Development of an Automated Gait Analysis Tool

by

Elham Dolatabadi

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Institute of Biomaterials and Biomedical Engineering
University of Toronto

Abstract

Development of an Automated Gait Analysis Tool

Elham Dolatabadi
Doctor of Philosophy
Institute of Biomaterials and Biomedical Engineering
University of Toronto
2016

Gait analysis involves the measurement of quantities that characterize human locomotion. This type of analysis helps clinicians detect gait disorder and therefore implement necessary therapeutic interventions. However, clinicians have been faced with numerous difficulties such as high dimensionality and nonlinear time dependencies among measured gait parameters. Moreover, the existing analysis of gait is compounded by long recording times in gait laboratories and delayed diagnosis of disorder. An effective analysis of gait could be facilitated if technology enabled frequent assessment in the home or in any clinic, i.e., without the need to visit a specialized gait clinic. The major goal of this thesis is therefore to investigate the use of easy-to-use sensing devices for home use, as well as novel algorithms capable of processing sensed data, e.g., machine learning methods to automatically analyze gait data. This thesis consists of four different studies toward addressing our major goal. In the first study, we examined the applicability of the Microsoft Kinect sensor as an affordable, portable and easy to use sensing technology for long-term monitoring of mobility after an orthopaedic surgical procedure. This study provided reasonable outcomes that confirmed using a single Kinect sensor for a gait and movement analysis. In the second study, we investigated the validity and reliability of the Kinect sensor for spatiotemporal gait analysis in clinical settings. The results presented in this study indicated that the Kinect sensor along with its tracking algorithm and our custom made algorithms provide excellent validity when compared to gait measures obtained from a pressure mat. In the third study, we explored the use of an unsupervised machine learning technique to separate individuals with gait disorder into homogenous groups based on selected spatiotemporal gait measures. Unsupervised learning methods for gait analysis have been shown to be very successful in (1) grouping the gait patterns with similar characteristics, and (2) generating a composite measure indicative of overall gait performance and sensitive to improvement or deterioration and rehabilitation over time. In the last study, we demonstrated the integration of an affordable and unobtrusive sensing technology and a machine learning approach (automated gait system) to distinguish pathological gait from healthy individuals. This study establishes the successful deployment of an automated gait system in a hospital clinic and during routine exams to (1) detect pathological gait pattern, and (2) identify which body part is mostly affected by this pathology. Together, the findings of this thesis improve the existing gait analysis tool to be more cost-effective and efficient in monitoring and diagnosing of pathological gait.

Acknowledgements

First, I would like to express my sincere gratitude and appreciation to my supervisor, Dr. Alex Mihailidis, for his fundamental role in my doctoral research. I appreciate his wisdom, vast interdisciplinary expertise in many areas (e.g., engineering, rehabilitation, aging, ethics and computer science), and technical guidance. His enthusiasm for the problems and sense of the "big picture" direction always inspired and motivated me. I would also like to express my gratitude to Dr. Babak Taati for his scientific advice and many insightful discussions at all levels of my doctoral endeavors. I also appreciate the guidance given by my supervisory committee members, Dr. Geoff Fernie and Dr. Karl Zabjek. Their comments and feedback were very valuable in the progress and completion of my thesis. I would also like to thank my examiners, Dr. Elaine Biddiss and Dr. Stephen Robinovitch who took time to read and examine this thesis and provided me with valuable suggestions.

Appreciation also goes out to my colleagues and friends in the Intelligent Assistive Technology Lab (IATSL) and iDAPT for their sharing of knowledge, collaboration, experience and friendship and making the lab a great place along the way. I am particularly thankful to Bing Ye, Marge Coahran, and Tizneem Jiancaro from IATSL and Simon Jones from iDAPT for their continuous assistance in collecting data, setting up the studies and editing my research papers despite the enormous work pressures they had.

This thesis would not have been possible without the support of a very special person, my loving husband, Ali. I consider myself fortunate to have his love and constant patience throughout the years. He was always around at times I thought that it is impossible to continue, he helped me to keep things in perspective. I greatly value his contribution and deeply appreciate his belief in me. I must acknowledge my dearest son, Borna, who was born before I finished my doctoral research. I owe my deepest gratitude to him for being a great baby during pregnancy and during his first year while I was collecting and processing my data until I finished this thesis. I am deeply sorry for the time we spent apart. I love you and I promise I will always be there for you. I am extremely grateful to my mom, Azam, my dad, Mohammad Ali, and my sister, Sepideh. Their support, encouragement and friendship have made me who I am and brought meaning to my endeavours. I am also thankful to my mother-in-law, Shahin and my father-in-law, Hamid, for looking after Borna at a moment's notice and for all their profound understanding. Words would never say how grateful I am to all of you. I consider myself the luckiest in the world to have such a lovely and caring family standing beside me with their love and unconditional support.

