EMG-Informed Estimation of Human Walking Dynamics for Assistive Robots

By

Homayoon Zarshenas

Supervised By Prof. Thor Besier

Co-supervised By Dr. Bryan Ruddy

Auckland Bioengineering Institute

University of Auckland

March 2022

A dissertation submitted in partial fulfilment of the requirements for the degree of Ph.D. in Bioengineering

Abstract

The demand for gait rehabilitation is increasing globally, and conventional rehabilitation practices cannot cope with this increase. Robotic-based rehabilitation and assistive robots are alternative solutions for gait rehabilitation, but challenges remain to bring this technology into the clinic. Assistive robots would ideally provide a personalized level of assistance based on an individual's physical and neurological condition, biomechanics, and muscular fatigue. An assistive robot should also produce a smooth movement based on the user's motion intention. Thus, a prediction of motion intention and corresponding adjustment of the robot actuator forces are the fundamental requirements for a controller of an assistive robot.

Electromyography (EMG) signals have been used widely for motion intention estimation. However, most EMG-based models are subject or task-specific, requiring complex calibration. Creating an accurate, EMG-based motion estimation model which is generalizable across individuals and experimental conditions is a major challenge and was the goal of this thesis.

The chosen application was to predict the motion and moments of the ankle joint during a range of different walking conditions. As such, a set of experiments was designed to collect motion-related data from 10 individuals during a wide range of activities. Initially, an artificial neural network was designed to predict ankle moment during constant speed walking based on a list of input time series, including the EMG signals of four muscles from each leg and ankle kinematics. The results helped find the list of most important input time series and the length of information required for ankle moment prediction at each step. Next, a machine learning approach was explored, including feature extraction and selection from the input time series. The selected list of features optimized the model training process and was generalizable across individuals to estimate the ankle moment during constant-speed walking. Exploring the influence of the training dataset on model predictions at various walking speeds was the focus of the next step. It was discovered that training the model on acceleration data from 0.5 m/s to 2.5 m/s enabled the model to predict ankle moment during walking at any speed in this range. Random forest, backpropagation neural networks, and linear regression were compared as potential predictive models, with the random forest having the best predictions across walking speeds.

In addition to making the model compatible with a range of activities, the desire was to update the model parameters based on the error between the model output and target value regardless of the training dataset. An adaptive model was developed and implemented to predict ankle angle during walking at four different speeds and three inclines to achieve this. The base model was initially trained on data from level ground walking on one participant at 1 m/s. The simplicity of the model

structure made it possible to update the parameters whenever there was an error between the predicted and actual ankle angle with less than a 30 ms time delay. The RMSE of the model for all of the test conditions was less than 5 degrees across the cohort of ten individuals (including nine unseen individuals). Continuous and accurate prediction of joint kinematics under different walking conditions and multiple individuals promises a stable and reliable control for wearable assistive robots, thus achieving the goal of the thesis.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my Ph.D. supervisors, Prof Thor Besier and Dr. Bryan Ruddy. They have provided me with the opportunity to work on such a prodigious project and helped me through this journey. They allowed me to explore my ideas, make mistakes and learn from them while patiently guiding me to find a solution. I also would like to extend my appreciation to The MedTech Centre of Research Excellence (MedTech CoRE) for providing financial support for the duration of my Ph.D. study. My special thanks go to Prof Thor Besier. It was an absolute honour to work with him. He taught me skills and supported me in the last four years, even outside office hours. I encountered many problems during my Ph.D., and his persistence and guidance led me to the end of the tunnel with solutions.

I'm incredibly grateful to Dr. Andreas Wilhelm Kempa-Liehr for his invaluable advice, continuous support, and guidance. His immense knowledge has taught me critical lessons during the past four years and helped me grow in my professional life. I also appreciate all the support from the Auckland Bioengineering Institute, especially the Bioinstrumentation lab.

Finally, my appreciation goes out to my mom, dad and my sister, who sacrificed everything for me and provided me with an opportunity to fly and follow my dreams. I always remember and cherish their love and support, which helped me to get this far in life. There is no word to describe my appreciation and gratefulness for the unconditional lover I have received from my family, I can just say Love You Forever.

Table of Contents

Table of Figures

List of tables

List of Abbreviations and Acronyms

Abbreviations and Acronyms	Descriptions
3D	3-Dimensions
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
AR	Auto Regressive Model
ECoG	Electrocorticogram
EEG	electroencephalogram
EMG	Electromyogram
FES	Functional Electrical Stimulation
FMD	Frequency Median
FR	Frequency Ratio
Gas-Med	Gastrocnemius Medialis muscle
Gas-Lat	Gastrocnemius Lateralis muscle
GRF	Ground Reaction Force
ID	Inverse Dynamics
IK	Inverse Kinematics
IMU	Inertial Measurement Unit
LDA	Linear Discriminant Analysis
LSTM	Long Short Term
MAV	Mean Absolute Value
MDS	Multi-Dimensional Scaling
ML	Machine learning
MVC	Maximum Voluntary Contraction
NARX	Nonlinear AutoRegressive network with eXogenous inputs
NMS	NeuroMusculoskeletal System
NN	Neural network
NRMSE	Normalized Root Mean Squared Error
PCA	Principal Component Analysis
PSD	Power Spectrum Density
R^2	Coefficient of Determination
RMSE	Root Mean Squared Error
RTT	Repetitive Task Training

Sol	Soleus muscle
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
TDANN	Time Delayed Artificial Neural Network
Tib	Tibialis anterior muscle
TSFRESH	Time Series FeatuRe Extraction on basis of Scalable Hypothesis
VEL	Velocity
WPT	Wavelet Packet Transform
WT	Wavelet Transform
ZC	Zero Crossing

1 Introduction

For people with physical disorders, such as stroke patients, it is necessary to perform rehabilitation exercises for a sufficient duration and intensity to retrieve their physical movement. As conventional rehabilitation methods are not coping with the increasing demand for rehabilitation, new emerging technologies come into play, and robotic rehabilitation is an alternative solution. In this chapter, the novel approaches in robotic rehabilitation and assistive robots are explained, and their benefits and limitations are discussed. The main goal of this thesis is defined and justified based on the gap of knowledge in the field of assistive robots. Finally, the steps necessary to fulfil the study's main goal and the driving research questions are described.

1.1 Motivation

One billion people, or 15% of the world's population, experience some form of disability, and disability prevalence is higher in developing countries [1]. Neurological disorders, such as stroke, are the leading cause of motor disabilities worldwide. Approximately 15 million people around the world suffer a stroke each year [2]. Lifestyle factors including high blood pressure, smoking, diabetes, high blood cholesterol level, heavy drinking, high fat diet, and lack of exercise have increased the rate of stroke by more than 40% among younger adults (under 45 years old) in the past decade [3]. Moreover, population ageing is another reason for the increased incidence of stroke. The population above 65 years of age will double by 2050. These people are at greater risk of stroke [4].

Stroke can result in varying disabilities, from sensual to motor control. Motor disabilities from a stroke might range from slight to severe. Some people make a speedy recovery and return to normal lives by undergoing a short-term rehabilitation process. Others have disabilities that may improve with time and can be managed over a longer time. For many, disabilities may last a lifetime with the need for rehabilitation and home care [5]. About 80% of stroke survivors suffer from significant motor impairments, and only 10% of them recover completely [5]. In New Zealand, stroke is the second single biggest killer and the leading cause of severe adult disability [6]. 18% of the population over 15 in New Zealand experiences some disability, of which 85% are caused by stroke [6].

One of the significant challenges that post-stroke patients have to deal with is limitations in movement. The neurological disorder of stroke degrades motor control and subsequently leads to lack of use, muscle atrophy and secondary health problems. Movement dysfunction occurs in both lower and upper limbs and can affect one side or both sides of the body, depending on the affected part of the brain and the severity of the damage [7]. Post-stroke patients are generally advised to start physiotherapy as soon as possible after the stroke to improve recovery of the lost movement. Stroke recovery can also be altered by the diversity and duration of treatment [8]. However, limitations in access and cost of attending a clinic and having trained clinicians make it difficult for some of the patients to have access to rehabilitation support, particularly considering the growing number of patients.

One solution to make rehabilitation accessible is to use assistive robots [9]. In the past decade, improvements in the design and manufacturing of assistive robots have made the interaction between robot and human safe, reliable, and comfortable. However, there is still room for improvement in terms of control of these robots. One of the main challenges of assistive robots is

providing an intuitive interaction between the wearer and the robot. An intuitive interaction means the robot needs to be aware of the user's motion intention and adjust the level of assistive force based on this knowledge.

Several approaches, including using mechanical feedback from body-robot interaction force [10, 11] and bio-signals such as electroencephalogram (EEG) [12, 13] and electromyogram (EMG) [14-16], have been suggested to predict the motion intention and adjust the robot's response correspondingly. Among all of the proposed methods to inform the robot about the movement interest of the wearer, muscle EMG signals are widely used. Two main characteristics of EMG signals make them a suitable option for robot control:

1. Electromechanical delay.

   There is a time delay between the onset of the electrical activity of muscles (EMG signals) and measurable tension in muscles, which is called electromechanical delay [17] (Figure 1.1). Thus, it is possible to predict the movement based on EMG signals before the actual motion occurs. This time delay ranges from 30 ms to 100 ms, depending upon the muscle size and location, muscle fibre type, and the intensity of the movement [17].

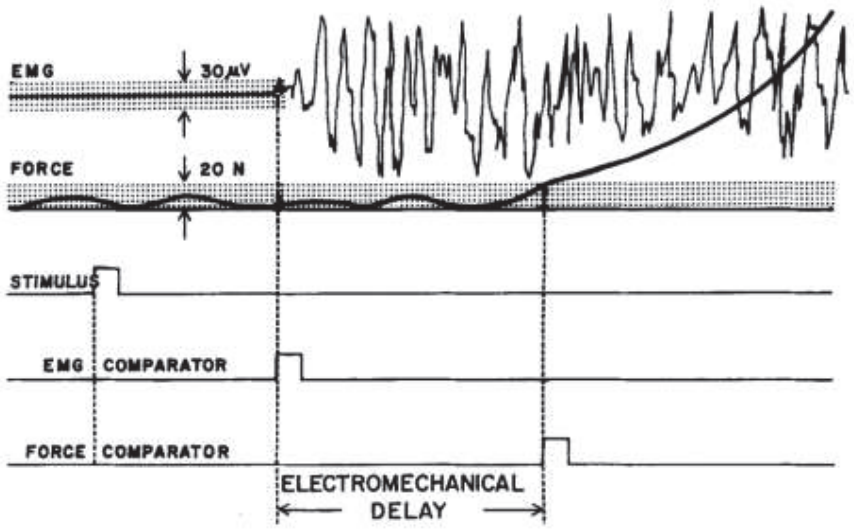


Figure 1.1 The demonstration of time delay between EMG signals and physical muscle contraction [17].

2. Ease of use and accurate data acquisition process.

   EMGs are the most reliable signals to detect motion intention. They offer a higher signal-to-noise ratio and need less expensive data acquisition equipment in comparison to Electroencephalogram (EEG) and Electrocorticogram (ECoG).

   Besides EMGs, the other electro-biological signal providing information about movement intention is Electroencephalogram (EEG). EEG is recorded directly from the brain via

electrodes located around the head. As the skull is between the emitted electrical activities of the brain and EEG receptors, the quality of the recordings is not high enough [18]. The solution is recording EEG signals directly from the exposed brain surface by locating receptors under the skull. This method is called Electrocorticogram (ECoG). ECoG is an invasive approach that involves surgery to implant the receptors. All the risks around body reaction to implanted receptors and brain infection because of the surgery limit the application of this method in assistive robot control. Limitations around the quality of EEG signals and the expensive equipment for EEG signal processing and enhancement have restricted the EEG-based approaches for robot control [18]. In comparison, EMG signals are easier to measure. Moreover, collecting EMG signals is non-invasive, and there is no health-related risk of inconvenience associated with EMG sensors.

The literature defines two main approaches to translating EMG signals to the joint moments. The first is based on the mechanical characteristics of the muscle-tendon system of the body, which can be called a "Neuromusculoskeletal" (NMS) model [19]. The second one is considered a "data-driven approach", which instead of the actual physiological correlation between EMG signals and muscle forces, uses a statistical model to correlate EMG signals to joint angle or joint moments as indicators of motion intention. Further explanation about these methods is provided in Chapter 2.

Even though NMS models have been shown to provide an accurate prediction of joint moments from EMG signals, the complexity of the model and the number of parameters in the model make the process of calibration complicated and time-consuming [20]. On the other hand, data-driven models have a more straightforward training process; however, these models are subject and task-specific [16, 21]. That is, data-driven models typically only predict joint moments or angles for the individual and particular task they were trained to. Herein lies the challenge that this thesis seeks to address. Can a data-driven approach be developed to predict human motion intention (i.e. forward in time) based on EMG signals with the capability to adapt to new individuals and new tasks (outside of the training dataset). Such a model requires "on-the-fly" updating with regard to every unique data point that is being collected, with an objective function based on the difference between the predicted and actual value of the output at each time step.

One of the movement-related disorders caused by stroke is foot-drop. Patients with this problem cannot lift their feet during walking and have to drag their feet on the ground [22]. Foot drop is the result of muscle deficit around the ankle joint. The four main movements of the ankle joints are plantar flexion, dorsiflexion in the sagittal plane, and eversion, inversion in the frontal plane [23] (Figure 1.3 and Figure 1.3). The dorsiflexion and plantar flexion are the main generated moments in the ankle joint during walking. In this thesis, the words torque and moment are used

interchangeably. Despite the subtle difference in the definition of these two words in mechanical engineering and material science, for the purpose of this thesis, I considered the same meaning for these two words, which is the source of ankle rotation in the sagittal plane.

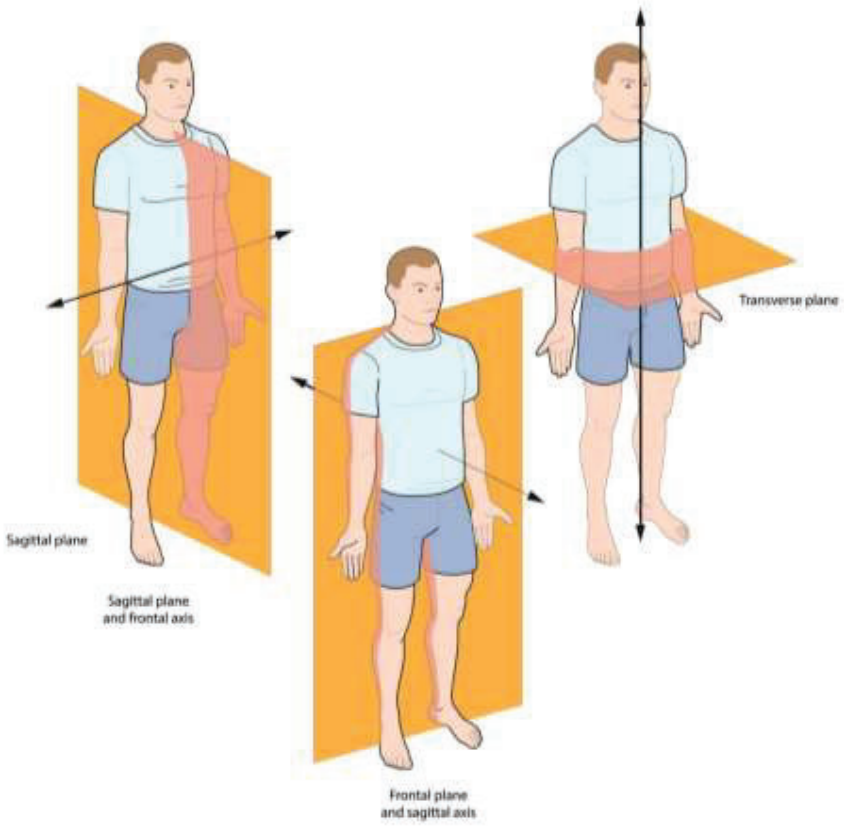


Figure 1.2 *The anatomical planes to define movements in 3D space [Photo ID: 2078502457, from www.shutterstock.com].*

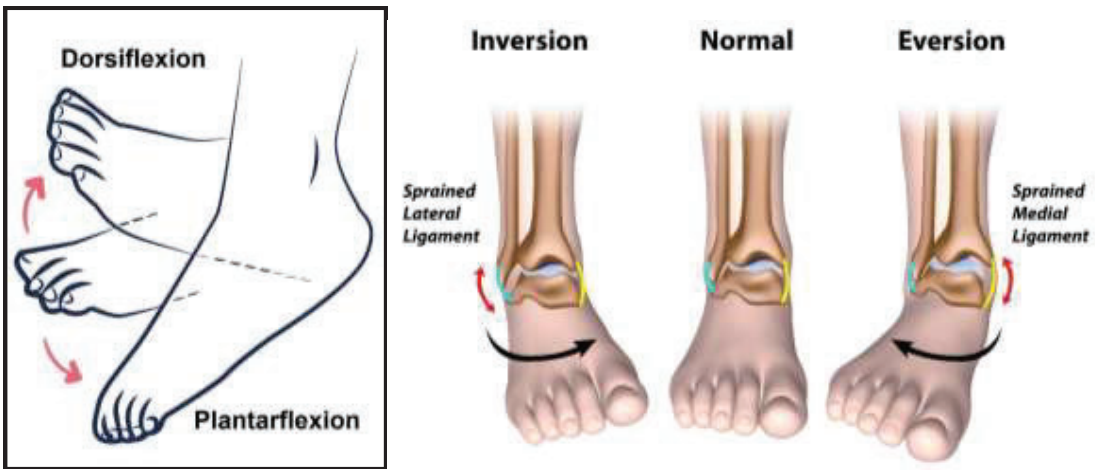


Figure 1.3 *Movements of the ankle joint [Photo ID: 228843235 and 1814883851, from www.shutterstock.com].*

Lack of plantarflexion moment in foot-drop patients during push-off limits their walking speed and stability. Because food drop can increase the risk of tripping and falling, using a lightweight brace or shoe inserts (orthotics) will be helpful in keeping the patients' foot in a fixed position. However, orthotics will not help with increasing muscle strength and ankle range of motion. In this case, rehabilitation is the solution to the patient's recovery.

The developed model in this thesis was considered to be implemented as part of the control method for an ankle assistive wearable robot to assist foot-drop patients and provide ankle plantar flexion moment using a spring with an active clutch mechanism. The active clutch is in charge of the engagement and disengagement of the spring to regulate the amount of the restored potential energy in the spring and release that energy based on the phase of gait. In this case, the robot will be able to provide enough force to help the patient lift their food with maximum patient involvement in the movement.

This thesis aimed to develop a data-driven model based on EMG signals to predict ankle motion intention. The main characteristic of the desired data-driven model in this thesis is its compatibility with different tasks and individuals beyond what was used for model training. Several research questions needed to be addressed to accomplish this aim:

1. What is the optimum length of input data for the model to predict motion intention in the ankle joint?
2. What are the most important features that can be extracted from the EMG time series to predict ankle moment?
3. Can a data-driven model predict a wide range of activities for a single individual?
4. Can a generic data-driven model be adapted "on-the-fly" to predict motion intention across a range of individuals and movement tasks?

These research questions constitute the chapters of the thesis, as outlined below.

1.2 Thesis outline

This thesis is constructed of 8 chapters, which are briefly explained in this section. Chapter 4 is an IEEE conference paper, and chapters 5, 6, and 7 are submitted for journal publication and appeared in this thesis in their as-submitted form.

1.2.1 Chapter 2 – assistive robots and interaction with humans – Background

A cursory summary of the background of stroke rehabilitation is provided, focusing on rehabilitation robots, investigating different control methods and how assistive robots can help make rehabilitation exercises more effective. Models to predict the motion intention of humans

are the central part of the assistive robot controllers, discussed in further detail, to critically evaluate the pros and cons of various approaches. Lastly, the main challenge of motion intention predictive models is described, justifying the knowledge gap that I have addressed in this thesis.

1.2.2 Chapter 3 – data collection and preparation.

Since the main focus of this thesis was to develop a data-driven model, the first step that was needed was to conduct motion capture experiments. This chapter provides the details regarding volunteer recruitment, experimental protocol, data collection processes, data preparation and filtering. These details are often overlooked in journal publications for the sake of brevity. However, since these data were critical to the subsequent modelling work, this chapter provides the necessary details to ensure others can reproduce the protocols and processes used to derive the input and output signals for model development.

1.2.3 Chapter 4 – Ankle moment forecasting using time-delayed neural networks.

On the way towards developing a data-driven model to predict ankle movements, there were some critical questions regarding the input data, including the optimum epoch of the input time series required to predict ankle moment forward in time. I developed a time-delayed artificial neural network (TDANN) and a feature extraction approach. To lead to accurate joint moment predictions, it was also necessary to determine the most informative EMG channels and kinematics input data. An additional question that was explored here was "how far into the future could this model accurately predict ankle moment, given a window of prior data points of EMGs and kinematics".

1.2.4 Chapter 5 – Forward prediction of ankle joint moment using a generic feature set.

Data-driven models are subject-specific. It means it is required to train the model for each individual separately. The model training process for each individual can be prolonged and computationally expensive. Extracting valuable features out of input time series for each individual is a time-consuming part of the training process. Chapter 5 tried to introduce a generic feature set applicable to different individuals. Using this generic feature set will reduce the required time for model training while the model accuracy remains high.

1.2.5 Chapter 6 – Ankle moment prediction at varying walking speeds using EMG and machine learning.

One of the limitations of the current data-driven models is that they are task-specific. It means it is not possible to use a model that is trained to predict joint moment based on EMG signals on a specific task for another task. Chapter 6 addressed this limitation and found a training dataset that makes the model compatible with a broad range of walking speeds. In this case, it doesn't need to train the model for each speed individually. This characteristic makes the training process faster and computationally less expensive.

1.2.6 Chapter 7 – An adaptive regression model to forecast ankle motion based on EMG signals during walking at different speeds and inclines.

Different muscle activation patterns around the ankle joint during various activities and diversity of muscle activation among populations degrade the accuracy of motion intention prediction based on EMGs during the test for unseen tasks and individuals. Having a model adaptive to different working situations and individuals makes the training process short and fast. The central part of the model adjustment happens during the testing process (on the fly). In chapter 7, the structure of such a model is explained, and its performance is evaluated to predict ankle angle during different moving tasks for ten other individuals based on EMGs and kinematics as input. The model initially trained on data from walking at a constant speed of one of the individuals, and its performance was evaluated on data from a broad range of activities of nine unseen individuals.

1.2.7 Chapter 8 – Conclusion

A summary of the novel contributions of this thesis is presented in Chapter 8. Potential improvements and future work is discussed to make the proposed model efficient for assistive robot control applications.

2 Assistive robots and interaction with human

Background and Literature Review

In this chapter, the importance of rehabilitation for post-stroke recovery is explained. Different physical rehabilitation methods are described and categorized as conventional and robotic assistive methods. The advantages of robotic rehabilitation over traditional methods are explained. Next, the control methods for assistive robots are described, divided into non-EMG-based and EMG-based methods. The EMG-based control methods construct a model to translate EMGs to muscle forces and moment and joint angles. Two main approaches for developing such a model are NMS and data-driven models and are explained in this chapter.

Moreover, EMG-based controllers are discussed in detail based on their applications and categorized as classifiers and regression models. Finally, the structure of data-driven regression models for estimating motion intention based on EMG signals is explained. The purpose of this study is justified based on the limitations of the state-of-the-art data-driven regression models.

2.1 Importance of rehabilitation for patients with movement disorders

A stroke is a sudden interruption of blood flow to a part of the brain, causing it to stop working and eventually damaging brain cells. During a stroke, the affected brain's cells start to die, and that part of the brain cannot function. This can affect a person's ability to walk, talk, eat, see, read, socialize or do things they were able to do before the stroke. The effects of a stroke depend on which part of the brain is damaged and how severe the damage is. Many people with stroke may also have fatigue or problems remembering, understanding, or thinking correctly [5].

The most effective rehabilitation is specific to the skills the patient needs with sufficient intensity and duration. Innervating the muscles and part of the nervous system affected by stroke can increase the chance of recovery from stroke [8, 24]. However, it is unclear what volume of training is optimal for recovery. Limitations related to the number of available trained therapists and the capacity of rehabilitation centres make it difficult to provide adequate rehabilitation for all patients, while the demand is growing year by year [9, 25].

The sooner the patient begins stroke rehabilitation, the more likely they will regain lost abilities and skills [24]. It's common for stroke rehabilitation to start as quickly as 24 to 48 hours after stroke [26]. In the initial phase, the main focus of rehabilitation is stabilizing the medical condition, controlling the life-threatening condition, and preventing another stroke. After getting over the initial stage, the direction of the rehabilitation program is toward regaining lost motor function [26].

2.1.1 What is post-stroke rehabilitation?

Rehabilitation can help stroke patients regain independence and improve their quality of life. The area of skills that rehabilitation can help patients depends on the part of the brain affected by the stroke. So the neurorehabilitation program must be customized to practice those skills impaired due to the stroke, such as weakness, lack of coordination, gait-related problems, or problems with hand grasp [27]. The severity of stroke complications and each person's recovery ability vary widely. Because of that, personalized and focused rehabilitation is needed for each individual after stroke [28].

The duration of stroke rehabilitation varies from person to person. Some stroke survivors recover their mobility over a few weeks. However, most need some form of long-term stroke rehabilitation, lasting possibly months or years following a stroke. Moreover, the rehabilitation plan will change during the recovery process as the patient re-learns skills [29]. Updating the rehabilitation program helps patients gain the most out of the rehabilitation program [29].

The success of rehabilitation program depends on multiple factors such as:

- Physical characteristics, including the severity of your stroke in terms of both cognitive and physical effects
- Emotional factors, such as the patient's motivation and mood, and ability to stick with rehabilitation activities outside of therapy sessions
- Social factors, such as the support of friends and family
- Therapeutic factors, including an early start to your rehabilitation and the skill of your stroke rehabilitation team.

As mentioned earlier, the sooner the patient starts the rehabilitation program, the higher the chance of recovery [24]. However, there is evidence that performance can improve even 12 to 18 months after a stroke [30].

2.2 Rehabilitation methods

We can divide the rehabilitation methods focusing on motor recovery into three main areas; physical therapy, Functional electrical stimulation (FES), and robotic therapy.

2.2.1 Physical therapy

Physical therapy is about the exercises that help stroke patients relearn movement skills and physical strength they need to get back to everyday life and do their daily tasks independently. Physical exercises simulate day-to-day activities such as reaching and grasping objects for the upper limb and walking and step climbing for the lower limb. Therapists mainly support patients during these exercises. Based on the level of disability, patients will be asked to participate in "group session" or "individual session." In group sessions, the patients can move their body limbs and go through physical activities such as walking, ascending/descending steps, and playing with a ball under the supervision of a therapist. As a more personalized solution, patients who have more severe problems with moving their body limbs are prescribed to participate in individual sessions. During individual session therapist help patients to move their body by providing the required force and helping them do some exercises that would be helpful in their daily life activities (Figure 2.1) [31]. In other words, in individual sessions, the therapists are more involved during the exercises and help patients complete tasks; however, in group sessions, patients go through the exercises mainly without physical help from clinicians, and clinicians only guide them through the exercises.

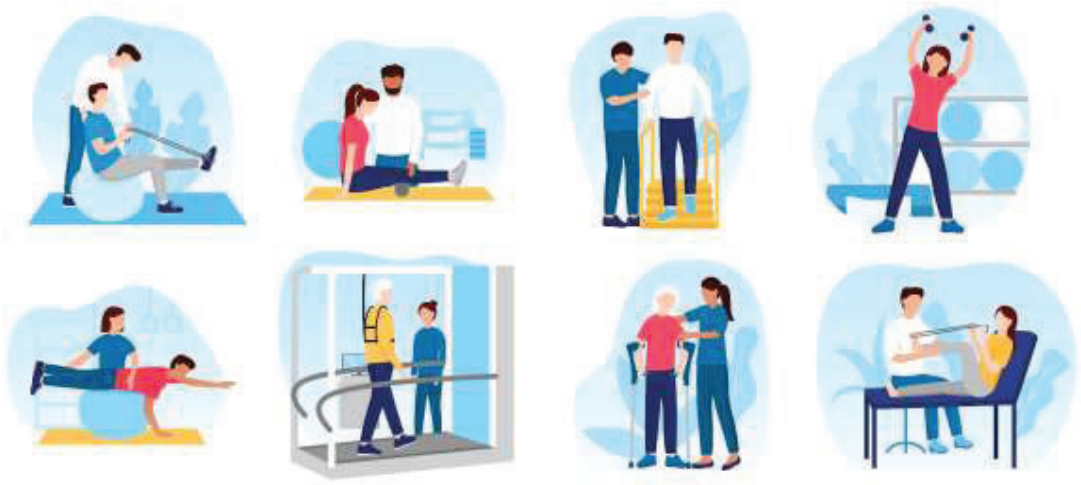


Figure 2.1 Sample of physical therapy for gait rehabilitation [Photo ID: 1909064194, from www.shutterstock.com].

Each exercise in physical therapy is designed for a particular purpose. These exercises are explained below.

2.2.1.1 Muscle strengthening and early mobilization

Strength training is defined as an intervention where a participant exercises a muscle or group of muscles against an external resistance. Applying external resistance helps muscles regain strength faster and helps the nervous system control the body during the movement. There are active and passive interventions during mobilizing exercises. Both are safe to perform for stroke patients and can induce functional improvement only if an expert therapist selects them based on the patient's physical status [32].

Passive exercises are defined so that the therapist moves the patient's limb without any effort from the patient. In progressive resistive exercises, the resistance against the movement or the weight the patient needs to carry increases during the rehabilitation process to improve muscle strength. The patients do most of the job during the active exercises, and the therapist guides the person to complete the task correctly. These are the three main types of muscle-strengthening exercises which therapists commonly use to improve motor power and functions, including walking, gait, and balance [33]. The passive exercises over limbs and early mobilization even during the acute stage of stroke leads to better recovery based on observation in studies [34-36]. Early mobilization also enhances neuroplasticity, apart from psychological benefits.

2.2.1.2 *Repetitive task training (RTT)*

These exercises are designed to improve brain control over multiple joints' movement simultaneously. During the rehabilitation session, the participants are asked to do task-specific motor activities repeatedly (e.g., lifting a cup or walking). This type of exercise strengthens the neural connectivity between the brain and muscles, which recovers neuroplasticity in damaged brain areas because of stroke. Some evidence proves the effectiveness of the RTT exercises in improving lower limb and upper limb functionality for patients who were doing these exercises up to 6 months after stroke [37, 38]. In a specific type of RTT designed for upper limb rehabilitation, the non-paretic side of the body is constrained by a strap, and the focus of the exercise is on the paretic side. During each training session, longing between 0.5 to 2 hours, the patients are asked to do task-specific motor activities with their affected limb while the non-affected one is constrained. This particular type of exercise has demonstrated benefits in motor function arm-hand activities in the long-term based on observations in studies [39, 40]. Despite the reports that show the impact of RTT on patients' recovery, the reported improvement is not significant. It is required to prescribe other types of exercises to prepare a stroke patient for daily activities. So other rehabilitation methods explained in this chapter will be necessary and RTT can play the role of a complementary exercise to accelerate the rehabilitation process.

2.2.1.3 *Gait rehabilitation*

Because of the importance of walking in terms of quality of life and independence of stroke patients, a particular class of physical exercises is focused on gait rehabilitation or gait retraining [41]. Standard devices to assist gait retraining include walking aids [42], stationary cycle [43], stepping machine [43, 44], and treadmill training [45] with or without support [46]. The training schedule includes specific gait symmetry, coordination, and walking speed enhancement exercises. A daily session of 30-min duration for four weeks has shown a significant improvement in gait speed and balance [47].

Physical therapy is the conventional approach for rehabilitation for post-stroke patients. It has been proved that physical therapy helps patients to recover their movements. The main advantages of physical therapy are included in voluntary activation in key group muscles, increasing muscle strength and coordination, maximizing moving skills, and increasing flexibility [47]. However, there are some drawbacks associated with this rehabilitation method, such as short session time, the nervousness of patients because of the clinic environment, schedule constraints of clinics, and lack of compatibility of exercises with daily basis tasks. These limitations harm the effectiveness of therapeutic exercises for post-stroke patients. Moreover, most physical therapy exercises are

labour-intensive for clinicians as they have to carry the body weight of patients or control the balance of patients during exercise by constantly applying force to them [48].

