TÜRK-ALMAN ÜNİVERSİTESİ
TÜRKISCH-DEUTSCHE UNIVERSITÄT

T.C.

TURKISH-GERMAN UNIVERSITY

INSTITUTE OF THE GRADUATE STUDIES

IN SCIENCE AND ENGINEERING

Prediction of Gait Kinetics from Joint Angles Using Machine Learning in Patients with Cerebral Palsy

Doctoral Thesis

Mustafa Erkam ÖZATEŞ

ISTANBUL 2023

TÜRK-ALMAN ÜNİVERSİTESİ
TÜRKISCH-DEUTSCHE UNIVERSITÄT

T.C.

TURKISH-GERMAN UNIVERSITY

INSTITUTE OF THE GRADUATE STUDIES

IN SCIENCE AND ENGINEERING

Prediction of Gait Kinetics from Joint Angles Using Machine Learning in Patients with Cerebral Palsy

Doctoral Thesis

Mustafa Erkam ÖZATEŞ

B.Sc., Control and Automation Engineering, Istanbul Technical University, 2014

M.Sc., Railway Systems Engineering, Istanbul Technical University, 2017

Advisor
Prof. Dr. Yunus Ziya ARSLAN

Co-Advisor
Apl. Prof. Dr. rer. nat. Sebastian Immanuel Wolf

Submitted to the Institute of the Graduate Studies in
Science and Engineering in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy

ISTANBUL 2023

Prediction of Gait Kinetics from Joint Angles Using Machine Learning in Patients with Cerebral Palsy

APPROVED BY:

Prof. Dr. Yunus Ziya ARSLAN (Thesis Advisor)	.....
Assoc. Prof. Dr. Yaşar Mahsut DİNÇEL Tekirdağ Namık Kemal University	.....
Assoc. Prof. Dr. Ali Fuat ERGENÇ Istanbul Technical University	.....
Assist. Prof. Dr. Mehmet Gökhan HABİBOĞLU Turkish German University	.....
Assist. Prof. Dr. Mehmet İPEKOĞLU Turkish German University	.....

DATE OF APPROVAL: 22 June 2023

ACKNOWLEDGMENTS

I would like to thank my supervisor Prof. Dr. Yunus Ziya Arslan for his big support at each step of my Ph.D. and for believing in my success from the very beginning. I also would like to thank my co-supervisor apl. Prof. Dr. rer. nat. Sebastian Wolf for his deep interest and support on my study.

I am also grateful to the motivation I have received from my dear professors in my department of Electrical and Electronics Engineering.

I can't repay my little but growing family's kindness. As an engineer, I won’t be trying to emphasize it, since it is inefficient to implement it hereby and far from being possible.

DECLARATION OF AUTHENTICITY

I declare that I completed the PhD thesis independently and used only the materials that are listed. All materials used, from published as well as unpublished sources, whether directly quoted or paraphrased, are duly reported. Furthermore, I declare that the master's / doctoral thesis, or any abridgment of it, was not used for any other degree-seeking purpose and give the publication rights of the thesis to the Institute of the Graduate Studies in Science and Engineering, Turkish-German University.

Mustafa Erkam ÖZATEŞ

ABSTRACT

Prediction of Gait Kinetics from Joint Angles Using Machine Learning in Patients with Cerebral Palsy

Joint moment and ground reaction force (GRF) during gait provide valuable information for clinical decision-making in patients with cerebral palsy (CP). Joint moments are calculated based on GRF using inverse dynamics models. Obtaining GRF from patients with CP is challenging. Typically developed (TD) individuals' joint moments and GRFs were predicted from joint angles using machine learning (ML), but no such study has been conducted on patients with CP. Accordingly, we aimed to predict the vertical GRF, dorsi-plantar flexion, knee flexion-extension, hip flexion-extension, and hip adduction-abduction moments based on the trunk, pelvis, hip, knee, and ankle kinematics during gait in patients with CP and TD individuals using one-dimensional convolutional neural networks. The anonymized retrospective gait data of CP and TD subjects were used. The data were collected in the course of patient care over the last two decades in the Department of Orthopedics and Traumatology of Heidelberg University. For broadening the ML study, we trained specific ML models of ridge regression, random forest, multilayer neural network, k-nearest neighbour, long short term memory neural network algorithms by using manually extracted time domain features and automatically generated features of gait kinematics. Their performances were evaluated and compared using isolated test subject groups based on normalized root mean square error (nRMSE) and Pearson correlation coefficient (PCC). Joint moments were predicted with nRMSE between 18.02% and 13.58% for the CP and between 12.55% and 8.58% for the TD groups, whereas with PCC between 0.90 and 0.96 for the CP and between 0.96 and 0.99 for the TD groups. GRF was predicted with an nRMSE of 7.47% for TD subjects and 11.75% for CP subjects, while with a PCC score of 0.98 for the TD and 0.94 for the CP group. ML algorithms using time domain features and automatically generated features showed similar performance. ML-based joint moment prediction from kinematics could replace conventional moment calculation in CP patients in the future, but the current level of prediction errors restricts its use for clinical decision-making today.

Keywords: *Machine Learning; Cerebral palsy; Gait; Ground reaction force; Joint moment.*

ÖZET

Serebral Palsili Hastalarda Makine Öğrenmesi Kullanarak Eklem Açılarından Yürüme Kinetiğinin Kestirilmesi

Serebral palsili (SP) hastalarda yürüme sırasında eklem momentleri ve yer reaksiyon kuvveti (YTK), klinik karar verme sürecinde değerli bilgiler sağlar. Eklem momentleri, ters kinematik modeller kullanılarak YTK'ne dayalı olarak hesaplanır. SP'li hastalardan YTK elde etmek zorlu bir süreçtir. Sağlıklı bireylerde eklem momentleri ve YTK'leri, makine öğrenmesi (MÖ) kullanılarak eklem açılarından tahmin edilmiştir, ancak SP'li hastalar üzerinde böyle bir çalışma henüz yapılmamıştır. Bu nedenle, bu çalışmada SP'li hastalarda ve sağlıklı bireylerde yürüme sırasında gövde, pelvis, kalça, diz ve ayak bileği kinematiği temel alınarak dikey YTK, dorsi-plantar fleksiyon, diz fleksiyon-ekstansiyon, kalça fleksiyon-ekstansiyon ve kalça addüksiyon-abdüksiyon momentlerinin bir boyutlu konvolüsyonel sinir ağları kullanılarak tahmin edilmesi amaçlandı. SP ve sağlıklı deneklerin anonimleştirilmiş geriye dönük yürüme verileri kullanıldı. Veriler, Heidelberg Üniversitesi, Ortopedi ve Travmatoloji Bölümü'nde son yirmi yıl içinde hastaların klinik çalışmaları sırasında toplanmıştır. MÖ çalışmasını genişletmek için, yürüme kinematiğinin zaman alanı özniteliklerinin manuel ve otomatik olarak çıkarılması yöntemleri kullanılarak ridge regresyon, rastgele karar ormanı, çok katmanlı sinir ağı, k-en yakın komşu, uzun-kısa süreli bellek sinir ağı algoritmalarının özelleştirilmiş MÖ modelleri eğitildi. Performansları, normalize edilmiş karesel ortalama hata (nKOH) ve Pearson korelasyon katsayısı (PCC) kullanılarak izole edilmiş test örneği gruplarında değerlendirildi ve karşılaştırıldı. Eklem momentleri, SP grubu için %18,02 ila %13,58 arasında nKOH ve sağlıklılar grubu için %12,55 ila %8,58 arasında nKOH ile tahmin edildi, SP grubu için 0,90 ila 0,96 arasında PCC ve sağlıklılar grubu için 0,96 ila 0,99 arasında PCC ile tahmin edildi. YTK, sağlıklı bireyler için %7,47 nKOH ve SP'li hastalar için %11,75 nKOH ile tahmin edildi, sağlıklılar grubu için 0,98 PCC ve SP grubu için 0,94 PCC ile tahmin edildi. Makine öğrenimi algoritmaları, manuel ve otomatik çıkarılmış öznitelikleri kullanarak benzer performans sergiledi. Makine öğrenimine dayalı eklem momenti tahmini, SP'li hastalarda geleneksel moment hesaplamanın yerine geçebilir, ancak mevcut düzeydeki tahmin hataları, klinik karar vermede bugün kullanımını sınırlamaktadır.

Anahtar Sözcükler: *Makine öğrenmesi; Serebral palsi; Yürüyüş; Yer tepki kuvveti; Eklem momenti.*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

LIST OF ABBREVIATIONS

CP	Cerebral palsy
TD	Typically developed
GRF	Ground reaction force
ML	Machine learning
1D-CNN	One dimensional convolutional neural network
LSTM	Long short term memory network
R-Regr	Ridge regression
KNN	K-nearest neighbor
RF	Random forest
MLNN	Multilayer neural network

1. INTRODUCTION

1.1 Background

Cerebral Palsy (CP), which is characterized by a variety of neurological and motor impairments, encompasses a diverse group of disorders that significantly impact an individual's neuromotor functions. The most obviously disordered neuromotor function of CP patients is their natural walking, namely gait.

Gait analysis, recognized as a valuable adjunctive tool in the realm of clinical decision-making, holds a crucial role in the assessment and management of various neurological and musculoskeletal conditions. Its primary objective is to identify and characterize gait abnormalities, providing healthcare professionals with comprehensive insights into an individual's walking pattern.. Additionally, gait analysis serves as a reliable means to continually monitor and evaluate the progress and effectiveness of treatment interventions implemented over time, facilitating a more personalized and targeted approach to patient care [1]-(Baker, 2013).

Motion capturing encompasses a comprehensive technology that enables to acquire a substantial volume of high-dimensional 3D kinematics and kinetics data during gait. This wealth of information is obtained through meticulous post-processing of images captured by precisely calibrated cameras and sensors. By harnessing the power of advanced imaging technology, gait analysis facilitates the acquisition of precise and detailed measurements, enabling healthcare professionals to gain deeper insights into the intricate biomechanics of an individual's walking pattern [2]-(Halilaj et al., 2018).

In clinical settings, the 3D data acquired through motion capturing is used for gait analysis in order to derive kinematic and kinetic data of the subject. Human body models like Plugin Gait Model (Oxford Metrics, Oxford, UK) is used for this purpose. The main kinematic information gathered in gait analysis is joint angles.

1.2 Problem and importance of the problem

In addition to the joint angles in evaluating, tracking, and managing CP, particular attention is also directed towards the ground reaction force (GRF) and joint moments. Joint moment is a significant gait parameter as it plays a pivotal role in the comprehensive assessment, ongoing monitoring, and effective therapeutic interventions for individuals affected by CP [3,4,5,6,7]-(Lai et al., 1988; Gage, 1994; Ounpuu et al., 1996; Lin et al., 2000; Novacheck and Gage, 2007).

Ground reaction force (GRF) stands as a distinctive and essential kinetic parameter within the realm of gait analysis. What sets GRF apart is its ability to be directly measured through experimental means. By employing force plates clinicians can precisely capture the forces exerted between the foot and the ground during walking or running. This direct experimental measurement of GRF plays a crucial role in unraveling the intricate biomechanics of human locomotion, providing valuable insights into the distribution, timing, and magnitude of forces acting on the body.

The joint moments offer tremendous potential in unraveling the intricate dynamics of muscle behaviors during joint motion exertion. Notably, research has demonstrated that the analysis of joint moments, both pre- and postoperatively, assumes a crucial role in

informing the decision-making process regarding the treatment of CP. Specifically, the examination of joint moments within the sagittal plane of the lower extremities, as well as the hip joint moment within the frontal plane, has been shown to exert a significant influence on treatment strategies [5,8,9,10]-(Ounpuu et al., 1996; De Luca et al., 1997; Kay et al., 2000, Rhodes et al., 2023). By incorporating these valuable insights, healthcare professionals are equipped with a more comprehensive understanding of the biomechanical intricacies underlying CP, enabling them to make informed and personalized decisions to optimize the treatment outcomes for individuals affected by this condition.

A prominent example of the significance of joint moments in clinical decision-making lies in the context of crouch gait observed in patients with CP. In this case, the insufficient strength of the quadriceps muscle group emerges as a primary contributing factor [11]-(Lenhart et al., 2017). Consequently, the magnitude and pattern of the knee extension moment play a pivotal role in reflecting the impact of quadriceps weakness on crouch gait. The careful analysis of these knee extension moments assumes critical importance when making surgical decisions for individuals presenting with CP and crouch gait. By thoroughly evaluating the knee extension moments, healthcare professionals gain crucial insights into the extent of quadriceps weakness, enabling them to make informed surgical choices aimed at addressing this specific issue and optimizing the gait biomechanics and functional outcomes for patients affected by CP [11,12]-(Lenhart et al., 2017; Karabulut et al., 2021).

Joint moments also provide crucial information about the joint mechanics that may be contributing to gait abnormalities in patients and for tailoring surgical interventions to address specific issues and improve gait mechanics. For instance, if a patient with CP exhibits excessive knee flexion during the stance phase of gait, surgical interventions such as lengthening the hamstrings or transferring the rectus femoris can be carried out to improve knee extension [10]-(Rhodes et al., 2023). Similarly, if a patient exhibits excessive hip adduction during the stance phase of gait, surgical interventions such as pelvic osteotomy or soft tissue release can be carried out to improve hip alignment and reduce the risk of hip dislocation [30]-(Miller, 2020).

Obtaining joint moments and ground reaction forces in clinical gait analysis leads to challenges. The measurement of ground reaction force (GRF) presents inherent challenges when attempting to capture it during natural walking [13]-(Caldas et al., 2020). Furthermore, these difficulties are compounded when dealing with deviated gaits, particularly in cases of CP [14]-(White et al., 1999). The complexity lies in accurately quantifying GRF in real-time while individuals exhibit variations in gait patterns, such as altered foot placements and asymmetrical weight distribution. Such deviations pose significant obstacles to obtaining precise and reliable GRF measurements.

On the other hand, the calculation of joint moments necessitates the utilization of GRF measurements. By employing the principles of inverse dynamics, the forces and moments acting on the body segments can be determined based on the measured GRF data [15,16]-(Winter, 2009; Whittle, 2014). This computational approach allows for the

computation of the joint moments at various joints throughout the body, providing valuable insights into the dynamic forces and torques experienced during different phases of the gait cycle. In addition to the problematic acquisition process of GRF measurements, these models are not perfect and especially far from being so for widely varying CP cases.

Machine Learning (ML) has emerged as a powerful tool in addressing various challenges, especially when dealing with tasks that involve missing measurements or the absence of comprehensive physical models. ML techniques excel in leveraging patterns, correlations, and underlying structures within data to make predictions, classifications, and decisions.

In situations where measurements are missing or incomplete, ML algorithms can effectively analyze the available data and uncover hidden relationships. By training on existing data, ML models can learn patterns and generalize from them, enabling them to make informed predictions or fill in missing information. This ability to handle missing measurements is particularly valuable when dealing with complex systems or scenarios where data collection might be limited or challenging.

Furthermore, ML can be applied when there is a lack of well-defined physical models. Instead of relying solely on explicit equations or models, ML algorithms can learn from data to infer relationships, capture nonlinearities, and make accurate predictions. This flexibility makes ML a versatile tool for solving problems across various domains, from image and speech recognition to natural language processing and complex system optimization.

By harnessing the power of ML, researchers and practitioners can overcome the limitations posed by missing measurements or the absence of physical models. ML enables them to leverage available data effectively, uncover hidden patterns, and make reliable predictions or decisions, thus contributing to advancements in fields where traditional approaches may fall short.

1.3 Aim and importance of the study

The exploration of predicting GRF during gait using (ML) techniques has been an active research area. One of the pioneering attempts in this field dates back to 2013, where conventional ML algorithms were employed for GRF prediction based on kinematics [17]-(Oh et al., 2013). Subsequently, other studies have further expanded on this research, investigating statistical approaches [18]-(Johnson et al., 2018) as well as deep learning algorithms utilizing motion capture data [19]-(Mundt et al., 2020a) and incorporating spatio-temporal information [20]-(Johnson et al., 2018).

ML algorithms have also been successfully applied to patients with CP, a condition characterized by non-uniform gait characteristics. Researchers have leveraged ML techniques for various tasks, including the detection of CP disease using video recordings [21]-(Ihlen et al., 2020) or gait kinematics [22]-(Zhang and Ye, 2019), as well as the classification of gait phases in CP patients using electromyography (EMG) [23]-(Morbidoni et al., 2021) or marker data [24]-(Kim et al., 2022). These applications demonstrate the potential of ML in assisting with the diagnosis, assessment, and classification of gait abnormalities in CP patients, providing valuable insights for clinical decision-making and personalized treatment interventions.