In conclusion, I recognize that this research would not have been possible without the financial assistance of Natural Sciences and Engineering Research Council CREATE CARE, Toronto Rehabilitation Institute graduate scholarship and my supervisor's research funding, and express my gratitude to those agencies.

Contents

List of Tables

List of Figures

Acronyms

ABI Acquired Brain Injury. xiii, 12, 71–75, 82–84, 87, 88, 90, 91, 95

AVHE Angular Velocity of Hip extension. 10, 25, 28, 30, 94

AVHF Angular Velocity of Hip flexion. 10, 25, 28, 30, 32, 94

BBS Berg Balance Scale. 53, 56, 58, 61, 64–66

BIC Baysian Information Criterion. 11, 14, 50, 52, 56

BMFClinic Balance Mobility and Falls Clinic. 36

CMSA Chedoke-McMaster Stroke Assessment. 53, 56, 58, 61, 64–66

CoM Center of Mass. 22–24, 38

CV cross validation. 12, 80, 81

DTW dynamic time warping. 71, 78, 82, 90, 91

EM Expectation Maximization. 51

GMM Gaussian Mixture Model. 50, 51

GPC central pattern generating. 8

GPLVM Gaussian Process Latent Variable Model. 70, 71, 78–84, 88, 90, 91, 93, 95, 96

HCoM Horizontal Velocity of CoM. 10, 25, 28, 30, 32, 94

HMM Hidden Markov Model. 70, 79, 82–84, 90

ICC Intraclass Correlation Coefficient. 11, 40–42, 44, 94

K4W Kinect for Windows. 5, 8, 10–14, 36–38, 40–42, 45, 46, 71, 75, 93–96

kNN k-Nearest Neighbor. 71, 78

LoA Limit of Agreement. 11, 40–42, 44

LOOCV Leave One Out Cross Validation. x, 12, 14, 50–52, 54–56, 63

ML Machine Learning. 3, 4, 6–9, 11, 12, 14, 69–71, 93, 95, 96

MS multiple sequences. 76, 80, 83

PCA principle component Analysis. 78, 79, 82–84

ROC eeceiver operating curve. 79

SCG scaled conjugate Gradient. 79, 83

SD standard deviation. 40, 44, 50, 52, 54, 56, 63, 83

SDK software development kit. 75, 94, 96

SRStance Symmetry Ratio for Stance Time. 54, 57, 60, 63, 64

SRStepL Symmetry Ratio for Step Length. 54, 57, 60, 64

SRSwing Symmetry Ratio for Swing Time. 54, 57, 60, 64, 65

SS single sequence. 76, 80, 83

StepLCV Coefficient of Variation for Step Length. 54, 57, 58, 60

StepTCV Coefficient of Variation for Step Time. 54, 57, 58, 60, 63

STS Sit to Stand. 10, 13, 18–24, 28, 31, 32, 94

THR Total Hip Replacement. 10, 13, 18, 20, 21, 28, 31, 32

TRI Toronto Rehabilitation Institute. 21, 36

VCoM Vertical Velocity of CoM. 10, 25, 28, 30, 32, 94

WD Dual Task Walking. 11–13, 34, 36, 37, 42, 45, 75, 82–84, 87, 88, 93

WFP Fast Pace Walking. 11–13, 36, 75, 82, 84, 87, 88, 93, 95

WSP Self Pace Walking. 11–13, 36, 37, 41, 75, 82–84, 87, 88, 93

Chapter 1

Introduction

1.1 Statement of Problem and Rationale

1.1.1 Gait Analysis

Gait is a dynamic rhythmic motion that involves a balance between various parts of the body (limbs, hands, head, arm and trunk) and the environment [4–7]. Problems with any of these areas can lead to gait disorders, which are on the rise these days. Studies have drawn attention to the significant impact of gait disorders on an individual's quality of life and independency as shown in Figure 1.1 [8, 9]. Some falls are a direct consequence of gait disorder and can lead to major health-related incidents including injuries, morbidity, fear of falling and immobility and sometimes mortality.