2.2.2 Functional electrical stimulation (FES)

In functional electrical stimulation (FES), low-energy electrical pulses induce movements in a muscle or group of weak muscles in the paralytic limb [49]. Electrical pulses are applied directly to the muscles via electrodes attached to the skin over the muscle belly. FES is one of the most widely used therapies with benefits including; spasticity reduction [50], improvements in range of motion [50] and proprioception [51], and pain reduction [51] (Figure 2.2).

(a) FES for upper limb *(b) FES for lower limb*

Figure 2.2 Functional electrical stimulation for a) upper limb and b) lower limb [Photo ID: 1320908207 and 236091475, from www.shutterstock.com].

FES aims to reconnect the afferent sensory information between muscles and the motor cortex to help the central nervous system relearn movement execution [52]. The FES treatments are delivered in 30- to 45-minute sessions. It is expected that electrical stimulations applied to muscle for short periods assist the user in performing motor functions such as walking, standing, reaching, and grasping [52]. FES is used in combination with physical rehabilitation to make therapy exercises more effective and efficient. By applying FES during the exercise, muscle engagement increases, and learning can be accelerated [53].

2.2.3 Robotic therapy

Despite the efficacy of physical therapy for post-stroke patients, it is challenging to provide a sufficient volume of exercise, particularly considering the increasing rate of stroke worldwide [54]. Moreover, most physical therapy exercises are labour-intensive, and it is difficult for clinicians to execute those exercises on adults, particularly overweight patients.

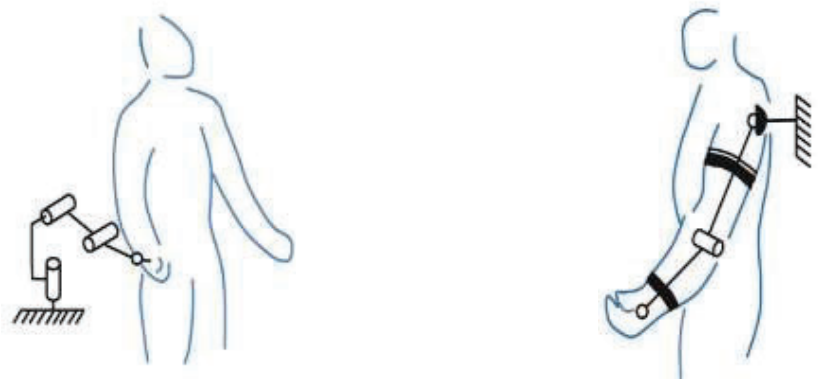
To overcome the drawbacks of physical therapy, robotic therapy is suggested as an alternative, which is expected to be an efficient and accessible solution for stroke patients who need therapy and assist clinicians during therapeutic exercises. In robotic therapy, mechanical devices are used to apply forces or torques to assist or resist the movement during the exercises and support the participants' weight in the rehabilitation exercises. Another advantage of robotic rehabilitation is providing opportunities for remote assessment and treatment. Portable solutions would enable stroke patients to perform their exercises at home instead of clinics [55]. In addition, robotic therapy makes it possible to measure and monitor the improvement in patient movement based on collected data via different types of sensors.

The robots commonly used for stroke rehabilitation are a subsection of robots called exoskeletons. Exoskeletons can be categorized into three subsections. The first category is "empowering exoskeletons," which are portable and wearable motorized robots that are mainly designed to make physical activities such as lifting and carrying heavy loads more manageable for individuals. These robots empower the wearer to carry loads beyond their physical capacity and reduce energy consumption during daily tasks. The second broad category is known as "assistive robots," designed for people with some level of disability, typically caused by neuromuscular problems such as stroke, spinal cord injury, or other neurological disorders that can cause difficulties in moving the upper or lower limbs. People with these problems need help to complete physical activities. In these cases, assistive robots will help to bring back physical performance to the patients. The main difference between empowering exoskeletons and assistive robots is the user's involvement. In empowering exoskeletons, the robot carries the load, and the person only controls the direction of movement. The robot provides assistive force to help the person complete a task but does not bear additional loads in assistive robots.

The third category is "therapeutic exoskeletons," used in rehabilitation to train the patient's muscles and/or nervous system. According to the desired rehabilitation exercise, these robots are designed to assist, perturb, or resist the user's movement. Therapeutic robots are designed to do repetitive rehabilitation exercises and are not necessarily portable. Based on the desired application, they are focused on end-effector movement or motion-control of various joints (Figure 2.3).

There are different ways to classify exoskeletons. The abovementioned categories are based on the application of exoskeletons regardless of their mechanical design or control methods. Robotic exoskeletons have attracted significant attention from the scientific and industrial community in recent years, and the number of publications in this field is increasing rapidly [56]. Over 30 years, from 1990 to 2019, ~4000 papers were published on robotic exoskeletons and rehabilitation

robots. From 2015 to 2019, the number of publications doubled, showing the increasing attention from scientists and engineers to this field, driven by increasing clinical demand [56].



(a) End-effector robot *(b) Exoskeleton*

Figure 2.3 *Schematic examples of rehabilitation robots. a) end-effector robot in which the person freely choose the movement in each joint to fulfil the required movement in hand, b) exoskeleton which controls the movement of all joints during the movement [57]. © 2013 IEEE.*

2.2.3.1 Promises and challenges of robotic rehabilitation

Robotic therapy promises to develop and expand home-based rehabilitation systems for stroke patients. The ability to provide home-based therapy can potentially improve the rate of recovery by enhancing the quality and quantity of movement therapy [58]. In addition, robotic therapy allows patients to start treatment in the early days following stroke, without frequent and prolonged visits to the clinic. In the comfort of their own homes, people could get specific training at the appropriate level of intensity while their progress is remotely monitored by a therapist [59]. In terms of efficacy, different studies showed that robotic rehabilitation provides effectiveness in the recovery of patients with spinal cord injuries and stroke [57][58]. Different studies compared the effectiveness of robotic rehabilitation with conventional therapy methods in upper limb and lower limb rehabilitation. In all cases, the reported improvement in range of motion and muscle strength for patients who underwent the robotic therapy was at the same level or better than those who only experienced the conventional therapy [58-60].

Despite the advantages that robotic rehabilitation can offer patients, high costs and limitations on the performance of the available systems are the main obstacles preventing the use of these systems in patients' homes [60]. Moreover, there are further limitations regarding the range of activities that robots can be used for long-term use due to power demands. However, emerging technologies in

mechanical and electrical systems reduce the cost and improve the functionality of assistive robots by reducing their size, weight and cost.

Despite the improvements made to the physical system, the greatest challenge in developing assistive robots is providing an intuitive interaction between man and machine. Man-machine interfaces have also seen an exponential rise in research interest over the last decade in the field of control and robotics. There is an infinite number of possible applications for which an intuitive controller is warranted. However, let us consider two fundamental approaches. The first approach is where a human directly provides the control scheme, and the robot follows the commands. In this approach, the robot does not understand the user's movement. Still, it delivers a pre-programmed response (e.g., a pre-defined kinematic trajectory or a pre-defined resistance through a range of motion). Ideally, for therapy applications, the robot would understand the user's motion intention to provide assistance or resist forces precisely based on the movement that the user wants to make. This describes the second approach, and this type of interaction between humans and robots is known as assist-as-needed control [61].

The mechanical design of assistive robots has evolved and improved in the past decade. The developments in the design of these robots improved their performance and versatility over the past years [62]. However, many studies focused on control strategies to increase assistive robots' precision, efficiency, and comfort [63]. Using mechanical and electrical sensors to measure the robot's kinematics and kinetics and estimate the muscle activity from electromyography (EMG) are common approaches to achieve assist-as-needed control. The following section explains the main approaches used to implement control methods for assistive robots.

2.3 Assistive robot control methods

The ideal control strategy for assistive robots is based on estimating the motion intention of the wearer to achieve a safe and smooth interaction between humans and assistive robots. Moreover, the reliability and effectiveness of therapy exercises provided by assistive robots depend on having a clear understanding of the motion intention of the wearer. Hence, no matter if the assistive robot is used for therapy purposes for post-stroke patients or for empowering healthy individuals, accurate prediction of the human motion intention is a critical factor to have an ideal human-robot interaction [64].

The control strategies implemented in prior work can be categorized into two groups, based on whether they use muscle activity estimated from electromyography; 1) non-EMG-based control methods, and 2) EMG-based control methods. Comparing the outcomes from existing methods

illustrates each approach's advantages and disadvantages and provides insight into which method is more reliable to use as an assistive robot control system.

2.3.1 Non-EMG-based controllers

In this case, the safety and reliability of robot movement are controlled based on the feedback of the interaction force between user and robot. In some studies, the interaction force is measured directly at the attachment points between body and robot. In some others, the interaction force is estimated based on deformation and tension in robot parts, measured by force or strain sensors [65, 66]. Flexible, distributed pressure sensors exemplify how the impact force between the user's limb and robot frame can be measured and used for control [67]. Flexible sensors, such as those used in [67], are ideal for assistive robot applications as the robot is in contact with soft tissue (human body) and can conform to different body parts and shapes.

The ground reaction force (GRF) provides valuable information for controlling gait-assisting robots and can be measured using force or pressure sensors at the soles of the foot [68]. The HAL-5 assistive robot, for example, uses the GRF vector measured from insole force sensors to calculate the position of the centre of gravity and detect three main phases of walking gait: swing phase, landing phase, and support phase [68]. This particular type of controller focused on distinguishing different movement phases is known as "state-based controller." In a state-based controller, the output from robot actuators changes with regard to the stage of the movement.

Besides the interaction force between ground and robot, a force sensor embedded in the exoskeleton structure can measure the interaction force between user and robot. Usually, a strain gauge is used in the design of the robot to measure very tiny deformations in the robot frame as a result of the interaction force with the human body. The measured force informs the robot's control system about the user's motion intention, and the controller adjusts the level of power in the robot actuator in favour of user motion intention to minimize the interaction force [69-71]. In this way, the robot helps the user complete the movement task in a hybrid mode as part of the required force to complete the task provided by the user.

Measuring interaction force is used to implement a particular branch of controllers for assistive robots and rehabilitation robots called impedance control [72]. When the users' movement deviates from the desired motion trajectory, the robot provides assistive force to return to the trajectory, termed 'assist-as-needed.' In this manner, the robot does not control the motion continuously. Still, it engages when required, and the reaction force between the body limb and the robot is critical in detecting the right time for robot activation [73].

In another approach, the dynamic stability of the exoskeleton is controlled, which is particularly relevant for assisting standing and walking robots (e.g., the BLEEX exoskeleton [74]). Instead of measuring the interaction force between the human and the robot, the applied force from robot actuators and the wearer are measured. The controller adjusts the robot actuators to maintain the wearer's stability [74].

To implement a non-EMG-based controller, it is required to have an accurate model of the exoskeleton dynamics and a clear understanding of human body dynamics. Providing such an exact model of the exoskeleton is challenging because of the redundancy and uncertainty embedded in the structure of these robots. On the other hand, the exoskeleton's dynamic model is usually too complex, causing noticeable time delay between wearer action and robot reaction. The delay is because of the long computation process, which imposes unexpected forces on the wearer and reduces the comfortability and reliability of the exoskeleton. Moreover, when the interaction force between the wearer and robot is used as input for the controller, there is an inherent delay embedded in the structure of the controller as the robot needs to first wait for the wearer to move before adjustments can be made. In other words, there is no way to predict the intention of movement using this control approach.

For this reason, researchers have investigated biological signals, such as EMG to obtain information regarding the intention of movement before the measurement of motion or forces that arise from the movement. The time delay between EMG signals and muscle movement varies based on the muscle size. It can be between 30 ms and 100 ms, providing a helpful time window for an assistive robot to provide a control signal.

2.3.2 EMG-based models

EMG-based control methods estimate the muscle activity from EMG electrodes (typically attached to the skin over a muscle belly) and perform some signal processing to provide an estimate of motion intention, which is used as input to a controller. The core of an EMG-based controller is a model that translates the EMG signals to kinematic information, including joint angle or moment, to inform the robot's controller about the movement that the wearer is about to make. EMG-based control methods can be classified as moment prediction or angle prediction models.

2.3.2.1 Moment prediction models

There are numerous ways of converting an EMG signal into a corresponding joint moment. The simplest method is to define a proportional relationship between the level of EMG signals and the joint's moment. This relation requires a calibration between EMG and the measured joint moment, which can be linear or non-linear [75]. For example, Ferris et al. used this strategy to adjust the

level of force in pneumatic actuators designed to simulate the soleus and tibialis anterior muscles in an ankle assistive orthosis. The implemented model was used in a clinical study on two participants with incomplete spinal cord injury, which improved muscle activation during walking at the self-selected speed [76]. In the proportional approach, one muscle is selected for each joint to represent the flexion moment and one muscle for the extension moment. During a calibration phase, the proportion gains that correlate the level of EMG in each of the selected muscles to the moment are defined [14, 77]. In the proportional approach, the function that correlates the level of EMG to joint torque is a simple linear regression [78].

Because of the complicated non-linear relationship between EMG and joint torque, more sophisticated models have been developed to estimate joint torque from EMG. These approaches are either based on the physiological behaviour of the muscle-tendon unit, which is often referred to as *musculoskeletal models* or relies upon large datasets of experimental data and a *data-driven model*.

The musculoskeletal model provides the most accurate model of the body in which we can correlate EMG signals to joint torque based on the muscles and tendons characteristics. The complexity associated with muscles and tendons modelling limits the functionality of this model in real-time applications such as assistive robot control. This model controls the robot moves around a single degree of freedom joints with low computational loads, such as the knee [79] and ankle [20].

The data-driven models are introduced to bypass the complexity of the musculoskeletal models while keeping the accuracy high in terms of defined correlation between EMG and joint torque. There are many types of data-driven approaches, including neural networks [80-83], fuzzy logic [84-86], regression models [75, 87, 88]. These are three main branches of data-driven models, but some other methods have been developed during the past few years based on a combination of main methods. For instance, a neuro-fuzzy model was implemented to detect the moments of the knee and hip joints [15]. A neural network-based approach was used to detect the fuzzy model parameters for joint moment prediction in this model. In another example, the neural network was used to determine the parameters of a regression model to predict hand force and wrist moments based on EMG signals [89]. A more complex version of neural network-based models known as "deep neural network" or "deep learning models" are also used for motion intention detection based on EMG signals [90]. The main difference between the deep neural network and a standard neural network is the number of hidden layers. A deep neural network can have as many as 150 hidden layers, while this number is 2 or 3 in a standard neural network.

The main papers that used explained methods in this section for joint moment prediction based on EMG signals are listed in Table 2-1.

2.3.2.2 *Angle prediction models*

It is not possible to calculate joint angle directly from EMG signal, as it is an indicator of muscle force. There are two main approaches to calculating the joint angle from EMG data. The first is to translate EMG to the joint force, calculate acceleration from the force, and then calculate the angular position of the joint by double integration from the acceleration [91]. The other approach uses data-driven models such as machine learning-based methods or other data-driven models to map the EMG signals to joint angles.

Table 2-1 List of papers in which EMG-based models were used to predict joint moments.

ref number	Authors	Joint	Number of muscles	Method	Number of participants
[76]	D. P. Ferris, et al	ankle	2	Proportional	2
[77]	H. Kawamoto, et al	Hip, knee	4	Proportional	1
[14]	P. N. Fernandes, et al	Knee	4	Proportional	2
[78]	Rong Song, et al	Ankle	2	Linear regression	8
[79]	N. Karavas, et al	Knee	6	Musculoskeletal model	1
[20]	D. Ao, et al	Ankle	2	Musculoskeletal model	8
[15]	H. He, et al	Hip, knee	8	Data driven (neuro-fuzzy)	1
[92]	K. Gui, et al	Hip, knee	2	Data driven (RBFNN)	4

Data-driven models such as third-degree polynomial [93], linear state space [94], and hierarchical projected regression method [95] were applied in the literature to predict elbow angle. Numerical models reported high accuracy in angular position prediction for single joints with one degree of freedom (DoF), such as the elbow. Neural network-based models were used for multi-joint position prediction and joints with more than one DoF. For instance, a multilayer perceptron was used for wrist angle prediction in [96]. Also, elbow and shoulder angles in 3D space are predicted through a multilayer perceptron neural network in [97, 98].

More sophisticated neural network-based models were introduced in the literature to provide higher accuracy in joint angle prediction. For instance, a time-delayed neural network (TDNN) was implemented to predict ankle angle based on EMGs from three muscles involved in this joint movement [99]. Or a long short-term memory (LSTM) network was used to predict knee angle

continuously [100]. The reported accuracy in angle prediction for both of these examples was over 90% which shows the effectiveness of sophisticated neural networks in joint angle prediction [100].

Linear and nonlinear regression models were also used for joint angle prediction based on EMG signals. An autoregressive integrated moving average with exogenous input (ARIMAX) was used as a regression model to predict ankle angle based on EMGs [101]. A nonlinear regression model based on the Kalman filter was implemented to predict knee and ankle angle [102].

More sophisticated approaches have been implemented for angle prediction in the past few years [103]. For example, a wavelet neural network was implemented to predict knee joints based on EMG [104]. In this model, the wavelet transform was used for feature extraction, and a convolution neural network was used to predict knee angle based on extracted features. Generalized regression neural network is another example in which the neural network was used to define the parameters of the regression model to predict knee angle during walking at the constant speed [105]. Deep learning-based models are another example of complex models used for accurate angle prediction based on EMG signals [106]. A list of papers that used data-driven models to predict joint angle based on EMG signals is represented in Table 2-2.

2.4 Numerical methods to predict motion intention based on EMG

Different types of numerical models are implemented in the literature to predict the motion intention of the wearer based on EMG signals to control assistive robots. Some of these models are designed to predict the joint angle, while others predict the joint moment to indicate the motion intention. In this section, the structure of these models is discussed in detail. Generally, numerical models for motion prediction based on EMGs are categorized into three main types: a) proportional methods, b) musculoskeletal models, and c) data-driven models. Each one of these groups is explained in the rest of this section.

2.4.1 Proportional methods

In the proportional approach, the main target is finding the gain that correlates the EMG signals to the joint torque or angle. To find this correlation, a calibration phase is required in which the amount of torque that each person needs to apply to maintain the joint at a fixed angle is measured. At the same time, the level of muscle activation is measured via EMG electrodes. Then the proportional gain that relates EMGs to joint torque was found. In [14, 20, 77], this approach was used to predict joint torque based on EMG signals. The results showed promising performance for proportional approaches in joint torque prediction, especially for single joint torque prediction during constant speed movement [14] or isokinetic acts [20].

Table 2-2 List of papers in which EMG-based models were used to predict joint angles.

ref number	authors	joint	Number of muscles	Method	Number of participants
[93]	H. Yu, et al	Elbow	2	third degree polynomial	1
[94]	Q. Zhang, et al	Elbow	2	linear state space	1
[95]	Y. Chen, et al	elbow	2	hierarchical projected regression	5
[96]	S. Muceli, et al	wrist	4	multilayer perceptron	1
[97]	Q. Zhang, et al	Elbow, shoulder	10	multilayer perceptron	1
[107]	Y. M. Aung, et al	Elbow, shoulder	10	back propagation neural network	4
[101]	H. Al Kouzbary, et al	Ankle	3	autoregressive integrated moving average with exogenous input	1
[102]	P. B. Jephil	Knee, ankle	9	non-linear extension of the Kalman filter	6
[104]	J. Wang	knee	5	Back propagation neural network	5
[108]	T. Anwar, et al	knee	5	generalized regression neural network	1
[21]	J. Chen, et al	Hip, knee, ankle	10	Back propagation neural network	1

2.4.2 Musculoskeletal models

To predict joint torque based on EMG signals and joint angles, a Hill-type model was proposed in [109] and developed in [110] as the first musculoskeletal model. This model was improved and evolved by others to make the calibration process faster and reduce the number of uncertainties. A list of muscle parameters is determined during the calibration process to construct the model. The greater the model fidelity, the longer the list of muscle parameters must be defined. So the calibration process is a trade-off between model accuracy/need and level of complexity [79].

The Hill-type muscle model was used for knee torque predicted in [19] based on EMG from thirteen muscles. Eighteen parameters were defined during the calibration process, which was important in muscle force calculation. These parameters were related to the muscle and tendon characteristics. The most important parameters were the pennation angle and muscle-tendon

length, which depends on the percentage of change in optimal fibre length (γ). The importance of γ in the model accuracy was investigated in this study by allocating different values to this parameter. The model's performance has been evaluated by testing the model for knee joint torque prediction for six healthy individuals during a wide range of activities. By changing the value of γ from 0 to 15%, the coefficient of determination (R^2) increased from 0.85 to 0.91. This shows how sensitive the model is to this variable. In another study, a musculoskeletal model of the lower limb, including 13 muscles developed to predict knee and ankle joint moment during walking [111]. Even though the model was accurate in moment prediction, its main drawback was its complicated and time-consuming calibration process.

Different works focused on making the process of training for musculoskeletal models faster and more efficient. In [112], knee joint torque was predicted based on EMG signals, joint angle and ground reaction forces. A complete lower limb model implemented in this study comprised of two legs with feet, shanks, thighs, and the torso. Only two parameters were trained during the calibration phase. The model performance was evaluated using data from one healthy individual during knee joint flexion/extension and stair climbing. The predicted knee joint torque was compared to the calculated values by inverse dynamics. A good correlation between predicted joint torque and calculated values was observed; however, the difference was significant in some cases.

In [113], the process of musculoskeletal model calibration shortened by neglecting the impact of tendon strain parameter in generated force in muscles. Despite ignoring this parameter, the model could predict knee torque during isometric movement with high accuracy (correlation coefficient (R) was 0.982). However, because of simplification in the model, calibration and the prediction time delay were significantly shorter.

To find a trade-off between model complexity and uncertainty, the number of parameters for each muscle during the calibration process was reduced to 8 in [114]. The reported 12.4% NRMSE for knee joint prediction during flexion and extension showed that this model was capable of accurate prediction. In another attempt to simplify the musculoskeletal model for joint torque prediction based on EMG signals, instead of taking all of the muscles involved in the motion generation around a joint into account for only the limited number of main muscles considered for model development. Such a simplified musculoskeletal model was used for knee joint moment prediction with an R^2 value of ~0.9, which proves the potential for the simplified model in moment prediction [115]. Another simplified musculoskeletal model for ankle moment prediction was implemented in literature which showed a high rate of accuracy [20]. In this model, only two muscles were used to determine the moment around the ankle. Each muscle only had three parameters, which made the process of calibration simple [20].

2.4.3 Data-driven models

As explained previously, musculoskeletal models have a computationally expensive and time-consuming calibration process. As an alternative solution to translate EMG signals to motion intention, data-driven models, such as neuro-fuzzy models [15, 16] and neural networks [92, 116, 117], have been presented in the literature. Data-driven models are used for two main purposes: classification and regression.

Data-driven models for classification use EMG signals as input to define the phase of the movement. For instance, a deep learning approach is introduced in [118], which used EMG signals as input to detect the swing and stand phases during level-ground walking. A multi-layer perceptron neural network was designed, which required EMGs from eight lower limb muscles to predict the gait phase during different walking patterns, including acceleration and deceleration. Recurrent neural networks and long short-term memory (LSTM) networks were implemented in [87, 119] to distinguish eight hand postures based on collected EMG signals from the forearm. The implemented models achieved about 95% accuracy in the evaluation process. The LSTM had a relatively shorter training process and needed a smaller training data set than these two models. A similar study compared support vector machine (SVM) and adaptive neuro-fuzzy inference system (ANFIS) methods to classify seven different hand postures based on EMG signals. In this study, the ANFIS was reported to have a more accurate performance than SVM by over 91% accuracy in classification [119]. In [120], a nonlinear regression model (random forest) was developed to distinguish among five movement types: level-ground walking, stair ascent, ramp ascent, stair descent, and ramp descent based on EMG signals recorded from gluteus medius muscle. The model had >99% accuracy in five movement modes recognition.

Besides classification, data-driven models can also be used as a regression model to continuously predict the value of joint angle or torque based on EMG signals instead of only predicting the movement phase. For instance, to predict the fingers' movement, a linear regression model was developed in [121]. In this study, a pattern recognition model was developed, which used an optimized number of EMG signals as input to predict the movement of each finger individually. Or in another example, a nonlinear regression model was implemented to predict knee joint angle continuously during walking at a constant speed based on collected EMG from ten muscles [21]. A neural network-based regression model was implemented to constantly decode the shoulder, elbow, and wrist movements during free arm reaching movements based on recorded EMG from six muscles involved in these joints' movement [122]. These are examples of data-driven regression models based on EMG signals for motion intention prediction. In the next section, the structure

of regression models is explained in detail, and novel approaches applied for regression model development are discussed based on the studies we found in the literature.

2.5 Structure of the regression model for time series prediction

A time-series prediction pipeline comprises three main blocks, feature extraction, feature selection, and regression model training as the predictive model (Figure 2.4).

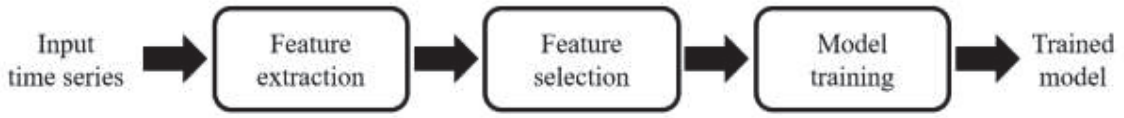


Figure 2.4 The pipeline of regression model training, representing the main steps required to develop a data-driven model.

The purpose of each block of the regression model structure and some examples of the most common methods implemented for each part are explained in this section.

2.5.1 Feature extraction

It is not practical to feed the prediction model directly by EMG signals because of the large number of inputs and the redundancy in EMG data. Therefore, it is essential to extract a set of features from input EMGs that represent the most important characteristics of them [123].

By extracting features from input time series, we emphasize the main characteristics of input data which will help to improve the classification or regression accuracy [124]. The success of any predictive model directly depends on the quality of the extracted features as input. According to the literature, extracted features from EMG signals fall into three categories: time domain, frequency domain, and time-frequency domain [125].

Table 2-3 shows the list of the most common feature extraction methods in each of the abovementioned categories.

Table 2-3 Feature extraction methods categorized in three main groups based on the type of the feature.

Time-domain	Frequency domain	Time-frequency domain
Mean absolute value		
Root mean square	Auto-regressive coefficients	Short-time Fourier transform
Variance	Frequency median	Wavelet transform
Waveform length	Frequency ratio	Wavelet packet transform
Zero crossing		

2.5.1.1 *Time-domain*

The time-domain features are commonly used in EMG-based models for motion prediction [126]. It is easy and quick to calculate these features as they don't need transformation. The time-domain features give us information about signal amplitude, the waveform, and some frequency-related characteristics of signal indirectly [127]. Some of the most important time-domain features applied in EMG signal processing are as follows:

- Mean absolute value (MAV)

Returns the mean of the absolute values of signal for each window of input data by adding the absolute value of all of the data points and dividing it by the length of the window [128]

$$MAV_k = \frac{1}{N}\sum_{i=1}^{N}|x_i| \tag{1}$$

k represents the number of the window of data, and N is the number of data points in each window (length of window).

- Root mean square (RMS)

The value of RMS for each window of the input signal is calculated based on equation 2 [129].

$$RMS_k = \sqrt{\frac{1}{N}\sum_{i=1}^{N}{x_i}^2} \tag{2}$$

In which x_i is the i^{th} data point in a window of input time series with the length of N.

- Variance (VAR)

It is given by [127]

$$VAR_k = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 \tag{3}$$

Where $\bar{x}$ is the mean value of the segment k.

- Waveform length (WL)

It is calculated based on the difference between two consecutive data points of time series in each window of data.

$$WL_k = \sum_{i=1}^{N-1}|x_{i+1} - x_i| \tag{4}$$

This reflects the amplitude, frequency, and waveform length for each time series window [128].

- Zero crossing (ZC)

Returns the number of times a time series crosses zero, equal to the times the waveform sign changes [128]. To mitigate the impact of signal noise on zero-crossing a narrow band around the zero line as a threshold for the zero-crossing event so that a ZC will happen if:

$$\{x_i > 0 \; and \; x_{i+1} < 0\} \; or \; \{x_i < 0 \; and \; x_{i+1} > 0\} \; and \; |x_i - x_{i+1}| > \epsilon \tag{5}$$

2.5.1.2 *Frequency domain*

These features are related to the signal's power spectrum density (PSD). Compared to time-domain features, the frequency-related ones are more complex and need more computational time. Here is a list of the most common frequency domain features used in EMG signal processing.

- Auto-regressive coefficients (AR)

This feature is about describing each data point in time series as a function of previous data points as [129]:

$$x_k = -\sum_{i=1}^{N} a_i x_{k-i} + e_k \tag{6}$$

The coefficients (a_i) in this equation are considered as a feature for EMG-based model development.

- Frequency median (FMD)

According to the following equation, it is calculated based on power spectral density [127].

$$F_{MD} = \frac{1}{2}\sum_{i=1}^{M} PSD_i \tag{7}$$

Where *M* is the length of the power spectrum density, PSD_i is the i^{th} line of the power spectrum density.

- Frequency ratio (FR)

It is given based on the Fourier transform of the EMG signal. The Fourier transform of low-frequency signal is divided by the Fourier transform of the high-frequency signal [130].

$$FR_j = \frac{|F(.)|_{jlowfreq}}{|F(.)|_{jhighfreq}} \tag{8}$$

The threshold for high and low frequency is defined through the experiment.

2.5.1.3 *Time-frequency domain*

These features map the energy of the signal in time and frequency dimensions. Even though these features provide valuable information about the signal, they require a computationally expensive transformation [131]. Here is a list of the most popular time-frequency-based features for EMG signal processing.

- Short-time Fourier transform (STFT)

This is an extension to the Fourier transform, in which the input signal is divided into small segments before applying the Fourier transform to it. The STFT transform is defined as:

$$STFT_x(t,\omega) = \int W * (\tau - t)x(\tau)e^{-j\omega\tau}d\tau \quad (9)$$

Where *W(t)* is the window function, * is the complex conjugate, τ represents time, and W stands for frequency [132].

- Wavelet transform (WT)

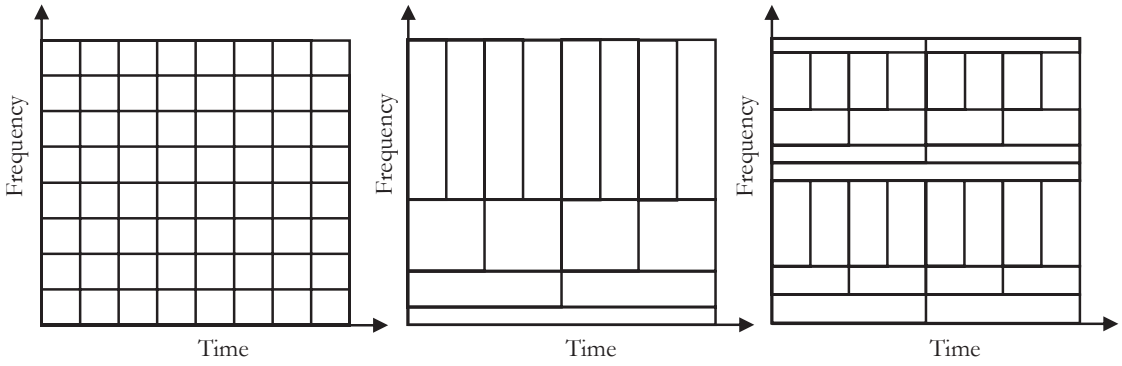
In this transform, the signal is compared with a master shape of signal (mother wavelet function), which is shifted and scaled [133].

$$W_x(a.b) = \int x(t)\left(\frac{1}{\sqrt{a}}\right)\Psi^*\left(\frac{t-b}{a}\right)dt \quad (10)$$

Where $x(t)$ is the function representing the input signal, Ψ^* is the complex conjugate of the mother wavelet function, and $\Psi((t-b)/a)$ is the shifted and scaled version of the wavelet at time b and scale a.

- Wavelet packet transform (WPT)

It is an extension of the continuous wavelet transform toward a discrete wavelet transform. Figure 2.5 compares the difference between SFTF, WT, and WPT in time and frequency domain resolution [134].