With the typically developed (TD) subjects, a few attempts have been made to predict joint moments using ML during gait. For example, the above mentioned study of Mundt et al. not only predicts the GRF but also the joint moments of TD subjects successfully using kinematic data from three-dimensional motion capture by employing feed-forward neural networks and long short-term memory neural networks [19]-(Mundt et al., 2020a). Another study published by Ardestani et al. introduced a wavelet neural network that considered frequency information to predict joint moments of TD subjects using both kinematic and EMG data [25]-(Ardestani et al., 2014).

This highlighted a research gap and presented an opportunity for the investigations in this thesis to apply ML techniques to predict GRF and joint moments in individuals with CP, which will enhance clinicians' understanding of CP's gait mechanics and contribute to personalized treatment interventions.

In this thesis, our main objective was to predict *i*) GRF and specific joint moments, including dorsi-plantar flexion, knee flexion-extension, hip flexion-extension, and hip adduction-abduction moments, using joint angles derived from marker data during gait in patients with CP.

Furthermore, we aimed to compare the performance of different input-algorithm settings for predicting these variables. The input sets comprised various types of kinematic information such as kinematic features and kinematic signals, and we carefully selected and developed appropriate conventional and deep ML algorithms accordingly. By exploring different combinations of input data and ML algorithms, we aimed to identify

the most effective approach for predicting the GRF and joint moments in patients with CP during gait.

Such a comprehensive analysis allowed us to evaluate and compare the predictive capabilities of different input variables and ML techniques. The findings from this study would contribute to advancing our understanding of the relationships between kinematic information, GRF, and joint moments in CP patients, and provide valuable insights into the optimal approaches for predicting these parameters.

1.4 Original contributions

To the best of our knowledge, prior to this thesis and the associated publications, no other study has specifically focused on predicting GRF or joint moments in patients with CP using ML algorithms. The prediction of such kinetic parameters in CP patients during gait represents a novel and important research area, as it can provide valuable insights into their biomechanical characteristics and aid in clinical decision-making.

By conducting this thesis and the associated publications, we addressed the existing research gap and contribute to the field by exploring the prediction of GRF and joint moments in CP patients. The unique challenges posed by CP, including non-uniform gait characteristics, require specific attention and tailored approaches for accurate prediction. Therefore, our work represents an important contribution to the understanding and application of ML techniques in the context of CP gait analysis.

The results and findings obtained from this thesis and the associated publications will help advance the knowledge and understanding of gait mechanics in CP patients, paving the way for future research and clinical applications in this area. The identification of effective prediction models and approaches for GRF and joint moments can have significant implications for treatment planning, rehabilitation strategies, and

personalized interventions for individuals with CP.

1.5 Organization of the Thesis

There are three concluded works within this thesis, namely

- Study I: Predicting ground reaction forces using one-dimensional convolutional neural network based on kinematics during gait in patients with cerebral palsy (*paper in preparation*).
- Study II: Deep learning-based prediction of joint moments based on kinematics in patients with cerebral palsy (*published: Ozates, Musta Erkam, et al. "Machine learning-based prediction of joint moments based on kinematics in patients with cerebral palsy."* ***Journal of Biomechanics*** *(2023): 111668*.)
- Study III: Predicting joint moments in patients with cerebral palsy using deep and various conventional machine learning methods (*paper in preparation*).

The thesis is organized in the same order both in methodology and results sections.

2. MATERIALS AND METHODS

2.1 Subjects

The studies involved in this thesis received ethical approval from the local ethical committee of the University Hospital of Heidelberg (S-227/2021), ensuring that it adhered to the necessary ethical guidelines and considerations. For the research, a comprehensive dataset of gait data was utilized, consisting of anonymized retrospective information. The dataset included 329 TD subjects with typical gait characteristics and 917 patients with CP. The TD subjects had an average age of 26 years (±14), a mass of 70kg (±15), and a height of 167cm (±89). On the other hand, the CP patients had an average age of 17 years (±9), a mass of 47kg (±19), and a height of 153cm (±36). The data was obtained from routine patient care, ensuring a real-world and clinically relevant context.

To capture the kinematic information, the Plugin Gait Model (Oxford Metrics, Oxford, UK) was utilized, which involved the placement of 19 markers on the subjects. The data was captured using a 12-camera motion capture system (Vicon Motion Systems Ltd., Oxfordshire, UK) while the subjects walked at their self-selected speed. Simultaneously, GRF data was collected using force plates (Kistler Instruments, Winterthur, Switzerland). To calculate joint moments, an inverse kinematics model, based on the work by Harrison et al. was employed [26]-(Harrison et al., 2012). These joint moments were then normalized by the body mass of the individuals, allowing for standardized comparisons.

The inclusion criteria for the TD and CP subjects did not involve specific age or gender

requirements. For the TD subjects, those who walked barefoot and had complete measurements were included in the study, as depicted in Figure 1 of the flowchart.

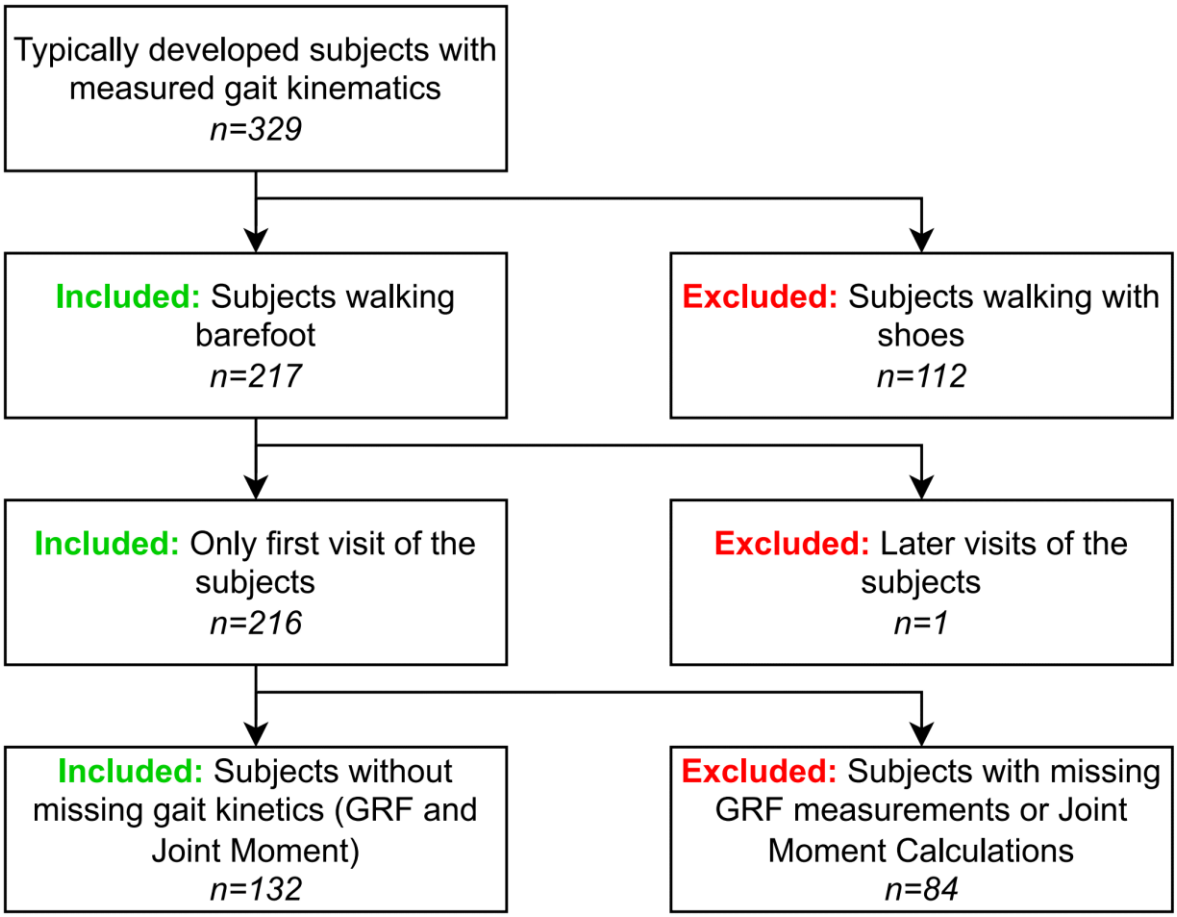


Figure 2.1: Inclusion/exclusion flow of the typically developed subjects. GRF: Ground reaction force

As for the CP subjects, their first visits were considered, and only those who were able to walk without assistive devices and had complete measurements were included, as illustrated in Figure 2 of the patient flowchart. Notably, the Gross Motor Function Classification System (GMFCS) levels of the included CP patients were limited to levels I and II.

After applying inclusion–exclusion criteria, 132 TD and 622 CP patients with spastic

diplegia were selected.

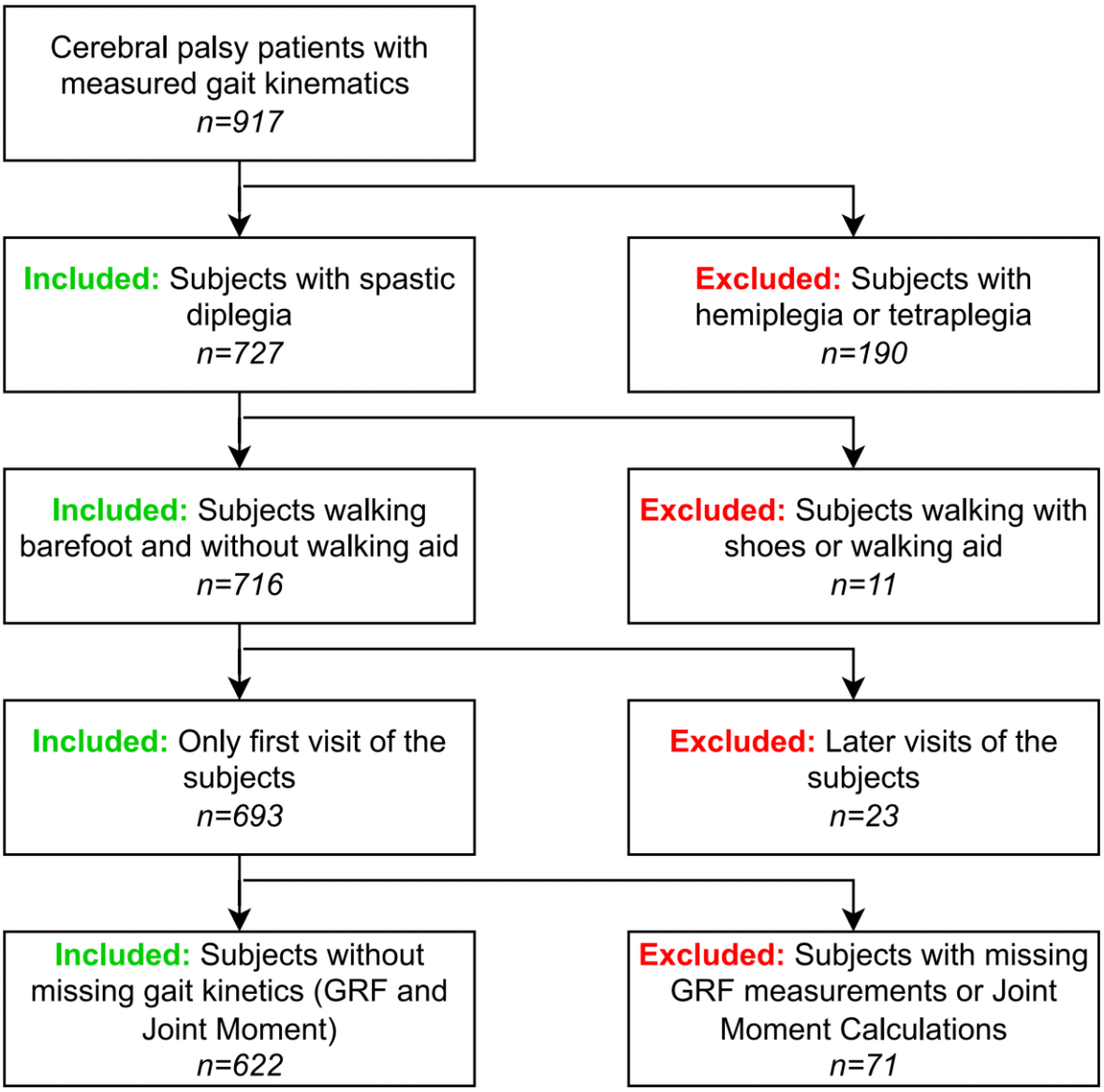


Figure 2.2: Inclusion-exclusion flow of the subjects with cerebral palsy. GRF: Ground reaction force

By utilizing this comprehensive dataset and implementing stringent inclusion criteria, the study aimed to ensure accurate and representative results, contributing valuable insights into the prediction of GRF and joint moments in both TD individuals and those with CP.

2.2 Data gathered from the subjects

A comprehensive set of kinematic and kinetic data was collected for analysis in this study. The kinematic data encompassed measurements from various body segments, including the trunk, pelvis, hip, knee, and ankle, in three planes of motion (sagittal, coronal, and transverse). This resulted in a total of 15 angles that were considered. Furthermore, the kinetic data included GRF, flexion-extension moments of the ankle, knee, and hip, as well as the adduction-abduction moment of the hip.

To ensure reliable and representative results, the data from each subject was averaged across 7-10 strides. This averaging process helped to minimize any potential variability and provided a more robust representation of the individual's gait pattern. Additionally, the data was normalized to a percentage gait cycle, where each time series consisted of 101-time points representing the entire gait cycle, ranging from 0% to 100%.

In addition to the averaged time series, the dataset also included the standard deviations of the time series throughout the individual strides. This information provided information of the variability of the kinematic and kinetic parameters within each gait cycle.

To further refine the analysis, the time series were segmented into stance and swing phases based on the temporal foot-off values. Since GRF data were not available during the swing phase, only the stance phases of the time series were utilized. This allowed for the inclusion of directly measured GRF and directly calculated moment data during

the periods when the foot was in contact with the ground.

By incorporating this comprehensive dataset of kinematic and kinetic measurements, along with the appropriate segmentation (stance and swing phases), the study is designed to use the most appropriate data in prediction of GRFs and joint moments during the gait cycle. These considerations ensured that the study focused on relevant and meaningful data, contributing to a more accurate learning of the biomechanical aspects of gait in both TD individuals and those with CP.

2.3 Predicting ground reaction forces using one-dimensional convolutional neural network

As mentioned above, this study of the thesis aimed to predict the GRF during gait without requiring force plates to measure it. The subjects and data explained above in Sections 2.1 and 2.2 are used for this study.

Figure 2.3: Force platform, embedded in the ground, used in measurement of ground reaction force [27]

2.3.1 Pre-processing

To ensure uniformity in the data size for training the ML algorithms, a standardized length of 60-time points was established for the stance segments of the time series. Since each subject may have a different duration of stance time, a standard interpolation technique was employed to adjust the time series to this standardized length. This interpolation process allowed for consistent data size across all subjects, facilitating the ML training process.

Furthermore, to ensure fair and unbiased learning, all time series values, regardless of their unit or magnitude, were normalized within the range of 0 to 1. This normalization technique prevented any particular time series with higher magnitudes from dominating the learning process. By scaling the values within a standardized range, the ML algorithms could effectively analyze and compare the patterns and relationships within the data.

Following the normalization step, the 15 kinematic time series, along with their corresponding standard deviations, were organized and stacked into a matrix format. This matrix, namely the input matrix, had 30 rows, representing 15 time series and their associated standard deviations, and 60 columns, corresponding to the standardized length of the stance segments. This matrix format allowed for a structured and consistent representation of the data.

In total, 132 matrices were created for training and testing the ML process in the case of TD subjects, while 622 matrices were generated for patients with CP. These matrices

served as the input data for the ML algorithms, enabling the learning and prediction of GRFs and joint moments based on the kinematic information.

By standardizing the data size, normalizing the values, and organizing the time series into matrices, the study aimed to establish a consistent and compatible format for the ML algorithms to process the gait data effectively. This approach ensured that each subject's data contributed equally to the training and testing processes, facilitating accurate predictions of GRFs and joint moments in both TD individuals and those with CP.

2.3.2 Machine learning approach

To effectively process the time series data and capture distinct features from each time series, a one-dimensional convolution neural network (1D-CNN) model was utilized. The 1D-CNN model employs convolutional layers specifically designed for processing sequential data, such as the joint angles in our case. These convolutional layers extract features from the time series data by considering different temporal ranges, enabling the model to capture valuable information relevant to predicting another time series' data, in our case, the GRFs and the joint moments [28,29]-(Hua et al., 2020; Malek et al., 2018).