There is a large proportion of the population suffering from gait disorders, including individuals with neurological conditions such as cerebral palsy, traumatic injury to the brain or cerebrovascular accident; overt diseases such as stroke or Parkinson's; and orthopedic surgeries such as hip replacement. Among these, ageing is the leading cause of gait dysfunction. Almost 15% of people 60 years and older suffer from deteriorated gait [8].

Gait analysis is the systematic study of human locomotion which involves characterizing and quantifying the trajectories of body parts over time [10]. Clinicians, therefore, rely extensively on gait analysis to diagnose gait disorders, identify balance factors, examine the effectiveness of the therapeutic treatment or surgical operations and make required therapeutic interventions [11, 12].

1.1.2 Measurement Techniques for Gait Analysis

Gait analysis is usually performed either in the clinic through observational assessment or in the laboratory through quantitative measuring devices. The former is a quick, easy and inexpensive method, which involves scoring all different elements of locomotion and balance

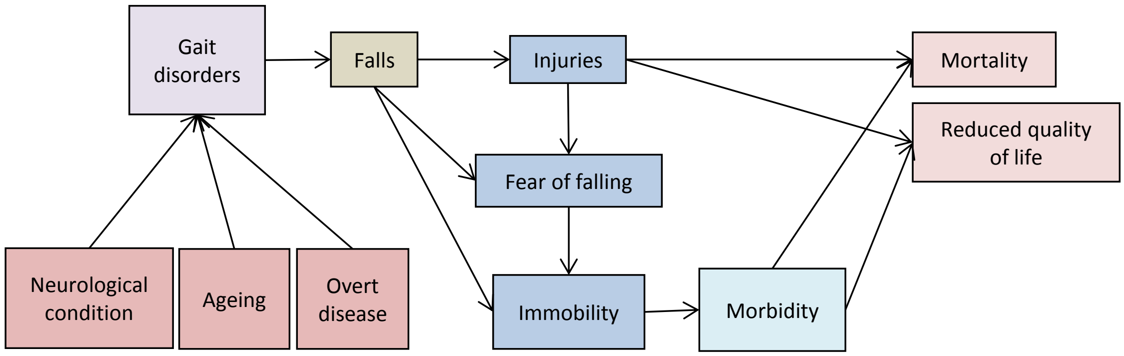

Figure 1.1: The causes of gait disorders and their impacts on individual's quality of life and independence.

through observations (e.g. Dynamic Gait Index [13], Berg Balance Scale [14], Timed up and Go [15]) [16]. However, the outcomes are hampered by poor validity, reliability, sensitivity, and specificity [17–21]. Moreover, physical examination rooms at clinics are too small and cannot reflect disturbed gait since patients must walk, turn and also do other tasks in tight quarters. That is why patients with gait disorders usually walk normally or indicate improvements in gait when walking in the examination room. Gait assessment using quantitative measuring devices gives the clinician objective information about a patient's gait with higher accuracy and reliability, and thus overcomes most of the limitations of the observational assessment in the clinic [22]. Despite this fact, their usability is limited in daily life applications and clinical settings because of high cost and time-consuming, elaborate and cumbersome setup requirements, i.e., the gait trials must be conducted by specialized expert staff [23, 24]. Due to these properties, use of existing quantitative measuring devices is limited outside laboratory settings. However, the underlying causes of gait abnormalities come to light when gait is examined in a real environment with natural settings and barriers, not a clinic or laboratory [8]. For instance, in gait laboratories, gait is assessed under constrained and highly unusual circumstances which affect patients' free movement.

In addition, both assessment techniques (observational assessment in the clinic and quantitative assessment in the laboratory) capture a snapshot of gait and assume that information obtained during the examination represents the patient's mobility status for a longer period of time prior to the assessment. Self-report questionnaires [25] are the only tools used by clinicians to follow-up patients during their stay at home. However, outcome measures derived from self-reports fail to provide reliable information as they depend on patient's personal impressions and feelings. Consequently, analysis of gait could be facilitated if technology enabled frequent examination of gait in the home, i.e., without the need to visit a specialized gait clinic.

Existing quantitative measuring devices for gait analysis generate lengthy measures in terms of the data produced. These measures indicate complex movement of body muscles and joints using computer-based pressure sensors and 3-D optical tracking which is hard to interpret [23]. The interpretation of this information is very challenging [26] due to its high-dimensionality, temporal dependence, high variability, and nonlinearity.