(a) STFT *(b) WT* *(c) WPT*

Figure 2.5 *The visual comparison of time-frequency tiling between a) the Short Term Fourier Transform (STFT), b) the Wavelet Transform (WT), and c) the Wavelet Packet Transform (WPT) [135].*

2.5.2 Feature selection

Feature extraction is one step toward optimizing the input information for the predictive model. However, part of the extracted features conveys redundant information, and some are related to noise. Therefore, it is required to only keep the most valuable information in the feature vector [136]. Feature selection helps to reduce the dimension of the feature vector by keeping only the most valuable information and removing redundant and noise-related information. The importance of extracted features depends on their impact on the accuracy of model prediction and the correlation between them.

In machine learning applications, feature selection is essential before training the model. Feature selection increases the model's accuracy and reduces the number of inputs, making the training process faster and less computationally expensive. When too many features are used for training a model, the model becomes too complex and tends to overfit during the training process. It is crucial to go through the feature selection phase to prevent this.

Several feature selection methods are designed for different types of data and different applications. Table 2-4 categorized these methods.

Table 2-4 *Feature selection methods are divided into two categories depending on the selection approach.*

Keep the informative features	**Find the new combination of features**	
	Linear methods	Nonlinear methods
Backward elimination Forward elimination Random forest	Principal component analysis linear discriminant analysis	Kernel PCA multidimensional scaling

As described in Table 2-4, there are two main feature selection methods. One type only selects the valuable features and removes the noise-related or redundant ones. Backward elimination, forward selection, and random forest are examples of that. In the other approach, features get transformed into new features to make them more distinguishable. The new set of features contains different values than the original feature set. Depending on the method used for transformation, these feature selection methods are divided into linear and non-linear methods. Principal Component Analysis (PCA) [137] and Linear Discriminant Analysis (LDA) [138] are examples of linear dimensionality reduction methods. Kernel PCA [139], Multidimensional Scaling (MDS) [140], and isometric mapping [141] are examples of non-linear dimensionality reduction methods. Here each of these methods is explained briefly.

2.5.2.1 Linear methods

These methods project the original data linearly to a low-dimensional space. These methods work perfectly on linear data but not very well on non-linear data. Principal component analysis (PCA) is one of the most common linear feature selection methods [142]. PCA translates the original correlated features to a set of smaller uncorrelated variables called principal components while retaining as much variation in the original features. Another famous linear method is linear discriminant analysis (LDA), similar to PCA. Instead of finding uncorrelated components, combine the input features to optimize the difference between inputs related to different classes [143].

2.5.2.2 Nonlinear methods

If we have nonlinear data as input, the nonlinear methods work better in feature selection. Kernel PCA is one of the most common nonlinear methods. This method transfers the original nonlinear features to a high dimensional space, turning them into linear data. Then, PCA is applied to them to select the uncorrelated values [144]. Another nonlinear method is multidimensional scaling (MDS), which tries to preserve the distances between instances while reducing the dimensionality of the input features [145]. Isometric mapping is another nonlinear method that is an extension to MDS [146]. In this method, the distance between features is the key factor to select the most valuable ones. A single value represents features closer to each other than a specific distance. In this way, similar features will be removed from the list of input features.

2.5.2.3 *Keep the informative features methods*

These methods only remove the redundant information from input features without changing the remaining ones. These methods improve the model accuracy. Here, three of the most popular ones are explained:

2.5.2.3.1 *Backward elimination*

This approach works based on a recursive feature elimination algorithm. In the first attempt, the algorithm trains the model based on the initial set of features. It calculates the performance of the model (usually based on the error between the model output and actual target values). Then, the algorithm drops one feature after each attempt, trains the model on the remaining features, and calculates the performance scores. The algorithm repeats eliminating features until it detects a small (or no) change in the performance score of the model and stops there. At the end of the process, features with minimum impact on model output are removed from the feature set [147, 148].

2.5.2.3.2 *Forward selection*

This method is the opposite process of backward elimination. The algorithm trains the model based on a single feature and measures the model performance. Then, new features are added to the list until no significant change in model performance happens. Similar to the backward elimination, at the end of the process, features that had the maximum impact on the model performance will make the new feature set [146, 149].

2.5.2.3.3 *Random forest*

The random forest is constructed of a set of decision trees. Each decision tree is a set of internal nodes and leaves. The internal nodes are used to decide how to divide an input feature set. At each node, the tree divides the dataset into two buckets, each hosting observations that are more similar among themselves and different from the ones in the other bucket [144]. Therefore, the importance of each feature is derived from how "pure" each of the buckets is. The more a feature decreases the impurity, the more influential the feature is. In random forests, the impurity decrease from each feature can be averaged across trees to determine the final importance of the variable [150]. In other words, features selected at the top of the trees are generally more important than those chosen at the trees' end nodes, as generally, the top splits lead to more significant information gains.

2.5.3 Feature extraction and selection in action

Starting from scratch and developing all of the abovementioned feature extraction methods and implementing the right method for feature selection is a tedious and time-consuming process. A

python-based library provides all the required functions for comprehensive feature extraction and feature selection for time series. Tsfresh (Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests) is a python package that is used to automatically calculate a wide range of time series features [151]. Further, the package contains methods to evaluate the importance of features for regression or classification tasks. The "feature_extraction" submodule covers a comprehensive range of features, including the time domain, frequency domain, and time-frequency domain. It provides 63-time series characterization methods, which compute a total of 794 time-series features. The "feature_selection" submodule offers the function "feature selector," which sorts the extracted features based on their P-values. The feature selector algorithm evaluates the influence of each of the input features on the target prediction based on the P-value. After that, the "Benjamini Hochberg" process decides which features are valuable and which ones are not [152].

2.5.4 Model training (for regression models)

The purpose of the training process is to define the coefficient of the regression model to provide the most accurate results. The training dataset includes all the input features and corresponding outputs (targets). During the training process, the model's parameters get defined so that the model output has the minimum difference with the targets. In this section, we take a brief look at two main training techniques for regression models.

2.5.4.1 Ordinary least square

This method is usually applied for linear regression model training. This method minimizes the squared error summation between model output and target for each data point in the training dataset [153]. To implement this method, it is required to have access to all training data points at once as the summation of the error is needed to calculate the cost function [154]. So we need to apply this method to an extensive training dataset. It will be computationally expensive. Because of that, this method is not practical in most real-world applications.

2.5.4.2 Gradient descent

The cost function in the gradient descent method is the same as the ordinary least square method. However, in this approach input dataset is divided into smaller sub-clusters, and the cost function is calculated for each cluster. It means in this method. It is not required to have access to all input data simultaneously. The training process starts by considering random values for regression model parameters and calculating the squared error summation between model output and target values for the first batch of inputs. Then during the training process, the model parameters get updated to minimize the cost function by feeding the model by other training batches one by one [155]. One crucial factor in this process is the learning rate that defines how fast the model parameters

adjust. A large value for learning rate may prevent the model from getting trained as in each iteration, the value of the parameters will change drastically, and the cost function would not converge to the minimum value. On the other hand, a small learning rate makes the training process very slow, as the parameters will not change appropriately after each iteration [82].

2.6 Discussion

This chapter explains the impact of stroke on human life and the importance of rehabilitation to help stroke patients get back to their normal lives. Different rehabilitation methods were discussed, including physical therapy (conventional therapy), the most common approach for the physical recovery of stroke patients. However, the limited number of clinicians and growing demand for rehabilitation (because of population ageing) make physical therapy difficult. Consequently, patients wouldn't receive the level of support they need. Robotic rehabilitation came into play to facilitate stroke patients' rehabilitation and help clinicians cope with increasing demand. Developments in mechanical design and electrical engineering improved the performance of rehabilitation robots during the past decade. Still, the field's remaining challenge is providing a reliable interaction between the user and robot, which needs an accurate understanding of user movement. This challenge is directly related to the control approach implemented for robots.

In this chapter, a review has been represented about all of the developed control methods for assistive robots, including those that used electromechanical feedback to control the robot and those that use biosignals (such as EMG) to provide more compatibility between human and robot. Three main categories for EMG-based control methods, including the proportional, musculoskeletal, and data-driven models for motion intention prediction, were discussed. The proportional approach is simple but only works for limited movements under a controlled environment. Musculoskeletal models provide an accurate understanding of motion intention based on EMG signals for a wide range of activities but have a long and computationally expensive calibration process. The data-driven methods, especially the regression models, have a more reasonable training process than the musculoskeletal models in terms of required time and computational costs. At the same time, they reportedly predict the joint angle and torque accurately.

Despite the accurate performance reported in the literature for data-driven models, there are a couple of limitations associated with these models. Data-driven models are subject-specific. Anatomical differences among the population and different patterns in muscles engagement during the same type of movement make it necessary to train the model for each new individual. Also, these models are task-specific as well. The results reported in the literature are for the situation in which the models were trained and tested on collected data during the same type of activity.

Different tasks cause different patterns in muscle activation. As long as those patterns were not covered in the training dataset, a data-driven model would not predict their corresponding motions accurately. In this situation, making a generic model works accurately for a wide range of activities requires a large training dataset, making the training process long. Therefore, there is still a need for a robust model compatible with different activities and subjects to understand the user's motion intention. This model should have an efficient training process and be computationally optimized for real-time applications such as assistive robot control.

3 Data collection

The first step toward model development involved data collection for model training. An experiment was designed to calculate lower limb kinematics and kinetics during a range of gait-related tasks. Experimental data, including EMG signals and motion capture data, went through a signal processing process to increase the signal-to-noise ratio (i.e. filter out the noise) and calculate ankle angle and ankle joint moment. The final step was to export the required inputs and outputs for model development. This chapter explains the process of participant recruitment, data collection, and data preparation for model development.

3.1 Introduction

The main purpose of this thesis was to develop an EMG-based model to predict the motion intention of the ankle. Ankle motion prediction is possible by predicting the ankle angle or ankle moment during movement. For this thesis, two scenarios were considered to develop the EMG-based model:

1. **INPUT:** EMG signals + ankle kinematics; **OUTPUT:** ankle moment
2. **INPUT:** EMG signals; **OUTPUT:** ankle angle.

These scenarios required the following data to be collected or calculated:

- EMG data (collected from quadriceps, hamstring, and ankle dorsi-plantar flexor muscles)
- Ground reaction forces (required for ankle moment calculation)
- Ankle flexion-extension angle (calculated by solving inverse kinematics)
- Ankle moments (calculated by solving inverse dynamics)

The abovementioned data were collected from ten volunteers who participated in the experiment. Each participant was asked to perform different tasks during the data collection process to provide a dataset for model development. The list of tasks is explained in section 3.3 of this chapter (experimental protocol). Ethics approval was obtained from the University of Auckland ethics committee before participant recruitment (ethics number: 022112). For the ethics approval, see Appendix A.1.

3.2 Lab equipment

Data collection was performed at the Sports Kinesiology laboratory located at AUT millennium institution of sport and health, Auckland, New Zealand. The lab was equipped with an 11-camera optical motion capture system (VICON, Oxford Metrics) to track retroreflective markers attached to the body landmarks of the participant. The markers were placed on the landmarks of the participant's body according to the UWA marker set [156] (Figure 3.1). The attached markers for tracking the motion include the static and tracking markers. The static markers were attached to the body landmarks (anatomical markers), and the tracking markers were attached to limbs (body segments). In our experiment, the tracking markers were attached to the thigh and shank (Figure 3.2). Each set of markers has different applications. The static markers were used to scale the body's anatomical model, and the tracking markers were used to translate the body movement to the anatomical model.

Figure 3.1 *The attached reflective static markers to the body landmarks of one of the participants.*

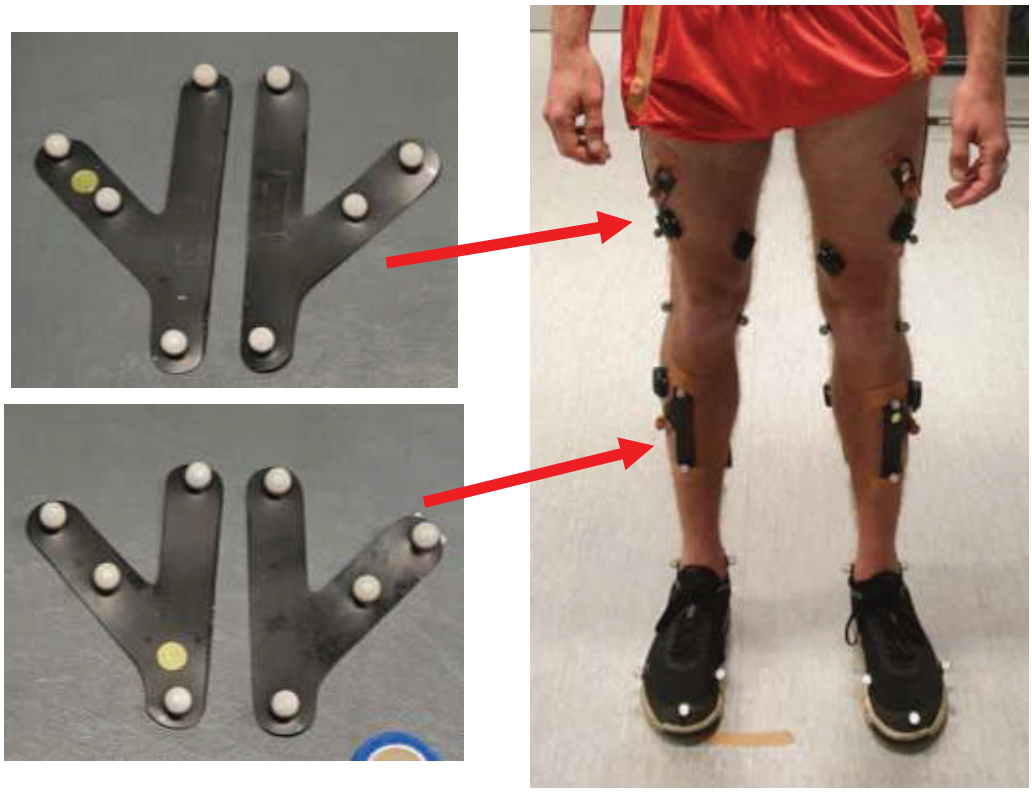


Figure 3.2 *The tracking markers attached to the lower limb of a participant.*

To collect the muscle activities during the experiment, a set of wireless EMG electrodes (Delsys Trigno, Natick USA) was attached to the muscles of interest. In this experiment, EMG was recorded from 14 muscles of the lower limb (7 muscles of each leg), including:

- Soleus
- Tibialis anterior
- Gastrocnemius lateral
- Gastrocnemius medial
- Vastus lateralis
- Vastus medialis
- Rectus femoris

EMG electrodes were attached to the belly of the muscles of interest in accordance with the SENIAM standard [157] (Figure 3.3).

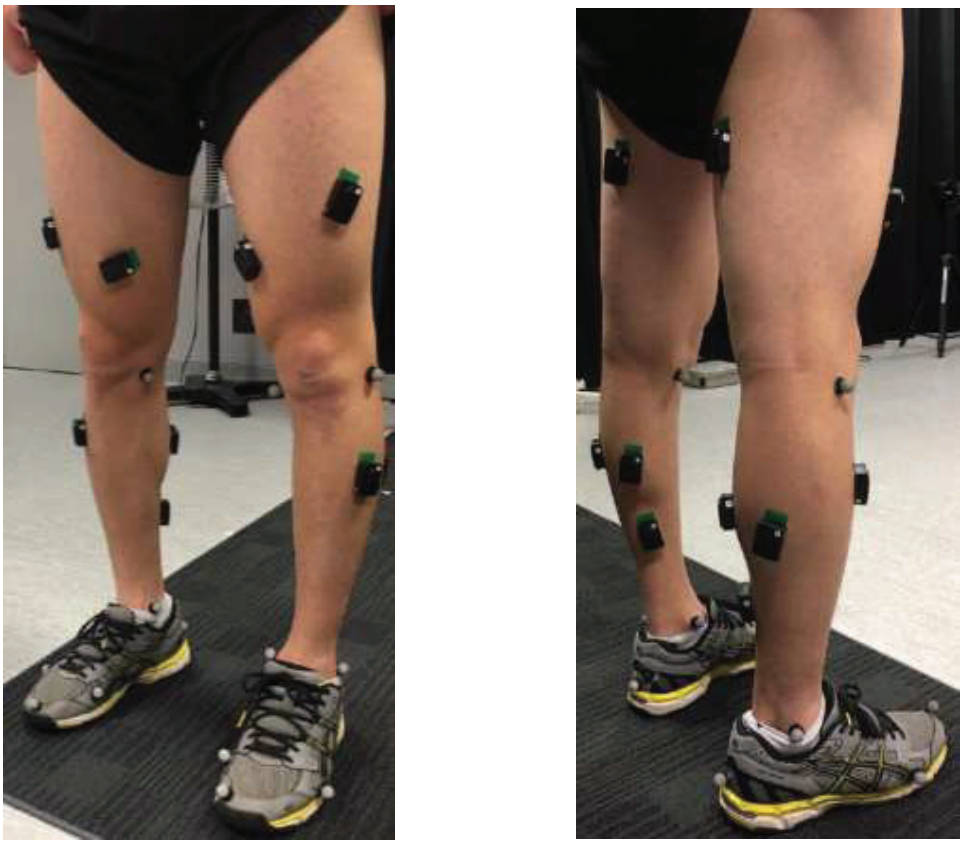


Figure 3.3 The attached EMG electrodes to the muscles for a participant in the experiment.

Ground reaction forces are required input for inverse dynamics calculations, and these were obtained from an instrumented split-belt treadmill (Bertec, OH). The treadmill was equipped with two force plates to measure left and right feet reaction forces separately during the experiment.

Moreover, the treadmill provided an incline control, enabling participants to walk up and down an incline (Figure 3.4).

(a) The back viewpoint

(b) The front viewpoint

Figure 3.4 *The motion capture set up and instrumented treadmill. The lab layout form (a) the back viewpoint and (b) the front viewpoint.*

3.3 Experiment protocol

The experiment protocol was designed to cover a range of tasks placing various demands on the ankle joint, including walking at different speeds, walking at different inclines, and ascending/descending steps. In each data collection session, after attaching all the markers and EMG electrodes, the participant was asked to perform the following tasks in order:

- Maximum Voluntary Contractions (MVCs) of lower limb muscles.
    - The MVC data was used after data collection during the signal processing process to normalize the recorded EMG signals of each muscle to their maximum value recorded during MVC. The MVC was recorded for muscles of interest (listed in section 3.2). During the MVC collections, the lab operator held the participant's leg in a fixed position and asked them to move their leg to activate each of the muscles of interest.
- Static pose.
    - The participant stood still at the centre of the motion capture lab while their arms were hanging on sides in a normal position. Motion capture data were collected in this position for one second. The recorded trace of anatomical markers was used to calibrate the anatomical model used for inverse kinematics and dynamics calculation at the next stage.
- Level ground walking at a constant speed.
    - The level ground walking on the instrumented treadmill was repeated at 1 m/s, 1.5 m/s, 2 m/s, and 2.5 m/s speeds (Figure 3.5). These speeds were selected to cover a range from slow walking at 1 m/s (Figure 3.5) to jogging/running at 2.5 m/s (Figure 3.6). The participants were asked to walk at each of these speeds for 30 seconds.

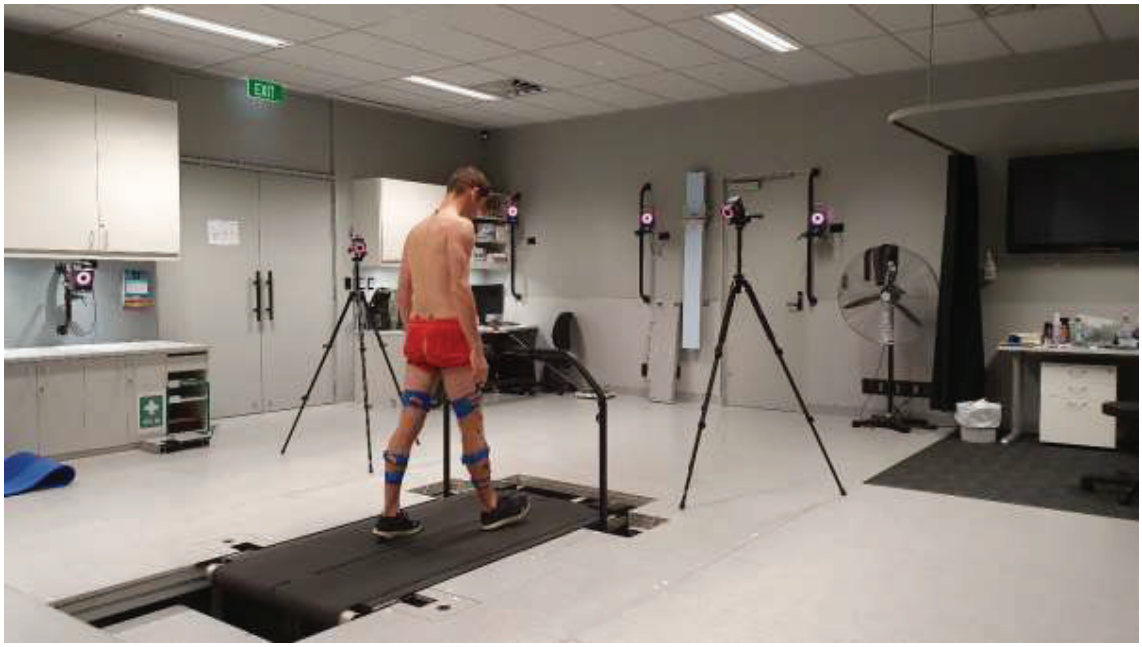


Figure 3.5 *Participant during walking on the instrumented treadmill. The EMG electrodes and reflected markers were attached to the body.*

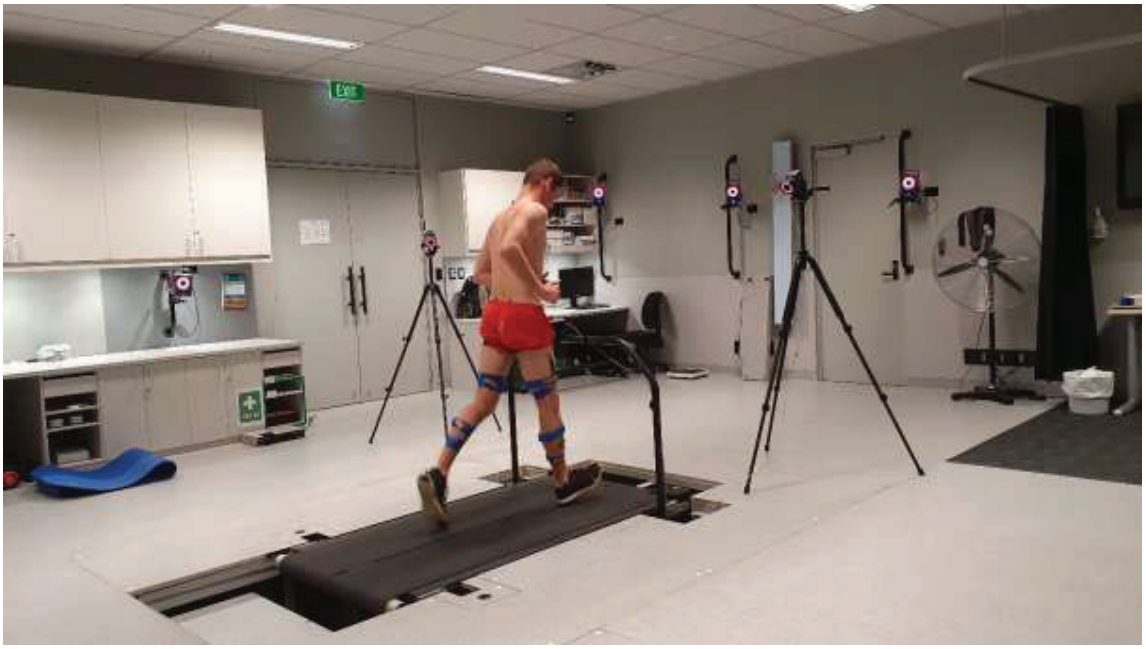


Figure 3.6 *Participant during jogging on the instrumented treadmill. The EMG electrodes and reflected markers were attached to the body.*

- 10% incline (5.71 deg) walking at a constant speed.
    - Repeated for walking at 1 m/s, 1.5 m/s, 2 m/s speeds (Figure 3.7). The participants walked for 30 seconds at each of these speeds. The incline angle was selected

considering the safety of the participants, as steeper slopes were not safe for fast walking.

Figure 3.7 Participant during walking on the treadmill at +10% slope (inclined walking).

- 10% (5.71 deg) decline walking at a constant speed.
    - Repeated for walking at 1 m/s, 1.5 m/s, 2 m/s speeds (Figure 3.8). The participants walked for 30 seconds at each of these speeds. The same slope as inclined ground walking was considered for declined walking. Maximum speed was determined based on the safety of the participants. During the inclined and declined walking, the 2.5 m/s was over the comfort zone of some of the participants, thus to have a consistent situation for everybody, the maximum speed was limited to 2 m/s.

Figure 3.8 Participant during walking on the treadmill at -10% slope (declined walking).

- Walking at self-selected speed for 60 seconds repeated on level ground, 10% incline, and -10% decline ground.
    - Repeating the experiment for self-selected speed gave a unique dataset for each individual as the selected speed was different across participants. The operator gradually increased the treadmill's speed, and the participants were asked to select their comfortable walking pace for level ground, inclined, and declined walking.
- Stationary squat (5 repeats)
    - The movement started from the complete standing position to the sitting position. The participant paused for 1 second at the bottom before returning to the starting position. A two-second pause was considered in between the iterations (Figure 3.9).

Figure 3.9 Example of a standard squat. Form standing position to sitting position [Photo ID: 1305075652, from www.shutterstock.com].

- Sit to stand (chair squat) (3 repeats)
    - This movement started from sitting position to standing position. The participant had a one-second pause in the standing position before returning to the sitting position. A three-second pause was considered in between repeats. The participant's hands were on their waist during the task (Figure 3.10).

Figure 3.10 Participant during the sit to stand experiment.

- Step ascending and descending (5 repeats)
    - The experiment started from the bottom to the top of the stairs. The force plates under the first and second boxes measured the ground reaction forces. Weights were added to the first and second boxes to improve stability. The force plates were zeroed before and after adding boxes and weights (Figure 3.11 and Figure 3.12). Further calculations were necessary to calculate the centre of pressure account for the padding on the force plates.

Figure 3.11 Participant during the steps ascending part of the experiment.

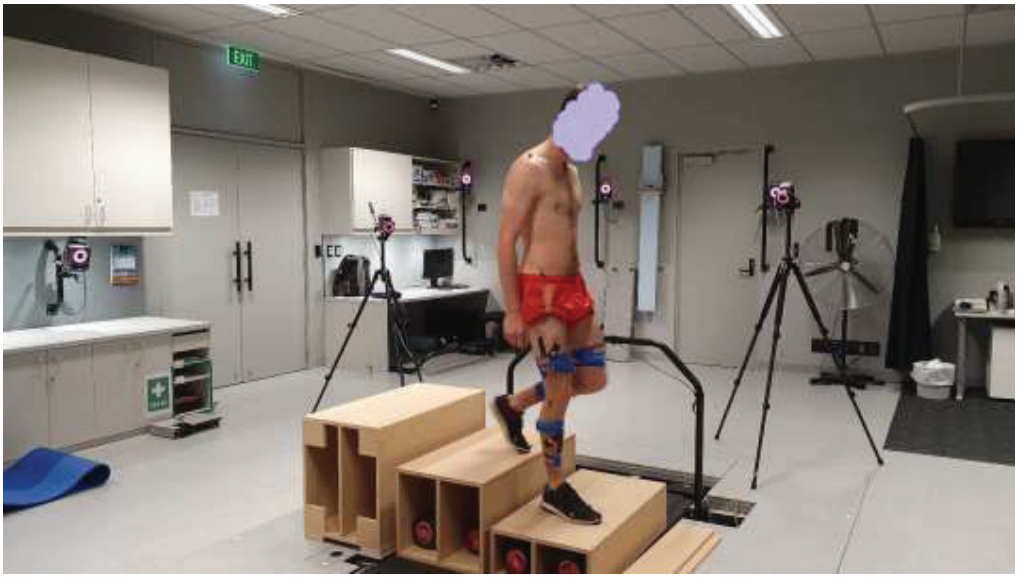


Figure 3.12 Participant during the steps descending part of the experiment.

Each of these activities was repeated more than once (at least three times) to check the consistency during the experiment and provide enough information for model development. The experiment run sheet is represented in Appendix A.2.

3.4 Data preparation

Collected data from the motion capture system, EMG system, and force plates were synchronised and recorded via VICON NEXUS software (Figure 3.13). Muscle activities and force plate data were collected at 1000 Hz in our experiment, whilst marker trajectories were captured at 100 Hz.

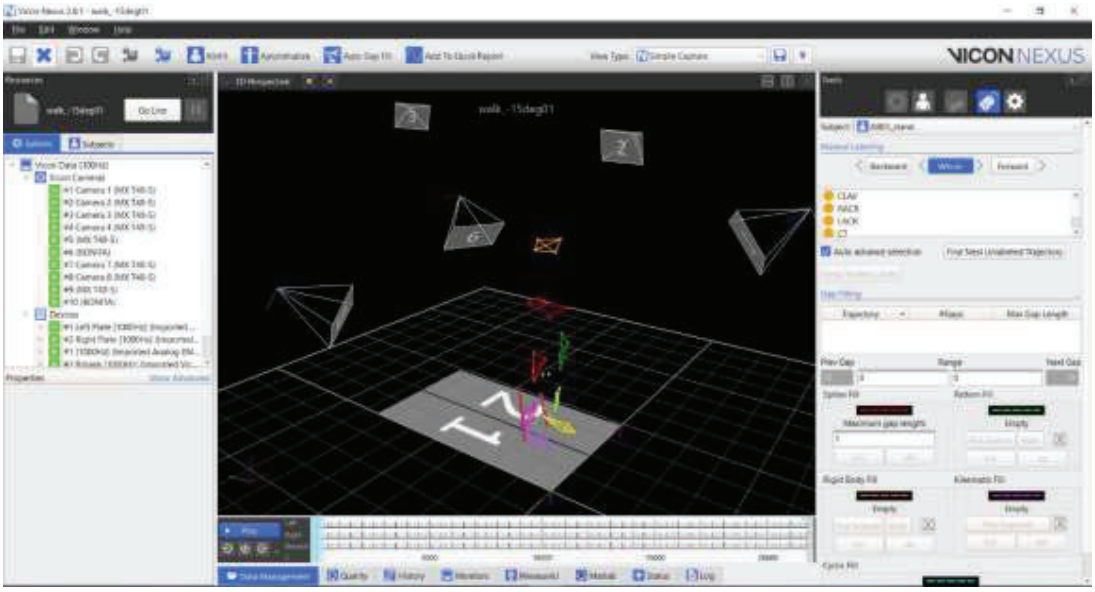


Figure 3.13 A view of the VICON NEXUS interface window during the data collection.

The collected data were processed in the following manner:

- **Motion capture data:** Markers were labelled and tracked using an 'auto label' function in NEXUS. However, manual intervention was required to fill any gaps and make sure marker trajectories were smooth and continuous. Gap-filling was performed by using a combination of automatic (for gaps <10 frames) and manual gap filling (for gaps >10 frames) using spline trajectories and 'Rigid body fill' options in NEXUS. Cleaned and labelled marker trajectories were exported as ".trc" files as input for joint angle calculation in OpenSim (Stanford CA).
- **Ground Reaction Force (GRF) data:** Raw forces and moments from both force plates were exported as a ".mot" file format (for input to OpenSim). Data were low-pass filtered with a Butterworth filter (6 Hz cut-off frequency) to ensure dynamics consistency with the joint kinematics data. Centre of pressure calculations was performed, and the GRF data were registered to the global coordinate system of the OpenSim.

Recorded GRF data from each force plate included force vectors in 3 dimensions (Fx, Fy, and Fz) and the moment components (Mx, My, and Mz). Based on the recorded force and moment vectors, the centre of pressure (COP) is definable, representing the contact point between the participant's feet and the ground. The COP is required for inverse dynamics calculation in OpenSim.

The coordinate system in OpenSim is different from the VICON system. For instance, in VICON, the z-axis is perpendicular to the ground, while in OpenSim, it is y-axis. Also, there is a translational misalignment between the origin of the coordinate system for the force plate, OpenSim and VICON's global coordinate system (i.e. marker trajectories). These differences were reconciled to ensure the ground reaction forces were registered to OpenSim's marker data and model.