During the initial experimentation phase, various network sizes were assessed using a separate development set. As the number of convolutional and densely connected layers increased from a basic architecture comprising only one of each, there was a gradual decrease in the calculated loss on the development set. The hyperparameters that best

suited the selected network size were determined. It became clear that further increasing the model's complexity did not yield significant improvements, but it considerably increased the computational demands. So, the complexity of the 1D-CNN model was manually defined to ensure optimal performance. The model's complexity was gradually increased until no significant decrease in loss, a measure of prediction error, was observed. This iterative process allowed for fine-tuning the model's architecture and finding the optimal balance between complexity and performance. To determine the model's effectiveness, a separate development set comprising 42 patients with CP was used to evaluate the model's performance, making adjustments until the desired accuracy and predictive power were achieved.

By incorporating the 1D-CNN model into the analysis, the study aimed to leverage its ability to extract meaningful features from the time series data, thereby enhancing the prediction accuracy of joint moments based on the corresponding joint angles. The iterative optimization process ensured that the model's complexity was tailored to the specific task, maximizing its effectiveness in capturing the intricate relationships and patterns within the gait data of patients with CP. The details of the designated 1D-CNN model as well as the evaluation metrics is explained below in Section 2.4 as similarly used for Study II, including Figure 2.5 showing the pipeline of the data processing and ML.

2.4 Deep learning-based prediction of joint moments based on kinematics in patients with cerebral palsy

As mentioned above, the thesis aimed to predict the dorsi-plantar flexion, knee flexion-

extension, hip flexion-extension, and hip adduction-abduction moments of patients with CP during gait without requiring inverse dynamics models, which is used to calculate joint moments based on GRFs measured by force plates. The subjects and data explained in Sections 2.1 and 2.2 were used in this study. The ML approach explained for Study I in Section 2.3.2. was followed. In the end, one 1D-CNN model was designed and used for both studies, namely Study I (Predicting ground reaction forces using one-dimensional convolutional neural network based on kinematics during gait in patients with cerebral palsy) and Study II (Deep learning-based prediction of joint moments based on kinematics in patients with cerebral palsy) explained in Sections 2.3 and 2.4.

2.4.1 Designed one dimensional convolutional neural network model

The designated 1D-CNN model utilized in the study consisted of five convolutional layers, each with a specific number of filters and 1D kernel sizes. These convolutional layers had the following numbers of filters: [128, 128, 512, 1024, 2048]. The corresponding 1D kernel sizes for these layers were [30, 15, 10, 5, 3]. This configuration was chosen to gradually extract features over decreasing time intervals, as the ascending number of filters and decreasing filter sizes allowed for capturing more detailed information at finer time resolutions.

After the output of the convolutional layers was flattened, ten densely connected layers were employed. These layers consisted of varying numbers of neurons, specifically [10000, 8000, 6000, 4000, 3000, 2000, 1000, 500, 250, 100]. The descending number of neurons in these layers was chosen to transform the information to the desired output size throughout the learning process. By using this architecture, the model was able to

effectively learn and extract meaningful features from the input data.

To introduce non-linearity and enhance the model's learning capabilities, a rectified linear unit (ReLU) activation function was applied to all layers. Additionally, a dropout layer with a 1% dropout fraction was attached to the output of each layer to prevent over-fitting. The final output layer, which was densely connected, employed a linear activation function and had a neuron size of 60, corresponding to the number of time points in the stance phase of the interpolated joint moment time series.

During the learning process, the stochastic gradient descent (SGD) algorithm was used as the optimization algorithm with a learning rate of 0.01. The loss criterion for evaluating the training performance was based on the root mean squared error (RMSE) and Pearson correlation coefficient (PCC) between the experimental and predicted time series. The implementation of the 1D-CNN algorithm was carried out using Keras on the Tensorflow framework.

To assess the model's performance and ensure robustness, a 10-fold cross-validation algorithm was employed. The dataset was divided into ten equal parts, with nine parts used for training and one part for testing in each fold. Range normalization was applied separately to the training and testing sets to prevent any information leakage between them. Each subject was included in only one of the ten subsets to avoid over-fitting the model to specific walking patterns.

2.4.2 Avoiding over-fitting

To monitor the training process and detect over-fitting, learning curves were plotted, comparing the decrease in loss on the training set to that of an isolated test set. The training process was limited to 500 epochs for each split, with batches of size 32. This approach allowed for effective training and evaluation of the model's performance while preventing over-fitting and ensuring generalizability.

Throughout the training phase, the model's parameters were adjusted via back propagation exclusively using the training set's loss, without any modifications to the hyper parameters. Consequently, the test set remained separate from the training process and served the sole purpose of evaluating potential over-fitting. The validation set, also known as the "development set," ceased to be utilized once the model was established. The learning curves demonstrated that notable over-fitting did not occur on the training set, as the loss on the test set decreased concurrently (though to a lesser extent) with the loss on the training set at each epoch.

Figure 2.4 shows the learning curve for the training of the model for ankle dorsi-plantar flexion moment as an example. The x-axis shows the number of epochs, while the light blue line represents the loss for the test set and the dark blue line shows the loss for the training set. The curve has been smoothed using the exponential moving average algorithm to provide an easy-to-interpret overall shape while retaining the actual loss values as a faded background.

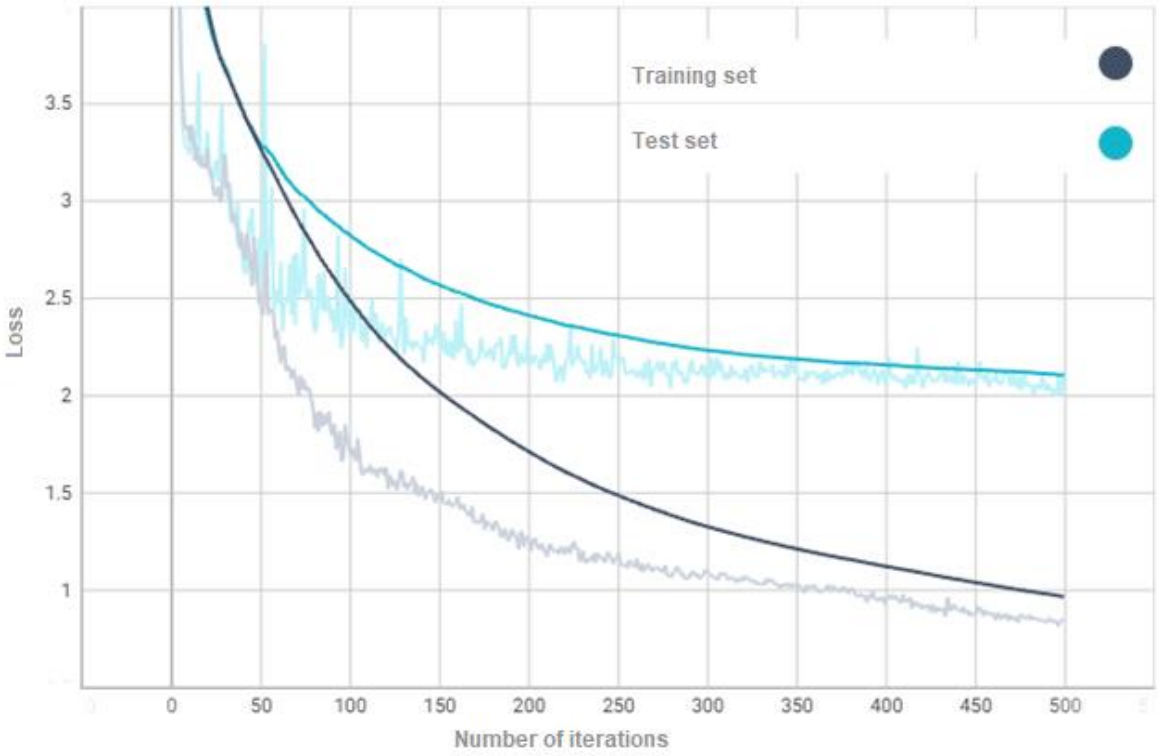


Figure 2.4: The learning curves plotted for the training of the model for the ankle dorsi-plan flexion moment

The detailed data processing and ML pipelines used in the study were given in Figure 2.5, providing a comprehensive overview of the methodology employed to predict GRFs and joint moments in patients with CP during gait.

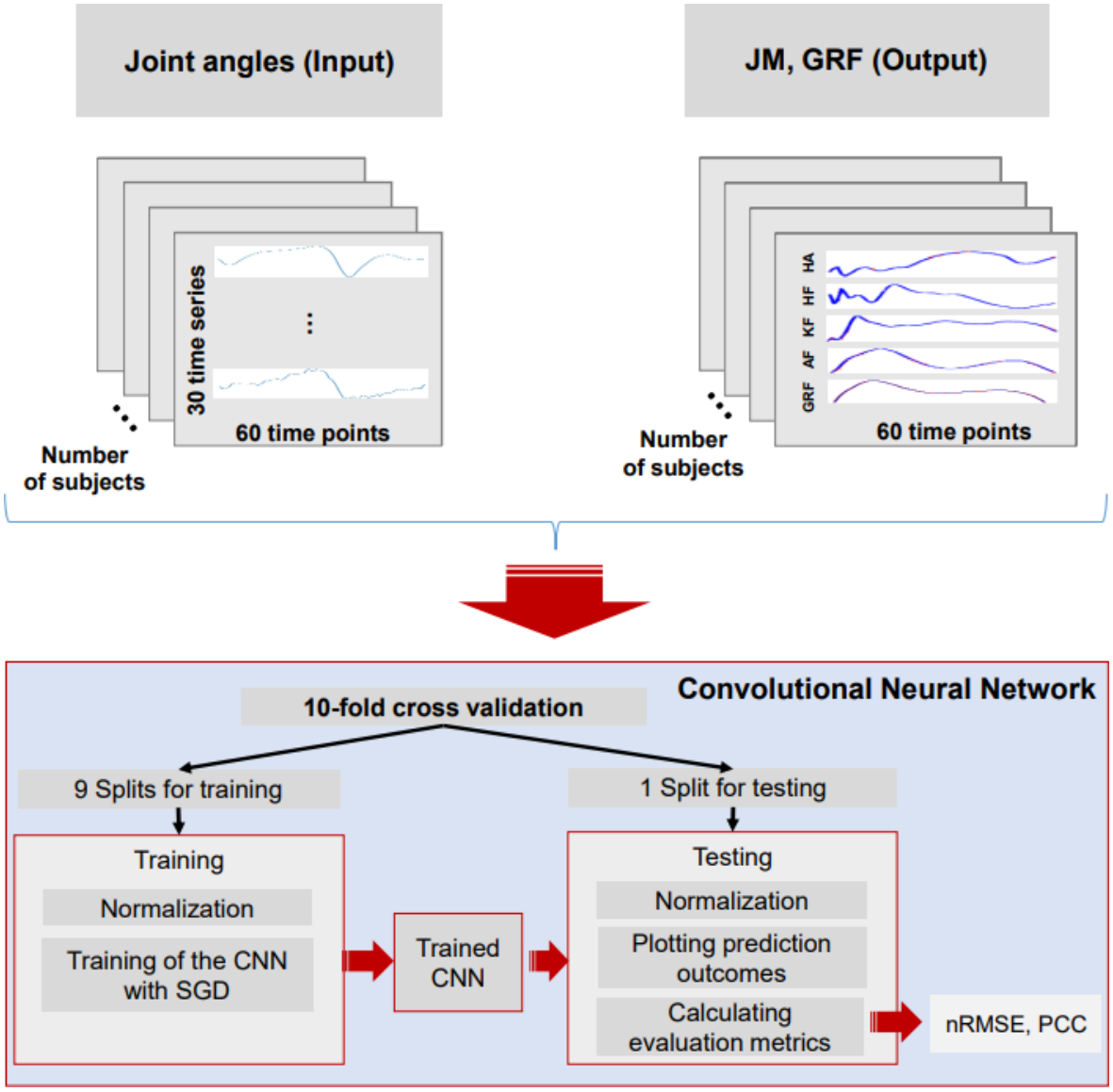


Figure 2.5: Data processing and machine learning pipeline. GRF: Ground reaction force, JM: Joint moment, AF: ankle dorsi-plantar flexion moment, KF: knee flexion-extension moment, HF: hip flexion-extension moment, HA: hip adduction abduction moment, CNN: convolutional neural network, SGD: stochastic gradient descent, nRMSE: normalized root mean square error, PCC: Pearson correlation coefficient.

2.4.3 Evaluation Metrics

The evaluation of the predicted joint moment time series involved the utilization of two widely recognized metrics: the normalized root mean square error (nRMSE (%)) and the Pearson correlation coefficient (PCC). These metrics have gained widespread acceptance in the field as reliable measures for assessing the accuracy and performance of ML algorithms in predicting joint moments, as evidenced by previous studies [19, 25,31,32] (Mundt et al., 2020a; Ardestani et al., 2014; Mundt et al., 2020b; Ripic et al., 2022).

The nRMSE metric quantifies the normalized magnitude difference between the predicted and experimental joint moment time series at each time point. nRMSE was calculated by dividing the RMSE value by the mean range of the experimental joint moment (μRoM) across all subjects of the same group as stated in Equation (1). In the equation, JM_P and JM_E denote predicted and experimental joint moments, respectively. Sub-indices *P* and *E* denote predicted and experimental quantity, respectively.

$$nRMSE = \sqrt{\frac{\Sigma_n (JM_P - JM_E)2}{n}} / \mu RoM \quad (1)$$

The Pearson correlation coefficient (PCC) metric calculates the degree of pattern similarity between the experimental and predicted joint moments, providing a quantitative measure of their correlation and alignment [33]-(Savelberg and Herzog, 1997), in which cross-covariance ($cov(E, P)$) of them and variance of each of them ($\boldsymbol{\sigma}_E, \boldsymbol{\sigma}_P$) respectively were used (Equation (2)).

$$PCC = \frac{cov(E,P)}{\sigma_E \sigma_P} \quad (2)$$

In order to address the issue of strong skewness in the distribution of the Pearson correlation coefficient (PCC) values, we employed Fisher's Z transformation. This transformation serves to normalize the PCC values, making the distribution more symmetric and suitable for statistical analysis. By applying the transformation, we were able to alleviate potential biases and achieve a more robust assessment of the correlation between the experimental and predicted joint moments. The transformed PCC values were then used to compute the mean, which was computed based on these normalized values. To interpret the results in the original PCC scale, we reversed the transformation by applying the inverse of Fisher's Z transformation, following a well-established approach outlined in the reference [34]-(Silver et al., 1987). This ensured that the final PCC values accurately reflected the pattern similarity between the experimental and predicted joint moments, providing a reliable measure of their association.

2.5 Predicting joint moments of patients with cerebral palsy using deep learning and various conventional machine learning methods

In this phase of the study, our objective was to design and evaluate various input-algorithm configurations to predict joint moments. We focused on constructing different input sets that encompassed the same range of kinematic information. Specifically, two sets of inputs

i) one for deep learning including row kinematic time series

ii) one for conventional ML algorithms that is composed of manually extracted features

were constructed for the prediction of the ankle dorsi-plantar flexion moment.

It is worth noting that conventional ML algorithms tend to be more straightforward to train when the provided information is relevant and representative. In other words, when the input features accurately capture the essential characteristics of the data, conventional algorithms are typically better equipped to learn and make accurate predictions. Therefore, in this study, careful consideration was given to selecting and constructing input sets that encompassed pertinent and meaningful information from the kinematic data, aiming to enhance the training process and optimize the performance of the ML models.

Deep learning, in contrast, has the ability to automatically create and extract its own features from the input data. It does not rely on handcrafted features like conventional ML approaches. By doing so, deep learning models can explore a much wider solution space, allowing them to capture complex patterns and relationships in the data.

However, deep learning algorithms typically require a larger amount of data to effectively learn these complex representations. With a substantial amount of data available for training, deep learning models can leverage their capacity to learn intricate patterns and achieve remarkable performance on more challenging and complex problems. The abundance of data enables the models to generalize well and make accurate predictions or classifications.

Therefore, deep learning's success is highly dependent on the availability of sufficient and representative data. When provided with enough data, deep learning models can

excel in tackling intricate and multifaceted problems, surpassing the capabilities of conventional ML methods.

2.5.1 Constructing Input Sets

In this study, two distinct types of input sets were created using the collected kinematic data, as depicted in Figure 2.6. The first type of input set involved stacking the extracted features from the kinematic time series, which were obtained following the Automated Feature Assessment Workflow for Instrumented Gait Analysis [35]-(Wolf et al., 2006). These extracted features captured relevant information from the time series to be used in conventional ML algorithms. On the other hand, the second type of input set comprised solely the stance segment of the raw kinematic time series themselves as explained for the Study II to be used in deep learning algorithms.