Recent advances in Machine Learning (ML) techniques have inspired research and development efforts toward intelligent data analysis techniques. These approaches have proposed algorithms capable of analysing patterns of walking [10, 27–32]. ML methods can potentially be used in a variety of gait analysis applications, such as to diagnose a gait disorder [27–29, 33, 34], to monitor rehabilitation recovery after surgical intervention [35], or to predict the risk of falls based on patterns of gait [36, 37] and rehabilitation assessment of stroke patients [38].

1.1.3 Gait Analysis within the context of smart homes

Long-term monitoring and screening of a change in gait over time is crucial for early diagnosis of gait disorder or movement dysfunction occurring slowly over time [39, 40]. This would enable implementation of effective strategies in a timely manner to prevent or reduce severe negative outcomes, as explained in section 1.1.1. Due to the significance of long-term monitoring of gait, the aim of this thesis is to turn an assessment of gait into an inherent component of the home. Development of a home-based gait assessment tool can potentially be achieved through (1) improving gait analysis with innovative devices that provide easy recording, portability and unobtrusiveness, and (2) an expansion of gait analysis with ML involvement.

A home-based gait analysis tool not only would enhance patient safety by monitoring their health conditions, but also reduce the used for expensive human intervention from a professional care provider. It also reduces the stress and anxiety in patients associated with controlled clinical gait studies [41].

1.2 Thesis Goal

The main goal of this research is to design and develop an automated gait analysis tool for people with gait disorder with the following characteristics:

- **Intelligence**: It capitalizes on the state-of-the-art ML algorithms and computational techniques for automated analysis of gait patterns.
- **Ubiquity**: It is portable and can be easily integrated into an individual's everyday environment.
- **Affordability:** It is cost effective.

- **Unobtrusiveness:** It is vision-based and markerless, so it requires no physical contact.

This thesis looks to lay the initial groundwork toward development of an automated gait analysis tool. Conducting feasibility analysis and prototyping of the concept design are critical steps in the development of a clinical product. Therefore, the purpose of the studies conducted in this thesis was to verify that certain technologies and methods have the potential for gait analysis. Accordingly, the specific objectives of the thesis are:

Investigation of the use of the Microsoft Kinect sensor as an appropriate sensing technology for gait analysis

O1 To determine whether we can capture the gait parameters using the Kinect sensor and detect changes in gait parameters over time using a custom made algorithm for a mobility impaired individual throughout rehabilitation recovery at home.

O2 To determine the concurrent validity and intra-reliability of the Kinect sensor and custom made algorithms in measuring gait parameters in comparison with the well-established clinical gait tools.

Explore the use of ML algorithms to classify a person's gait

O3 To investigate the use of an unsupervised ML algorithm for retrieving underlying gait patterns through clustering.

O4 To investigate the combination of an affordable sensing technology (the Kinect sensor) and supervised ML algorithm to (a) automatically discriminate between healthy and pathological gait, modelled here as a (binary) classification task, and (b) to identify the role of each body segments (upper limb, trunk and lower limb) in separating the pathological patterns from the standard healthy pattern.

The long-term plan of this research is to use the results of this thesis to develop an automated gait assessment tool, which will provide continuous monitoring of a person's gait within a smart home environment.

1.3 Literature Review

1.3.1 Sensor Technologies for Gait Data Acquisition

Recent advances in sensor technology benefit gait analysis through utilizing different types of sensors including force sensors, motion sensors such as accelerometers and gyroscopes, magnetoresistive sensors, and electromagnetic tracking systems [10, 42, 42, 43, 43–48]. Based on these sensors, a single type or multiple types of sensors may be worn or attached to various parts of the participant's body, called wearable sensors, or embedded into the environment, called ambient sensors [10, 42, 43]. Motion sensors, including accelerometers and gyroscopes, are usually attached to the foot and waist and have been used in several studies to measure motion related to human gait including the angular accelerations and velocity of the feet [42, 43]. Magnetoresistive sensors are used to estimate changes in the orientation of a body segment in relation to a determined origin [44–46]. Force sensors are embedded into footwear or embedded into the surface on which the subject walks. The electromagnetic sensing system includes a tracker and a transmitter worn by the subject. It provides 3-D positions and orientations of the subject in relation to the transmitter worn by the subject [47, 48].