- **EMG data:** The recorded EMG signals were also exported in the same ".mot" file format as the GRF data. Raw EMG data were band-pass filtered using a zero-lag Butterworth filter (20 Hz - 500 Hz), full-wave rectified and then filtered using a low-pass Butterworth filter (6 Hz) to create a time series linear envelope. The filtered EMG for each muscle was normalized to the maximum value of the MVC trial for each corresponding muscle (which was processed in the same manner). The normalized EMG data were then down-sampled to 100 Hz to match the kinetic data.

3.4.1 Ankle joint angle and moment calculation

Exported marker trajectories and GRF data were used as input to OpenSim software (Stanford CA) to calculate lower limb kinematics and joint moments by solving inverse kinematics and inverse dynamics equations, respectively [158]. The process of scaling a musculoskeletal model for each participant in the experiment to calculate ankle angle and moments comprises three main phases represented in Figure 3.14.

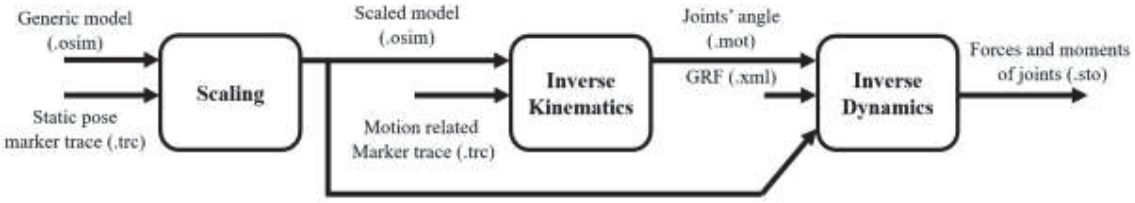


Figure 3.14 The block diagram of the OpenSim pipeline for data preparation. The format of information at each stage is represented in round brackets.

3.4.1.1 Scaling

A musculoskeletal model of the human body, including 23 degrees of freedom and 54 muscle-tendon units (Gait-2354), was used as a generic template model for scaling [159]. Scaling involves adjusting the body segment lengths to match anatomical marker lengths placed on the participant during a static trial. A statistical shape model-based approach was used to scale the pelvis, femur, and tibia. This process has been done via the "MAP Client" scale tool [160].

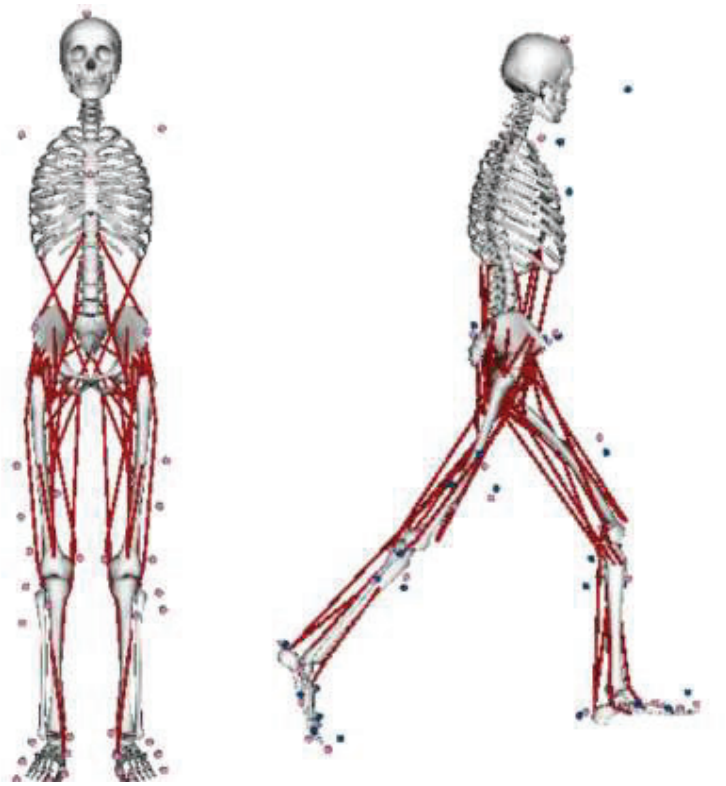


Figure 3.15 The anatomical model of the human body in OpenSim. The model includes 23 degrees of freedom and 54 muscle-tendon units (named Gait-2354). The pink markers are the model-defined markers, and the blue markers are recorded during the experiment.

For the torso and feet, the standard scaling tool of OpenSim was used. During the scaling process in OpenSim, the anatomical markers (static markers) were used to adjust the length of the segments of the anatomical model based on the recorded data during the static pose. The position of the model-based tracking markers was registered to match the location of the tracking markers obtained from optical motion capture (Figure 3.15).

3.4.1.2 Inverse kinematics (IK)

In this phase, the scaled OpenSim model was used to track the marker trajectories recorded during each movement using inverse kinematics (Figure 3.14). This involves an optimisation to find the 'best match' between experimental markers and model markers in a least-squares sense. Joint

constraints are imposed in this process, and weights can be applied to each marker to define what level of confidence you have in the experimental marker trajectory. This enables the user to account for soft tissue artefact, for example, by reducing the weight on the thigh markers, which are prone to large movements due to muscle tissue (Figure 3.14). The output of the IK tool is a motion file (.mot) containing the generalized coordinate trajectories (joint angles and/or translations) computed by IK.

3.4.1.3 Inverse dynamics (ID)

To solve the equations of motion using an inverse dynamics approach, OpenSim requires three inputs:

1. The scaled musculoskeletal model
2. The kinematics of the motion (IK output)
3. Ground reaction force (GRF) during the movement.

It is important at this stage to ensure that the kinematic data from IK are filtered, as the optimisation routine solves the pose-estimation problem frame-by-frame. As such, there is no guarantee that the kinematic output is smooth and continuous, which generates noise when double differentiating (as part of the inverse dynamics calculations). Furthermore, it is essential to filter the kinematic data at the same frequency as the GRF data to ensure dynamics consistency [161].

3.5 Time series segmentation (windowing)

The pipeline above converts the raw data recorded during the experiment to the time series applicable for data-driven model development. Depending on the prediction scenario, the desired model takes input time series and predicts the output time series, either ankle angle or ankle moment. The predictive model took information from a window of input time series and predicted a value at each time step to continuously estimate output. The process of output prediction was repeated in high frequency (at least 10 times per second), which makes a semi-continuous estimation of the ankle moment or angle. At each step, the model takes a limited length of input time series called a window of data. The length of the window directly impacts the model performance and the accuracy of the prediction [155]. To find the optimal length of input data, a backpropagation neural network was designed to predict the moment of the right ankle based on EMGs and kinematics inputs. The model had three hidden layers with rectified linear unit (ReLU) activation function. The model was trained and tested on data corresponding to 30 seconds of walking at 1 m/s from one of the participants in the experiment. 70% of data was used for model training, and 30% was used for testing. The input data included EMGs from the tibialis anterior, gastrocnemius lateral and medial, and soleus from the left and right leg and angle and angular

velocity of left and right ankle. The output was the right ankle moment. The process of train and testing was repeated for different lengths of windowed time series. Figure 3.16 shows the model's accuracy in ankle angle prediction for different input window lengths. The model accuracy is reported based on the root mean square error (RMSE) between the predicted value and actual ankle moment. The RMSE was calculated and reported as a percentage of the actual ankle moment based on the following equation:

$$\sqrt{\left(\frac{\sum_{i=1}^{n}\left(\frac{predicted\ moment - actual\ moment}{actual\ moment}\right)^2}{n}\right)} \quad (1)$$

In which n is the number of data points in test dataset.

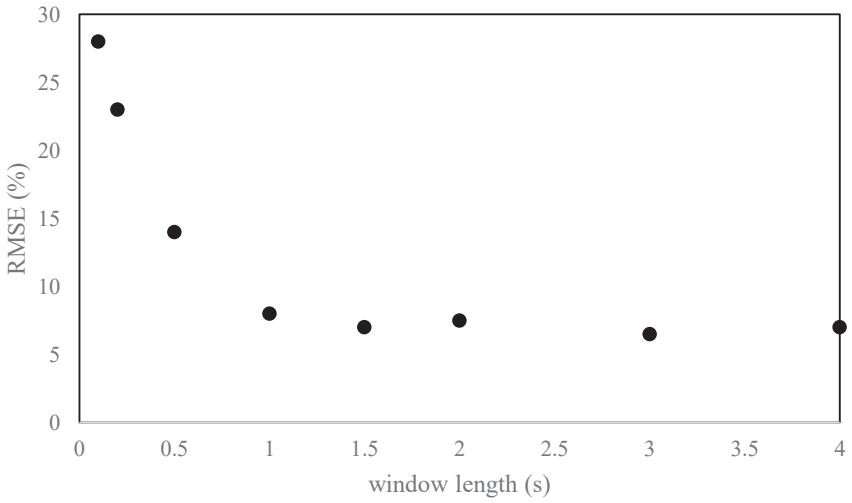


Figure 3.16 *The impact of input data window length on a data-driven model accuracy for ankle moment prediction based on EMGs and kinematics.*

As shown in Figure 3.16, by increasing the length of the input windowed time series, the model error decreased; however, for windows longer than one second, there was no significant improvement in model performance. So for the model developed in chapters 5, 6, and 7, the length of the input time series for model development was considered one second. Considering 100 Hz as the data collection frequency, one second of each input time series corresponded to 100 data points.

4 Ankle moment forecasting using time-delayed neural networks

Homayoon Zarshenas, Bryan P. Ruddy, *Member, IEEE*, Andreas W. Kempa-Liehr, *Member, IEEE* Thor F. Besier

A method for ankle torque prediction ahead of the current time is proposed in this paper. The mean average value of EMG signals from four muscles, alongside the right ankle's joint angle and angular velocity, were used as input parameters to train a time-delayed artificial neural network. Data collected from five healthy subjects were used to generate the dataset to train and test the model. The model predicted ankle torque for five different future times from zero to 2 seconds. Model predictions were compared to torque calculated from inverse dynamics for each subject. The model predicted ankle torque up to 1 second ahead of time with a normalized root mean squared error of less than 15 per cent while the coefficient of determination was over 0.85.

Clinical Relevance— the model's potential for predicting joint torque ahead of time is helpful to establish an intuitive interaction between humans and assistive robots. This model has application to assist patients with neurological disorders.

4.1 Introduction

Joint torque prediction is one of the most important research fields in biomechanics. It explains muscle-actuated human movement, which can be used for purposes ranging from athletes' exercise optimization [162]to assistive robots designed to help people who suffer from movement disabilities [163]. Joint torque is typically estimated by solving multi-body inverse dynamics equations based on body movement data measured with a 3D motion capture (MoCap) system in addition to external forces [19]. However, this approach is time-consuming and requires special facilities, limiting its practicality in real-time applications.

As an alternative solution, machine learning has been proposed to understand human body dynamics, as reviewed by He et al. [164]. In terms of joint moment prediction, a feed-forward artificial neural network has been used to predict elbow joint toque based on electromyography (EMG) signals [82]. Another approach uses EMG and kinematic features of the elbow joint to create input vectors for a recurrent artificial neural network (RANN) to predict joint torque [165]. A combination of fuzzy logic and neural networks [166] and wavelet-based artificial neural networks have also been proposed as solutions to predict joint torques [167].

The above-mentioned neural network methods have shown promising results in predicting joint torques based on EMG signals. In real-time applications such as teleoperations and assistive robot control, however, the time delay between motion intention detection by the level of EMG activity and actuator activity is of critical importance. Ideally, the required joint torque would be predicted to minimize the latency between motion intention detection and actuator response and reduce the unexpected forces between humans and robots. This paper presents the design and development of a time-delayed artificial neural network (TDANN) to predict ankle torques forward in time. The model was trained and tested to predict ankle joint torque up to 2 seconds ahead of time during level walking at a self-selected speed. We compared two approaches to evaluate the importance of kinematic features on torque prediction accuracy. Our first approach used only EMG data as input (EMG-alone); our second approach included ankle angle and angular velocity to the EMG data (EMG + kinematics).

4.2 Method

4.2.1 Experiment and dataset preparation

Five healthy young subjects (age: 29.2±1.7 yr; height: 174.5±10.3 cm; weight: 72.7±12.2 kg) volunteered to participate in this study. The experiment was approved by the University of Auckland human participants ethics committee (approval #022112). All participants gave their informed consent prior to the experiment, which involved measuring body movements, ground reaction

forces, and EMG signals whilst walking on a force-instrumented treadmill at a self-selected speed. Retroreflective markers were placed on anatomical landmarks and tracked at 100 Hz using a 3D motion capture system (VICON Nexus, v. 2.8.1). Ground reaction forces were captured synchronously (1000 Hz) by the force plates embedded in the treadmill for 30 seconds of steady-state walking. The EMG signals of the following four muscles of the right limb were recorded at 1000 Hz: tibialis anterior, medial gastrocnemius, lateral gastrocnemius, and soleus. The location for each EMG electrode on each muscle was determined based on the SENIAMS standard [168], and maximum voluntary contractions (MVCs) were recorded prior to the walking trial. Collected EMGs were band-pass filtered to remove movement artefact (450 Hz to 30 Hz), then full-wave rectified and low-pass filtered using a second-order Butterworth filter (cutoff frequency 5 Hz) to extract the muscle activity envelope [169]. Processed EMG signals of each muscle were then normalized to their corresponding MVC activation.

The marker trajectories for each subject were used as input to OpenSim (Stanford, CA) to scale a generic musculoskeletal model with 23 degrees of freedom and 54 muscle actuators. OpenSim was then used to solve the inverse kinematics (IK) problem and calculate the ankle angle and angular velocity [158]. The kinematic features alongside the ground reaction forces were used as input for an inverse dynamics (ID) analysis in OpenSim to calculate the ankle joint torque of the right limb during normal walking. The calculated ankle torque via inverse dynamics was considered as the desired value for training the neural network-based model. For the convenience of discussion, the results of the inverse dynamic are referred to as *actual* torque, and the torque predicted by the neural network-based model is named as *predicted* torque.

4.2.2 Model design and development

As a model to predict joint torque based on EMG signals, a time-delayed artificial neural network (TDANN) was designed and implemented using MATLAB software (v. 9.5.0). The model included four hidden layers with 40 neurons per layer (4 by 40) with two time-step delays. The structure of the hidden layers was determined by a systematic investigation over different combinations of numbers of neurons and hidden layers. The error between actual torque and predicted values and the consistency of the network's outcome over time were two main parameters for selecting the network structure. The output layer of the model included five nodes, which represented torque values for 0, 0.5, 1, 1.5, and 2 seconds ahead of the time frame of the input vector. A sigmoidal transfer function was used as the activation function for all neurons in hidden layers, and the output layer used a linear transfer function. A Levenberg-Marquardt algorithm was used to update the weights and biases in each hidden layer, which reduced the root mean squared error between the predicted torque value and the actual torque from inverse dynamics [170].

To make the input vector for this model, the EMG signals were down-sampled to 100 Hz by calculating the mean absolute value (MAV) as the representing feature for each EMG channel [136]. The performance of the designed TDANN was assessed under two scenarios. In the first scenario, the MAV of EMG channels with ankle angle and angular velocity was used as input vector, while in the second scenario, only the MAV of EMGs was considered as input for the network. Because of the delays in the network, the input features regarding the current time frame (*t*) and two time-steps before that (*t-1* and *t-2*) were used to generate the input vector for the model (Figure 4.1). So the feature vector represented information corresponding to a window of input data with 20 ms length moving forward 10 ms in each time step.

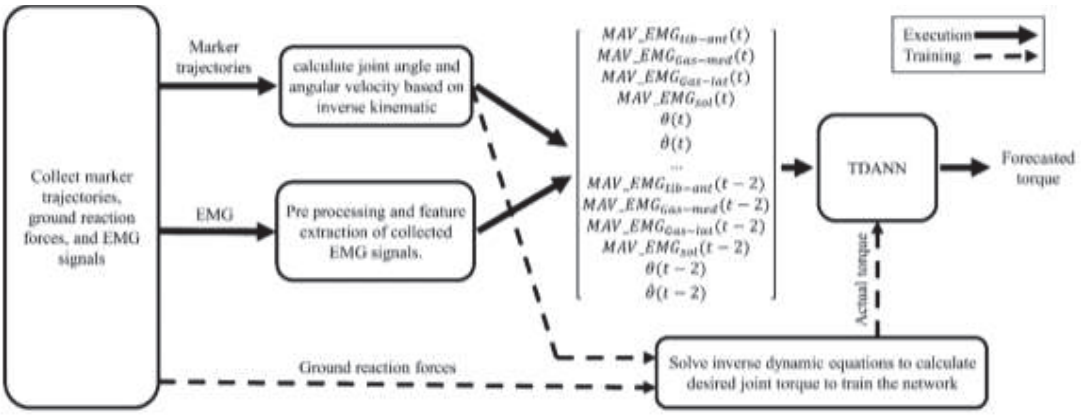


Figure 4.1 *The pipeline for developing the time-delayed artificial neural network for joint torque prediction. The solid lines represent the process of creating the required feature set from input time series (EMG and marker trajectories). The dashed line is the process of generating the desired output used for model training. © 2020 IEEE.*

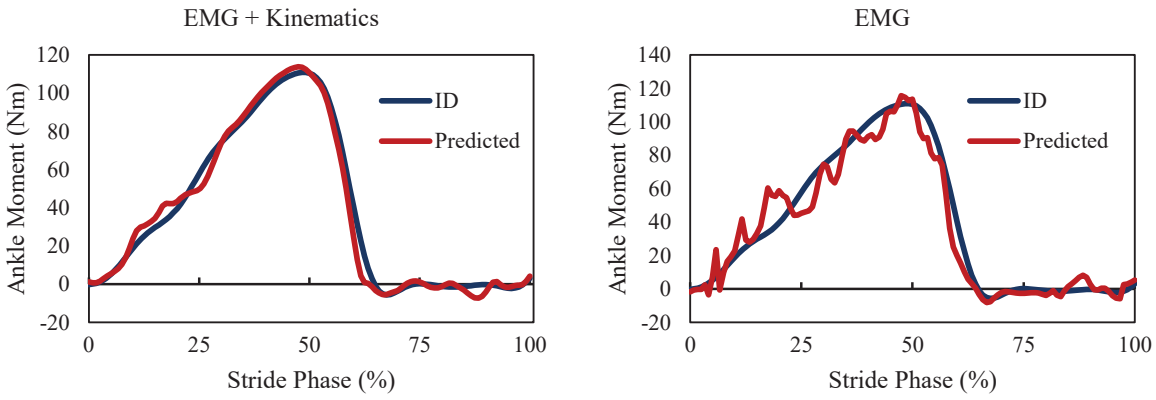
To compare model performance over different training scenarios and different future time horizons, the root means square error (RMSE) was calculated based on the difference between the predicted values and the actual torque for five subjects and normalized to the range of actual torque over the test dataset. We compared the curvature of the predicted vs actual joint torque waveform using the coefficient of determination (square of the correlation coefficient between actual torques and estimated torques – R^2).

The future torque prediction capability of the designed and developed TDANN was investigated in two different input vector structures; (i) "EMG + kinematics" and (ii) "EMG-alone." The training dataset for the TDANN included input vectors from six strides of the right leg for each subject, and the next five strides were used to test the network performance.

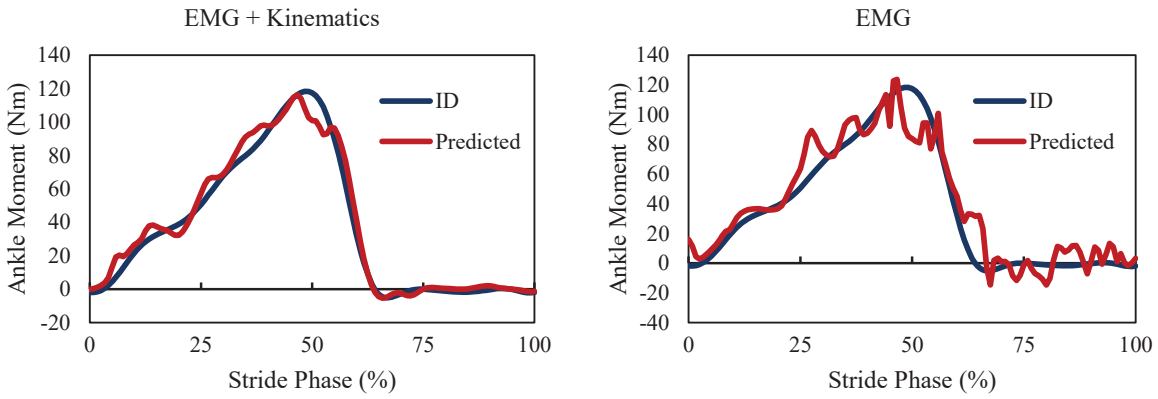
4.3 Results

More fluctuation in torque prediction over time steps in future prediction was observed in the EMG-alone model than in the "EMG + kinematics" model (Fig. 2). Model predictions for joint torques up to 1 sec beyond the current time point were reasonable, but predictions were less accurate from one to two seconds forward in time (Figure 4.2). The capability of the "EMG + kinematics" model for predicting the torque for a long time interval was assessed by testing its outcome over the next 6 seconds while predicting torque in different time steps ahead. The model outcome remained accurate, and there was no unusual under/over torque prediction for zero up to one second ahead of time, which shows model robustness over deviation in input signals over time (Figure 4.3).

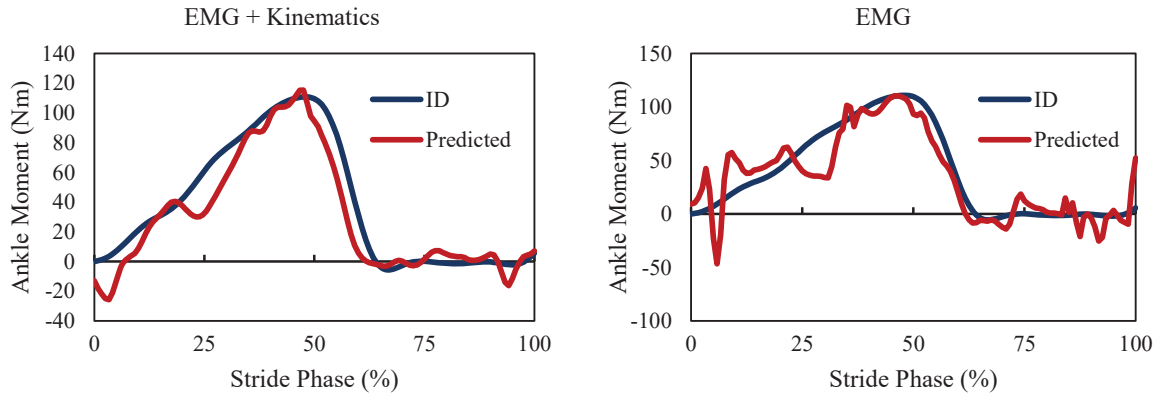
A quantitative comparison based on the value of normalized RMSE over five subjects showed the lower error rate for the "EMG + kinematics" approach rather than the "EMG-alone" approach overall future time-steps forecasting. Also, in "EMG + kinematics", the average of normalized RMSE was 8.5% for zero-time ahead and remained less than 15% until 1 s ahead prediction (Figure 4.4). The level of R^2 value for different time gaps in the future showed over 0.85 compatibility between actual torque and predicted values for up to one second ahead of time. The best R^2 value was in zero seconds ahead torque prediction, which was 0.95 on average for five subjects (Figure 4.5).



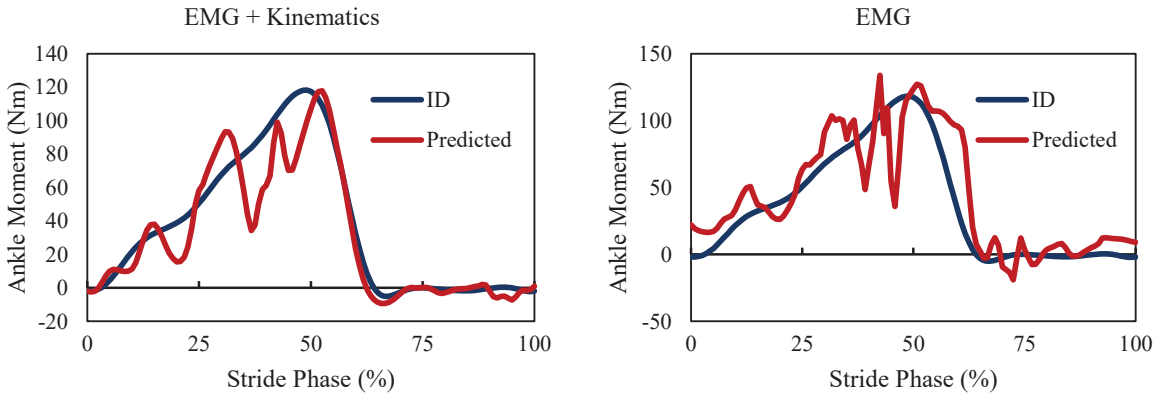
(a) 0 s forward



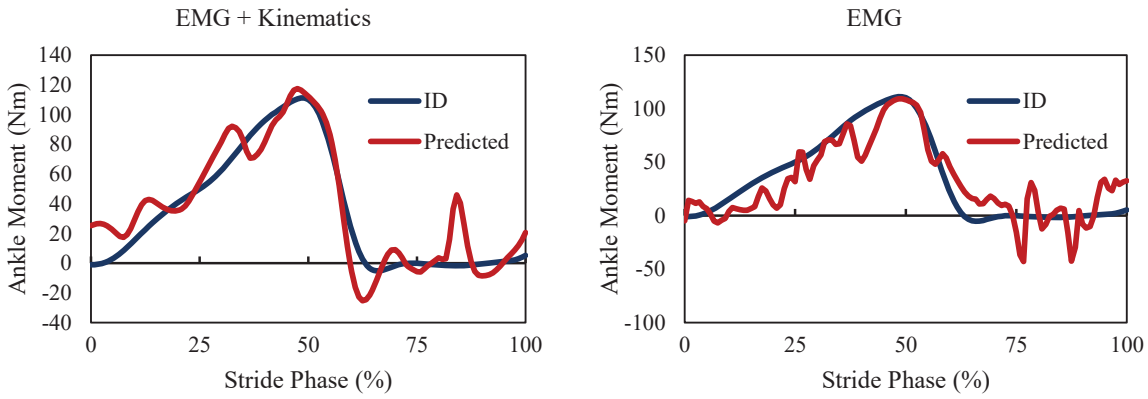
(b) 0.5 s forward



(c) 1.0 s forward

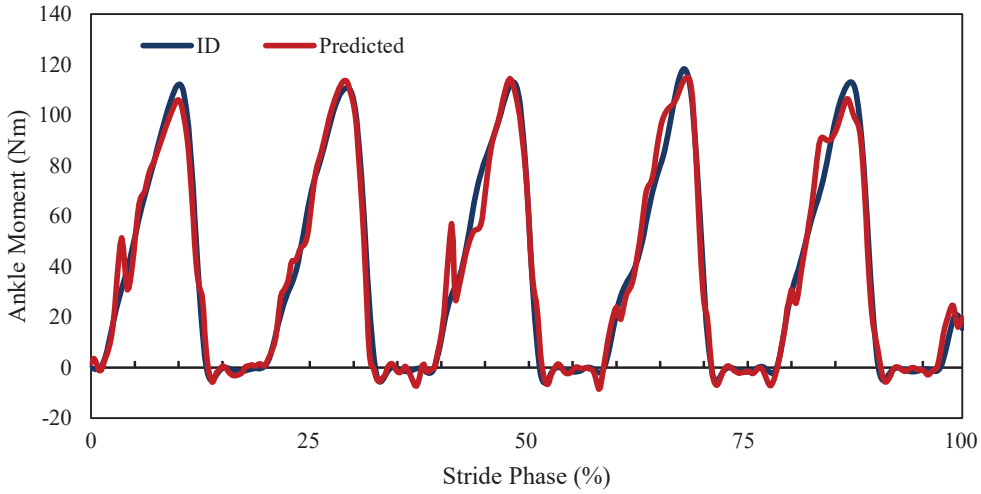


(d) 1.5 s forward

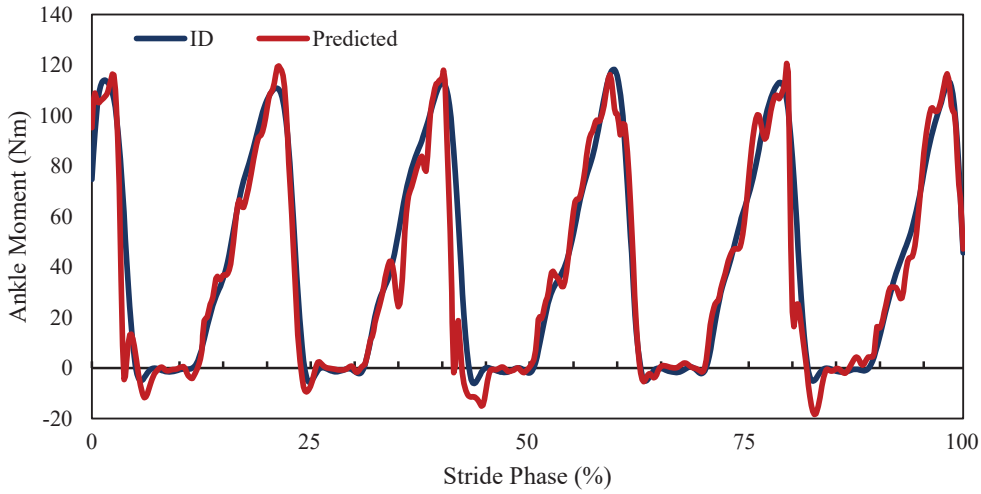


(e) 2.0 s forward

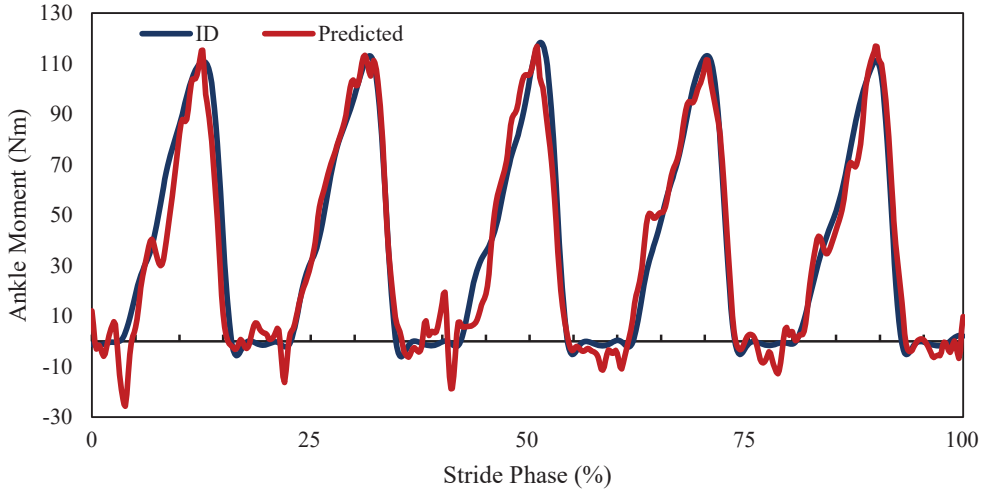
Figure 4.2 *Ankle torque prediction comparison between two model training approaches: "EMG + kinematics" (left column) and EMG alone (right column) over one heel-strike for one subject. The blue line is the actual moment, and the red line is the predicted moment.*