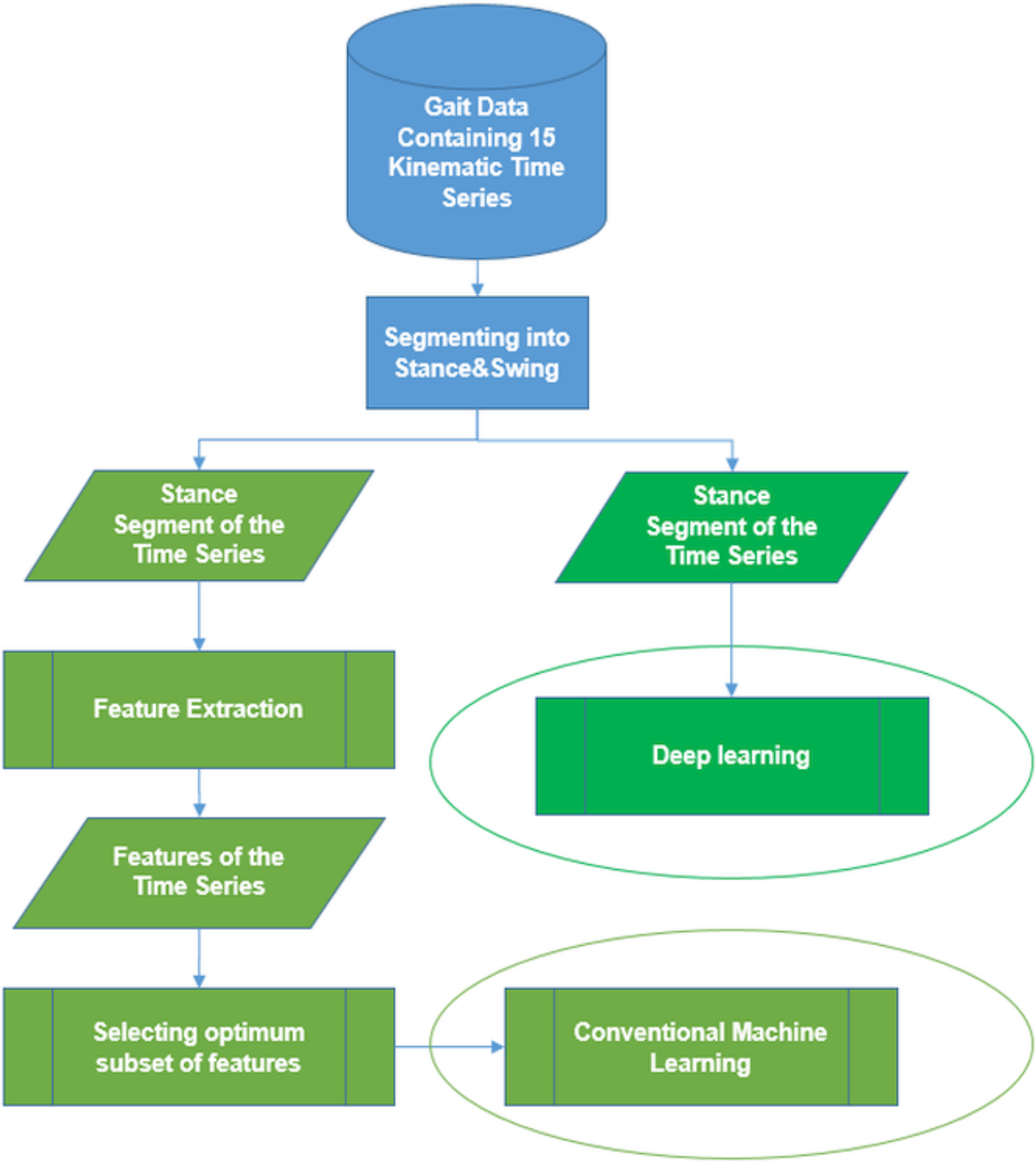


Figure 2.6: The flowchart for constructing two input subsets.

2.5.2 Extracting Features for Conventional ML Models

As conducted in the study of Wolf et al. [35]-(Wolf et al., 2006), two new time series were derived from the mean time series for each joint angle: The first gradient time series and the difference from the normative time series. The first gradient "*V*" of the mean time series "*U*" was calculated according to the Formula (3), where "*k*" is the data point index within a gait cycle. A discrete derivation of the mean time series must be done since the time series are in a discrete domain.

$$V[k] = \frac{1}{2}(U[k+1] - U[k-1]) \quad (3)$$

The difference relative to a reference considered to be normal U_{norm} namely difference from normative "DN" was calculated according to the Formula (4):

$$DN[k] = |U[k] - U_{norm}[k]| \quad (4)$$

For each joint angle, reference normal time series U_{norm} was calculated by averaging the corresponding time series across all TD subjects.

For each joint angle, the computed scalar features from both derived and original time series were as follows: Minimum and maximum values and their timings (i.e. temporal position in the gait cycle; x-axis in Figure 2.7). Note that, for the standard deviation time series, only the maximum value and its timing were considered as features since it makes sense that strongly varying gait of subjects with CP may cause more differences in the gait patterns among the strides. For the first gradient time series, the average

difference from normative was additionally considered because the difference from the normal gait pattern may contain meaningful information.

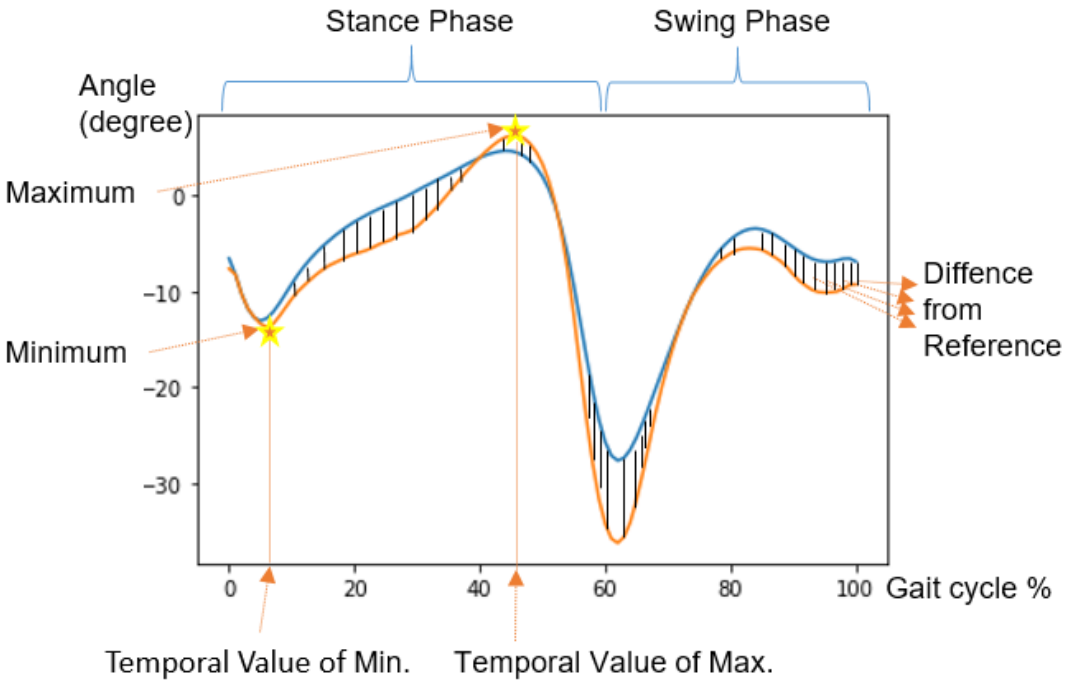


Figure 2.7: Computed features of an example time series. The orange line shows the ankle dorsi plantar flexion of the subject. The blue line shows the reference normal time series for the ankle dorsi plantar flexion.

2.5.3 Machine Learning Algorithms

Two deep learning models, namely 1D-CNN (the same model as used in Study II), long short term memory network (LSTM) and four conventional ML models, namely ridge regression (R-Regr), k-nearest neighbor (KNN), random forest (RF), and multilayer neural network (MLNN) were developed and compared regarding the same evaluation metrics described in Section 2.4.3 for Study II.

Based on the manual preliminary trials conducted on the development set, similarly executed as described for the Study I in Section 2.3.2, the hyper-parameters of the conventional ML algorithms, namely, R-Regr, KNN, and RF were carefully selected to ensure their suitability for our specific case of predicting joint moments. The deep learning model 1D-CNN is used same as developed in the Study II, explained in Section 2.4. The deep learning model LSTM is used with the same densely connected structure as in the 1D-CNN and the hyper parameters of the LSTM layers are selected with the same approach described for the Study I in Section 2.3.2. The MLNN model is designed to have the same structure as the densely connected layers of the 1D-CNN model. Table 2.1 below shows the hyper-parameters of the models.

Table 2.1: Hyper-parameters of the machine learning models

Model	**List of hyper-parameters**
LSTM	Number of LSTM layers, number of LSTM units at each layer, activation function, drop-out ratio, number of densely connected layers
1D-CNN	Number of convolutional layers, number of filter per each layer, size of the filters in each layer, activation function, drop-out ratio, number of densely connected layers
R-Regr	The regularization strength, the type of solver algorithm
KNN	The type of distance metric, the type of weight function, the type of algorithm for computing nearest neighbor, the size of leaf (for the requiring algorithms)
RF	The number of trees in the forest, split quality criterion, the minimum required samples for splitting a node, the minimum required samples for being a leaf node, maximum depth of trees, the maximum number of features for splitting, the existence of bootstrapping
MLNN	Number of densely connected layers, activation function, drop-out ratio

These architectures serve not only the best learning possible but also the purpose of ensuring a fair and meaningful comparison between the deep learning approach (based-on autonomously extracted features) and the conventional ML method (based-on manually extracted features). By aligning the architecture of the MLNN model with the densely connected layers of the 1D-CNN, we create a consistent framework that allows us to directly evaluate the impact of feature extraction on the performance of the models.

This approach enables a comprehensive analysis of the benefits and limitations associated with both approaches, shedding light on the effectiveness of feature extraction methods and their influence on the predictive capabilities of the models. The models are trained following the same pipeline as previously described in Figure 2.5 for the prediction of the ankle dorsi-plantar flexion moment.

2.6. Statistical Analysis

The resulting evaluation metrics for predicting GRF (Study 1) not only provided valuable insights but also necessitated a comprehensive statistical analysis for a more in-depth discussion and interpretation. This was particularly crucial given the presence of two distinct groups in the study, namely TD individuals and those with CP. By conducting statistical analyses on the evaluation metrics, we were able to explore and uncover potential differences, patterns, and trends between these two groups in terms of the predicted GRF.

Similarly, the same principle applies to the prediction study involving various ML algorithms (Study III). In this study, we explored the performance and effectiveness of six different models for predicting the desired outcomes. With the presence of multiple models to compare, it became necessary to employ statistical analysis techniques to assess and compare their predictive capabilities.

We hypothesized that the performance metrics differ significantly, between each group in Study I, and between the models in Study III. Statistical analysis was conducted using

SPSS software (Version 21.0; SPSS; Chicago, IL, USA) and the level of significance was set at 0.05. The resulting evaluation metrics were checked with Kolmogorov-Smirnov test whether they are normally distributed or not. They were both found not to be normally distributed. The statistical significance between two subject groups in Study I was checked with Mann-Whitney U test. The statistical significance between the models in Study III was checked with Friedman's Anova test. A Bonferroni correction was applied to adjust the p-value for multiple comparisons (p<0.016). A Bonferroni correction was implemented to account for multiple comparisons, resulting in an adjusted p-value threshold of 0.016.

The resulting evaluation metrics for predicting joint moments (Study II) required statistical analysis for further discussion, since there are four predicted joint moments of two subject groups, namely TD and CP. Statistical analysis was conducted using SPSS software (Version 21.0; SPSS; Chicago, IL, USA). We performed both inter-comparisons (between patient with CP and TD subjects) and intra-comparisons (within each group of subjects). In the inter-comparison, we hypothesized that the prediction success rates for the joint moments differ significantly between the groups, namely patients with CP and TD subjects. In the intra-comparison, we hypothesized that the prediction success rates of the aforementioned joint moments differ significantly within each group of subjects. The level of significance was set at 0.05. The Kolmogorov-Smirnov test was used to test the normality of the data, which was found not to be normally distributed. The predicted joint moments of the TD subjects and patients with CP were statistically analyzed using the Mann–Whitney U-test. For the intra-comparisons, Friedman's ANOVA test was used. The Mann–Whitney U-test was used

to identify significant differences between the methods. A Bonferroni correction was applied to adjust the *p*-value for multiple comparisons ($p<0.016$). Please refer to Appendix A for details on the intra- and inter-comparison groups and the type of analysis.

3. RESULTS

3.1. Results of predicting ground reaction forces using one-dimensional convolutional neural network

For TD subjects, the mean normalized root mean square error (nRMSE) value for predicting GRF was found 7.47%±3.53 (Figure 3.1), indicating a relatively low level of prediction error. On the other hand, patients with CP had a higher mean nRMSE value of 11.75%±6.88 (Figure 3.1), showing a higher level of prediction error in their case.

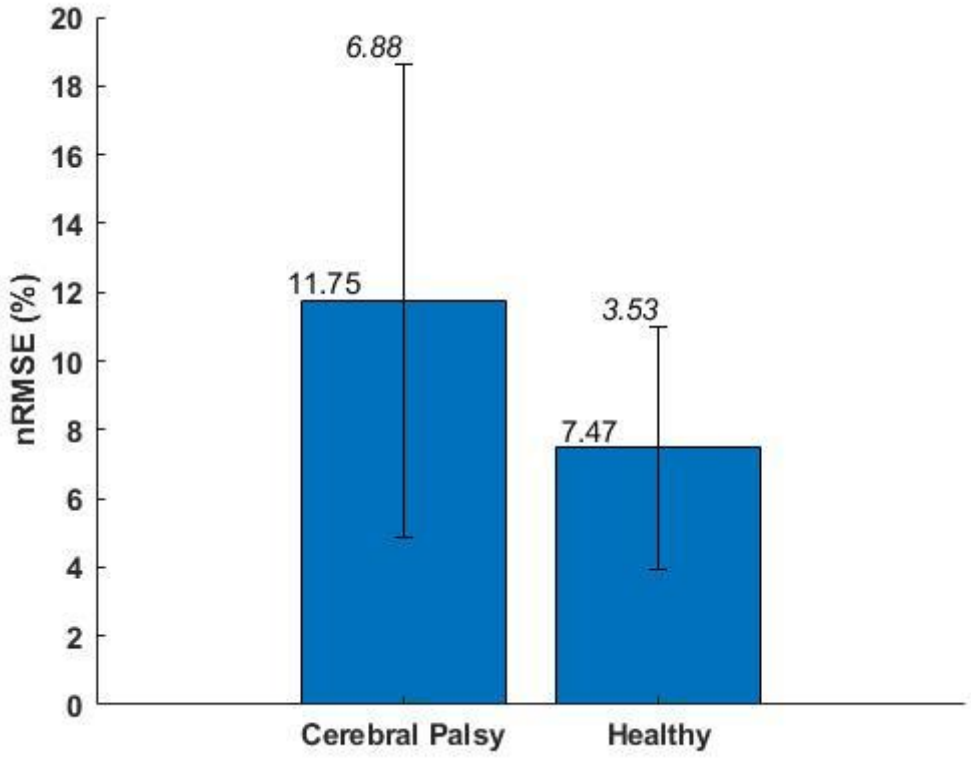


Figure 3.1: Normalized root mean square error (nRMSE) scores for ground reaction force predictions of TD subjects and patients with CP.

In terms of the Pearson correlation coefficient (PCC), TD subjects exhibited a mean PCC value of 0.98 (Figure 3.2), indicating a strong pattern similarity between the predicted and experimental GRF values. Conversely, patients with CP had a slightly lower mean PCC value of 0.94 (Figure 3.2), indicating a slightly weaker pattern similarity between the predicted and experimental GRF values in their case.