Although wearable sensors have made great progress and have shown good applications in various clinical conditions and monitoring physical activity, they suffer from limitations, including short battery life, trouble capturing information about the environment and context of activities, and sensitivity to sensor placement (e.g., foot versus ankle). Moreover, using multiple wearable sensors can cause individuals discomfort. Demiris et al. [49] has raised the issue of older adult's refusal to wear sensors all the time and to use them in actual emergency cases. In addition, wearable sensors sometimes affect user's mobility and gait as some of them would function only within a certain area. For instance, footwear gait analysis equipment can hinder the gait of mobility impaired-individuals [30].

Ambient sensors such as pressure mats are more comfortable and easier to use than wearable sensors; however, they only measure gait parameters based on foot placements and/or ground contact forces and cannot provide any information about the motion of the foot above the surface. Moreover, they constrain each subject to walk within the narrow width of the mat, which is challenging for populations with vision impairment, stroke or brain injuries (since they show walking deviation, and have difficulty avoiding the sensors on the mat or may step off the mat [50]).

A new application domain of computer vision has emerged over the past few years dealing with the analysis of human gait from RGB images and videos. This work seeks to retrieve the human pose and motion in the form of 3-D joint positions, orientations and forces during walking, running and playing games [51–55]. Recovery of human body parts is typically obtained at low cost and unobtrusively since there is no need to place markers on or around the subject's body. However, their applicability is limited to scenarios with a controlled environment. Moreover, they require complicated algorithms in order to track people, find their body parts, and analyze their behavior.

Recently, researchers have begun to address these issues through utilizing depth sensors for human motion analysis. Depth sensors such as the Kinect for Windows (K4W)s offer several advantages over RGB cameras, wearable sensors and ambient sensors. The depth sensor in K4W consists of an infrared laser projector combined with a semiconductor sensor, which captures video data in 3-D under any ambient light condition. It also simplifies silhouette extraction and human tracking. Moreover, it is more affordable and convenient than RGB cameras to use and tracks the human body without the need to wear any markers. The Kinect provides real-time tracking of people within its field of view through fitting a skeletal model to each person to track his or her major body parts (head, shoulders, elbows, hands, etc.) in real time, i.e., 30 frames per second [56]. As such, it has attracted considerable attention in gait analysis [57–64].

1.3.2 Machine Learning Techniques for Gait Data Processing

Machine learning has proven to be a promising and powerful approach for characterizing high dimensional (lengthy) gait data which is not simply achievable with traditional statistical methods [26, 65]. ML methodologies for gait analysis can be divided into three major categories. The first category, gait classification, is the first step toward the development of a gait analysis tool. Gait classification is the main focus of this thesis and will be explained in detail in the next paragraphs. The second category of using ML in gait analysis is gait prediction. The goal of a gait prediction tool is to estimate secondary gait parameters which are difficult to collect outside clinical settings (e.g. EMG of the semitendinosus and vastus medialis muscles from kinematic data) from primary gait measures (e.g. hip and knee angles) [66–71]. Gait generation is the third category and aims to develop artificial gait systems such as neural prostheses and robot navigation [72–74].

Gait Classification

A primary purpose of gait classification is to describe the hidden structure of the gait patterns using unsupervised learning methods [75–78] - known as the clustering task - or to automatically discriminate between healthy and pathological gait using supervised learning methods [27, 34, 79–82] - known as the binary classification task. In supervised learning methods, the gait measures are first labeled by a clinician and the objective is to design a model which predicts the labels for the recorded gait measures. Unlike the supervised method, there are no predefined labels that guide the classification in an unsupervised approach, and the model tries to learn the hidden structure of data through the optimization of a metric.

1. Unsupervised Gait Clustering

In gait analysis, unsupervised methods find natural groupings (or clusters) among multidimensional gait measures and then assign a new subject with unknown gait conditions to the discovered groups. Unsupervised learning techniques include k-means, fuzzy [75], hierarchical [76] and self organizing map [77, 78] clustering methods. Among the aforementioned approaches, fuzzy clustering is the most prominent method since it performs soft clustering on data. Soft clustering methods assign a membership index to a data point for each cluster. The value of the membership index indicates the strength of the association of the data point to each cluster.

Hence, it is very suitable for gait analysis; and the changes in gait during rehabilitation recovery can be detected and quantified as changes in cluster membership [83].

