galarraga C., N. Khouri, V. Vigneron, B. Dorizzi, and E. Desailly, "Predicting kinematic outcome of multi-level surgery in cerebral palsy," *Gait & Posture*, vol. 49, p. 1, Sept. 2016.

[174] R. Caruana, "Multitask learning," *Machine learning*, vol. 28, no. 1, pp. 41–75, 1997.

[175] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 1440–1448, 2015.

[176] L. Deng, G. Hinton, and B. Kingsbury, "New types of deep neural network learning for speech recognition and related applications: An overview," in *2013 IEEE international conference on acoustics, speech and signal processing*, pp. 8599–8603, IEEE, 2013.

[177] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in *Proceedings of the 25th international conference on Machine learning*, pp. 160–167, 2008.

[178] B. Ramsundar, S. Kearnes, P. Riley, D. Webster, D. Konerding, and V. Pande, "Massively multitask networks for drug discovery," *arXiv preprint arXiv:1502.02072*, 2015.

[179] R. Caruana, "Multitask learning: A knowledge-based source of inductive bias1," in *Proceedings of the Tenth International Conference on Machine Learning*, pp. 41–48, Citeseer, 1993.

[180] J. Baxter, "A bayesian/information theoretic model of learning to learn via multiple task sampling," *Machine learning*, vol. 28, pp. 7–39, 1997.

Titre: Apprentissage profond pour la simulation prédictive de la marche et les bénéfices fonctionnels post-traitement dans les maladies neurologiques

Mots clés: Apprentissage profond, mémoire à long terme, apprentissage multitâche, analyse quantifiée de la marche, rééducation de la marche et mécanisme d'attention

Résumé: Les affections neurologiques se manifestent souvent par des troubles de la marche, fréquemment liés à la spasticité. Injections de toxine botulique de type A (BTX-A) traitent généralement les problèmes de démarche liés à la spasticité. Il reste crucial d'obtenir des résultats thérapeutiques optimaux avec un rapport bénéfice-risque favorable. Les améliorations cinématiques obtenues par ce traitement sont parfois très efficaces, mais elles restent pour l'instant difficilement prévisibles. L'objectif de cette thèse est d'utiliser des techniques d'apprentissage profond (DL) pour simuler l'impact du traitement BTX-A sur les paramètres de marche. Le simulateur vise à afficher le résultat de marche le plus probable, améliorant ainsi le processus de prise de décision lors du traitement par BTX-A. La base de données comprenait 43 adultes diagnostiqués avec diverses maladies, notamment la CP, la SEP, le TBI, la LME et l'accident vasculaire cérébral. Chaque participant avait subi au moins une analyse clinique de la marche (CGA) avant et après avoir reçu le traitement. Le laboratoire UGCEAM a obtenu et traité les données cinématiques de la marche. Diverses techniques de régression ont été utilisées, notamment LSTM, BiLSTM, le mécanisme d'attention, l'apprentissage d'ensemble et l'apprentissage multitâche (MTL). Les méthodes évaluées et leur efficacité ont été comparées entre elles et à des approches alternatives documentées dans la littérature. Cette étude est la première à simuler quantitativement l'impact du traitement au BTX-A sur la démarche d'adultes atteints de diverses maladies. Il explore un large éventail de combinaisons de traitements et différents modèles de démarche.

Title: Deep Learning for Predictive Simulation of the Gait and Post-Treatment Functional Benefits in Neurological Diseases

Keywords: Deep Learning, Long-Short Term Memory, Multi-task Learning, Quantified Gait Analysis, Gait Rehabilitation, and Attention Mechanism

Abstract: Neurological conditions often manifest as gait disorders, frequently linked to spasticity. Botulinum Toxin Type A (BTX-A) injections commonly treat spasticity-related gait issues. Achieving optimal treatment outcomes with a favourable benefit-risk ratio remains crucial. Kinematic improvements obtained by this treatment are sometimes very efficient, but at this moment they remain difficultly predictable. The aim of this thesis is to employ deep learning (DL) techniques to simulate the impact of BTX-A treatment on gait parameters. The simulator aims to display the most probable gait result, enhancing the process of decision-making in BTX-A treatment. The database consisted of 43 adults diagnosed with various diseases, including CP, MS, TBI, SCI, and stroke. Each participant had undergone at least one clinical gait analysis (CGA) both before and after receiving treatment.The UGCEAM laboratory obtained and processed kinematic gait data. Various regression techniques were employed, including LSTM, BiLSTM, attention mechanism, ensemble learning, and Multi-task learning (MTL). The evaluated methods and their efficacy were compared both amongst themselves and to alternative approaches documented in the literature. This study is the first to quantitatively simulate the impact of BTX-A treatment on the gait of adults with various diseases. It explores a wide range of treatment combinations and different gait patterns.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France