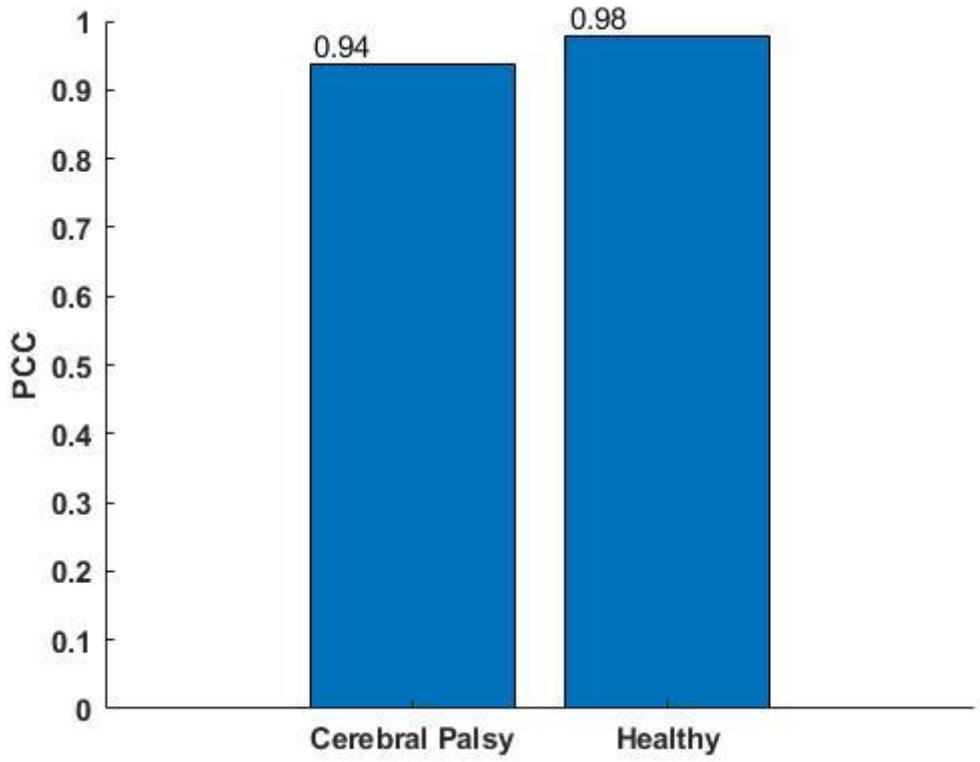
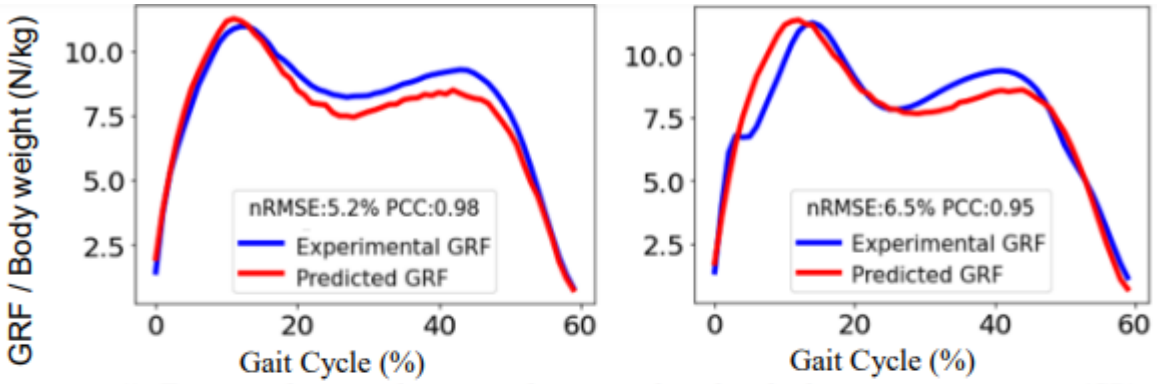


Figure 3.2: Pearson correlation coefficient (PCC) scores for ground reaction force predictions of TD subjects and patients with CP.

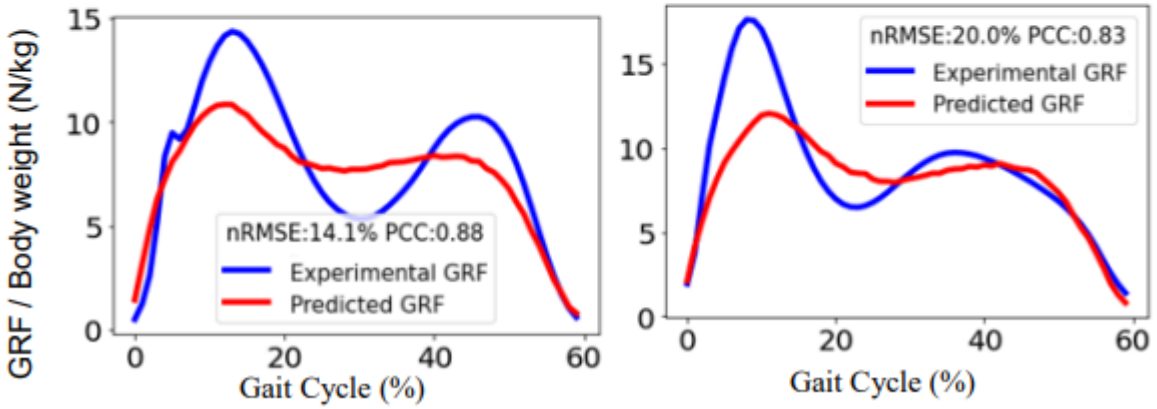
Statistical analysis, as described in Section 2.6, was conducted to compare the CP and TD subject groups. The resulting *p*-values for nRMSE and PCC values were found to be 0.032 and 0.027, respectively. These p-values indicate significant differences between the subject groups in terms of both magnitude (nRMSE) and pattern similarity (PCC).

Figure 3.3 provides a collection of representative predicted and experimental GRFs for patients with CP, offering valuable insights into the predictive capabilities of the trained models across different scenarios. The inclusion of these diverse examples aims to enhance the understanding of how well the models perform in predicting GRFs with

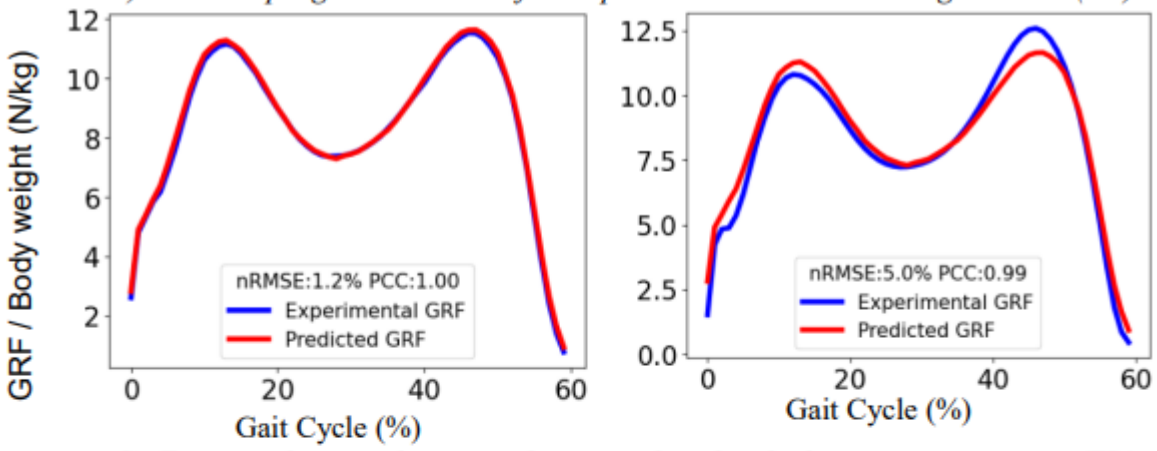
varying patterns. The upper row of the figure showcases predictions that have achieved relatively high levels of success, evaluated with lower normalized root mean square error (nRMSE) values and higher Pearson correlation coefficient (PCC) values compared to the average performance of the model. These predictions demonstrate the models' proficiency in successfully capturing the intricate dynamics of the GRFs. On the other hand, the lower row of the figure presents predictions that are comparatively less successful, exhibiting higher nRMSE values and lower PCC values than the average. These examples highlight the challenges faced by the models in accurately reproducing certain complex patterns within the GRF data. By including both successful and less successful predictions, the figure provides a comprehensive representation of the models' performance, enabling a nuanced evaluation of their overall predictive capabilities.



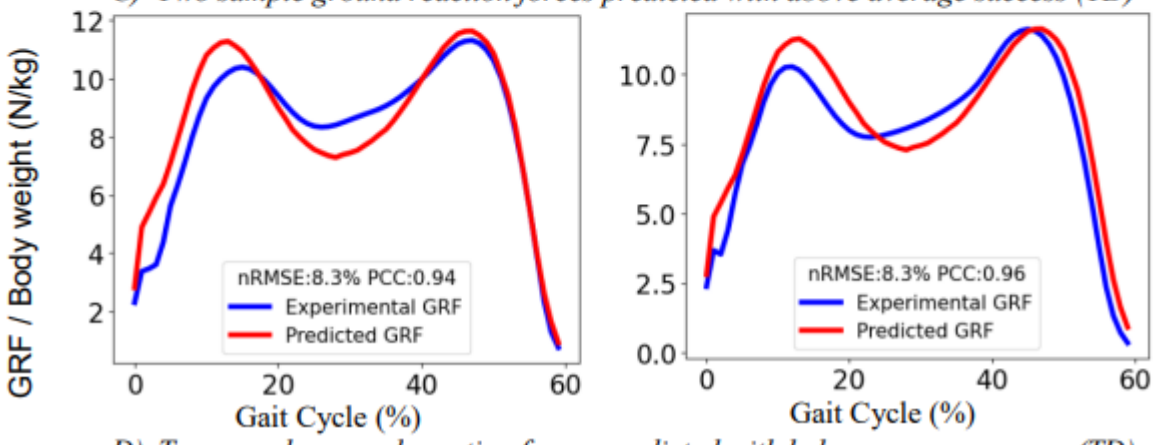
A) Two sample ground reaction forces predicted with above average success (CP)



B) Two sample ground reaction forces predicted with below average success (CP)



C) Two sample ground reaction forces predicted with above average success (TD)



D) Two sample ground reaction forces predicted with below average success (TD)

Figure 3.3: Secondary, representative results for aiding interpretations of Figure 3.1 and Figure 3.2. Predictions of ground reaction force for representative typically developed subjects and patients with cerebral palsy. CP: Cerebral palsy, TD: Typically developed. (A) corresponds to above-average prediction success for patients with CP, (B) corresponds to below-average prediction success for patients with CP, (C) corresponds to above-average prediction success for TD subjects, (D) corresponds to below-average prediction success for TD subjects. The blue line represents the experimental ground reaction force, while the red line represents the predicted ground reaction force.

3.2. Results of deep learning-based prediction of joint moments based on kinematics in patients with cerebral palsy

For TD subjects, all joint moments were predicted with mean nRMSE values less than 12.55%±5.08 (Figure 3.4). The knee flexion-extension moment is the least successfully predicted joint moment in terms of nRMSE score (12.55%±5.08). The dorsi-plantar flexion is the most successfully predicted joint moment (8.58%±3.87). The hip adduction-abduction and hip flexion-extension moments were predicted with an nRMSE value of 11.89%±4.72 and 10%±3.66 for TD subjects, respectively.

For patients with CP, all joint moments were predicted with mean nRMSE values less than 18.02%±9.14 (Figure 3.4). The knee flexion-extension moment is the least successfully predicted joint moment in terms of nRMSE (18.02%±9.14), while the hip flexion-extension is the most successfully predicted joint moment (13.58%±5.36). The hip adduction-abduction and dorsi-plantar flexion moments were predicted with an nRMSE value of 17.2%±6.53 and 14.78%±7.17 for the CP group, respectively.

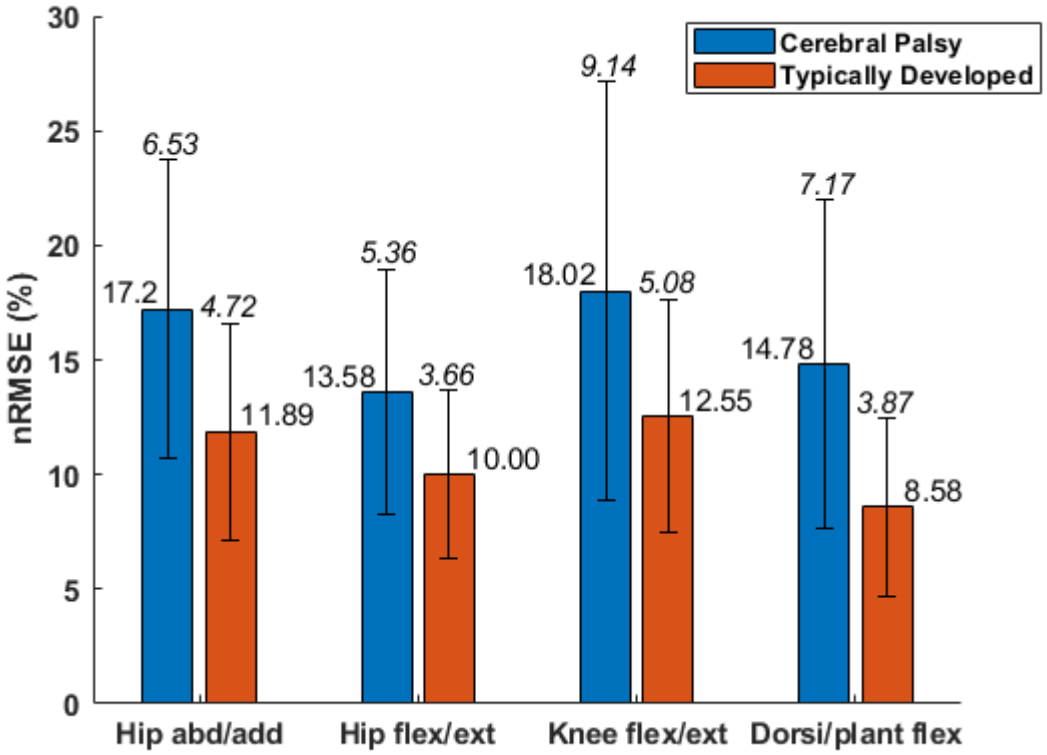


Figure 3.4: Normalized root mean square error (nRMSE) scores for joint moment predictions of TD subjects (red) and patients with CP (blue). Hip abd/add: hip adduction abduction, Hip flex/ext: hip flexion extension, Knee flex/ext: knee flexion extension, Dorsi/plant flex: dorsi plantar flexion.

For TD subjects, all joint moments were predicted with mean PCC scores higher than 0.96 (Figure 3.6). The dorsi-plantar flexion is the most successfully predicted joint moment in terms of PCC score (0.99), while the others have the same PCC (0.96).

For the patient group, all joint moments were predicted with mean PCC scores higher than 0.89 (Figure 3.6). The hip adduction-abduction moment is the least successfully predicted moment in terms of PCC score (0.89), while the dorsi-plantar flexion is the most successfully predicted one (0.96).

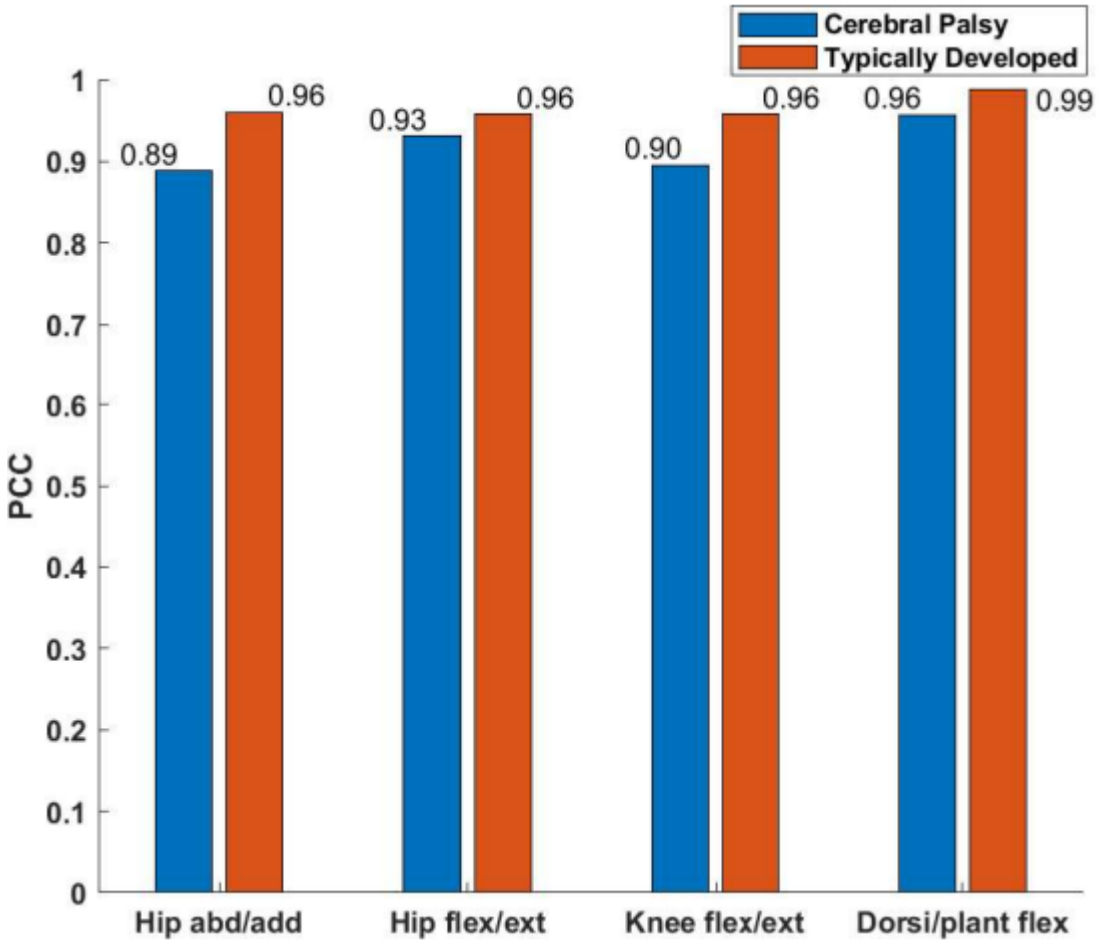


Figure 3.5: Pearson correlation coefficient (PCC) scores for joint moment predictions of TD subjects (red) and patients with CP (blue). Hip abd/add: hip adduction-abduction, Hip flex/ext: hip flexion-extension, Knee flex/ext: knee flexion-extension, Dorsi/plant flex: dorsi plantar-flexion.