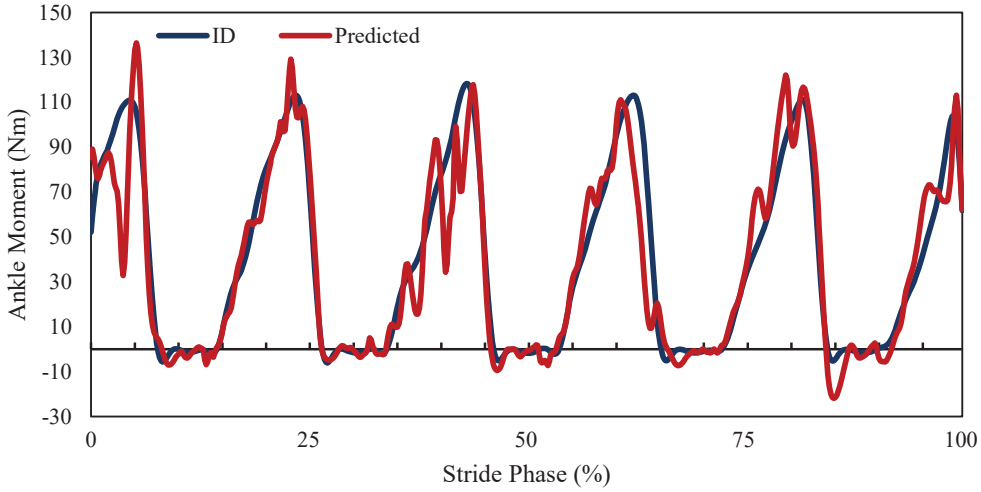
(a) 0 s forward



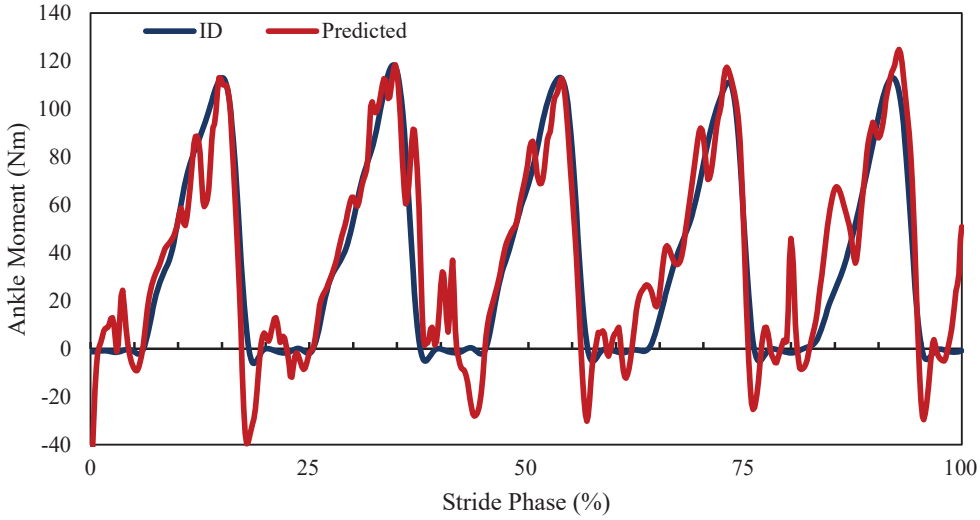
(b) 0.5 s forward



(c) 1.0 s forward



(d) 1.5 s forward



(e) 2.0 s forward

Figure 4.3 *Ankle torque prediction when EMG and kinematics are both used as input for the model. The results represent the model performance on 6 seconds of test data for one subject. The model predicted ankle torque (moment) for multiple time spans in future (a. 0.0s, b. 0.5s, c. 1s, d. 1.5s, and e. 2.0s). The blue line is actual torque, and the red line is predicted torque. © 2020 IEEE.*

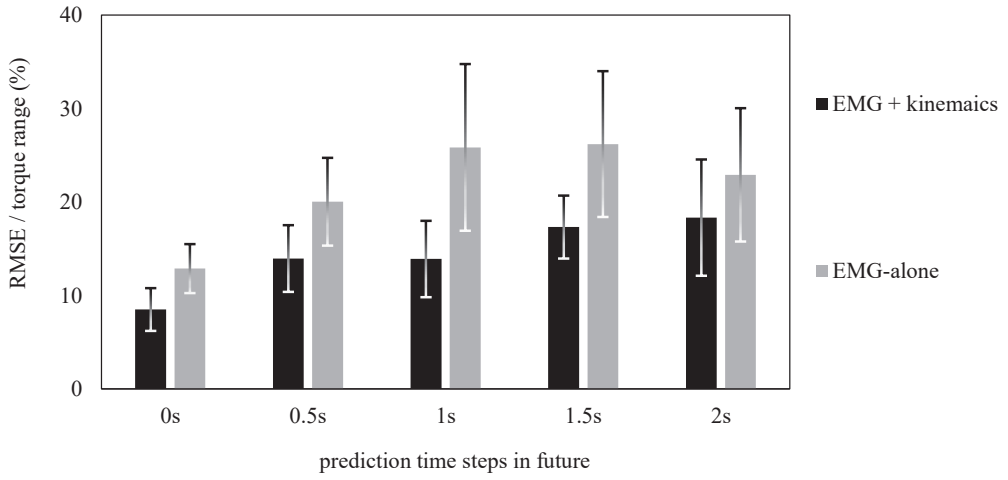


Figure 4.4 *The average and the standard deviation of RMSE (normalized to the torque range) calculated for five subjects for different future time spans. © 2020 IEEE.*

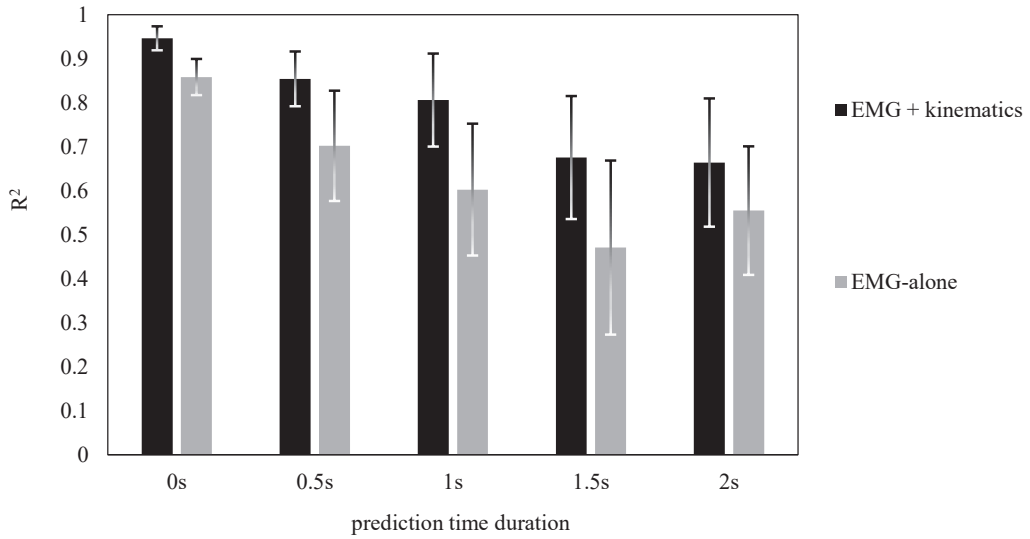


Figure 4.5 *The average of coefficient of determination (R^2) over five subjects for ankle torque prediction in five different time durations in the future. © 2020 IEEE.*

4.4 Discussion

This study assessed the ability of a time-delayed multi-layer artificial neural network model to predict ankle torque during walking at normal speed from EMGs of four muscles and kinematic features as the input parameters. The "4 by 40" network structure presented here produced a low RMSE and was less likely to be 'trapped' in local minima during the training phase. This structure also showed greater robustness to the time-dependent deviation in input signals than other network

structures combinations. More complex networks suffered from over-fitting, degrading their prediction accuracy over test data.

The quality of the predicted torque directly relies on the quality of collected EMG signals. This was particularly noticeable in the EMG-alone model, where noisy EMG signals resulted in fluctuations in predicted torques. The inclusion of kinematic features improved the torque predictions, which might result from the dependency of joint torque on joint posture and velocity and muscle activation. As expected, the prediction of joint torques forward in time was reduced as the prediction time duration increased. However, the reasonable performance of the model up to 1.0 s ahead of the input vector gives us confidence that this approach could be useful within a control scheme for controlling a robotic exoskeleton. There are some fluctuations in the predicted values for ankle moment, which are related to the noise of input signals. A predictive model with a more complex structure would be more robust to high-frequency noise and provide smoother output. The structure of the model in this study was determined as a trade-off between the computational cost and desired accuracy. As predicting ankle moment for larger time spans into the future needs more complex networks and will be computationally expensive, models that can predict ankle moment beyond one second into the future were not explored.

EMG is a short latency signal, meaning there is ~ 30 to 100 ms delay between signal onset and the physical movement, so there is no physiological correlation between the joint movement at each time step with EMG signals from one second in the past. However, as in this study, the focus was on a cyclic movement (gait), it was possible to mathematically correlate input signals with the output for time steps longer than the electromechanical gap between EMG and muscle contraction.

In robot control applications, the prediction time gap will be adjusted to a shorter duration, according to the size of the muscles involved in the movement and the required time for signal processing for the controller. As future work, the next step is assessing the model performance on ankle torque prediction over other lower limb activities such as step climbing and walking at different speeds. Also, the model performance will be examined over a more significant number of healthy subjects.

4.5 Conclusion

This work proposed a TDANN model developed for the ankle torque prediction based on the EMG signals and kinematics features. This model could predict joint torque up to 1 second ahead of time with less than 15 per cent error (normalized RMSE), an acceptable level of error for potential applications in the control of assistive robots.

5 Forward Prediction of Ankle Joint Moment Using a Generic Feature Set

Homayoon Zarshenas, Thor F. Besier, Bryan P. Ruddy, Alexander Woodall, Andreas W. Kempa-Liehr

Providing an adaptive interaction between the human and robot is essential to maximize the benefits of using a robotic exoskeleton for rehabilitation and empowering people in daily life movements. Machine learning (ML) based models are extensively used to predict motion intention. However, they are mainly task and subject-specific. A non-subject-specific model will fasten each individual's training process and make the ML models more practical in exoskeletons control. This study developed an ML-based model to predict ankle joint moment during walking at a constant speed. The model is based on the random forest regression method and uses electromyography (EMG) signals and ankle joint kinematics as input time series. A systematic feature engineering pipeline was implemented to find the generic feature set applicable to train the predictive model for an unseen individual, based on the data collected from 10 individuals walking at a constant speed. A ten-fold cross-validation method was used to find the list of most informative features in the feature engineering process. After that, the list of 16 features repeated in all permutations of the cross-validation was selected as the list of generic features. It has been proved the generic feature set introduced in this study is applicable to multiple individuals for ankle moment prediction, to speed up the model training process due to the reduction of computational complexity and provided the same level of accuracy as when a personalized set of features was used for model training.

Note to Practitioners— Machine learning-based models are commonly used to predict motion intention based on EMG signals. Typically, these models are subject-specific due to differences in muscle activities among different people and the many different possibilities a machine learning algorithm can learn relevant information from the measured EMG signals. This characteristic makes the process of model training for each individual time-consuming and challenging. We simplify the process by introducing a systematic time-series feature engineering process, which generates a small informative set of time-series features. In the presented case study, the feature engineering process was used to predict ankle joint moment 30 ms ahead of the current time step based on EMG signals and kinematics for a constant walking speed of 1 m/s. The identified generic set of 16 features makes the model training faster and easier for a new individual. The generic feature set generalizes well across a cohort of ten volunteers with a generalization error of one per cent compared to an individualized set of time-series features. Although an individual calibration step is still advised for the regression model, we conclude that the proposed systematic time-series feature engineering process reduces the dimensionality of the machine learning problem significantly and has the potential to accelerate the development of robotic exoskeletons.

Index Terms— EMG-based moment predictive model, feature engineering, motion intention detection, time-series prediction, assistive robot controller.

5.1 Introduction

Neurological disorders cause physical changes such as muscle weakness, poor coordination, and loss of sensation. These physical disorders affect daily activities, including locomotion in a large number of people. Providing targeted rehabilitation is the most important priority for people with movement disorders. Powered exoskeletons show much promise to provide locomotion rehabilitation for individuals with movement disorders such as spinal cord injury and stroke patients. One essential factor in exoskeleton-based rehabilitation is providing active participation for the patient.

The exoskeleton needs to adapt to the user movements to have active participation, which requires an understanding of motion intention [171]. Despite the progress of exoskeleton mechanical design, few control methods can provide active assistance to the user [172, 173]. The desired control algorithm must decode the motor functions accurately and robustly. Electrophysiological signals, such as EEG and EMG, can interpret the movement intention from the central nervous system, with EMG being the most common due to its ease of use. Surface EMG represents the summation of action potentials generated by the underlying motor units, containing useful information regarding muscle activities and human movement.

To understand motion intention based on EMGs, a model is necessary to correlate muscle activity to the corresponding joint moment. Neuromusculoskeletal (NMS) models based on the modified Hill-type muscle model have been used extensively to determine muscle forces and the joint moment [19, 174, 175]. Because of the nonlinear relationship between muscle activation and joint moment, the NMS model includes multiple parameters, which requires a complex calibration process and computational power. As an alternative, machine learning can be used to obtain a surrogate model to relate the measured EMG signals to joint moment without the need to model muscle mechanics and neurophysiology [176]. Artificial neural networks with an adequate number of hidden layers can predict joint kinematics based on EMG in relatively complex tasks, including lower limb kinetics [177], knee moment prediction during running at different speeds [178], isokinetic ankle joint moment estimation [176, 179], ankle moment prediction during walking [180], and elbow kinematics prediction based on EMGs from biceps and triceps EMGs [181].

Machine learning approaches show promising performance in joint moments and kinetics estimation. However, these models generally offer accurate performance only in isokinetic tasks, and they are mostly subject-specific. This means these models cannot estimate a broad range of tasks and must be trained to each individual using a lengthy calibration procedure.

The process of generating a machine learning model on sensor data comprises three stages; 1) signal preprocessing, 2) time-series feature extraction/selection, and 3) model training. Although the extraction and selection of time-series features can be highly automated [151, 182], it is still a time-consuming process that is subjective to different people and various tasks. This means the selected feature set for two individuals to predict a specific joint moment for the identical movement is not the same. Moreover, different sets of features are required for one person to predict joint moments during various activities.

This study aims to find a generic set of features to predict ankle joint moment during walking at a constant speed, which applies to multiple individuals. An investigation was conducted over a range of features to find a set of features out of EMG signals and kinematics to satisfy two main goals; a) minimize the number of required features for a precise moment estimation, and b) define a set of features applicable to a new individual (outside of the training set) to make the process of model training faster. The selected feature set was used as input for a regression model to predict ankle joint moment 30 ms ahead of the current time-step for ten individuals to test the method's robustness.

5.2 Method

5.2.1 Population

Ten volunteers participated in this study (eight males and two females, age: 27±3 yrs, height: 174.6±8.5 cm, mass: 73±11 kg). Selected participants had no history of walking-related disorders, lower limb surgery, or pain affecting their walking abilities. All participants were familiar with treadmill walking and gave their written, informed consent prior to participation. The study was approved by the University of Auckland human participants ethics committee (UAHPEC), ref. Number 022112 (date of approval: 09-Nov-2018).

5.2.2 Experiment process

Reflective markers were attached to the participant in accordance with the UWA marker set [183]. The trajectory of these markers was recorded at 100 Hz during the experiment using a VICON 11-camera motion capture system (VICON, Oxford UK). Wireless surface EMG electrodes (Delsys Trigno, Natick USA) were placed above the medial and lateral gastrocnemius, soleus, and tibialis anterior muscles of the left and right leg in accordance with SENIAM [168], and muscle activations were recorded at 1000 Hz.

Maximum voluntary contractions (MVC) of the muscles of interest were recorded before a 2 to 4-minutes familiarization session of walking at a self-selected speed on a split-belt force-instrumented

treadmill (Bertec, OH). Ground reaction forces (GRF) were captured at 1000 Hz during the data collection section. Participants walked steadily for 90 seconds at 1 m/s speed, and the last 30 seconds were used as input to the model.

5.2.3 Signal processing

Collected EMG signals were band-pass filtered using a zero-lag Butterworth filter (20 -500 Hz), full-wave rectified and filtered using a low-pass Butterworth filter (6 Hz). The filtered EMG for each muscle was normalized to the maximum value of the MVC trial for that muscle. The normalized EMG data were then downsampled to 100 Hz to match the kinetic data. Ground reaction forces were low-pass filtered at 6 Hz to be dynamically consistent with the kinematic data.

5.2.4 Model development

OpenSim 3.3 (Stanford, CA) was used to calculate the body's kinematics and kinetics based on the collected data in the experiment. A generic musculoskeletal model of the body (OpenSim gait 2354) was scaled to each individual's anthropometry using marker trajectories recorded in a static pose. It has been shown that the shaped-model scaling is more accurate than the linear scaling [160] [18], so a statistical shape model-based approach was used to scale the pelvis, femur, and tibia. This process has been done via the "MAP Client" scale tool [184]. Linear scaling was used for the foot and torso. Marker trajectories were then tracked using inverse kinematics and filtered using a Butterworth low-pass filter at 6 Hz to calculate the kinematics of the lower limb. Inverse kinematics (IK) and GRFs were used to calculate the ankle joint moment during walking by solving the inverse dynamics equations. The calculated moment was considered the 'desired value' for the surrogate model training to correlate muscle activities to the ankle's moment. The goal of the surrogate model was to predict ankle joint moment up to 30 ms ahead of the current time step. The actual target for the surrogate model was ankle moment normalized to the height times weight of each individual.

A specific length of input time series was required as input for the surrogate model to predict each data point of the moment time series. Our previous findings indicated that a 1s window was enough for moment prediction at the constant walking speed (including 100 data points at 100 Hz sampling rate) [185]. The data window slides forward 10 ms at each time (Figure 5.1). Windowed signals and their corresponding joint moment output were used for feature extraction.

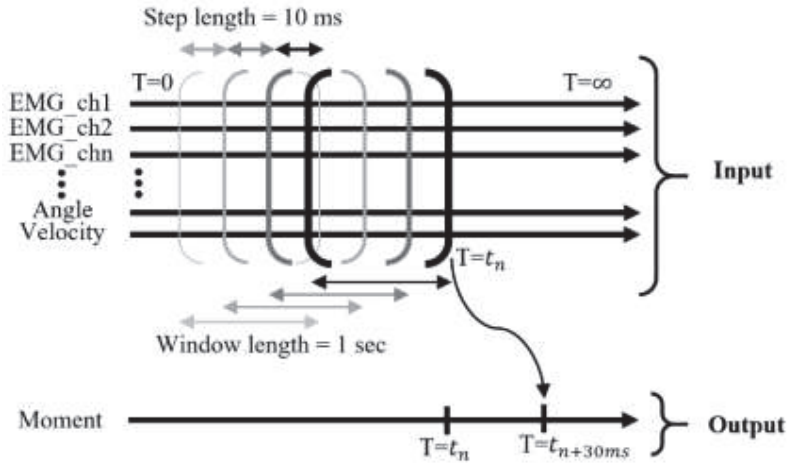


Figure 5.1 *The schematic view of input time series windowing for output prediction. Each window comprises one second of information from the input time series. At each time step, the model takes information from each window and predict the joint moment 30 ms in future. This process is repeated every 10 ms by sliding the window forward.*

The pipeline of data preparation and model development is represented in Figure 5.2. In the following sessions, each of the blocks in Figure 5.2 that correlate input data to the output is explained extensively.

5.2.5 Feature extraction

One of the essential characteristics of assistive robot control is minimizing the latency between robot and human; minimizing the required time to predict moment helps with that. The smaller the volume of input data, the faster the surrogate model can estimate the output. However, shortening the length of windows of input time series means missing potentially helpful information. The alternative solution is to extract informative characteristics of the input time series, including time-based, frequency-based, and time-frequency-based information, and use them as input to train the model.

By feature extraction, it is possible to keep the length of windows of the input data as long as required. At the same time, instead of all data points, only use the most important characteristics of the input time series for prediction. In this way, the computational cost of prediction will decrease. Also, it is possible to explain the importance of input time series based on the correlation between the extracted features and the desired output.

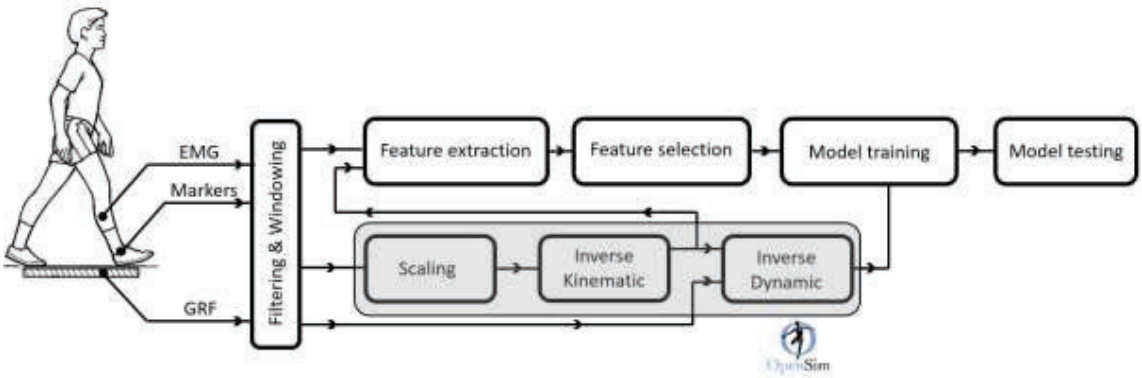


Figure 5.2 *The schematic pipeline of joint moment prediction model implementation based on MEG and kinematics. The shaded part of the graph in grey is happening in OpenSim, which is related to desired output generation for the model training.*

In this study, feature extraction was performed using the open-source Python package tsfresh (Time Series FeatuRe Extraction on the basis of Scalable Hypothesis tests), which applies 63 time-series characterization methods and generates 794 features from each one of the input time series [151]. As a rule of thumb, the more information we extract from the input time series, the more accurately the model can predict the output. However, this is not always true, particularly when dealing with noisy signals and a redundant mechanical system. Moreover, using all of these features to train the regression model is time-consuming and inefficient. Therefore, it is essential to find information with maximum relevance to the output.

Tsfresh filters the available features before using them for model training to limit irrelevant information. This process consists of three phases:

- Collect a comprehensive set of features from each window of input time series and aggregate them as a feature vector.
- Estimate the relevance of each feature for predicting the target variable on the basis of univariate hypothesis tests. The smaller the p-value, the more relevant is the feature. These evaluation results are a vector of p-values with smaller values indicating relevant features.
- Select relevant (statistically significant) features while controlling the false discovery rate [186].

An important question is how many features are required to predict the output. The key factor in answering this question is finding the minimum number of features without a significant drop in prediction accuracy.

5.2.6 Regression model

A random forest nonlinear regression model was selected to correlate the selected input features to the ankle moment. Random forests are constructed of many individual regression models known as decision trees at training. Each decision tree is fed with input data and independently predicts the output. Predictions from all trees are pooled to make the final prediction based on the mean prediction of each decision tree.

One of the main problems in machine learning is overfitting, which becomes more problematic if many colinear features are presented. However, it has been found that random forest models are robust estimators under these circumstances both for classification [182] and regression problems [187].

5.2.7 Performance evaluation

The regression model performance was assessed under different scenarios. The first scenario is finding the feature set for each participant and training and testing the regression model for each individual separately. The number of selected features for model training and the combination of different input time series were two items whose impact on model performance was assessed.

In the second scenario, the personalized feature set for one participant was used to train and test the regression model on the data of another participant to see how a personalized feature set works on unseen data. 10-fold cross-validation was conducted in the third scenario to find the most informative features among the population. To achieve this, 9 participants were used for feature extraction, and the list of top 50 features based on the p-values was used to train and test the regression model for an unseen participant.

In the fourth scenario, the similarity between all of the feature sets extracted based on 10-fold cross-validation in the previous scenario was checked, and the list of common features among all of the feature sets was considered as the generic feature set and used to train and test the regression model for each individual. In all of the above-mentioned scenarios, the training data was not used in the test phase. Even in the first scenario in which one individual's data was used to train and test the model, 70% of data was used for training, and the other 30% was used for testing.

5.3 Results

Several important decisions were made on the way to developing a generic feature set that works accurately for all individuals. In this section, the impact of each decision is explained.

5.3.1 Additional information based on the same number of inputs from sensors

One possible way to improve the model accuracy is to use recoded data from physical sensors to create more input time series called virtual sensors data [182]. In other words, combine existing time series to derive more information and use them for the feature extraction phase. The advantage of this time-series engineering is that it doesn't need extra input channels, which means no more hardware and sensor requirements for EMG measurement and less signal preparation. However, it is crucial to use a suitable combination of physical sensors.

In this study, the original input time series (physical sensors) included the EMG from four muscles of each leg plus the ankle angle for the left and right leg. We considered the difference between measured EMG from the same muscle on the left and right leg and the difference between EMGs of antagonistic muscles on the same leg as two approaches to make new time series out of existing EMG channels. The logic behind that was calculating the difference between EMGs activated at different stride phases on one leg or opposite legs to magnify the diversity among the input EMG channels.

Also, adding the angular velocity of the ankle and the difference between the left and right ankle angle and angular velocity was another stream of additional information derived from kinematics. The list of the additional time series is represented in Table 5-1.

Additional time series increases the number of inputs to 22. The impact of this additional information on the model performance was assessed, and the results were compared under different circumstances in Table 5-2 in terms of normalized root mean square of error (NRMSE) and R-squared. The reported RMSE was normalized with respect to the height and weight of the participants.

As shown in Table 5-2, additional information from EMGs slightly improved the model accuracy, but kinematics-related features caused a significant improvement in model performance. Fifty features were selected to represent each window of the input time series in each of the conditions in Table 5-2. The number of features for different input combinations, before and after feature selection, is represented in Table 5-3.

Table 5-1 Name of the additional time-series generated based on collected data during the experiment (virtual sensors).

EMG-related additional time series. Each row represent the EMG difference between the declared muscles	
Time series	**Description**
Sol_tib_r	Tibialis anterior and soleus muscles. (right leg)
Gasmed_tib_r	Gastrocnemius medial and tibialis anterior. (right leg)
Gaslat_tib_r	Gastrocnemius lateral and tibialis anterior. (right leg)
Sol_tib_l	Soleus and tibialis anterior. (left leg)
Gasmed_r_l	gastrocnemius medial of right and left leg
Gaslat_tib_l	Gastrocnemius lateral and tibialis anterior. (left leg)
Sol_r_l	Soleus of right and left leg
Tib_r_l	Tibialis anterior of right and left leg
Gaslat_r_l	Gastrocnemius lateral of right and left leg
Gasmed_r_l	Gastrocnemius medial of right and left leg
IK related additional time series	
Angle_R_L	Ankle angle [right – left]
Vel_R_L	Angular velocity of the ankle [right - left]

Table 5-2 Model performance under different sets of input time series. The performance is reported in terms of the value of the coefficient of determination and the normalized RMSE.

Input time series	R^2	NRMSE
EMG & IK	0.945 ± 0.014	6.94 ± 1.02
EMG, IK, additional EMG	0.96 ± 0.0145	7.01 ± 0.43
EMG, IK, additional EMG & IK	0.985 ± 0.086	6.38 ± 0.17

Table 5-3 Number of features before and after feature selection for different combinations of input time series.

Input time series	Number of input time series	Number of extracted features	Number of selected features
EMG & IK	10	7940	50
EMG, IK, additional EMG	20	15880	50
EMG, IK, additional EMG & IK	22	17468	50

5.3.2 Regression model performance for different number of features

One of the important factors in the model's performance is the number of selected features for prediction. Choosing the most important features confines the negative impact of redundant data and noise on the model performance. However, by reducing the number of features, it is likely to miss some potentially valuable information.

Considering the 22 input time series from physical and virtual sensors [182], each window of input time series was converted to an array of over 18,000 features by tsfresh in this study. Using all of these features for model training is not reasonable. It is essential to find the optimum number of features. Figure 5.3 shows the relation between the number of selected features and the model accuracy. Features were sorted from high to low impact rate, then a different number of features were selected from the top of the list to train the model. By increasing the number of features to 20, the model's accuracy increased rapidly (Figure 5.3). However, the model accuracy only slightly increased in the range of 20 up to 200 features.

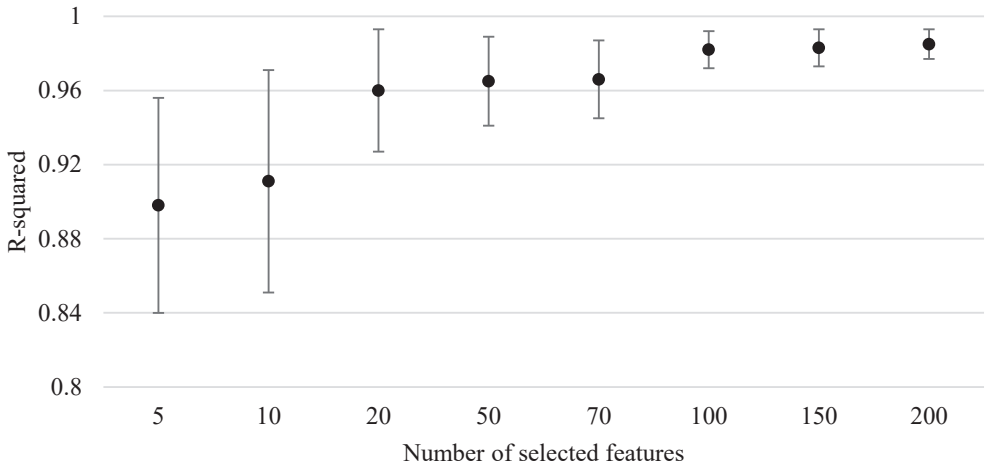


Figure 5.3 Model accuracy in terms of R^2 regarding different numbers of input features.

To keep the computational cost of the model low, the minimum number of features with an acceptable accuracy rate was selected to develop the model in this study. Based on the results in Figure 5.3, the number of selected features was fixed to 50 in this study.

5.3.3 Find a generic feature set based on the cross-validation approach.

A generic feature set means having a set of features applicable to all individuals. As we had collected data from 10 individuals, to implement the cross-validation, a set of top 50 features were selected based on data from nine participants and used to train and test the regression model for the left

out (unseen) individual as the "test participant." In this way, the test participant did not contribute to generated feature sets. This process was repeated over all ten participants to cover all possible permutations.

The impact of using a generic feature set on moment prediction is represented in Table 5-4. The "*Personalized feature set*" means the feature selection, model training, and testing were done based on the data of one participant. The "*Mismatched feature set*" is when a personalized feature set for one participant was used to train and test the regression model on the data of another participant. The "*Cross-validation feature set*" represents the average model accuracy for all permutations of the cross-validation approach for 10 participants.

Table 5-4 *Comparison of prediction accuracy for the same predictive model trained based on three different feature sets.*

Type of feature set	R^2	NRMSE
Personalized feature set	0.985 ± 0.009	6.38 ± 0.17
Mismatched feature set	0.92 ± 0.011	13.2 ± 1.08
Cross-validation feature set	0.98 ± 0.010	8.97 ± 0.86

Based on the results in Table 5-4, even though the test individual was not contributing in the feature set selection phase in the cross-validation-based approach, the model accuracy was similar to the personalized feature set. Extending the list of selected features based on the cross-validation to an unseen individual did not decrease the regression model performance.

5.3.4 Find the common features in the list of generic feature sets.

The results of the cross-validation approach showed that the selected feature set based on data of 9 participants performed better than the mismatched feature set in the prediction of ankle moment of an unseen individual. However, the selected feature sets based on cross-validation were not the same for all of the permutations. Figure 5.4 shows how many features were shared across different groups of individuals.

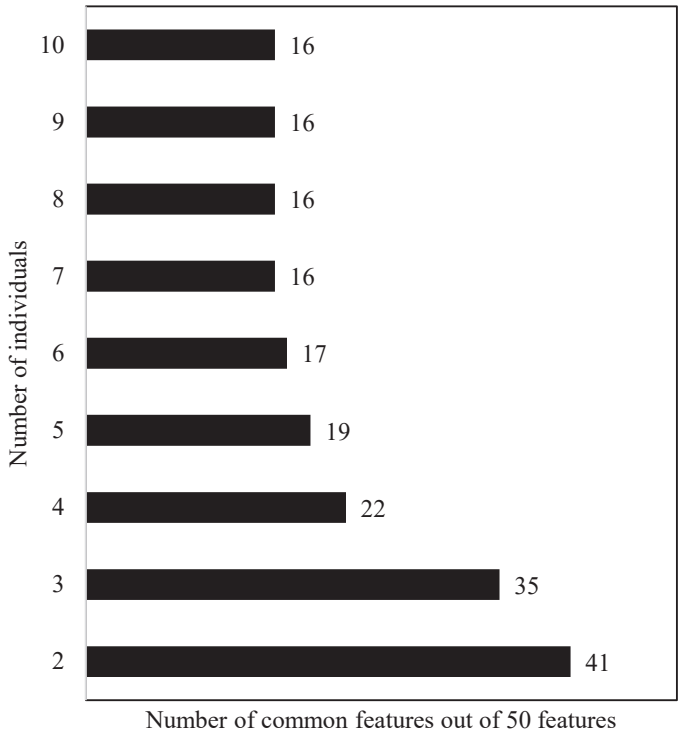


Figure 5.4 The number of common features among different numbers of individuals.