Despite the proven benefits of clustering gait data with fuzzy approaches, there is a major issue regarding their use in gait analysis [83]. This shortcoming arises from the fact that all clustering approaches used in gait analysis rely solely on different measures of similarity (e.g. distance, connectivity, and intensity) to form the clusters. This exclusive dependency make the clustering algorithm suffer from noise and outliers which are commonly present in gait measures [84].

On the basis of this literature, the use of a novel clustering approach which can overcome the aforementioned issues was studied in this thesis. The mixture model clustering approach [85] as an alternative to aforementioned clustering techniques is a more statistically formalized method, which not only handles noisy gait data and outliers but also benefits from soft clustering. This approach makes the problems of choosing the best model and number of clusters an optimal numerical choice problem. This clustering analysis is a powerful numerical methodology which enables the use of an approximate Bayes factor to compare different mixture models and choose the model with the best clustering algorithm and optimum number of clusters. Moreover, mixture model clustering can accommodate clusters that have different sizes, shapes and orientations.

2. Supervised Gait Classification

An automated binary classification of gait using supervised ML approaches has two important applications:

- Binary classification of gait has practical applications in early identification of abnormality in gait pattern. Early diagnosis of gait disorders specifically in the older adult population prompts appropriate intervention, which further results in preventing dysfunction and loss of independence [39].
- Successful binary classification of gait will demonstrate a reasonable capacity and provide the basis for automated clinical classification of gait disorder into different classes. Manually classifying gait disorder into multiple groups (hierarchical, anatomical, aetiological, or phenomenological) is now used in clinics [8]. For instance, in anatomical classification, the gait is divided by clinicians into two groups of frontal and cerebellar which are common neurological gait abnormalities in older adults. The frontal gait disorder (apraxia) as a result of frontal lobe pathology causes variable combinations of disequilibrium with a wide stance base, increased body sway and falls, loss of control of truncal motion, locomotor disability with gait ignition failure, start hesitation, shuffling, and freezing [86]. Cerebellar disorder is most commonly seen in cerebellar disease and its typical features are reduced cadence with increased balance related variables and an almost normal range

of motion (with increased variability) in the joints of the lower limb. The tandem gait paradigm accentuates all the features of gait ataxia and is the most sensitive clinical test [87]. On the basis of clinical observation, it may be difficult to distinguish these gaits as they share common clinical characteristics such as unsteadiness, slowing, and shuffling. However, **automated** and **quantitative** classification of gait disorder leads to quantitative monitoring of gait improvement/deterioration as a **unified** language among health professionals.

Although supervised learning approaches have offered a variety of novel solutions on successfully classifying a person's gait through different approaches [10, 23, 27–30, 42, 88–91], they are not without shortcomings. The key issue is that in several recent studies, automated classification of gait was mostly achieved using standard classification methods such as artificial neural networks and support vector machines [92–94]. However, the standard ML methods do not capture temporal dependence in gait data. Moreover, some of previous studies only used features extracted from the movement of the lower body for their gait analysis [95].

Human gait exhibits complex and rich dynamic behavior that consists of cyclic events controlled by central pattern generating (GPC) circuitry [96, 97]. GPC circuitry acting on each limb during gait output complex patterns of muscle activity and regulate phase relations among different muscle groups. Studies have shown that three component factors can account for a human gait including those of upper limb, trunk and lower limb [97–99]. For these three components, changes in muscle activity, joint angles and torques vary across time, subjects and speed. The results of these studies suggest that synergies of upper limb, truck and lower limb should be defined in the state-space of the orientation angles of each limb segment with respect to the direction of the gravity and that of forward progression.

It is therefore useful to explore ML methods to incorporate the dynamics (temporal dependence) of orientation angles for the upper limb, trunk and lower limb in order to better classify the gait sequences. In visual tracking applications, where the input of the classifier is sequential body joint orientations or positions (e.g. when using a K4W), dynamics can provide a powerful model that captures the essential structure of the sequential data in the presence of noise and bias of the samples. Another issue arises from the fact that most of the ML approaches developed so far are presented as black boxes which restrict clinicians to graphically visualize the relationships between the gait disorder and the learned ML model [12, 100]. That explains why clinicians are still not satisfied with the use of these technologies in their assessments. Therefore, much work remains to be done in improving the interpretation of gait analysis using ML approaches.

From this review, the last study of this thesis focuses on (1) utilizing a supervised ML approach that is more powerful in modeling high dimensional and complex human gait in com-