Tables 3.1 and 3.2 present the statistical significance of the nRMSE and PCC scores obtained for the joints of TD and patient groups, respectively. Within the TD group, the dorsi-plantar flexion and hip flexion-extension moments exhibited significantly better predictions than the hip abduction-adduction and knee flexion-extension moments in terms of nRMSE (Table 3.1). When considering the PCC scores, the prediction rate for the dorsi-plantar flexion moment was significantly higher than that for the knee flexion-extension joint moment (Table 3.1).

Table 3.1. P-values obtained for the nRMSE and PCC values of joint moment predictions for healthy subjects. Significant differences were marked bold. Hipaddabd: hip adduction-abduction, hipflexext: hip flexion-extension, kneeflexext: knee flexion-extension, dorsiplanflex: dorsi-plantar flexion

		nRMSE	PCC
Hip add/abd vs.	Hip flex/ext	**0.014**	0.021
	Knee flex/ext	0.019	0.020
	Dorsi/planflex	**0.015**	0.018
Hip flex/ext vs.	Hip add/abd	**0.014**	0.021
	Knee flex/ext	**0.012**	0.019
	Dorsi/planflex	0.020	0.016
Knee flex/ext vs.	Hip add/abd	0.019	0.020
	Hip flex/ext	**0.012**	0.019
	Dorsi/planflex	**0.014**	**0.014**
Dorsi/plan flex vs.	Hip add/abd	**0.015**	0.018
	Hip flex/ext	0.020	0.016
	Knee flex/ext	**0.014**	**0.014**

Within the patient group, the dorsi-plantar flexion moment was significantly better predicted than the hip abduction-adduction and knee flexion-extension moments in terms of nRMSE (Table 3.2). Furthermore, when taking the PCC scores into account, the prediction rates for the dorsi-plantar flexion and hip flexion-extension moments were significantly higher than those for the knee flexion-extension and hip abduction-adduction moments (Table 3.2).

Table 3.2. P-values obtained for the nRMSE and PCC values of joint moment predictions for the patients with CP. Significant differences were marked bold. Hipaddabd: hip adduction-abduction, hipflexext: hip flexion-extension, kneeflexext: knee flexion-extension, dorsiplanflex: dorsi-plantar flexion

		nRMSE	PCC
Hip add/abd vs.	Hip flex/ext	0.016	**0.015**
	Knee flex/ext	0.018	0.032
	Dorsi/planflex	**0.014**	**0.011**
Hip flex/ext vs.	Hip add/abd	0.016	**0.015**
	Knee flex/ext	**0.014**	**0.013**
	Dorsi/planflex	0.018	0.028
Knee flex/ext vs.	Hip add/abd	0.018	0.032
	Hip flex/ext	**0.014**	**0.013**
	Dorsi/planflex	**0.012**	**0.011**
Dorsi/plan flex vs.	Hip add/abd	**0.014**	**0.011**
	Hip flex/ext	0.018	0.028
	Knee flex/ext	**0.012**	**0.011**

Table 3.3 demonstrates the statistical significance of the scores between the TD individuals and patient groups. In terms of nRMSE, all four joint moments were predicted significantly higher in the TD group than in the CP group.

Table 3.3. P-values obtained for the comparison of the nRMSE and PCC values of joint moment predictions for the healthy subjects and patients with CP. Significant differences were marked in bold. Hipaddabd: hip adduction-abduction, hipflexext: hip flexion-extension, kneeflexext: knee flexion-extension, dorsiplanflex: dorsi-plantar flexion

		nRMSE	PCC
Healthy vs. Patients with CP	Hip add/abd	**0.041**	**0.033**
	Hip flex/ext	**0.047**	0.051
	Knee flex/ext	**0.038**	**0.037**
	Dorsi/plan flex	**0.034**	0.055

Figures 3.6 and 3.7 show some representative predicted and experimental joint moments of TD subjects and patients with CP, respectively. These figures are provided for a better understanding of the trained models' capability of predicting joint moments with varying patterns. To ensure the representativeness of the models' capability in predicting joint moments, the figures in the left column show relatively successful predictions (with lower nRMSE and higher PCC values than the average), while the figures in the right column show relatively less successful predictions (with higher nRMSE and lower PCC values than the average).

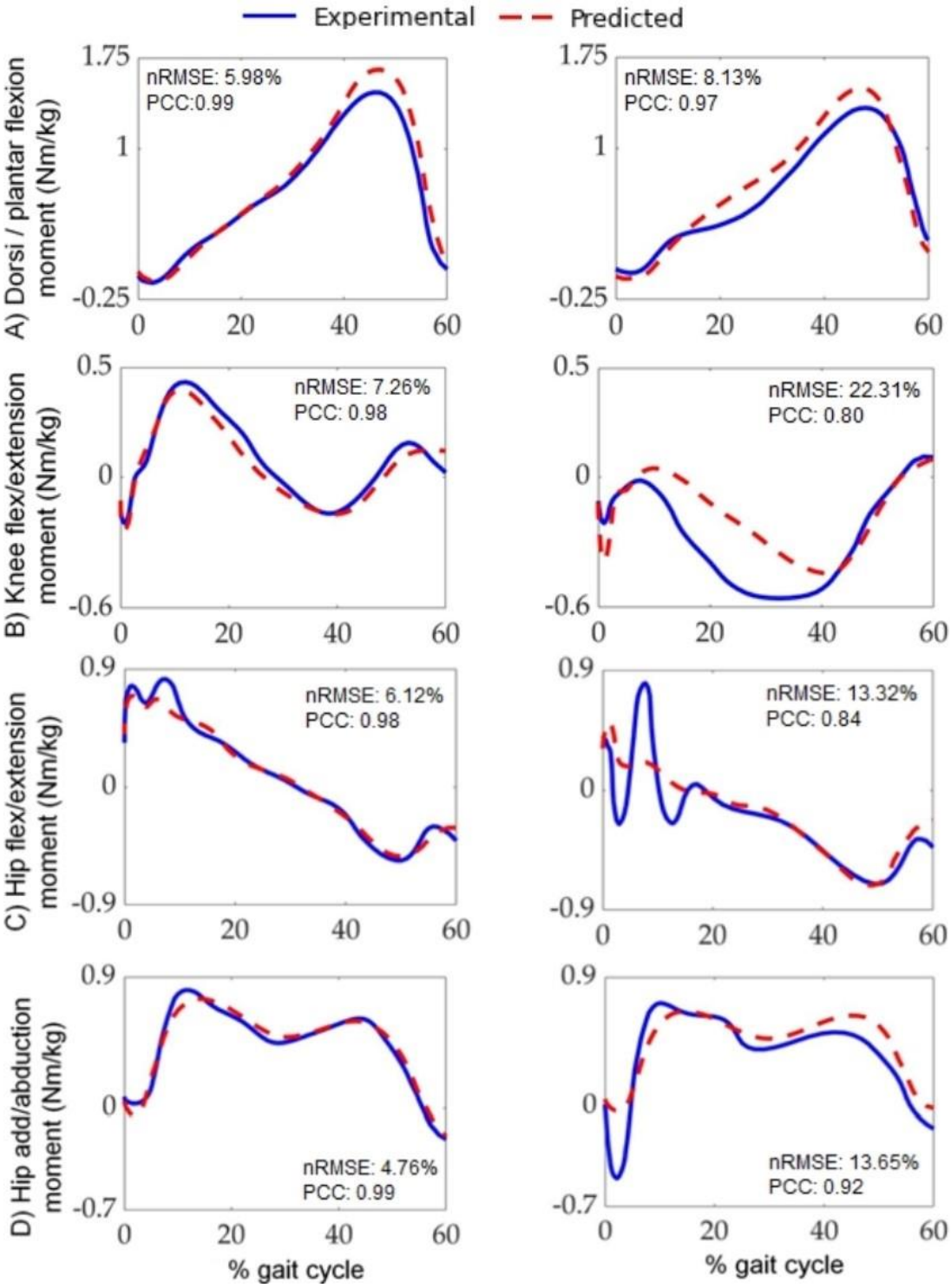


Figure 3.6: Secondary, representative results for aiding interpretations of Figure 3.4 and Figure 3.5. Joint moments of a) dorsi-plantar flexion, b) knee flexion-extension, c) hip flexion-extension, d) hip adduction-abduction for representative typically developed subjects. The predictions on the left column correspond to above-average success rates, while those on the right column correspond to below-average success rates. The blue line represents the experimental joint moment, while the red dashed line represents the predicted joint moment.

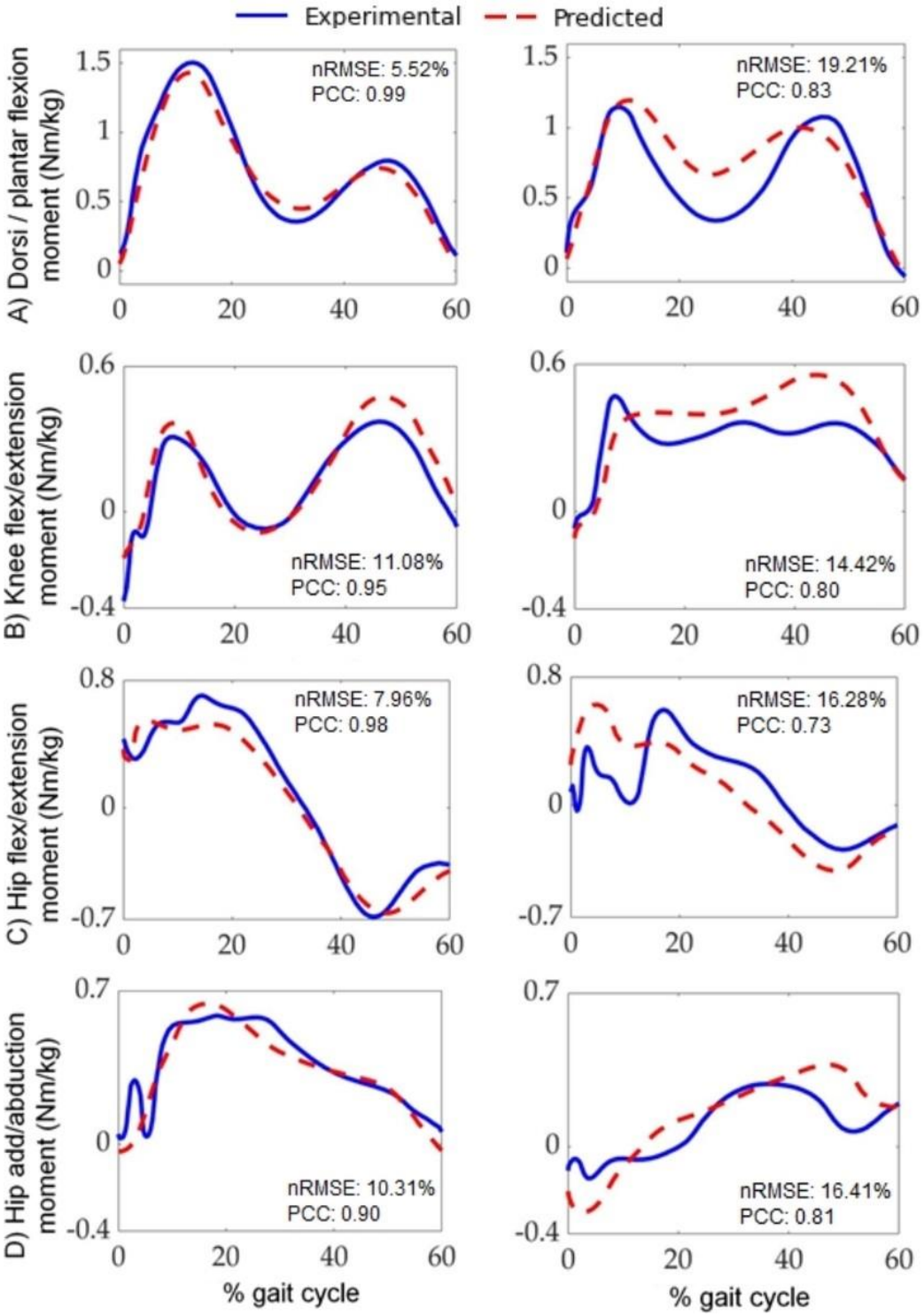


Figure 3.7: Secondary, representative results for aiding interpretations of Figure 3.4 and Figure 3.5. Joint moment predictions of a) dorsi-plantar flexion, b) knee flexion-extension, c) hip flexion-extension, d) hip adduction abduction for representative patients with cerebral palsy. The predictions on the left column correspond to above-average success rates, while those on the right column correspond to below-average success rates. The blue line represents the experimental joint moment, while the red dashed line represents the predicted joint moment.

3.3 Results of predicting joint moments of patients with cerebral palsy using deep and various conventional machine learning methods

For the random forest algorithm, the mean nRMSE value was found 14.93% (±6.81%). This suggests that, on average, the predicted ankle dorsi-plantar flexion values deviated from the experimental values by approximately 14.93%. Similarly, the multilayer neural network achieved a mean nRMSE value of 14.35% (±6.15%). The k-nearest neighbour algorithm exhibited a slightly higher mean nRMSE value of 17.94% (±7.96%). The ridge regression model achieved a mean nRMSE value of 14.19% (±6.34%). The one-dimensional convolutional neural network achieved a mean nRMSE value of 14.78% (±7.17%), while the long short-term memory network obtained a mean nRMSE value of 14.73% (±7.04%). See Figure 3.8 for a bar representation of the models' successes in terms of nRMSE values.