As illustrated in Figure 5.4, out of the selected 50 features, a total of 41 features were in common between two randomly selected participants. However, the number of common features decreased by increasing the number of individuals. The number of common features among seven individuals and beyond remained constant at 16. This means 16 features were repeated among seven individuals, and by increasing the number of individuals, these 16 features were still repeating. We can refer to these features as the generic feature set as they are consistently among the list of selected features. The list of generic features is represented in Table 5-5. The first column in Table 5-5 represents the time series name from which the feature was extracted. The column "TSFRESH algorithm" indicates which type of feature was extracted from each time series. The column "parameters" describes the parameters applied to extract each feature.

Table 5-5 List of final 16 features selected as the generic feature set.

Time series name	Description	Tsfresh algorithm	Parameters	Imp.
vel_L	Angular velocity of left ankle	energy_ratio_by_chunks	num_segments = 10, segment_focus = 9	0.46
gasmed_l	EMG of left leg gastroc nemius medial muscle	energy_ratio_by_chunks	num_segments = 10, segment_focus = 9	0.32
gasmed_r_l	Refer to table I	energy_ratio_by_chunks	num_segments = 10, segment_focus = 9	0.24
gasmed_l	EMG of left leg gastrocnemius medial muscle	index_mass_quantile	q = 0.9	0.11
angle_R	Angle of the right ankle	agg_linear_trend	attr = "slope", chunk_len = 50, f_agg = "min"	0.086
vel_R	Angular velocity of right ankle	fft_coefficient	Attr = "real", coeff = 1	0.078
angle_L	Angle of the left ankle	first_location_of_maximum		0.058
gasmed_l	EMG of left leg gastrocnemius medial muscle	index_mass_quantile	q = 0.8	0.055
vel_R_L	Refer to Table 5-1	time_reversal_asymmetry_statistic	lag = 2	0.039
angle_L	Angle of the left ankle	last_location_of_maximum		0.035
vel_L	Angular velocity of left ankle	energy_ratio_by_chunks	num_segments = 10, segment_focus = 3	0.029
vel_R	Angular velocity of right ankle	energy_ratio_by_chunks	num_segments = 10, segment_focus = 9	0.019
angle_L	Angle of the left ankle	change_quantiles	f_agg = "var", isabs = False, qh = 1.0, ql = 0.8	0.017
sol_r_l	Refer to Table 5-1	fft_coefficient	attr = "angle", coeff = 5	0.013
angle_R_L	Refer to Table 5-1	agg_linear_trend	attr = "slope", chunk_len = 50, f_agg = "min"	0.011
sol_r_l	Refer to Table 5-1	fft_coefficient	attr = "angle", coeff = 1	0.011

The last column represents the importance of each feature in prediction. Each of the features algorithms is described in Table 5-6.

Table 5-6 Description of each of the feature extraction functions of tsfresh, used in the generic feature set.

No.	Function name	Description
1	energy_ratio_by_chunks	Chunks the input time-series into 10 segments (num_segments=10) of approximately equal size and computes the sum of squared time-series values for each segment. This results in a list of 10 segment-specific "energy" measures. Return the energy ratio from the last segment (segment_focus=9) and the sum of squared time-series values for the complete time series.
2	index_mass_quantile	Given $x(t_i)$ as input time series, this function calculates the relative index i of time series x where 90% (q = 0.9) of the of x lies left of index i.
3	agg_linear_trend	Chunk the input time-series into segments of 50 measurements (chunk_len = 50), fit a linear regression to each segment. Return the minimum (f_agg = "min") of the slopes of the fitted linear regressions (attr = "slope").
4	fft_coefficient	Calculates the Fourier coefficients of the one-dimensional discrete Fast Fourier Transform and returns either the amplitude (attr=real) or phase (attr=angle) for a specific coefficient (coeff).
5	first_location_of_maximum	Given $x(t_i)$ as input time series, return the relative location of the first local maximum x.
6	time_reversal_asymmetry_statistic	Returns the time reversal asymmetry statistic [188].
7	last_location_of_maximum	Given $x(t_i)$ as input time series, this function returns the relative location of the last local maximum of x.
8	change_quantiles	Compute the conditional dynamics: Select values larger than the 80th percentile (ql=0.8) and smaller than the maximum (qh=1.0=). Subtract the consecutive values, do not compute the magnitude of the difference (isabs = False), and return the variance of these differences (f_agg = "var").

The performance of the regression model based on the generic set of 16 features is compared to a personalized set of 50 features in Table 5-7.

Table 5-7 Predictive model accuracy under personalized and generic features sets. The model went through the same train and test data set for both of these feature sets. The model accuracy is reported in terms of coefficient of determination (R^2) and the normalized RMSE.

Type of feature set	R^2	NRMSE
Personalized feature set	0.985 ± 0.0086	6.38 ± 0.17
Generic feature set	0.977 ± 0.0175	7.99 ± 0.66

By applying the generic feature set, the required training time for the regression model was reduced by about 70% (Table 5-8) in comparison to applying a personalized feature set for training. But it is not the only impact of using a generic feature set. Most importantly, the generic feature set saves

time during the feature engineering process. In the generic feature set method, as the list of important features is a known factor for an unseen individual, there is no need for comprehensive feature extraction from input time series, which means saving a significant time during the process of model preparation (about 27 mins based on Table 5-8). Also, the feature selection is shorter as instead of 50 features in the personalized feature set, only 16 features were extracted in the generic feature set approach. The timing for each phase is represented in Table 5-8 repeated ten times, and the average of consumed time is reported. The system used for timing in Table 5-8 had intel Core i7-7700 at 3.60 GHz CPU, 32 GB of RAM, and Windows 10 Enterprise – Version 1803. Python 3.6 was used to run the codes.

Table 5-8 Processing time comparison between model training based on the generic feature set versus the personalized feature set.

	Personalized feature set	**Generic feature set**
Comprehensive feature extraction	1620(s)	-
Feature selection	1040(s)	21.46(s)
Regression model training	4.53(s)	1.41(s)

A regression model trained on the generic features showed ~1 % reduction in the accuracy. However, the regression model had less computational cost than the trained model based on the personalized feature set because of decreased feature numbers.

A visualized comparison of the model performance with different feature sets is represented in Figure 5.5 and Figure 5.6. The average and standard deviation of ankle moment over six strides are illustrated in these figures.

Using the personalized feature set to train the predictive model (green dashed line in Figure 5.6) provided the highest accuracy as the predicted moment was close to the ID results (solid blue line in Figure 5.6). However, when the mismatched feature set was used for model training (dashed red line in Figure 5.5), the model accuracy in moment prediction degraded (solid line in Figure 5.5). However, using the cross-validation-based feature set (dashed blue line in Figure 5.5) increased the prediction accuracy due to having a more diverse data set during the feature selection phase. As the next step toward a generic feature set, a list of 16 features in common among all the permutations of the cross-validation approach was selected to train and test the model for an unseen individual, the results (red dashed line in Figure 5.6) didn't show any significant deterioration in prediction accuracy despite the limited number of features, compare to the situation in which train and test were done based on the personalized feature set.

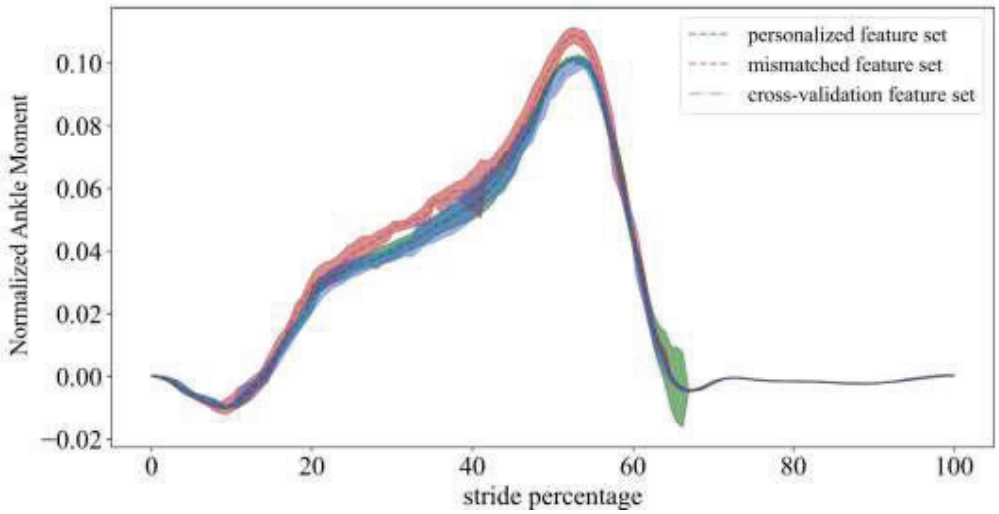


Figure 5.5 *The average and standard deviation of ankle moment over six strides for one individual during walking at a constant speed. Comparison between predicted values from a model trained on the personalized feature set, mismatched feature set, and cross-validation feature set.*

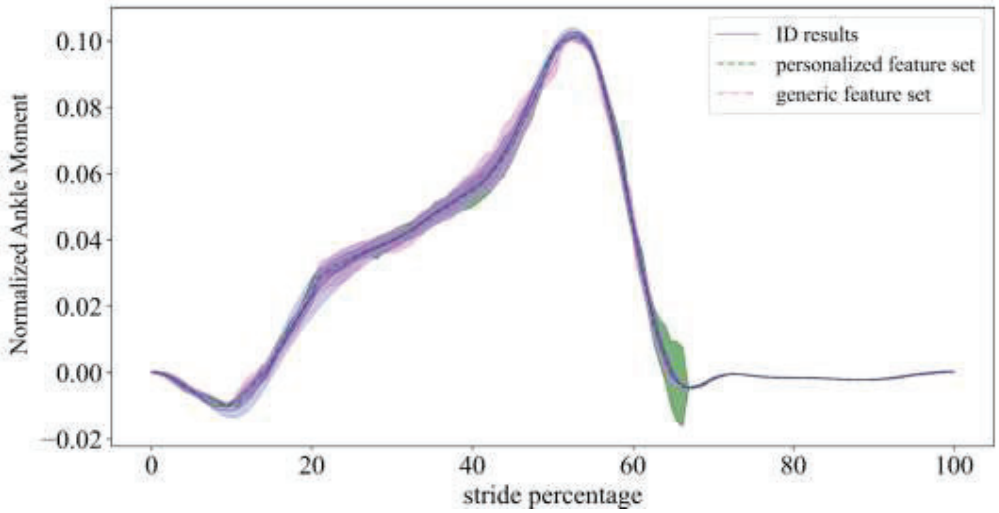


Figure 5.6 *The average and standard deviation of ankle moment over six strides for one individual during walking at a constant speed. Comparison between inverse dynamics results versus predicted values from a model trained on the personalized and generic feature sets.*

5.4 Discussion

The main expectation of a model for joint moment prediction is achieving high accuracy. Among all existing models, MNS models provide reasonable accuracy, but they are computationally

expensive and not practical in real-time applications [175]. As an alternative solution, machine learning-based models have been introduced to predict joint moments, especially when there is a nonlinear relation between input and output [176]. Minimizing the required time in the prediction process depends on various parameters, such as the computational power of the controller, predictive model structure, and the size of input data.

In this study, all of the muscles involved in ankle moment generation in the left and right leg (including 8 EMG channels) were initially used to implement the model. However, to minimize the computational cost, only 4 EMG channels were selected during the feature selection process to predict the ankle moment without compromise in the prediction accuracy. Minimizing the number of EMG channels helped save time during the signal preparation stage and eventually made the prediction process faster. Moreover, less physical equipment is required for signal acquisition, making the model implementation easier as a controller for assistive robots.

Apart from minimizing the required input time series for ankle moment prediction, the other outcome of this study was finding a generic set of features applicable to all individuals. Having a generic feature set makes model training faster as it is not required to select features for each individual separately.

A comprehensive investigation was conducted over the list of selected generic features to understand the physiological reasons behind selected features. As represented in Table 5-8, five features out of 16 came from virtual sensors. These virtual sensors-related features illustrate that the mathematical combination of data from physical sensors provides valuable information for model training. Ten features were related to joint angle and angular velocity, which indicates EMGs alone are not enough to predict joint moment. It is not surprising to have kinematic-related features as predictors since muscle force is a function of activation, length, and velocity [19].

Another interesting finding is that from all of the muscles involved in moment generation in the ankle joint, only recorded EMGs from gastrocnemius lateral and soleus (from the left and right leg) are among the time series used in the list of generic features, which means only 4 EMG channels is enough to provide information for ankle moment prediction. There are two explanations for it: First, the human error in electrode attachment to other muscles degraded the quality of collected signals from them in our experiment; as a result, extracted features from those channels were not in the list of informative features. The second reason backs to the presence of kinematics, which provided more informative features than what was potentially achievable from EMGs of other muscles.

Compared to EMGs, the kinematics-related inputs were less susceptible to human error, making them more reliable and repeatable. As a result, the corresponding features related to the kinematics dominated some muscles' features. The left leg contributed to 13 features in the list of generic features that show even though our model predicted the right ankle moment, both left and right leg EMGs and kinematics were necessary.

The list of algorithms used for feature extraction in the set of generic features is represented in Table 5-5. Part of these algorithms are related to the magnitude of input time series (algorithm numbers 1, 2, 5, and 7), some of them relate to the rate of variation in signals (algorithm numbers 3, 6, and 8). One of them is related to the frequency of input signals (algorithm number 4). Magnitude and rate of variation of IK-related time series were used in the list of the final feature list, which physiologically makes sense as the level of the generated moment in the ankle joint corresponding to the walking speed. The variation-related features from EMGs were not part of the generic features as these signals are noisy, and the variation rate may not provide clear information. The frequency-related features of both EMGs and IK-related signals contained valuable information, and because of that, they are part of the generic feature set.

It is important to note that the generic feature set implemented in this study tested for ankle moment prediction during walking at a constant speed. Implementing the same feature set to predict ankle moment in different walking speeds will not show the same accuracy because the ankle joint's generated moment corresponds to the walking speed. In that case, the training set used in the feature selection phase needs to cover a wide range of speeds.

Moreover, the continuous prediction of ankle moment was achieved at a capture rate of 100 Hz. This frequency matches the sampling rate of the input time series. Although 100 Hz is sufficient for ankle moment prediction during walking at a constant speed (which has fundamental frequencies of 8Hz or less), a higher sampling rate might be required during faster movements such as jumping and running to capture higher frequency components in joint moments.

5.5 Conclusion

This study introduced a generic set of features extracted from EMGs and ankle kinematics to predict ankle moment during walking at a constant speed. Due to selecting only a limited number of features to make the feature set, model training will not be computationally expensive as when all of the features were used for training. The generic model performance was tested on unseen individuals whose data was not used in the feature selection phase to evaluate the model performance. The results showed that the model prediction accuracy when the training was done on the generic feature set was as good as when the model trained on the personalized feature set.

This is a promising result showing the generic feature set is working for unseen individuals. The predictive model still needs to get calibrated for each individual; however, it is not required to find the feature set for each person separately. In this way, model training will be easier for new individuals.

5.6 Acknowledgements

We would like to thank the AUT Millennium motion capture laboratory for the use of their facility and the NZ Centre of Medical Research Technologies (MedTech CoRE) for funding.

6 Ankle moment prediction at varying walking speeds using EMG and machine learning

Homayoon Zarshenas, Andreas W. Kempa-Liehr, Bryan P. Ruddy, Thor F. Besier

A challenge for exoskeleton control is providing assistive force to augment muscle forces to complete a task. Electromyography (EMG) signals are ideally suited to predict motion intention related to activation and force production of muscles. Here we present a nonlinear regression model that uses EMG and ankle kinematics as input to predict ankle joint moment 30ms ahead of the current time step. To make model training efficient, a comprehensive feature extraction and selection process was conducted through a python-based package (tsfresh) to find the most valuable information for model training.

Our model is sensitive to the training dataset, similar to other data-driven methods. We tried different training datasets to address this limitation and make the model compatible with a range of walking speeds. We found that training the model on data related to acceleration from 0.5 m/s to 2.5 m/s enables the model to predict ankle moment accurately for all of the walking speeds within the range of acceleration (R^2 between 0.86 and 0.96 for ankle moment prediction during walking at 1 m/s, 1.5 m/s, 2 m/s, and 2.5 m/s). The train and test of the model were repeated for five individuals to check the consistency of the results. Also, to find the best data-driven model to predict ankle moment based on inputs in our study, a comparison across three predictive models

was conducted. We found the nonlinear regression (random forest) was compatible with input diversity and had a higher rate of accuracy in comparison to a neural network or linear regression.

Index Terms— human-robot interaction, EMG, moment prediction, regression model, tsfresh.

6.1 Introduction

Neurological disorders result in muscle weakness, coordination problems, and loss of sensation, negatively affecting activities of daily living and locomotion [189]. Restoring normal locomotion is essential for patients living with neurological disorders [190]. Conventional rehabilitation for people recovering from spinal cord injury or stroke includes hands-on physiotherapy such as joint mobilization, joint manipulation, muscle stretching, and soft tissue massaging [191]. This work is labour-intensive for therapists [192], and differences in the range of motion and level of muscle force generation create challenges for therapists to deliver a personalized rehabilitation plan.

Powered exoskeletons offer an alternative solution to assist therapists and deliver consistent, targeted rehabilitation to patients with neurological disorders [193]. First-generation assistive robots provided forces or joint moments to complete a specific movement task without considering the force generated from the user [194]. However, active participation from the patient is necessary to facilitate neuromuscular recovery [195, 196]. Despite improvements in mechanical and electrical aspects of assistive robots in the past decade, an intuitive interaction to provide maximum neurological recovery for patients is lacking.

Understanding the user's motion intention is critical to achieving intuitive control and adjusting the required level of assistance [176]. Controlling assistive robots based on the user's motion intention requires accurate and robust decoding of motor function [197-200]. It is important to predict motion intention far enough in the future to compensate for the associated delay to the decision-making process in the controller and movement generation in the actuators of an assistive robot. Electromyography (EMG) provides a valuable signal to achieve this, as it conveys information of muscle activation that precedes the corresponding joint moment by 20 ms to 100 ms [17]. However, relating EMG signals to muscle forces and joint moments is challenging and requires a mapping model.

Neuromusculoskeletal (NMS) models based on Hill-type muscle models of muscle mechanics have been used extensively, using EMG and joint kinematics as input [19, 175, 201]. However, musculoskeletal modelling is complex, requiring parameter estimation and comprehensive data to represent muscle activation, muscle-tendon dynamics and geometry, motion data, reaction forces, and joint moments. Calibrating these models to an individual requires considerable expertise, is time-consuming, and computationally expensive.

Machine learning has also been used to correlate EMG signals to joint moments [102]. These include adaptive neuro-fuzzy models [202, 203] to link EMG to movements; fuzzy logic models to calculate joint force based on EMG [204], convolutional neural networks to control assistive robots

based on EMG [205]; and recurrent neural networks for limb motion estimation [155]. Machine learning models have an acceptable performance when the training and testing dataset is from the same movement (e.g., predicting joint moments during a specified walking speed). With respect to walking gait, it remains to be seen whether a generalized machine learning approach can be used to predict joint moments across a range of walking speeds.

This study developed a regression model to predict ankle joint moment across four walking speeds using minimal data for model training. The ankle was selected as it produces over 40% of the total leg moment required to walk on level ground at different speeds [206]. Different methods for regression model development were investigated, including; recurrent neural network, linear regression, and a nonlinear random forest regression model.

The data-driven models are sensitive to the training dataset, so they work well in motion prediction for the trained movements. However, a slight change in movement, such as speed variation, can deteriorate the accuracy of the model [207]. Training a data-driven model on all possible activities is infeasible as it would require a very large training dataset [88], and the process of training would take too long. Here, we tested different combinations of training and testing datasets to find the minimum dataset required for model training to enable good prediction of ankle joint moment across a range of four walking speeds.

6.2 Method

6.2.1 Participants

Five volunteers participated in this study (three males and two females, age: 27±2 yr, height: 171.6±9.5 cm, mass: 72.2±8.5 kg). Participants had no history of walking-related disorders or no record of lower limb surgery or pain that affected their walking abilities. All participants were familiar with treadmill walking and gave their written, informed consent prior to participation. The study was approved by the University of Auckland Human Participants Ethics Committee (approval 022112).

6.2.2 Experimental protocol

Experimental data were collected from participants walking at various speeds on a force-instrumented split-belt treadmill (Bertec, OH). Ground reaction forces were recorded at 100 Hz via embedded force plates in the treadmill. To record the body movements during the experiment, reflective markers were attached to the participant in accordance with the UWA marker set [156]. The trajectories of these markers were recorded at 100 Hz using a VICON 11-camera motion capture system (VICON, Oxford UK). Wireless surface EMG electrodes (Delsys Trigno, Natick

USA) were placed above the medial and lateral gastrocnemius, soleus, and tibialis anterior muscles of the left and right leg following SENIAM guidelines [175]. EMG data were collected at 1000 Hz. Maximum voluntary contractions (MVC) of the muscles of interest were recorded before the data collection session and were used in the data processing phase to normalize the EMG signals.

Participants performed a familiarization trial of two to four minutes duration prior to the experiment, walking at a self-selected speed until they felt comfortable. They then completed level-walking trials of 30 seconds duration at constant speeds of 1 m/s, 1.5 m/s, 2 m/s, and 2.5 m/s. Finally, they completed a trial that started at rest while the speed increased to 2.5 m/s, at a 0.02 m/s^2 acceleration rate.

6.2.3 Signal processing

EMG signals were filtered using a zero-lag Butterworth band-pass filter (20 Hz - 500 Hz), full-wave rectified and filtered using a zero-lag low-pass Butterworth filter (6 Hz). The filtered EMG for each muscle was normalized to the maximum value of the MVC trial for each corresponding muscle. The normalized EMG data were then down-sampled to 100 Hz to match the kinetic data. Ground reaction forces were low-pass filtered at 6 Hz to be dynamically consistent with the kinematic data.

6.2.4 Surrogate model development

First, we defined the input and desired output for the model based on the experimental data. The desired output for the model was the net ankle joint moment, calculated from inverse dynamics (see below). The model was required to predict the ankle moment at each time step based on the recorded data, including EMGs and ankle joint kinematics, over one second period. The model development is described in the following five steps:

6.2.4.1 Desired output calculation

Ankle joint moment was calculated using an inverse dynamics analysis in OpenSim 3.3 (Stanford, CA). A generic musculoskeletal model of the body (gait 2354) was scaled to each individual's anthropometry using marker trajectories recorded in a static pose. Scaling was performed using the MAP Client scale tool, which relies on the lower limb bones [160]. Marker trajectories were used in inverse kinematics to calculate lower limb joints angle and then angles filtered using a Butterworth low-pass filter at 6 Hz. Inverse dynamics were then performed to calculate the ankle joint moment during walking based on inverse kinematics results and filtered ground reaction forces. The ankle joint moment was then normalized to the participants' height × weight.

6.2.4.2 *Surrogate model input data*

To predict the normalized ankle joint moment at each time step, the surrogate model received a one-second window of input time series data (Figure 6.1). The input time series for this model included a combination of 'primary' and 'derived' data. Primary data refer to both limbs' normalized EMG recordings and ankle joint kinematics. Derived time series refers to mathematical combinations of the preliminary data. Our previous study showed that combining EMG signals from antagonistic muscles and contra-lateral limbs provides additional features for prediction [185]. This resulted in five primary time series and 12 derived time series inputs (Table 6-1).

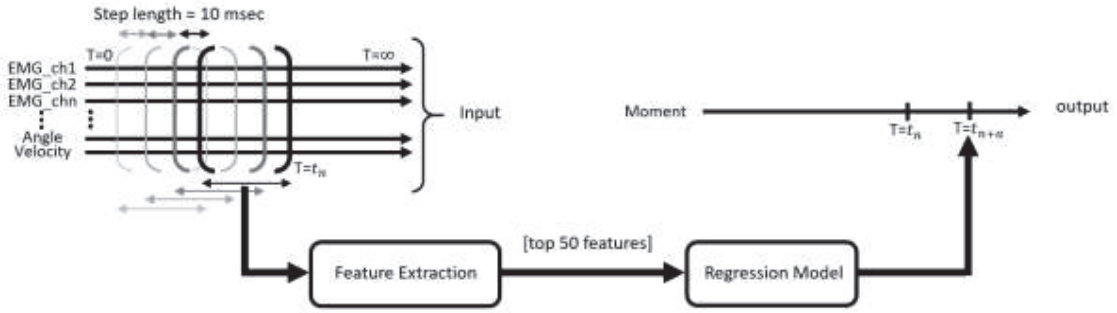


Figure 6.1 *The schematic view of the process of the ankle moment prediction based on a window of input time series.*

6.2.4.3 *Feature extraction*

Each window of the input time series (n=17) went through a feature extraction process. A comprehensive list of features extracted from each of the time series, including time domain, frequency domain, and time-frequency domain features via a python-based toolbox named "tsfresh" (Time Series Feature extraction based on scalable hypothesis tests) [151]. By going through the feature extraction phase, a window of one second of each input time series was replaced by 749 different features. The high dimensionality of input features makes the prediction process slow and impractical, especially for real-time applications [152]. Therefore, instead of using the extracted features from all time-series as input for model training, extracted features were sorted based on their importance for prediction (using a corresponding p-value). Only the most important features were selected for the training of the prediction model [208]. Feature selection was also performed using tsfresh [151], which sorted the features based on their statistical significance to predict the outcome variable. Selecting the number of features for training was a tradeoff between model accuracy and computational cost. Here, we selected the top 50 features based on our previous findings, in which model accuracy beyond 50 features was not significantly improved [185].

Table 6-1 List of primary and derived time-series used as input for the data-driven model training

Time series	Description
Angle_R / L	the right (R) and left (L) ankle angle
Sol_R / L	EMG of the right (R) and left (L) soleus
Tib_R / L	EMG of the right (R) and left (L) tibialis anterior
Gasmed_R /L	EMG of the right (R) and left (L) gastrocnemius medial
Gaslat_R / L	EMG of the right (R) and left (L) gastrocnemius lateral
Angle_R_L	Difference between right and left ankle angle
Vel_R_L	Difference between right and left ankle angular velocity
Sol_tib_r	EMG difference between tibialis anterior and soleus muscles of the right leg
Gasmed_tib_r	EMG difference between gastrocnemius medial and tibialis anterior of the right leg
Gaslat_tib_r	EMG difference between gastrocnemius lateral and tibialis anterior of the right leg
Sol_tib_l	EMG difference between soleus and tibialis anterior of the left leg
Gasmed_r_l	difference between EMG of gastrocnemius medial of right and left leg
Gaslat_tib_l	Difference between EMG of gastrocnemius lateral and tibialis anterior of the left leg
Sol_r_l	Difference between EMG of soleus of right and left leg
Tib_r_l	Difference between EMG of tibialis anterior of right and left leg
Gaslat_r_l	Difference between EMG of gastrocnemius lateral of right and left leg
Gasmed_r_l	Difference between EMG of gastrocnemius medial of right and left leg

6.2.4.4 Prediction model

We explored three data-driven prediction models to correlate our feature set with the corresponding desired ankle joint moment:

6.2.4.4.1 Recurrent Neural Network

Recurrent neural networks (RNNs) have been used previously for joint moment prediction and estimation of movement trajectories with reliable performance [155]. The RNN model we implemented in this study included three hidden layers; the first hidden layer was defined as a Long Short-Term Memory (LSTM) layer with 100 nodes. The other two layers were defined as fully connected layers of 50 nodes with a sigmoid activation function. The number of layers and nodes on each layer was determined through a trial and error process toward achieving the highest accuracy in model prediction.

The model was compiled with the Adam optimizer, a stochastic gradient descent-based method to minimize loss [209]. The loss function in our study was defined based on the difference between the predicted ankle joint moment and the actual value, which was minimized during the training

process by adjusting the model parameters (weights). A python library (Keras) was used to implement the model [210].

6.2.4.4.2 *Linear regression*

Linear regression considers a linear relationship between the inputs and outputs. In this study, a vector of 50 features (generated through the feature extraction stage) represented the required information at each data point for the regression model. The model was defined as:

$$\boldsymbol{y_t = \beta_0 + \beta_1 x_{1.t} + \beta_2 x_{2.t} + \cdots + \beta_k x_{k.t}} \quad (1)$$

in which $\beta_{1\ldots k}$ $(1 \leq k \leq 50)$ was the impact factor of each feature $(x_{k.t})$ at each time step (t) and β_0 was the linear offset between input and output. The values for β_0 to β_k were determined in a training process by minimizing the cost function defined based on the difference between the predicted value and desired output. The cost function in this study was defined as:

$$\sum_{t=1}^{T} \boldsymbol{\varepsilon_t^2} = \sum_{t=1}^{T} (\boldsymbol{Y_t - \beta_0 - \beta_1 x_{1.t} - \beta_2 x_{2.t} - \cdots - \beta_k x_{k.t}})^2 \quad (2)$$

in which ε_t^2 is the error function and Y_t is the desired value at each time step. A python-based package (sklearn) was used to develop the linear model [211]. The desired linear model comprised 51 parameters (50 impact factors for features plus one parameter as the offset) defined during the training phase.

6.2.4.4.3 *Nonlinear regression*

Random forest is a popular nonlinear regression model applied for time series prediction [212]. It comprises multiple independent decision-making units (trees) and returns the mean of prediction from the individual trees as the final decision (Figure 6.2). Random forest performance depends on the value of multiple hyperparameters. Among all of the parameters, the number of decision trees and the number of features used by each decision tree as inputs are the most important ones, considered 100 and 50 respectively in this study. The random forest model in this study was developed in python using the "ensemble" library from the "sklearn" toolbox.

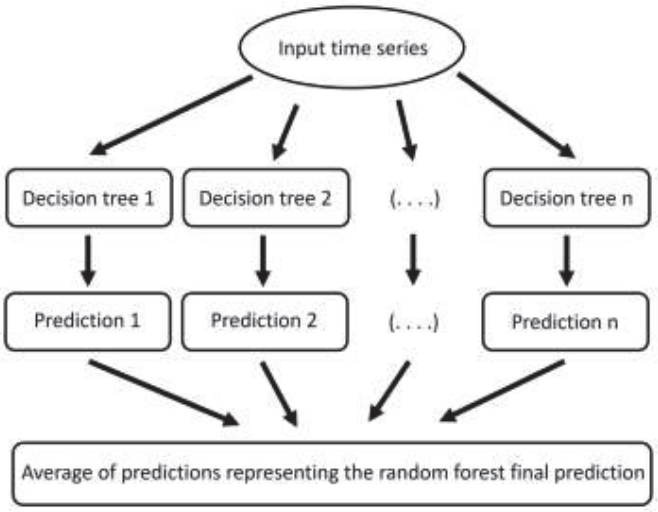


Figure 6.2 Decision-making process in the random forest regression model based on the input information.

6.2.4.5 Train and test datasets

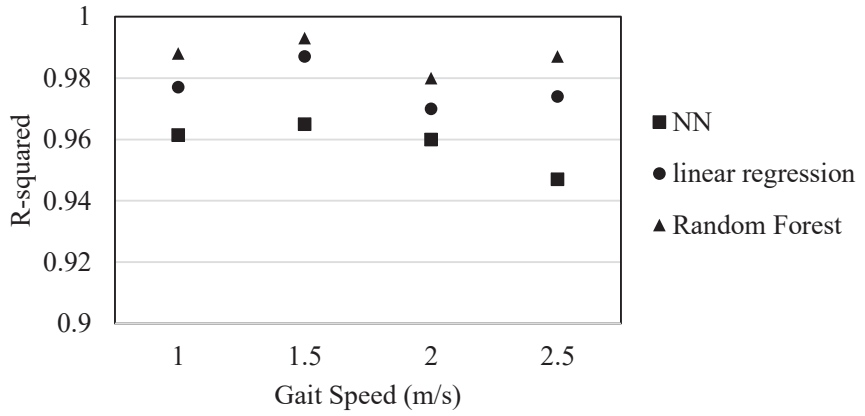
The performance of each of the three data-driven models was evaluated under the different train and test scenarios, including:

- **Train on one walking speed, test on different walking speeds.**
  In this case, the predictive model was trained on data from one walking speed for each individual and tested across walking at different speeds for the same individual. This approach was repeated for the four different walking speeds.
- **Train on acceleration walking trial, test on different walking speeds**
  Walking data obtained from the accelerated walking trial (from 0.5 m/s to 2.5 m/s) was used as training data. The model was tested using data from the four constant walking speed trials (1 m/s, 1.5 m/s, 2 m/s, and 2.5 m/s, respectively).
- **Train on the combination of slow and fast walking speed, test on other walking speeds.**
  In the third scenario, we combined the slowest (1 m/s) and fastest (2.5 m/s) walking trials for model training and predicted ankle joint moment for the 1.5 m/s and 2 m/s walking trials.

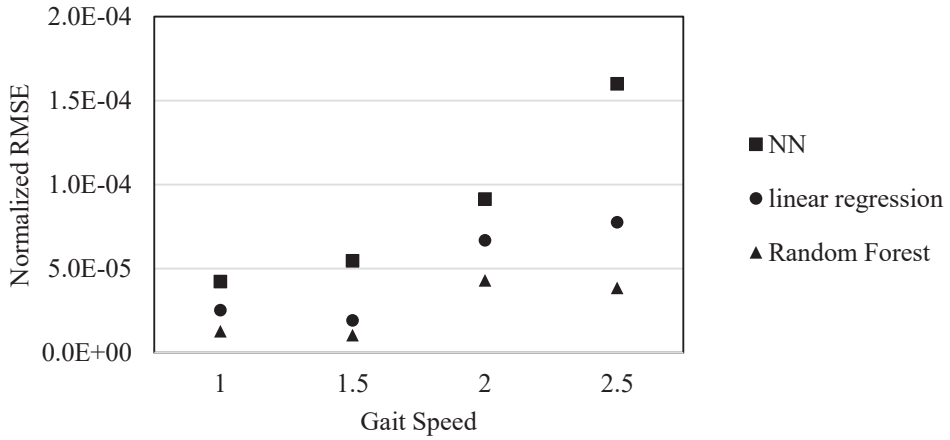
The model performance for each train-test dataset combination was reported based on the root mean square of error (RMSE) between normalized ankle moment and predicted one. Also, the coefficient of determination (R^2) was calculated based on the model output and actual ankle moment to assess how well the model could track the variation in the pattern of ankle moment.

6.3 Results

A comparison among three selected models in this study was performed under two train-test conditions to find the best data-driven model for ankle moment prediction. In the first scenario, the train and test were executed on the same walking speed data and repeated for 1 m/s, 1.5 m/s, 2 m/s, and 2.5 m/s (Figure 6.3). In the second approach, the models were trained on data from 1 m/s walking speeds. However, they were tested on data from multiple walking speeds (1.5 m/s, 2 m/s, and 2.5 m/s) (Figure 6.4). The values in these figures are averaged over five participants.

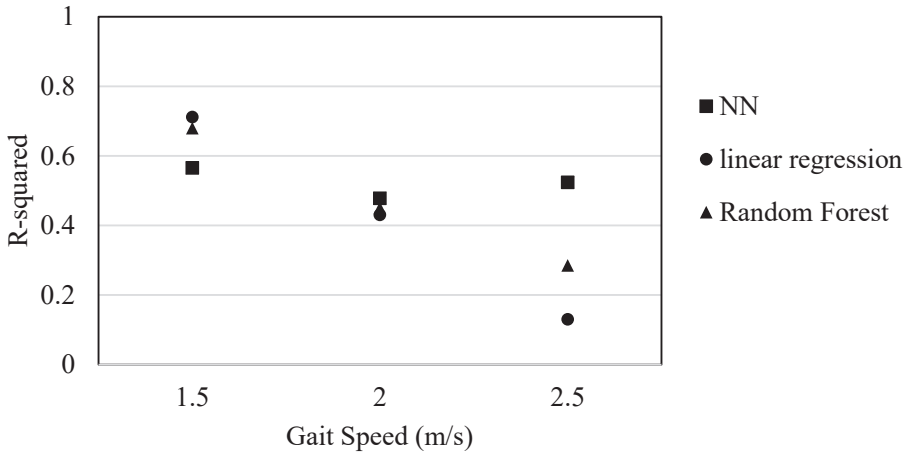


(a) The value of R^2

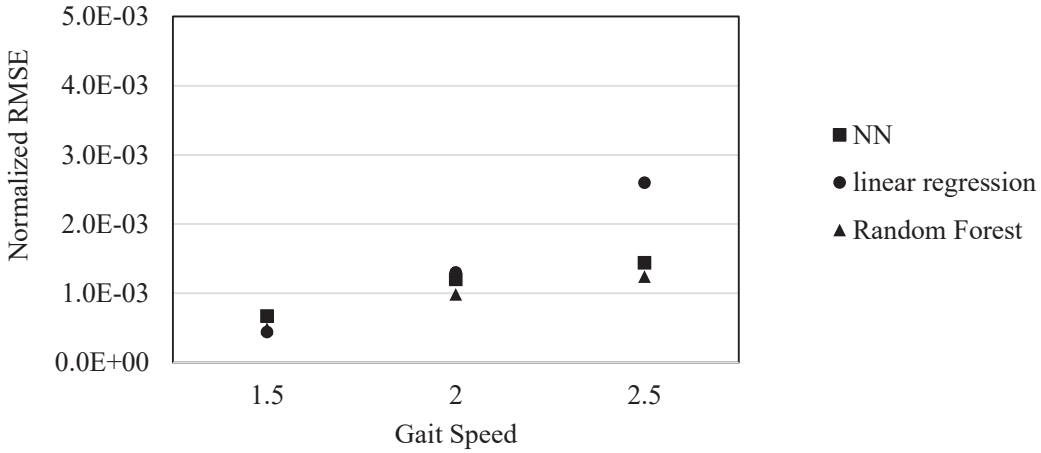


(b) The value of RMSE

Figure 6.3 *Comparison among three predictive models in terms of R^2 and RMSE. The train and test have been done at the same speed, and the process of train and test repeated for four different walking speeds.*



(a) The value of R^2

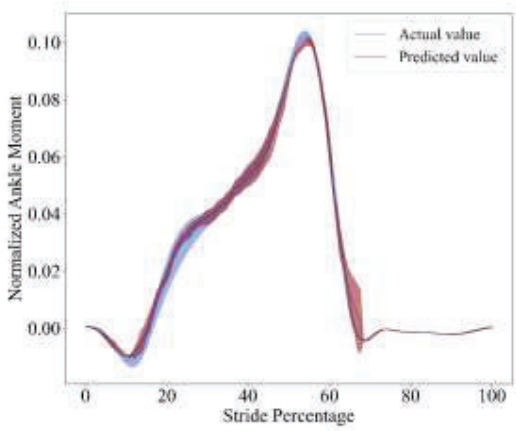


(b) The value of RMSE

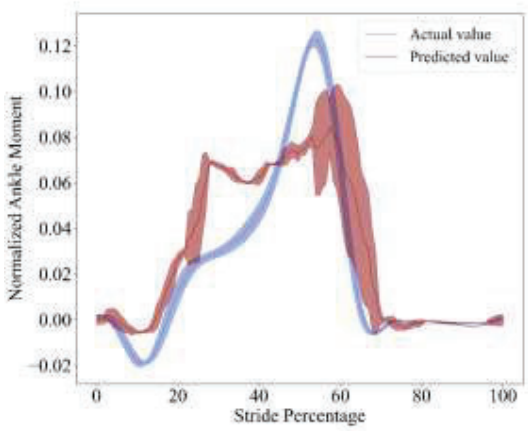
Figure 6.4 Comparison among three predictive models in terms of R^2 and RMSE for ankle moment prediction. Models were trained on data from 1 m/s speed and tested on three other walking speeds (1.5 m/s, 2 m/s, and 2.5 m/s).

Random forest had the minimum RMSE and highest R^2 compared to two other models in the first scenario. In the second scenario, the model accuracy decreased compared to the first scenario because of the speed difference between the training and test datasets.

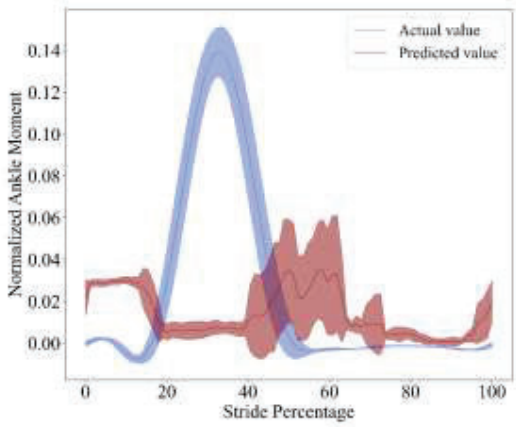
The random forest had the best performance among all of the models. It was selected for the second part of the study, which focused on finding the minimum dataset to train a compatible model with different walking speeds. During this process, an investigation on the impact of varying training datasets on model accuracy for ankle moment prediction during walking at different speeds has been conducted. Figure 6.5 shows the random forest performance under each of the train-test combinations explained in the method section.



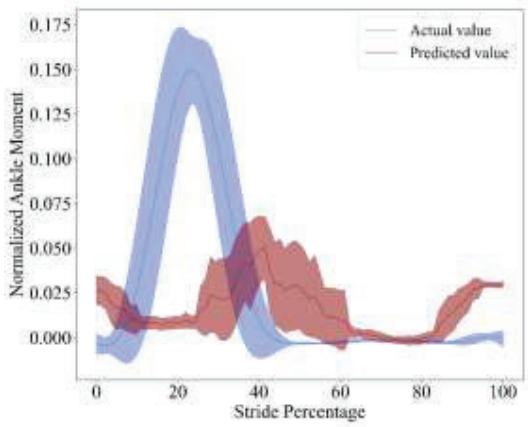
a.1) Test on 1 m/s



a.2) Test on 1.5 m/s

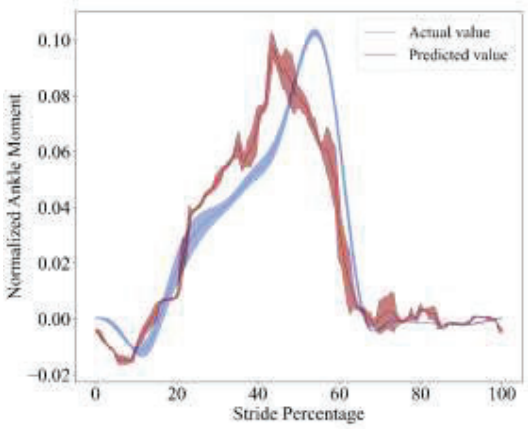


a.3) Test on 2 m/s

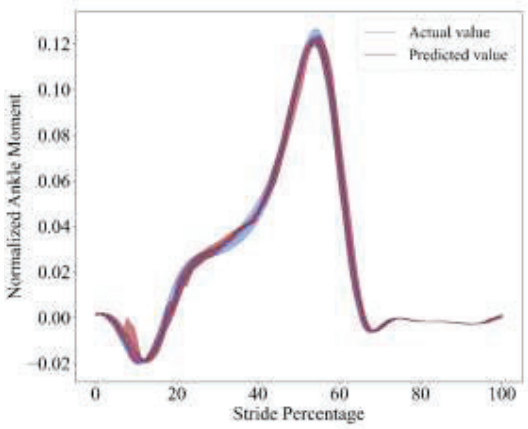


a.4) Test on 2.5 m/s

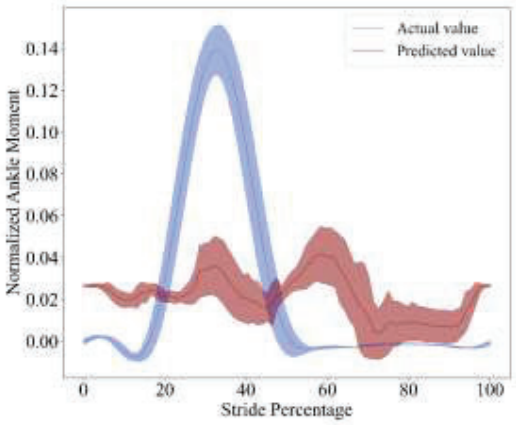
(a) Train on 1 m/s



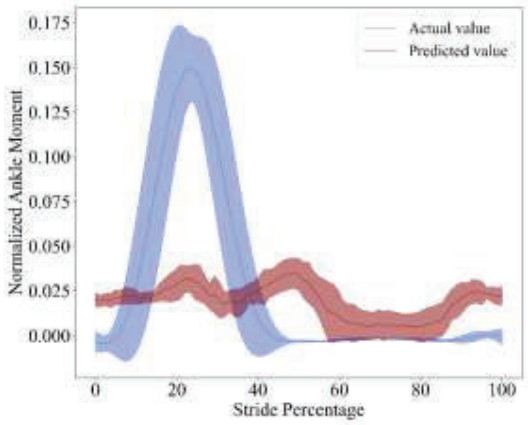
b.1) Test on 1 m/s



b.2) Test on 1.5 m/s

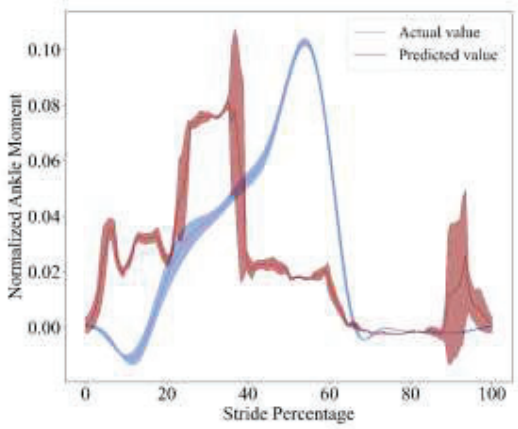


b.3) Test on 2 m/s

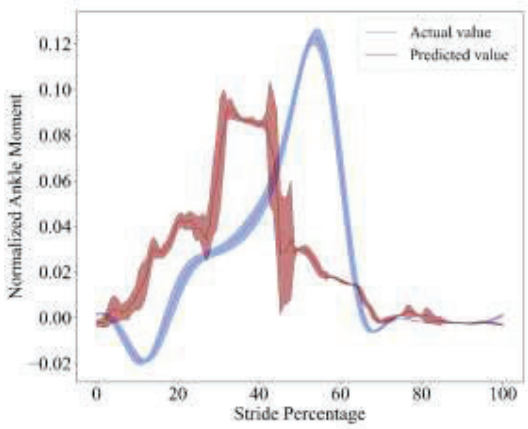


b.4) Test on 2.5 m/s

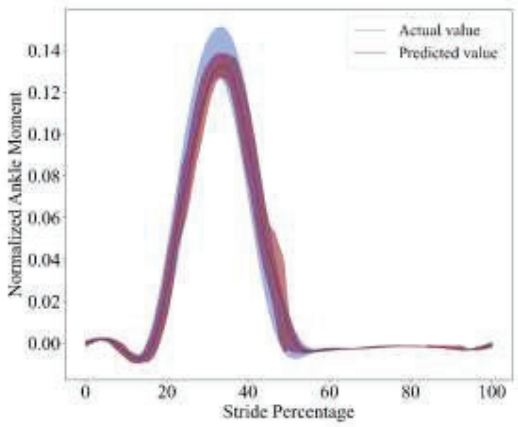
(b) Train on 1.5 m/s



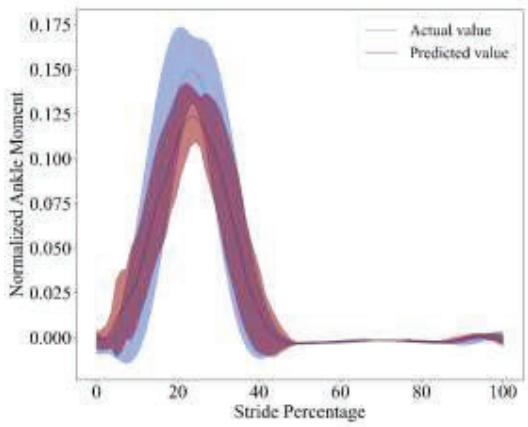
c.1) Test on 1 m/s



c.2) Test on 1.5 m/s

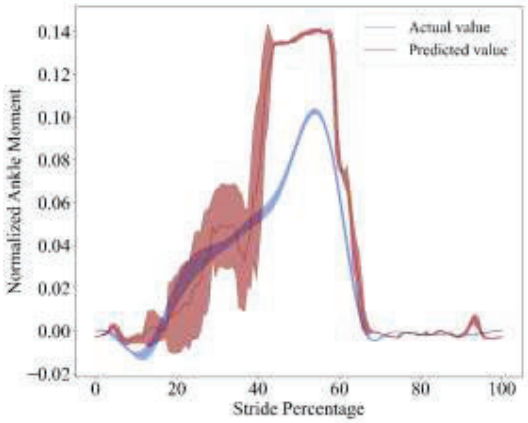


c.3) Test on 2 m/s

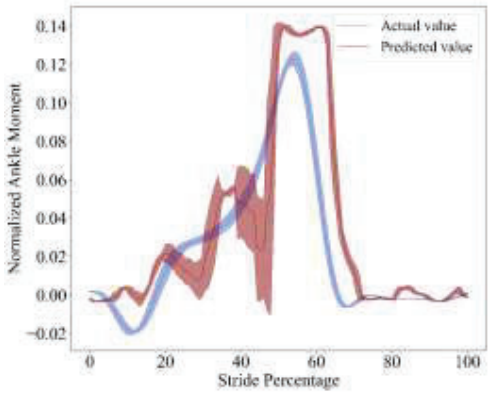


c.4) Test on 2.5 m/s

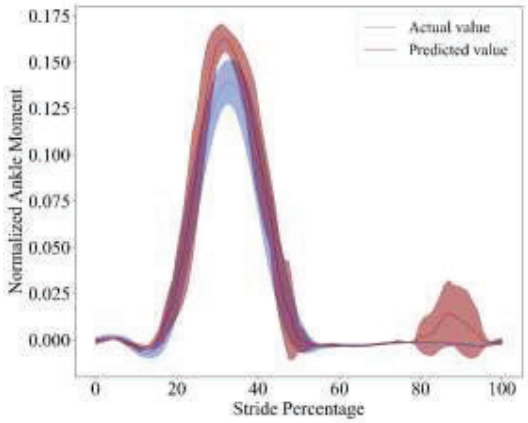
(c) Train on 2 m/s



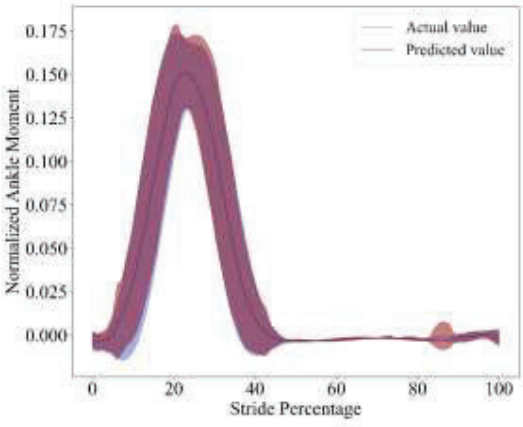
d.1) Test on 1 m/s



d.2) Test on 1.5 m/s

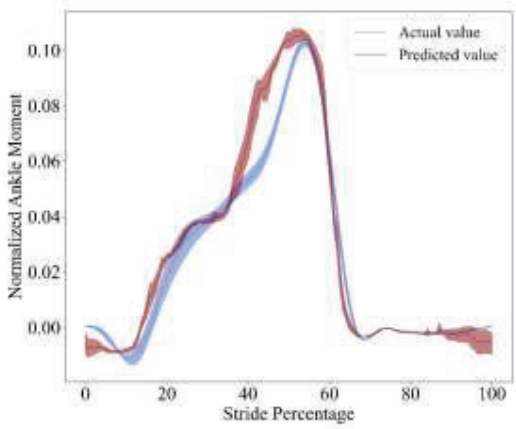


d.3) Test on 2 m/s

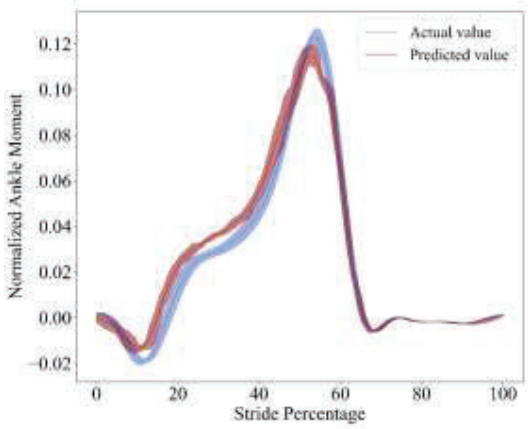


d.4) Test on 2.5 m/s

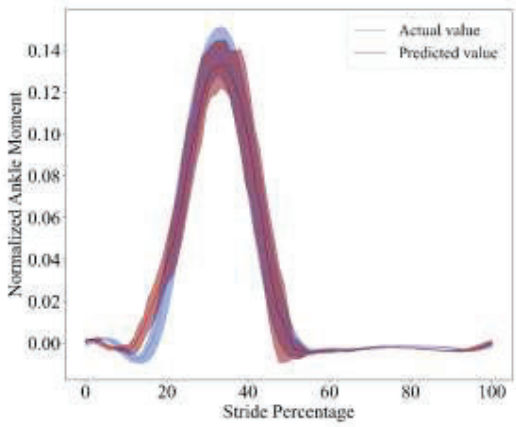
(d) Train on 2.5 m/s



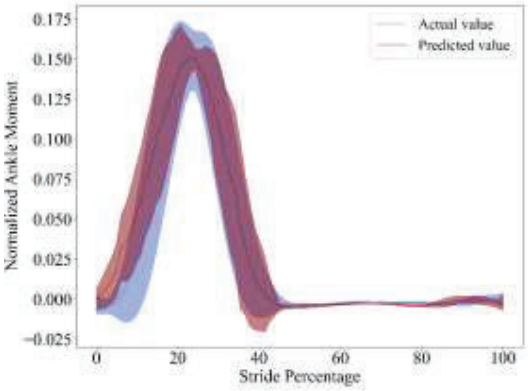
e.1) Test on 1 m/s



e.2) Test on 1.5 m/s

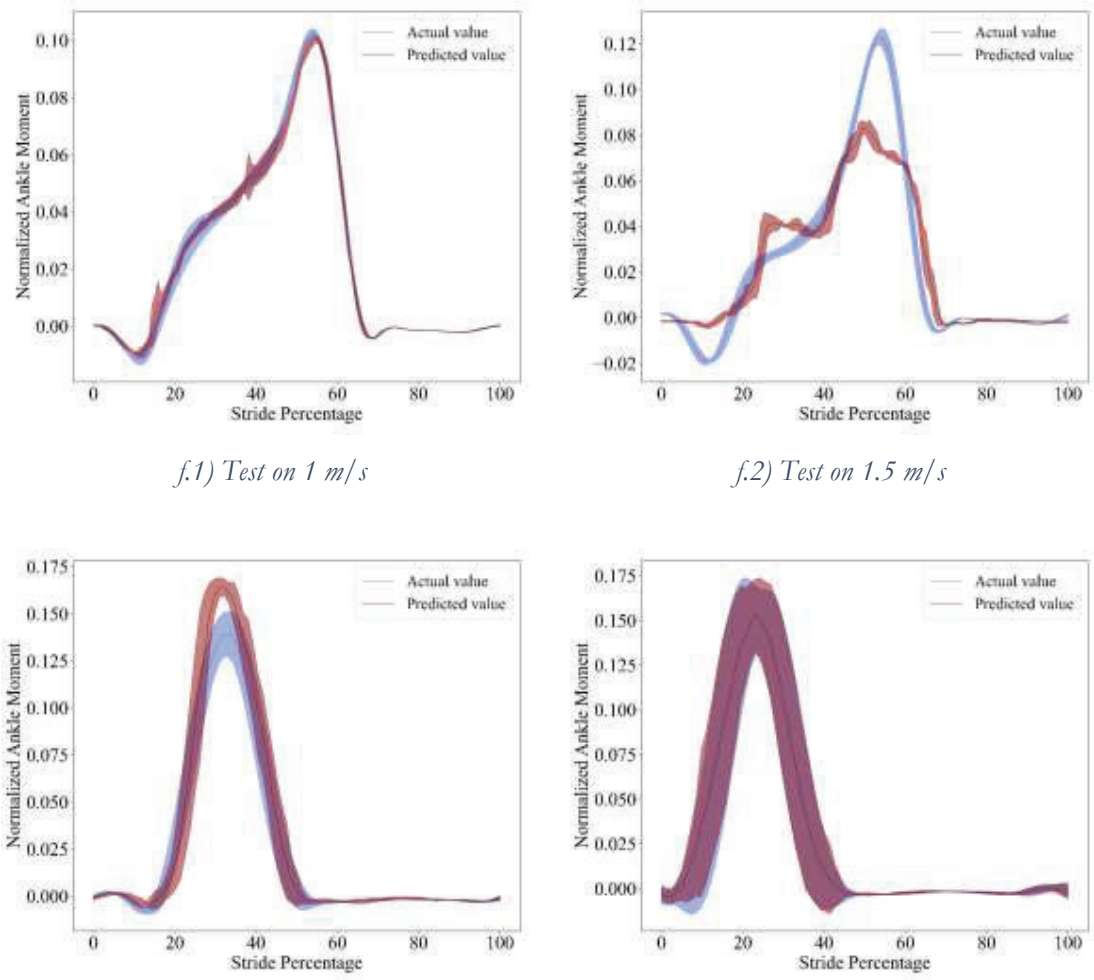


e.3) Test on 2 m/s



e.4) Test on 2.5 m/s

(e) Train on acceleration



f.1) Test on 1 m/s *f.2) Test on 1.5 m/s*

f.3) Test on 2 m/s *f.4) Test on 2.5 m/s*

(f) Train on 1 m/s and 2.5 m/s

Figure 6.5 *Comparison among different train and test combinations for the same predictive model. Each row represents the model trained on a specific training dataset. Each column shows the model performance for ankle moment prediction 30 ms ahead of the current time step during walking at different speeds.*

Each row in Figure 6.5 represents the speed used for model training, and each column is related to the testing speed. Each graph in Figure 6.5 compares the average and standard deviation of predicted ankle moment (red line) and actual ankle moment (blue line) over ten strides. Figure 6.5 represents the model performance only for one individual; the experiment over all of the explained

combinations of train/test datasets was repeated for five participants to check the consistency of the results. The average of R^2 and RMSE for five individuals is represented in Table 6-2.

Table 6-2 *The average of (a) R^2 and (b) RMSE of ankle moment prediction for different train-test configurations for the random forest regression model, each configuration repeated for five individuals. Green blocks show more similarity between model output and the actual value of ankle moment. As the colour shifts toward red, the accuracy of the model decreases.*

(a) The value of R^2

		Test on			
		1 m/s	1.5 m/s	2 m/s	2.5 m/s
Train on	1 m/s	0.982	0.505	0.271	0.162
	1.5 m/s	0.510	0.988	0.504	0.369
	2 m/s	0.110	0.675	0.986	0.802
	2.5 m/s	0.230	0.470	0.853	0.980
	acceleration	0.871	0.961	0.892	0.864
	1 m/s & 2.5 m/s	0.976	0.631	0.783	0.981

(b) The value of RMSE

		Test on			
		1 m/s	1.5 m/s	2 m/s	2.5 m/s
Train on	1 m/s	0.002	0.090	0.191	0.299
	1.5 m/s	0.058	0.002	0.118	0.196
	2 m/s	0.116	0.069	0.004	0.024
	2.5 m/s	0.177	0.120	0.030	0.005
	acceleration	0.012	0.006	0.025	0.021
	1 m/s & 2.5 m/s	0.003	0.142	0.090	0.007

6.4 Discussion

This study developed three different data-driven models based on EMG and kinematics as input to predict ankle moment 30 ms ahead of the current time step. For all of these models, the results showed better agreement in ankle moment prediction when the train and test were from walking at the same speed. The model accuracy reported in terms of RMSE and R^2 in Figure 6.3 and Figure 6.4 illustrates that data-driven models are not capable of predicting the joint moment during activities other than those used for training, as about 40% drop in R^2 occurred when the random forest model trained on data from 1 m/s walking speed and tested on 1.5 m/s. This decrease happened by only 0.5 m/s speed difference between train and test dataset. By increasing the speed difference, the accuracy of the model decreased.

In [176], a similar behaviour for an artificial neural network (ANN) model for moment prediction based on EMG data was reported. It was concluded that NMS models had better performance than data-driven models (ANN in this case) in moment prediction when the training and testing were not related to the same movement [176]. In another example, a comparison between NMS models, linear models, and ANN models for wrist and elbow angle prediction based on EMG during arbitrary movement in 3D space showed that linear models and ANN were not capable of angle prediction for activities that were not trained on [207].

The data-driven models applied for motion prediction in the literature are mainly used to predict joint moments during isometric tasks or constant speed movements [102, 202-204]. In each model, the training and test datasets were based on the same type of movement [205]. For instance, a comparison between ANN and RF for knee angle prediction during walking at constant speed showed that the RF predicted knee angle based on EMG signal more accurately [213]. However, the data used for training and testing the model were from the same walking speed [213]. In another study, a cascade linear model was developed with promising upper limb angles prediction results during arbitrary arm movement on a horizontal surface. The reported results showed a high level of accuracy during a five-fold cross-validation process when four out of five segments of data in each trial were used for training, and the remaining segment was used for validation [88]. Even though the movement was arbitrary in this study, the model was trained for each activity before prediction by going through five-fold cross-validation.

Making a data-driven model for motion prediction based on EMGs compatible with a wide range of movements requires a large dataset covering a range of activities [88]. This study attempted to find the minimum required input dataset that makes the model compatible with four different walking speeds.

According to reported values for R^2 and RMSE in Table 6-2, when the training of the random forest model had been performed on data from constant speed walking, the model was able to predict ankle moment accurately for that speed. The model performance degraded by increasing the speed difference between train and test datasets (first four rows in Table 6-2a and b). Using the acceleration data for training improved the model compatibility with different speeds. The error between the predicted ankle moment and the actual values reduced (row five in table 2a and b). By training the model on a combination of data from minimum and maximum walking speeds, the model became compatible only with speeds used for training. It could not predict ankle moment for speeds in between (last row of table 2 a and b). Figure 6.5 showed that the trained model on low speed could not reproduce the peak of ankle moment pattern at faster speeds. The model trained on fast speed predicted higher values for the maximum ankle moment in lower speeds. In

contrast, in both cases, the model struggled to follow the pattern of moment variation because of the speed difference between the train and the test dataset. This shows data-driven models of walking gait are speed-sensitive, and by testing them on speeds other than what is used for training, they cannot reproduce the ankle moment.

Even using a combination of data from walking at the minimum and maximum speeds could not make the model compatible with speeds in between. Ankle moment prediction was not as accurate. The reason for this behaviour is the nonlinear relationship between the walking speed and input EMG signals; because of this nonlinear correlation, even though the model was trained on EMGs collected during walking at the maximum and the minimum speed, it was not able to interpolate the ankle moment for walking speeds in between.

As an alternative approach, we tried a continuous acceleration from 0.5 m/s to the maximum walking speed in our experiment (2.5 m/s) to train the RF model. This dataset was used for feature selection and regression model training. In this situation, the model could predict ankle moment during walking at different speeds, not as well as a model trained and tested on the same walking speed, but better than any model trained and tested at different speeds (Figure 6.5). Thus, we can conclude that using acceleration data for model training enhanced the model accuracy in ankle moment prediction in a range of walking speeds from slow walking to running.

By increasing the walking speed to 2.5 m/s, ankle moment prediction became more challenging as the pattern of ankle moment variation during a stride became less consistent. The reason for ankle moment diversity at 2.5 m/s is that this speed approaches the transition between jogging and running. In other words, to keep up with pace during the experiment, some participants preferred to take long strides in jogging fashion while others chose to take short strides and started running.

One limitation of our approach, when considering a real-time application, is the post-processing of EMG data. The zero-lag filters used here are not applicable for real-time signal processing, so the lag time associated with the filtering method should be considered in the prediction time for real-life applications.

Even though the result in our study showed the accuracy of the data-driven model improved by using acceleration data for training, it is worth noting that the model was tested on speeds that were in the range of the speed variation of the training dataset. In other words, these results did not show if the model can predict ankle moment prediction for speeds outside the range of speed variation in the training dataset.