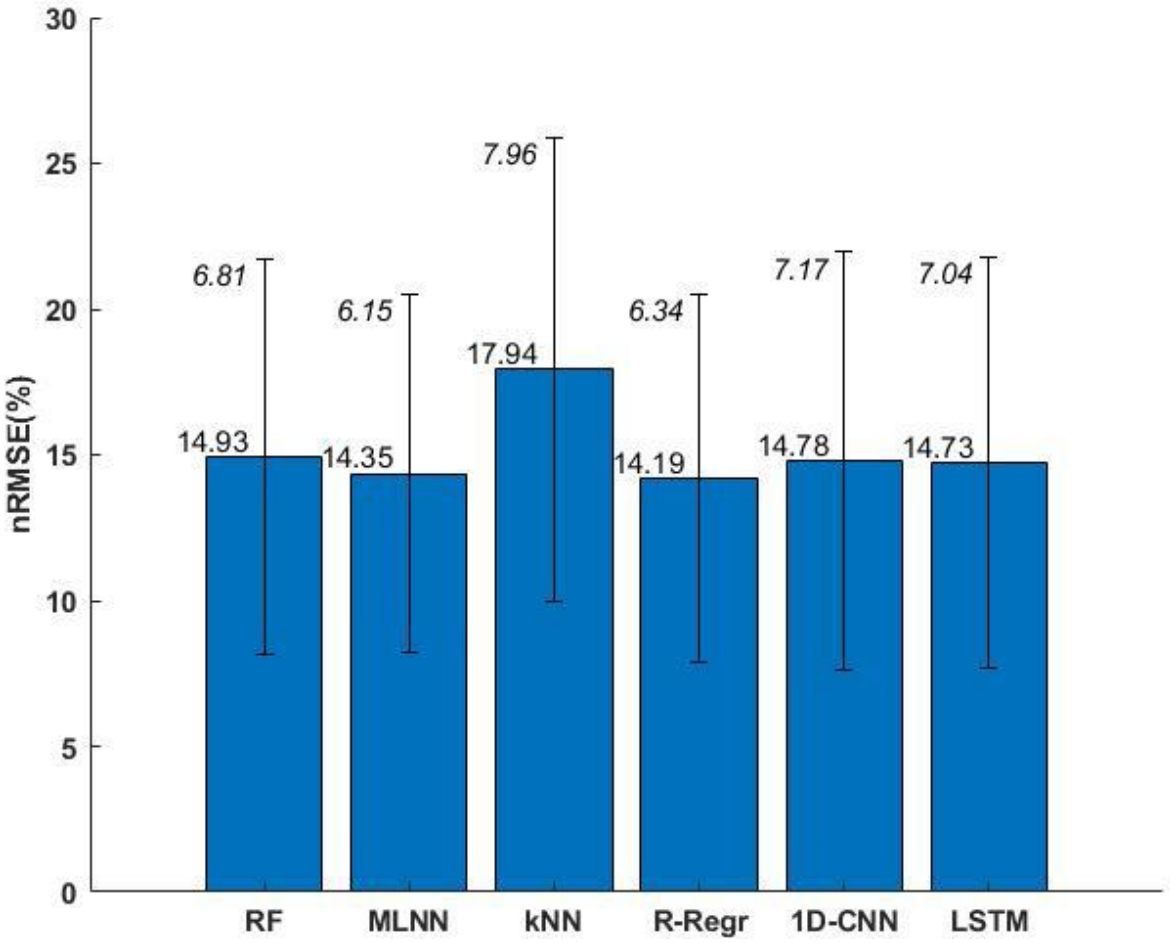


Figure 3.8: Normalized root mean square error (nRMSE) scores for predicting ankle dorsi-plantar flexion moment of patient with CP. RF: Random forest, MLNN: Multilayer neural network, kNN: K-nearest neighbor, CNN: one dimensional convolutional neural network, LSTM: Long short term memory neural network.

For ankle dorsi-plantar flexion prediction, the random forest, multilayer neural network, and ridge regression models achieved high PCC values of 0.95. This indicates a strong positive correlation between the predicted and experimental values, suggesting that these models were successful in capturing the underlying patterns in the ankle dorsi-plantar flexion data. The k-nearest neighbour algorithm achieved a slightly lower PCC value of 0.92, indicating a slightly weaker correlation. The one-dimensional convolutional neural network and long short-term memory network performed well, with PCC values of 0.96 and 0.95, respectively, indicating strong correlations between

the predicted and experimental ankle dorsi-plantar flexion values (Figure 3.9).

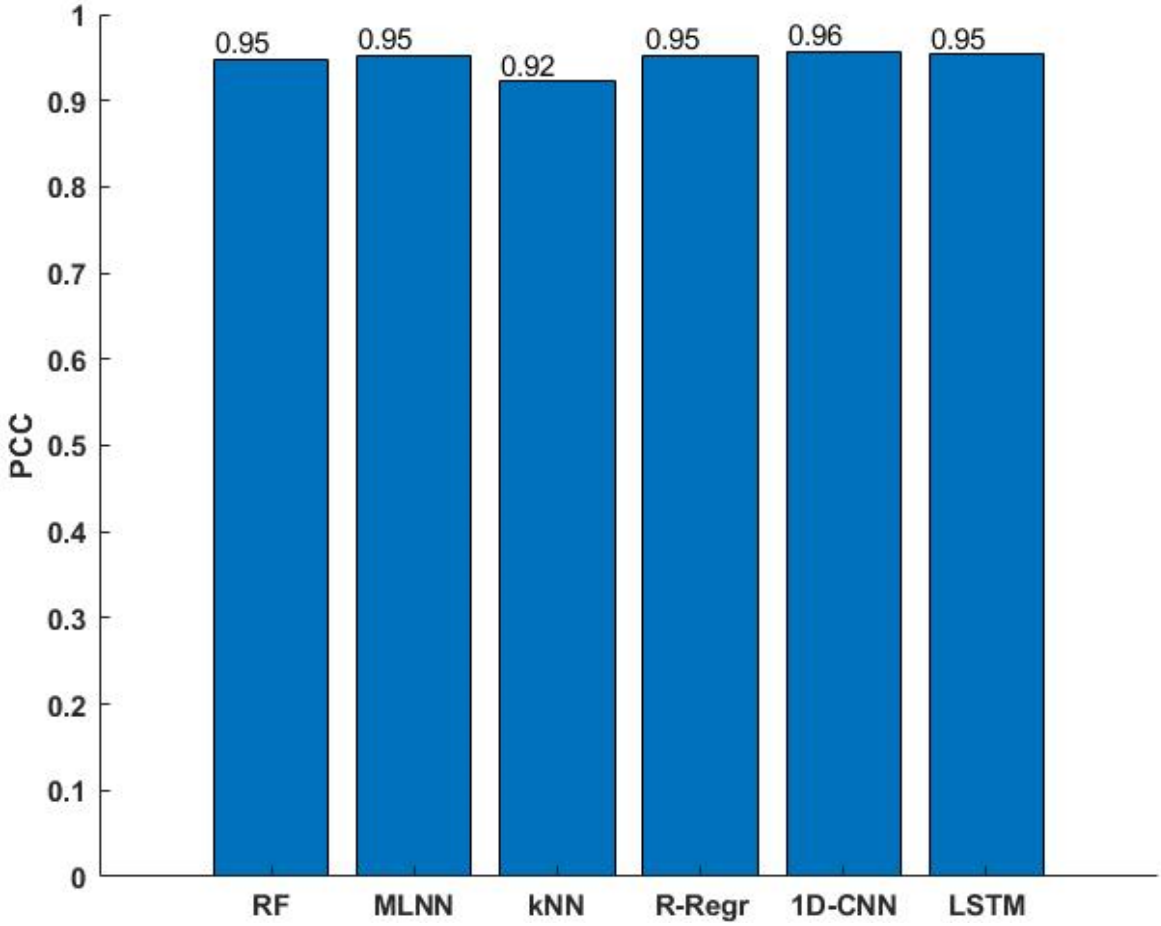


Figure 3.9: Pearson correlation coefficient (PCC) scores for predicting ankle dorsi-plantar flexion moment of patient with CP. RF: Random forest, MLNN: Multilayer neural network, kNN: K-nearest neighbor, CNN: one dimensional convolutional neural network, LSTM: Long short term memory neural network.

These results demonstrate the capabilities of the different ML algorithms in predicting ankle dorsi-plantar flexion for patients with CP. The nRMSE values provide insights into the magnitude of prediction errors, while the PCC values reflect the degree of pattern similarity. Overall, the algorithms exhibited relatively low prediction errors and strong correlations, indicating their potential usefulness in accurately predicting ankle dorsi-plantar flexion for patients with CP.

Table 3.4 presents the p-values resulting from the statistical analysis approach described in Section 2.6. These p-values provide a quantitative measure of the statistical significance of the observed differences and enable a rigorous evaluation of the obtained performance results of the six aforementioned ML models. Each row of the table represents a specific model (KNN, LSTM, MLNN, RF, R-Regr, 1DCNN), and each column represents the model being compared against (KNN, LSTM, MLNN, RF, R-Regr, 1DCNN). For example, the first row indicates the p-values for comparing KNN with other models in terms of nRMSE and PCC.

The obvious outcomes of the statistical significance results regarding the Table 3.4 should be noted as following. RF generally shows lower p-values compared to other models, indicating significant differences in performance. MLNN and LSTM often have higher p-values, suggesting less significant differences compared to other models. 1DCNN shows relatively higher p-values in some comparisons, implying less significant differences compared to certain models.

Table 3.4. *p*-values obtained for the comparison of the nRMSE and PCC values of ankle dorsi-plantar flexion moment predictions for the patients with CP. Significant differences were marked bold. RF: Random forest, MLNN: Multilayer neural network, kNN: K-nearest neighbour, CNN: one dimensional convolutional neural network, LSTM: Long short term memory neural network

		nRMSE	PCC
KNN vs	LSTM	**0.0139**	**0.0131**
	MLNN	**0.0132**	**0.0137**
	RF	**0.0126**	**0.0131**
	R-Regr	**0.0128**	**0.0136**
	1DCNN	**0.0132**	**0.0127**
LSTM vs	KNN	**0.0139**	**0.0131**
	MLNN	0.0169	**0.0159**
	RF	**0.0157**	0.0172
	R-Regr	**0.0157**	0.0161
	1DCNN	0.0169	**0.0158**
MLNN vs	KNN	**0.0132**	**0.0137**
	LSTM	0.0169	**0.0159**
	RF	0.0175	**0.0152**
	R-Regr	0.0175	0.0173
	1DCNN	0.0522	**0.0149**
RF vs	KNN	**0.0126**	**0.0131**
	LSTM	**0.0157**	0.0172
	MLNN	0.0175	**0.0152**
	R-Regr	0.0518	**0.0148**
	1DCNN	0.0497	**0.0151**
R-Regr vs	KNN	**0.0128**	**0.0136**
	LSTM	**0.0157**	0.0161
	MLNN	0.0175	0.0173
	RF	0.0518	**0.0148**
	1DCNN	0.0492	**0.0142**
1DCNN vs	KNN	**0.0132**	**0.0127**
	LSTM	0.0169	**0.0158**
	MLNN	0.0522	**0.0149**
	RF	0.0497	**0.0151**
	R-Regr	0.0492	**0.0142**

$p<0.016$

4. DISCUSSION

Since joint moments and GRFs are valuable assessment parameters in the management of CP [3]-(Lai et al., 1988) and hard to capture experimentally, we predicted vertical GRF and, dorsi-plantar flexion, knee flexion-extension, hip flexion-extension, and the hip adduction-abduction moments of patients with CP during gait from joint angles using 1D CNN in our study. We found that the joint moments of patients could be predicted with nRMSE values less than 18.02% and PCC scores higher than 0.85 and GRF of the patients with CP could be predicted with an average nRMSE value of 11.75% and an average PCC value of 0.94. In the TD group, all joint moments were predicted with nRMSE values less than 12.55% and PCC scores higher than 0.94, whereas GRF was predicted with an nRMSE of 7.47% and a PCC of 0.98. The predictions mostly captured the patterns and magnitudes of the experimentally obtained joint moments and GRF.

Mundt et al. predicted joint moments from joint angles of TD subjects using a densely connected feed-forward and an LSTM neural network achieved nRMSE scores between 12.14% to 15.00% and PCC scores between 0.92 to 0.97 on cross validation splits [19]-(Mundt et al., 2020a), whereas in our study the CNN model achieved nRMSE scores between 8.58% to 12.55% and PCC scores between 0.94 to 0.98 for TD subjects (Figures 4 and 5). There is another study predicting joint moments of TD subjects based on EMG and GRF components using wavelet neural networks, which achieved higher success in terms of nRMSE (lower than 5.69%) and PCC (above 0.99) (Ardestani et al., 2014). Using GRF as input information would increase the prediction success since GRF and joint moments are biomechanically coupled. Thus, GRF that was leaked into

the calculated joint moments was considered the golden standard in this study. In our study, only the joint angles were used as input which were separately measured and easily accessible information in routine gait analysis, hence there is no further need for costly equipment like force plates.

4.1 Discussion for Study I

Comparing the prediction of GRF between TD subjects and those with CP, TD subjects exhibited significantly higher success in terms of nRMSE and PCC (Section 3.1). The presence of diverse gait deviations in CP cases poses challenges for the learning process of CNN models, resulting in lower performance in predicting joint moments for patients compared to TD subjects. This outcome was anticipated due to the increased complexity of the coupled relationship between joint angles and GRFs in patients with CP. Despite the TD group having a relatively smaller number of subjects in comparison to the CP group, the models for TD subjects exhibited higher success rates in predicting joint moments. This finding is noteworthy considering the commonly recognized disadvantage of training machine learning models with a limited sample size, showing that strongly varying gait characteristics of CP patients won't allow ML models to gain advantage from larger sample size when compared with TD prediction successes. The GRF patterns depicted in Figure 3.3 serve as evidence that the models successfully predicted GRFs with different characteristics. To assess the models' performance, blind testing was conducted using randomly selected test splits across all participants, indicating their potential for accurately predicting gait kinetics across a range of gait patterns. This study presents a significant advancement by offering the potential to enable motion analysis of patients with CP without the need for force plates, thus

eliminating the reliance on costly equipment and simplifying the assessment process.

4.2 Discussion for Study II

The prediction of joint moments for TD subjects was achieved with a significantly higher success regarding nRMSE within all considered joint moments, whereby with a significantly higher PCC within hip adduction-abduction and knee flexion-extension moments (Table 3) when compared to those for subjects with CP. The varying deviation of gait in CP cases makes the learning process of the CNN models harder, which caused less moment prediction performance in the patient group compared to TD subjects. This was expected due to the coupled relation between joint angles and joint moments becoming more complex in patients with CP. The models for TD subjects have achieved higher success rates despite having a relatively smaller number of subjects than the CP group, which is a commonly recognized disadvantage when training ML models. The sub-classification of CP groups based on altered gait patterns, such as crouch gait and tip-toe, and training separate ML models for each subgroup could improve the prediction accuracy. We consider this attempt as the next step in gait kinetics prediction studies for CP patients.

One could argue that the prediction of moments in the joints that are closer to the ground (distal joints) would be more successful than those that are further from the ground (proximal joints) because the calculation of joint moments that is based on inverse dynamics is performed in a stepwise fashion from bottom to the top resulting in accumulating errors in calculations [16]-(Whittle et al., 2014). This was not totally observable in our results however, the joint moment with the highest prediction success

for TD subjects was the ankle dorsi-plantar flexion, which fits that expectation.

The representative joint moments presented in Figure 3.6 and Figure 3.7 indicate that the models were able to predict joint moments with different profiles. The models were blindly tested with randomly selected test splits across all included subjects, hence the performance of the models is promising for predicting gait kinetics of varying gaits. Although the results are promising, the fact that the gait analysis is used for surgical decision-making in CP makes the use of ML-based joint moment predictions limited, since the obtained error rates might still be critical for surgical decision-making. The accuracy of obtaining kinematics data from markers directly affects the correctness of joint moment prediction. Moreover, inaccurate recording of kinematics data, caused by marker misplacement or soft tissue artifacts, can result in biomechanically inaccurate joint moments (Fonseca et al, 2020). However, the successful application of this workflow would facilitate the gait analysis of patients with CP by reducing laboratory effort and eliminating the need for complex musculoskeletal models for calculating joint moments. Furthermore, this workflow can help clinicians with the treatment protocol by providing joint moments of the patients with CP, whose GRFs could not be correctly measured at all due to using assistive devices or very short stride length.

4.3 Discussion for Study III

In the part of the study aiming to predict joint moments of patients with CP using various ML methods, six ML models were developed and evaluated for their prediction success in ankle dorsi-plantar flexion. The models included kNN, LSTM, MLNN, RF, R-Regr, and 1DCNN. This part of the study also focused on comparing two types of kinematic

input data: manually extracted time domain features for conventional ML models and automatically extracted features within the deep learning models.

Interestingly, all models, except for kNN, achieved similar levels of success with relatively low standard deviations. This finding suggests that regardless of the specific form of kinematic data representation, accurate predictions of gait kinetics can be achieved if the models are trained with suitable ML algorithms. This implies that both manually extracted time domain features and automatically extracted features within deep learning models can provide sufficient information for predicting joint moments in patients with CP.

The results of this study highlight the versatility and effectiveness of different ML approaches in predicting gait kinetics. It emphasizes the importance of selecting the appropriate algorithm based on the specific dataset and research objective. While conventional ML models and deep learning models demonstrated comparable success in this study, it is crucial to consider factors such as computational efficiency, interpretability, and generalizability when choosing the most suitable ML approach for a particular application.

Moreover, the finding that different forms of kinematic data representation yielded similar prediction performance suggests that the choice between manual feature extraction and automatic feature extraction can be based on practical considerations and the availability of data. Manual feature extraction requires domain expertise and prior knowledge of relevant features, whereas automatic feature extraction allows the model

to learn complex patterns and representations directly from the raw input data. Both approaches have their advantages and limitations, and the decision should be based on the specific requirements and constraints of the study.

This part of the study demonstrates that accurate predictions can be achieved regardless of the specific data representation method. This highlights the flexibility of ML approaches in gait analysis and provides valuable insights for future research and clinical applications in the field of CP management.

4.4 Limitations

Limitations of this study should be considered. Firstly, the models were limited to the aforementioned vertical GRF and four joint moments, which are major kinetic parameters for the management of CP, however additional joint moments like hip internal/external rotation and ankle inversion/eversion may also be taken into account in monitoring CP. Secondly, the kinematics data included only the trunk from the upper body, however further kinematics data from upper extremities like arms may provide valuable information, thereby improving the ML models' prediction success rates. Thirdly, it is ambiguous if the model would be able to predict a marginal GRF or joint moment from a CP patient with a novel form of deviation, which did not show up in our subject dataset. Although the used dataset is large and has been collected over two decades, the ML algorithm should always be further developed with potential new cases' data. For example, we did not include hemiplegic and tetraplegic subjects in the study. The implementation of ML algorithms on such patients would improve the applicability of the proposed joint kinetics prediction procedure.

5. CONCLUSION AND FUTURE WORK

In conclusion, the findings from this study highlight the potential of ML-based prediction of joint moments and GRF using kinematics as an alternative technique to conventional joint moment calculation in the gait analysis of patients with CP in the near future. The results demonstrate that ML models can successfully estimate joint moments and GRFs based on kinematic data, offering a promising avenue for capturing important kinetic parameters in an accessible manner.

However, it is important to acknowledge that the current level of prediction errors may still limit the immediate use of ML-based techniques for clinical decision-making today. While the models achieved favourable prediction success rates, there is a limitation that need to be addressed before these techniques can be seamlessly integrated into clinical practice. This consideration is the clinical significance of prediction errors. Even though the models showed promising performance regarding the well-accepted evaluation metrics, the accuracy and reliability of the predicted joint moments may not meet the critical thresholds required for making surgical decisions or implementing specific treatment protocols.

On the other hand, to the best of our knowledge, the clinicians did not explicitly define a specific threshold in their academic literature, indicating a gap in the existing knowledge. This observation underscores the pressing need for stronger and more collaborative partnerships between researchers and clinicians. By fostering closer collaborations, we can bridge this gap and facilitate the exchange of expertise and insights, ultimately enhancing the integration of scientific findings into clinical practice.

Such collaborations would enable researchers and clinicians to collectively establish meaningful thresholds and guidelines that can effectively inform decision-making processes and improve patient care.

The limitations identified in this study open up possibilities for future research and development in the field of predicting joint moments in patients with CP. Here are some potential avenues for future work based on the limitations mentioned:

A potential avenue for future work in joint moment prediction studies for CP patients involves the sub-classification of CP groups according to specific altered gait patterns, such as crouch gait and tip-toe. By creating distinct subgroups based on these variations, and subsequently training separate ML models for each subgroup, it is anticipated that the prediction accuracy can be significantly improved. This approach represents a logical progression in the field, and its implementation holds promise for advancing our understanding of CP biomechanics and optimizing the accuracy of joint moment predictions in clinical settings.

Expansion of joint moment analysis is possible. While this study focused on vertical GRFs and four major joint moments, there is room for including additional joint moments such as hip internal/external rotation and ankle inversion/eversion. Considering these additional joint moments could provide a more comprehensive understanding of the biomechanics of patients with CP and further improve the monitoring and management of their condition.

It is also a good idea to incorporate the upper extremity kinematics. The current study primarily focused on kinematics data from the trunk and lower extremities. Future research could explore the inclusion of kinematic data from the upper extremities, particularly the arms. This additional information could provide valuable insights into the overall movement patterns and contribute to improving the prediction success rates of ML models.

It remains uncertain whether the developed models would be capable of predicting marginal GRF or joint moments in patients with CP exhibiting novel forms of deviation that were not present in the subject dataset used in this study. To address this limitation, future work could involve collecting data from a broader range of patients, including those with unique gait deviations or specific subtypes of CP, such as hemiplegia or tetraplegia. By including diverse cases and continuously updating the ML algorithms with new data, the applicability and robustness of the joint kinetics prediction procedure can be enhanced.

The dataset used in this study was collected over two decades, providing a substantial amount of information. However, to further improve the ML models and their generalizability, it would be valuable to gather longitudinal data from patients with CP. Long-term follow-up studies can help capture the progression of the condition, assess treatment effectiveness, and refine the predictive capabilities of the models over time.

This study focused specifically on patients with CP. Future research could explore the application of ML algorithms to other patient groups, such as individuals with different

neuromuscular disorders or orthopaedic conditions. By adapting and fine-tuning the ML models for specific patient populations, the joint kinetics prediction procedure can be extended to a broader range of clinical scenarios, enabling personalized treatment planning and evaluation.

In conclusion, future work in this field could involve sub classifying the CP groups, expanding the analysis to include additional joint moments, incorporating upper extremity kinematics, addressing novel forms of deviation, collecting longitudinal data, and applying ML algorithms to diverse patient groups. These advancements would contribute to a more comprehensive understanding of gait kinetics in various clinical populations and further enhance the clinical utility of predictive models for joint moment analysis.

6. APPENDIX A: Details of the statistical analysis for the resulting evaluation metrics of predicting joint moments (Study II)

Table. Details of the statistical analysis

Type of analysis	• Friedman's ANOVA *(for the intra-comparison)* • Mann-Whitney U test *(for the inter-comparison)*
Intra-comparison	• nRMSE values calculated between the predicted and experimental hip abduction-adduction moment vs. nRMSE values calculated between the predicted and experimental hip flexion-extension moment, • nRMSE values calculated between the predicted and experimental hip abduction-adduction moment vs. nRMSE values calculated between the predicted and experimental knee flexion-extension moment, • nRMSE values calculated between the predicted and experimental hip abduction-adduction moment vs. nRMSE values calculated between the predicted and experimental dorsi-plantar flexion moment, • nRMSE values calculated between the predicted and experimental hip flexion-extension moment vs. nRMSE values calculated between the predicted and experimental knee flexion-extension moment, • nRMSE values calculated between the predicted and experimental hip flexion-extension moment vs. nRMSE values calculated between the predicted and experimental dorsi-plantar flexion moment, • nRMSE values calculated between the predicted and experimental knee flexion-extension moment vs. nRMSE values calculated between the predicted and experimental dorsi-plantar flexion moment, • PCC values calculated between the predicted and experimental hip abduction-adduction moment vs. PCC values calculated between the predicted and experimental hip flexion-extension moment, • PCC values calculated between the predicted and experimental hip abduction-adduction moment vs. PCC values calculated between the predicted and experimental knee flexion-extension moment, • PCC values calculated between the predicted and experimental hip abduction-adduction moment vs. PCC values calculated between the predicted and experimental dorsi-plantar flexion moment, • PCC values calculated between the predicted and experimental hip flexion-extension moment vs. PCC values calculated between the predicted and experimental knee flexion-extension moment,

	• PCC values calculated between the predicted and experimental hip flexion-extension moment vs. PCC values calculated between the predicted and experimental dorsi plantar flexion moment, • PCC values calculated between the predicted and experimental knee flexion-extension moment vs. PCC values calculated between the predicted and experimental dorsi-plantar flexion moment,
Inter-comparison	• nRMSE values calculated between the predicted and experimental hip abduction-adduction moment of the CP patients vs. those calculated healthy subjects, • nRMSE values calculated between the predicted and experimental hip flexion-extension moment of the CP patients vs. those calculated healthy subjects, • nRMSE values calculated between the predicted and experimental knee flexion-extension moment of the CP patients vs. those calculated healthy subjects, • nRMSE values calculated between the predicted and experimental dorsi-plantar flexion moment of the CP patients vs. those calculated healthy subjects, • PCC values calculated between the predicted and experimental hip abduction-adduction moment of the CP patients vs. those calculated healthy subjects, • PCC values calculated between the predicted and experimental hip flexion-extension moment of the CP patients vs. those calculated healthy subjects, • PCC values calculated between the predicted and experimental knee flexion-extension moment of the CP patients vs. those calculated healthy subjects, • PCC values calculated between the predicted and experimental dorsi -plantar flexion moment of the CP patients vs. those calculated healthy subjects.

CP: Cerebral palsy, nRMSE: normalized root-mean-square error, PCC: Pearson cross-correlation coefficient.

7. REFERENCES

[1] Baker, R. (2013). Measuring walking: a handbook of clinical gait analysis. Mac Keith Press.

[2] Halilaj, E., Rajagopal, A., Fiterau, M., Hicks, J. L., Hastie, T. J., & Delp, S. L. (2018). Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities. Journal of biomechanics, 81, 1-11.

[3] Lai, K. A. Kuo, K. N. Andriacchi, T. P., 1988. Relationship between dynamic deformities and joint moments in children with cerebral palsy. Journal of Pediatric Orthopedics 8(6), 690-695.

[4] Gage, J. R., 1994. The clinical use of kinetics for evaluation of pathological gait in cerebral palsy. Journal of Bone and Joint Surgery 76(4), 622-631.

[5] Ounpuu, S. Davis, R. B. Deluca, P. A., 1996. Joint kinetics: methods, interpretation and treatment decision-making in children with cerebral palsy and myelomeningocele. Gait & posture 4(1), 62-78.

[6] Lin, C. J. Guo, L. Y. Su, F. C. Chou, Y. L. Cherng, R. J., 2000. Common abnormal kinetic patterns of the knee in gait in spastic diplegia of cerebral palsy. Gait & Posture 11(3), 224-232.

[7] Novacheck, T. F. Gage, J. R., 2007. Orthopedic management of spasticity in cerebral palsy. Child's Nervous System 23(9), 1015-1031.

[8] DeLuca, P. Davis, R. Õunpuu, S. Rose, S.; Sirkin, R., 1997. Alterations in surgical decision making in patients with cerebral palsy based on three-dimensional gait analysis. Journal of Pediatric Orthopaedics 17(5): 608-614.

[9] Kay, R. M., Dennis, S., Rethlefsen, S., Reynolds, R. A., Skaggs, D. L., & Tolo, V. T., 2000. The effect of preoperative gait analysis on orthopaedic decision making. Clinical Orthopaedics and Related Research (1976-2007), 372, 217-222.

[10] Rhodes, Jason T., et al., 2023. "Rectus femoris transfers with and without a hamstring lengthening will not change hip kinematics in children with cerebral palsy." Gait & Posture 99: 119-123.

[11] Lenhart, R. L. Brandon, S. C. Smith, C. R. Novacheck, T. F. Schwartz, M. H. Thelen, D. G., 2017. Influence of patellar position on the knee extensor mechanism in normal and crouched walking. Journal of Biomechanics 51, 1-7.

[12] Karabulut, D. Arslan, Y. Z. Götze, M. Wolf, S. I., 2021. The Impact of Patellar Tendon Advancement on Knee Joint Moment and Muscle Forces in Patients with Cerebral Palsy. Life 11(9), 944.

[13] Caldas, R. Fadel, T. Buarque, F. Markert, B., 2020. Adaptive predictive systems applied to gait analysis: A systematic review. Gait & Posture 77, 75-82.

[14] White, R. Agouris, I. Selbie, R. D. Kirkpatrick, M., 1999. The variability of force platform data in normal and cerebral palsy gait. Clinical Biomechanics 14(3), 185-192.

[15] Winter, D. A., 2009. Biomechanics and motor control of human movement. John Wiley & Sons.

[16] Whittle, M. W., 2014. Gait analysis: an introduction. Butterworth-Heinemann.

[17] Oh, S. E., Choi, A., & Mun, J. H., 2013. Prediction of ground reaction forces during gait based on kinematics and a neural network model. Journal of biomechanics, 46(14), 2372-2380.

[18] Johnson, W. R., Mian, A., Donnelly, C. J., Lloyd, D., & Alderson, J., 2018. Predicting athlete ground reaction forces and moments from motion capture. Medical & biological engineering & computing, 56, 1781-1792.

[19] Mundt, M. Koeppe, A. David, S. Bamer, F. Potthast, W. Markert, B., 2020a. Prediction of ground reaction force and joint moments based on optical motion capture data during gait. Medical Engineering & Physics 86, 29-34.

[20] Johnson, W. R., Alderson, J., Lloyd, D., & Mian, A, 2018. Predicting athlete ground reaction forces and moments from spatio-temporal driven CNN models. IEEE Transactions on Biomedical Engineering, 66(3), 689-694.

[21] Ihlen, E. A. Støen, R. Boswell, L. de Regnier, R. A. Fjørtoft, T. Gaebler-Spira, D., Labori, C. Loennecken, M. C. Msall, M. E. Möinichen, U. I. Peyton, C. Schreiber, M. D. Silberg, I. E. Songstad, N. T. Vågen, R. T. Øberg, G. K. Adde, L., 2019. Machine learning of infant spontaneous movements for the early prediction of cerebral palsy: A multi-site cohort study. Journal of Clinical Medicine 9(1), 5.

[22] Zhang, Y. Ye, M., 2019. 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 106, 33-39.

[23] Morbidoni, C. Cucchiarelli, A. Agostini, V. Knaflitz, M. Fioretti, S. Di Nardo, F., 2021. Machine-learning-based prediction of gait events from EMG in cerebral palsy

children. IEEE Transactions on Neural Systems and Rehabilitation Engineering 29, 819-830.

[24] Kim, Y. K. Visscher, R. M. Viehweger, E. Singh, N. B. Taylor, W. R. Vogl, F., 2022. A deep-learning approach for automatically detecting gait-events based on foot-marker kinematics in children with cerebral palsy—Which markers work best for which gait patterns?. Plos One 17(10), e0275878.

[25] Ardestani, M. M. Zhang, X. Wang, L. Lian, Q. Liu, Y. He, J. Li, D. Jin, Z., 2014. Human lower extremity joint moment prediction: A wavelet neural network approach. Expert Systems with Applications 41(9), 4422-4433.

[26] Harrison, S.M. Whitton, R.C. King, M. Haussler, K. K. Kawcak, C. E. Stover, S. M. Pandy, M.G., 2012. Forelimb muscle activity during equine locomotion. Journal of Experimenal Biology 215, 2980–2991.

[27] Wikimedia Foundation. (2022, December 6). Force platform. Wikipedia. Retrieved February 9, 2023, from https://en.wikipedia.org/wiki/Force_platform

[28] Hua, X. Han, J. Zhao, C. Tang, H. He, Z. Chen, Q. Tang, S. Tang, J. Zhou, W., 2020. A novel method for ECG signal classification via one-dimensional convolutional neural network. Multimedia Systems, 1-13.

[29] Malek, S. Melgani, F. Bazi, Y., 2018. One-dimensional convolutional neural networks for spectroscopic signal regression. Journal of Chemometrics 32(5), e2977.

[30] Miller, F. (2020). Hip and Pelvic Kinematic Pathology in Cerebral Palsy Gait. Cerebral Palsy, 1471-1487.

[31] Mundt, M. Koeppe, A. David, S. Witter, T. Bamer, F. Potthast, W. Markert, B., 2020b. Estimation of gait mechanics based on simulated and measured IMU data using an artificial neural network. Frontiers in Bioengineering and Biotechnology, 8, 41.

[32] Ripic, Z. Kuenze, C. Andersen, M. S. Theodorakos, I. Signorile, J. Eltoukhy, M., 2022. Ground reaction force and joint moment estimation during gait using an Azure Kinect-driven musculoskeletal modeling approach. Gait & Posture 95, 49-55.

[33] Savelberg, H.H. and Herzog, W., 1997. Prediction of dynamic tendon forces from electromyographic signals: An artificial neural network approach. Journal of neuroscience methods, 78(1-2), pp.65-74.

[34] Silver, N. Clayton, and William P. Dunlap, 1987. "Averaging correlation coefficients: should Fisher's z transformation be used?." Journal of applied psychology 72.1: 146.

[35] Wolf S, Loose T, Schablowski M, Döderlein L, Rupp R, Gerner HJ, et al., 2006. Automated feature assessment in instrumented gait analysis. Gait Posture. 23(3):331-8.

8. LIST OF PUBLICATIONS

- *Özates, Musta Erkam, et al. "Machine learning-based prediction of joint moments based on kinematics in patients with cerebral palsy."* ***Journal of Biomechanics*** *(2023): 111668.)*
- *Özates, Mustafa Erkam, et al. "Predicting Joint Moments from Kinematics in Patients with Cerebral Palsy using Deep Learning."* ***Gait & Posture*** *100 (2023): 28-29.*
- *Özates, Mustafa Erkam., et al. "Vorhersage der Bodenreaktionskräfte während Gangzyklus aus den Gelenkwinkeln durch maschinelles Lernen.",* ***12. Kongress der Deutschen Gesellschaft für Biomechanik****, (September 2022).*