The other consideration about training a model based on acceleration data is that, as the speed was constantly changing in the training dataset, there was no complete stride at any of the constant

speeds the model tested on. However, the model showed the potential to predict ankle moment at those constant speeds. The reason that explains this accurate performance is that the model was responsive to the speed variation. When the model was trained on acceleration data, the speed variation became indirectly one of the model's inputs as it caused changes in the level of EMG and kinematics, which were the direct inputs for the model. This means to get the model compatible with different walking speeds, it is not required to have a set of strides at each one of the speeds. Instead, it is needed to make the model compatible with the variation in speed, which occurred when using acceleration data for training. This brings another question to attention regarding the "ideal" acceleration profile for the training dataset. In this study, we trained the model using increments of 0.02 m/s^2. However, higher or lower acceleration might lead to changes in prediction accuracy, which remains a question for future work.

6.5 Conclusion

Predicting joint moments forward in time across a range of tasks remains a significant challenge for intuitive robotic rehabilitation. In this paper, we compared the performance of three regression models to predict ankle moment during four different walking speeds. A nonlinear random forest model showed the most promising results. The integration of EMG signals and ankle-related kinematics were used as input for these models to predict ankle moment 30 ms ahead of the current time step. As expected, the accuracy of all of the models was dependent upon the training dataset. To have a comprehensive predictive model that can accurately predict the ankle moment over a wide range of speeds, we propose using data from a walking trial that accelerates from a standstill to a maximum walking speed.

6.6 Acknowledgement

We would like to thank Kelly Sheerin and the AUT Millennium motion capture laboratory for the use of their facility and the NZ Centre of Medical Research Technologies (MedTech CoRE) for funding.

7 An adaptive regression model to forecast ankle motion based on EMG signals during walking at different speeds and inclines

Homayoon Zarshenas, Andreas W. Kempa-Liehr, Bryan P. Ruddy, Thor F. Besier

Predicting the user's motion intention accurately and reliably is critical in human-machine interface applications. In this study, an adaptive regression model was constructed using time series information from surface electromyography (EMG) signals and kinematics to forecast right ankle flexion/extension angle 30 ms in the future. Model performance was compared to a non-adaptive regression model to illustrate the predictive capability across different walking speeds and incline conditions. Three-dimensional lower limb motion dynamics was reconstructed using an optical motion capture system, and EMG signals were collected from the tibialis anterior, soleus, gastrocnemius lateral and medial of the left and right leg from 10 participants at four different walking speeds during incline, decline, and level walking on a treadmill plus acceleration from slow walking to jogging on a treadmill at different slopes. The model predicted ankle angle at 100 Hz frequency. At each time step, a one-second window of input time series leading up to the current data point was used to predict ankle angle 30 ms in the future. The input time series went through a feature engineering process to optimize the size of input information for the model. The model was constructed of polynomial linear regression to correlate the input data to ankle angle and closed-loop feedback control that adjusted the model parameters whenever the error between

predicted and actual ankle angle was greater than three degrees. The model was initially trained on one participant walking at 1 m/s with no incline and tested across three other walking speeds (1.5 m/s, 2 m/s, and 2.5 m/s) from 10 individuals (including nine unseen individuals and one individual whose data was used for training) with incline and decline conditions. The adaptive model was able to predict ankle angle for unseen individuals during unseen walking speeds with 0.82 to 0.964 coefficient of determination (R^2) and with the root mean square of error (RMSE) between 1.66 to 5.06 deg. On another test on acceleration data (from 0.5 m/s to 2.5 m/s), the adaptive model followed the variation in ankle angle with less than 3 deg RMSE.

The adaptive model outperformed the conventional non-adaptive model whenever the test dataset was not part of the initial training dataset. The adaptive model reduced the root mean square error (RMSE) between the estimated and actual ankle angle by ~70% during walking at different speeds. The proposed model was sensitive to variation in EMG patterns caused by speed changes, anatomical differences among people, and terrain diversity. Continuous and accurate prediction of joint kinematics under different walking circumstances and for multiple individuals by the proposed model promises a stable and reliable control for wearable assistive robots for the ankle joint.

7.1 Introduction

Robot development in recent decades makes it possible to use robots in close contact with humans in different applications, such as empowering individuals to perform heavy-duty tasks or assisting disabled individuals with locomotion tasks (robotic rehabilitation) [214, 215]. However, providing interactive communication between humans and robots remains a challenging research area. In robotic rehabilitation, to provide intuitive and effective assistance, the robot needs to understand the user movement intention and work in synchronization with the user's sensorimotor abilities [92]. An intuitive controller adaptable to different conditions is critical for assistive robots. For example, in terms of lower limb assistive robots, a key desirable feature is the ability to forecast intended joint motion or joint torques across a range of different walking speeds and inclines.

In a conventional approach, assistive robots are controlled passively. A binary on/off command (from some state-prediction algorithm) triggers the robot to go through a predefined path [216]. In this approach, the robot leads the motion regardless of the wearer's induced assistive or resistive force. Such a controlling approach is far from an intuitive control strategy as there is no contribution required from the wearer to generate the final movement. Moreover, these schemes are challenged by different environments. For example, incline or decline walking and walking at different speeds for lower limb assistive robots or moving objects in a 3D space at different speeds and paths for upper limb robots causes misalignment between robot movement and wearer desired motion [155, 213, 217]. To accommodate these variations, the predefined trajectory for the robot needs to change for every new working environment.

One approach toward an intuitive controller for assistive robots is relying on kinetics/kinematics information from the person, such as joint angle or joint torque, to understand motion intention for robot control. However, in this approach, the assistive robot only considers the current state of the wearer instead of understanding the movement intention. By applying this control method, the robot's movement will not be natural and smooth as there is a delay between user motion and robot movement. This delay comes from the required time for processing input signals, generating the actuators' command, and the actuators' response time [218, 219].

The second approach toward intuitive robot control is using electromyography (EMG) signals to provide an insight into the movement intention and minimize the time delay between robot and human movement [220, 221]. Neural information from EMG signals offers valuable information to interpret motion intention [222, 223]. As electrical impulses in muscles are triggered before the actual muscle contraction (electromechanical delay), they give information about movements 30 ms to 100 ms before the physical movement for different joints depending on the size and location

of the muscles involved in movement generation [128]. EMG is widely used for motion prediction because it is easy to collect and has a higher signal-to-noise ratio than other neurological signals such as electroencephalogram (EEG) [224-227]. However, EMG signals are still noisy and require post-processing and an appropriate model for interpretation.

EMG-based continuous motion prediction models can be broadly categorized into *mechanistic* or *data-driven* approaches. A mechanistic approach involves a neuromusculoskeletal (NMS) model, which uses the EMG signals as input to estimate the muscle activation and subsequent muscle mechanics (typically a Hill-type muscle model) to predict muscle forces and joint torques. NMS model was used to predict flexion/extension moment of knee joint with an average of R^2 of 0.91 [19]. NMS model also was used for motion prediction of joints with higher degrees of freedom. For instance, it has been implemented for wrist angle estimation based on EMG signals with 10 deg RMSE [228]. Although these approaches have been shown to predict joint contact forces accurately, they suffer from a drawback; having a prolonged and comprehensive calibration phase for each individual. A wide range of variables corresponding to the muscle and tendon properties correlating the EMG signals to muscle force is determined during a calibration process [19, 229]. The calibration process needs to be performed for each individual, so it is not easy to generalize this approach. Moreover, musculoskeletal models typically require kinematics data and EMG as input, so their ability to predict kinematics from EMG is limited.

The second approach of EMG-based continuous motion prediction involves data-driven models, including regression-based, artificial neural networks, and statistical models. These models are less complex than neuromusculoskeletal models and can be computationally attractive for real-time control. Extensive research in data-driven motion prediction models has improved the performance of these models in assistive robot control. Some examples of models developed to predict the pattern of motion trajectory include multi-layer perceptron neural networks [230, 231], recurrent neural networks [232, 233], adaptive neuro-fuzzy neural networks [124]. The reported results in these studies proved the effectiveness of neural network-based models in motion intention prediction based on EMG signals. Nonlinear autoregressive with exogenous inputs neural network (NARX-NN) has been used in literature to predict elbow, wrist, and ankle angle or torque [122, 234]. It showed promising results in terms of compatibility with different activities (incline/decline walking & ascending and descending steps) [235].

Regression models have also been used to determine the nonlinear translation between EMG signals and joint angle. Lower limb joint angles were predicted based on EMG signals during walking at a constant speed with less than 9 deg RMSE by implementing a nonlinear regression model [21]. Support vector regression is another model implemented in literature for joint angle

prediction based on EMG, which could predict knee angle with less than 3 deg RMSE [236]. Hidden Markov models [237, 238] Kalman filters [10, 11, 239, 240] are other nonlinear methods used for motion intention prediction based on EMG signals.

Despite these algorithms' reported accurate joint moment and angle prediction, they suffer from limitations. Typically, data-driven models only work well on data they have been trained on, which creates a challenge for training a new model to match a new activity [117, 241, 242]. The training dataset size would be prohibitive in the logistical application and computational cost. Moreover, often these models work well on one individual but do not transfer across to another person and, as such, are non-generalizable [103, 231].

Some efforts have been made to make data-driven models compatible with different individuals and activities. For instance, a recurrent neural network has been developed to predict knee angle during various activities, including standing, sitting, going up and down the stairs, and walking on the level ground [233]. The model was trained on data from all of the mentioned activities from two individuals and tested on data from two new individuals doing the same movements. The reported RMSE for the test dataset was less than 3 deg [233]. An adaptive prediction network was developed for knee angle prediction during walking at a constant speed tested over a group of 10 people. The model parameters were updated online whenever inputs were disturbed because of muscle fatigue or individual differences. The model predicted knee joint with less than 3 deg RMSE across 10 participants during walking at a constant speed [92]. The developed model was robust to disturbance of EMG signals during walking at a single constant speed, but it was not evaluated during walking at multiple speeds or different inclines [92].

This chapter aimed to develop a new motion intention prediction method using EMG signals as input based on a polynomial regression model. The model was designed for ankle angle prediction 30 ms in the future based on a short history of EMG signals and measured ankle angle at each time step. The main feature of this model is its compatibility with different walking speeds on different inclines across a range of individuals. Data from 10 individuals walking at multiple speeds on a treadmill at three different slopes were recorded for training and testing the model. During the initial training process, the model was only exposed to data related to one of the individuals walking at one constant speed. However, it is required to be compatible with other walking conditions for different individuals beyond the training dataset. The model performance was assessed by testing ankle angle prediction during walking at various speeds on inclined and declined surfaces for ten individuals.

7.2 Method

The implemented model in this study took a one-second history of EMG signals and kinematics as input at each time step and forecasted the right ankle flexion angle 30 ms in the future across a range of different walking conditions and people. Four main steps have been taken to create the adaptive model, including A) Data acquisition and preparation, B) signal processing, C) kinematic modelling, and D) model development. Each of these steps is explained in this section.

7.2.1 Data acquisition and preparation

The data collection process provided data across a range of walking speeds covering slow walking to fast walking (jogging) on different slopes. The data collection experiment was repeated for multiple participants to check the model's predictive capability across a larger cohort. Ten volunteers with no history of walking-related disorders or no record of lower limb surgery or pain that affects their walking pattern participated in this study (eight males and two females, age: 27±3 yo, height: 174.6±8.5 cm, mass: 73±11 kg). Ethics approval for this study was obtained from the University of Auckland human participants ethics committee (UAHPEC, ref no: 022112).

To record the body movement during the experiment, reflective markers were attached to the participant's body landmarks in accordance with the UWA marker set [243]. A VICON 11-camera motion capture system (VICON, Oxford USA) recorded marker trajectories at 100 Hz during the experiment. Wireless surface EMG electrodes (Delsys Trigno, Natick USA) were placed over the medial and lateral gastrocnemius (Gasmed and Gaslat, respectively), soleus (Sol), and Tibialis Anterior (Tib) muscles of the left and right leg following SENIAM protocols [175] to record muscle activities at 1000 Hz. These muscles were selected as they influence ankle moment generation. Prior to data collection, participants walked on the treadmill at a self-selected speed for three minutes to become familiar with the split-belt force-instrumented treadmill (Bertec, OH). Also, before starting the experiment, the maximum voluntary contraction (MVC) of the muscles of interest was recorded to normalize the collected EMGs.

Following the familiarization session, participants were asked to walk at various speeds and slopes (Table 7-1) to create the required dataset for model development.

Table 7-1 List of activities each individual was asked to complete during the data collection experiment.

activity	Repeat at slope			Duration (s)
	-10%	**0%**	**+10%**	
Walking at 1 m/s	*	*	*	30
Walking at 1.5 m/s	*	*	*	30
Walking at 2 m/s	*	*	*	30
Walking at 2.5 m/s		*		30
Acceleration from 0.5 m/s to 2.5 m/s at 0.02 m/s²		*		50

The slopes for inclined and declined walking were selected based on the treadmill's limitation and the participants' safety. Also, the range of walking speeds was chosen to cover a wide range of rates from slow walking to jogging. Also, collecting data during acceleration and walking on the inclined and declined treadmill covered a much greater range of conditions each individual might experience during walking outside the lab. The acceleration condition was included to test the predictive capability of the model at various walking speeds.

7.2.1 Signal processing

Collected EMG signals were band-pass filtered using a zero-lag Butterworth filter (20 Hz - 500 Hz), full-wave rectified and then filtered using a low-pass Butterworth filter (6 Hz). The filtered EMG for each muscle was normalized to the maximum value of the MVC trial for that muscle. The normalized EMG data were then down-sampled to 100 Hz to match the motion data frequency.

7.2.2 Kinematic modelling

As mentioned previously, the right ankle angle was the model's output. 'Ground truth' lower limb kinematics were calculated using a scaled musculoskeletal model and inverse kinematics algorithm using OpenSim (Stanford, CA). Briefly, a 23 degree-of-freedom lower limb musculoskeletal model was scaled to match the anthropometry of each participant. The MAPClient scaling tool was used during the scaling process, which relies on the statistical shape model of the lower limb bones [160]. The scaled model was used to reconstruct the kinematics of each activity by adjusting the pose of the scaled kinematic model to 'best-match' recorded marker trajectories. Ankle flexion-extension kinematics were exported and low-pass filtered with a Butterworth filter and a cut-off frequency of 6 Hz. The first derivative of the ankle angle was also calculated to provide time-series input of ankle angle and ankle angular velocity (Figure 7.1).

7.2.3 Model development

7.2.3.1 Model structure

According to Figure 7.1, a window of one second of the input time series was used for the feature extraction process. A set of most important features was extracted and used as input for the predictive model. At each time step, the model predicted the value of the ankle angle 30 ms in advance ($t_n + 30$ ms) and at the same time the predicted value at 30 ms in the past ($t_n - 30$ ms) compared to the actual ankle angle at the current time step (t_n) to check the magnitude of the error. If the error was greater than 3 deg, the updating phase was triggered, which was responsible for changing the model parameters to minimize the error rate (updating loop).

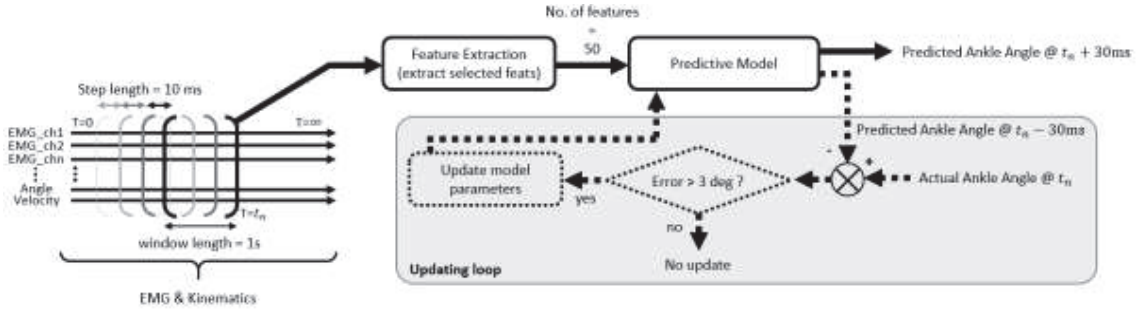


Figure 7.1 The block diagram of the adaptive regression model.

The number and list of features extracted from the input time series at the feature extraction phase were decided during a feature engineering process as part of the model training. The predictive model in Figure 7.1 was a linear regression trained initially on a limited part of the collected data during the experiment and evaluated on a wide range of datasets related to different activities and people beyond training (see "model training" below).

7.2.3.2 Model training

7.2.3.2.1 Feature engineering

Feature engineering optimizes the predictive model performance by finding the most valuable information to predict the model output. A mathematical combination of recorded data from EMG electrodes or kinematic information makes the inputs more distinguishable during different activities, which helps to improve the model accuracy. For instance, the difference between left and right ankle angle and EMGs from the same muscles on opposite legs or antagonistic muscles on the same leg increases the accuracy of ankle moment predictions [185].

This study used a mathematical combination of recorded data as input. The recorded data during the experiment collected from the EMG sensors and motion capture system are called "original

time series," and the mathematical combination of them is called "derived time series". The original and derived time series lists are represented in Table 7-2 and Table 7-3, respectively.

Table 7-2 List of original input time series.

EMG-related and kinematics related collected signals	
Time series	**Description**
Sol_r / l	EMG from soleus muscle (right leg / left leg)
Gasmed_r / l	EMG from medial gastrocnemius (right leg / left leg)
Gaslat_r / l	EMG from lateral gastrocnemius (right leg / left leg)
Tib_r / l	EMG from tibialis anterior (right leg / left leg)
Angle_R / L	Ankle angle [right leg / left leg]
Vel_R / L	Angular velocity of the ankle [right leg / left leg]

* R / L (for kinematics) or r / l (for EMGs) in this table shows the input signal recorded for the left and right leg separately.

Table 7-3 Derived time series from original input time series.

EMG-related derived time series. Each row represents the EMG difference between the declared muscles	
Time series	**Description**
Sol_tib_r	[tibialis anterior – soleus] (right leg)
Gasmed_tib_r	[medial gastrocnemius - tibialis anterior] (right leg)
Gaslat_tib_r	[lateral gastrocnemius - tibialis anterior] (right leg)
Sol_tib_l	[soleus - tibialis anterior] (left leg)
Gasmed_r_l	[medial gastrocnemius] (right leg - left leg)
Gaslat_tib_l	[lateral gastrocnemius - tibialis anterior] (left leg)
Sol_r_l	[soleus] (right leg - left leg)
Tib_r_l	[tibialis anterior] (right leg - left leg)
Gaslat_r_l	[lateral gastrocnemius] (right leg - left leg)
Gasmed_r_l	[medial gastrocnemius] (right leg - left leg)
Kinematics-related additional time series	
Angle_R_L	Ankle angle [right – left]
Vel_R_L	Angular velocity of the ankle [right - left]

The continuous input time series were divided into windows of 1-second length to provide input to the model. As such, the model used data from each 1-sec window ending to the current time step to predict the ankle angle 30 ms ahead of the present time step. To continuously predict ankle angle, the data window moved forward 10 ms at each step. Therefore, the model predicted the ankle angle every 10 ms (100 Hz) (Figure 7.1).

The total number of input time series (original + derived) was 24 (Table 7-2 and Table 7-3). Considering the 100 Hz sampling rate, a 1-second window of the input dataset contained 2400 data points. This is a large input for a predictive model, making prediction computationally expensive and time-consuming. To optimize the list of input data for the predictive model, a feature engineering pipeline was performed using the "tsfresh" python package to collect a list of the most informative characteristics of the time series for model training (Figure 7.2).

A set of data related to walking at different speeds (1 m/s, 1.5 m/s, 2 m/s, and 2.5 m/s) from one of the individuals (selected randomly) was used in the feature engineering process to find the list of the most important features. Using data from multiple speeds helped to find features corresponding to ankle angle prediction based on EMG signals across a wide range of speeds.

In the first step of the feature engineering process, a list of 794 features was extracted from each windowed time series, including time-based, frequency-based, and time-frequency-based characteristics of the time series. In the second step, all extracted features were sorted based on their impact on model prediction accuracy. It has been shown that model accuracy does not necessarily improve when using all the extracted features [21]. Some of the features are noise-related, and others are redundant because they convey similar information. The number of features for model training is a tradeoff between model accuracy and computational cost. Our previous work found that model prediction accuracy was not improved by using more than the top 50 features [185]. Here, we selected the top 50 features for model training based on the list of importance obtained from the feature engineering process (Figure 7.2).

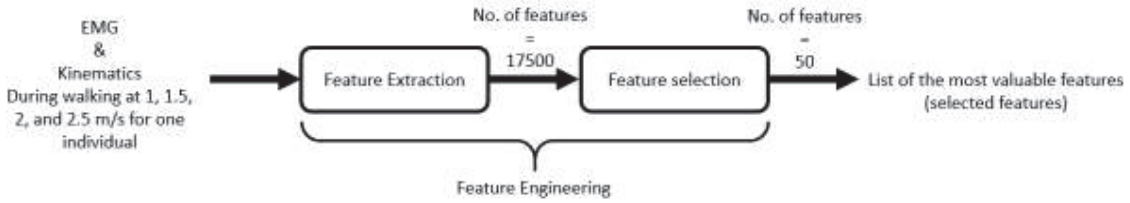


Figure 7.2 Feature engineering process.

7.2.3.2.2 Model fitting

A regression model was developed to predict ankle angle based on extracted features (Figure 7.1). The *predictive model* was a linear regression, in the form of a basis function:

$$y(x, w) = \omega_0 + \sum_{j=1}^{M} \omega_j \phi(x_j) \quad (1)$$

The parameter ω_0 defines a fixed offset in the data, or "bias" parameter, and $\phi(x_j)$ represents a basis function based on the input features x_j ($1 < j < M = number\ of\ parameters\ in\ \phi(x)$). ω_j is representing the impact factor of each basis function. w in eq. 1 represents model parameters including all of the impact factors plus ω_0, so $w = [\omega_0, \omega_1, \dots, \omega_M]$.

We can generate the basis function by taking the powers of inputs as $\phi(x) = x^i$. By increasing the value of "i", the level of complexity of the equation increases by the power of the i. Based on trial

and error, we set $i = 2$, as this improved model accuracy compared to $i = 1$ while keeping the complexity of the model at an acceptable level. Considering $i = 2$ the $\phi(x)$ is:

$$\phi(x) = x_1 + x_2 + \cdots + x_n + x_1x_2 + \cdots + x_nx_m + {x_1}^2 + \cdots + {x_n}^2 \quad (2)$$

While x is the input feature and $1 < m.n < 50$ $(number\ of\ input\ features)$ $and\ m \neq n$. Model parameters in (1) were determined by fitting a polynomial to the training dataset. During the training process, the impact factors were adjusted to best-fit the model output with the training data using the following cost function ($E(w)$).

$$E(w) = \frac{1}{2}\sum_{n=1}^{M}(y(x_n, w) - Y_n)^2 \quad (3)$$

In which $y(x_n, w)$ is the predicted value for each data point (n), and Y_n is the actual value corresponding to each one of the input data points ($1 < n < N = number\ of\ input\ data\ points$).

To control over-fitting, the regularization term was added to the cost function, so the total cost function was defined as:

$$E_D(w) + \lambda E_w(w) \quad (4)$$

In which $E_w(w)$ is the regularization term and λ is the regularization coefficient that controls the relative impact of the data-dependent error $E_D(w)$ (represented in eq. 2).

$$E_w(w) = \frac{1}{2}w^T w \quad (5)$$

This regularizer is known as "weight decay" because it encourages the value of weights to decay towards zero. Considering the sum of the squares error function as:

$$E(w) = \frac{1}{2}\sum_{n=1}^{N}(t_n - w^T\phi(x_n))^2 \quad (6)$$

the total cost function was:

$$\frac{1}{2}\sum_{n=1}^{M}(Y_n - w^T\phi(x_n))^2 + \frac{\lambda}{2}w^T w \quad (7)$$

By setting the gradient of the cost function to zero with respect to w, and solving it for w, we obtain:

$$w = (\lambda I + \Phi^T\Phi)^{-1}\Phi^T t \quad (8)$$

This function correlates the regression model weights (w) to the basis functions, $\Phi = \phi(x)$. In this equation, λ is the training rate.

After the training process, the regression model parameters were adjusted to predict the ankle angle based on input features. Data relating to walking at 1 m/s on level ground from one participant was selected randomly to train the model (Figure 7.3).

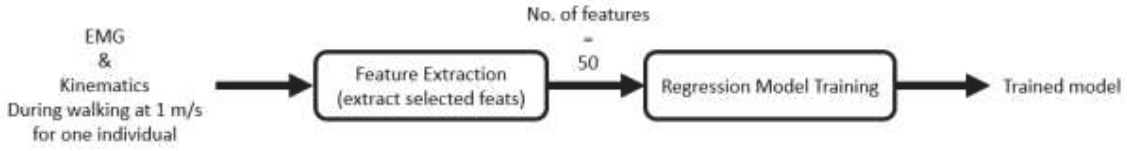


Figure 7.3 Regression model training.

7.2.3.3 Updating phase

The updating phase adjusted the model parameters (w) whenever the difference between the forecasted ankle angle and actual value was more than 3 deg (Figure 7.1). This threshold was determined through a trial and error process. For higher thresholds, the model was less sensitive to variation in output, and for lower thresholds, the model output was noisy with poor performance.

Providing adjustable parameters for the model is called "sequential learning," which is the critical characteristic of our adaptive model. Sequential learning is appropriate for real-time applications in which the data observations arrive in a continuous stream. The model needs adaptation during the prediction process based on the difference between the predicted value and the actual value of the ankle angle. We applied sequential gradient descent to update the model parameters based on the difference between the predicted value and actual value at each data point. The *gradient descent* algorithm updated the parameters of the model (w) using:

$$w_{\tau+1} = w_\tau - \eta \nabla E_\tau \quad (9)$$

In which E_n is the error summation between model output and actual output over the past n data points. Also, T denotes the time step, and η represents the learning rate. Based on the error function defined in (5), the new value of the weight vector was determined as:

$$w_{\tau+1} = w_\tau + \eta(Y_\tau - {w_\tau}^T \phi_\tau)\phi_\tau \quad (10)$$

Where $\phi_\tau = \phi(x_\tau)$ at data point τ. This is known as the least-mean-square or LMS algorithm. In this equation, the value of η is important as it defines the update rate, determined based on the trial and error approach in this study.

A weighted average of errors in the past five time-steps was used to update the regression model coefficients to prevent large changes in weights during the updating process. In this case, the update function for coefficients was:

$$w_{\tau+1} = w_\tau + (\sum_{i=1}^{5} \eta_i (Y_i - {w_i}^T \phi_i))\phi_i \quad (11)$$

In which the Y_i is the actual value of joint angle and ${w_i}^T \phi_i$ is the model prediction, η_i is the impact factor for each error, and index i represents the number of time steps in the past used in the summation. $i = 1$ means one time step back, and $i = 5$ means five time-steps back. The value of η_i and i were defined via trial-and-error to minimize the updating delay, the fluctuations in predicted values, and the difference between predicted and actual ankle angles.

7.2.4 Model evaluation

The performance of the adaptive model was evaluated under different walking conditions across ten individuals to check the robustness of adaptation to input variation. To compare the model prediction with actual ankle angle from inverse kinematics, the value of root mean square of error (RMSE) and coefficient of determination (R^2) were reported.

To show the role of updating on the adaptive model performance, the results from the adaptive model were compared to the results of the predictive model without the updating phase. As the model parameters without the updating phase were fixed, this model will be referred to the 'static model' for the rest of this paper.

The list of activities used for model evaluation included:

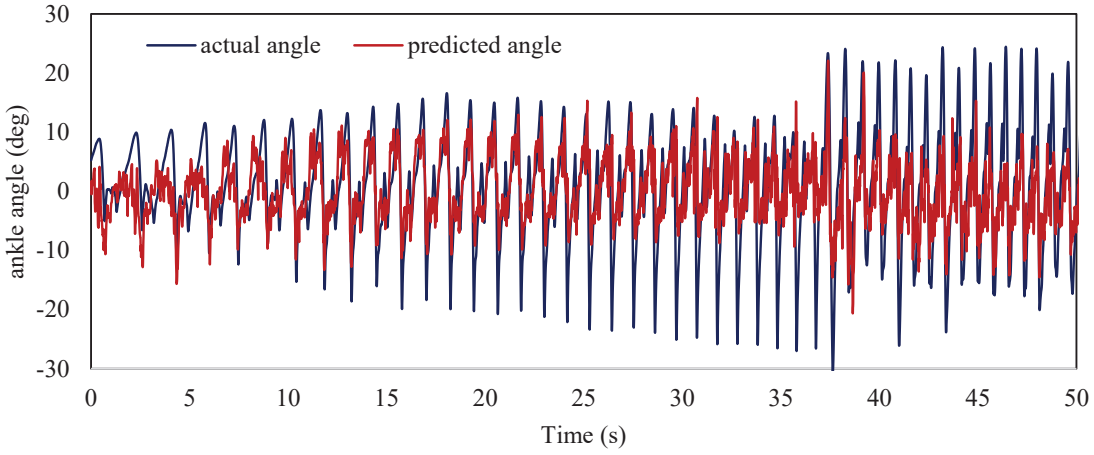
- Walking at variable speed (acceleration data)
- Level, inclined and declined ground walking at different constant speeds.

Moreover, a visual comparison between the predicted ankle angle via the adaptive model and static model versus the calculated angle via inverse kinematics was reported to show the accuracy of these models in ankle angle prediction during one stride.

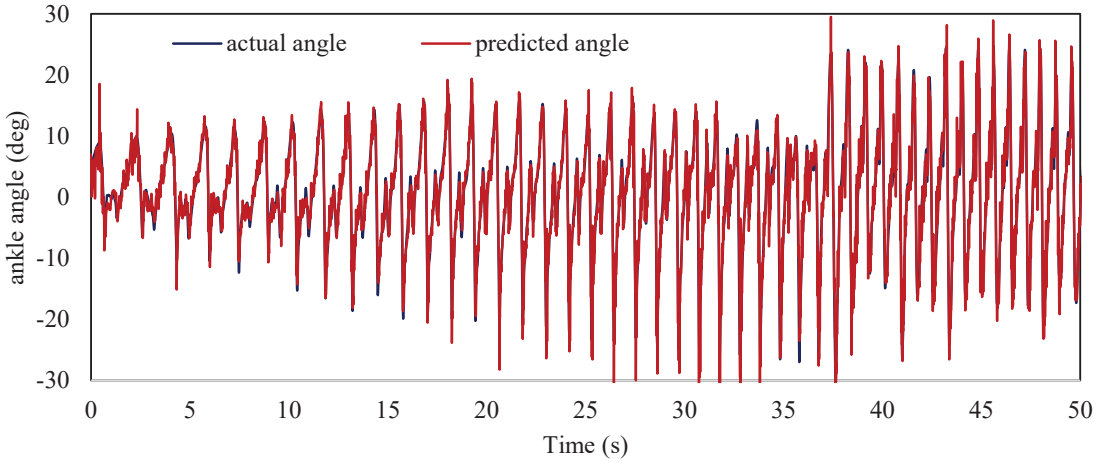
7.3 Results

The adaptive model's main advantage is its ability to adjust to EMG signal patterns variations beyond the training dataset. One reason for the change in EMG signal patterns is variation in walking speed. Testing the adaptive model on acceleration data showed that it was compatible with speed variation (Figure 7.4a), even though it was initially trained on data from constant speed walking (1 m/s). Results in Figure 7.4a illustrate the static model performance in the acceleration condition, predicting ankle angle for a participant whose data were not used for initial model training. The same dataset using the adaptive model shows the ability of the model to rapidly adjust to the varying speed condition and predict a wide range of ankle angles (Figure 7.4b).

The difference between the adaptive and static models is captured in the updating loop. Figure 7.5 shows how quickly the model changes from static to the adaptive mode by activating the updating loop. Here the updating mode was activated at second 4, and the model quickly began to follow the trajectory of the ankle angle. Five data points were required to run the optimization process in the updating phase, which corresponds to a 20 ms delay plus the required execution time for updating phase, which was less than 10 ms on average (9.4±0.7 ms) (using intel Core i7-7700 at 3.60 GHz CPU, 32 GB of RAM, and Windows 10 Enterprise – Version 1803). So there is a 30 ms delay between the updating phase onset and the generation of the new parameters.

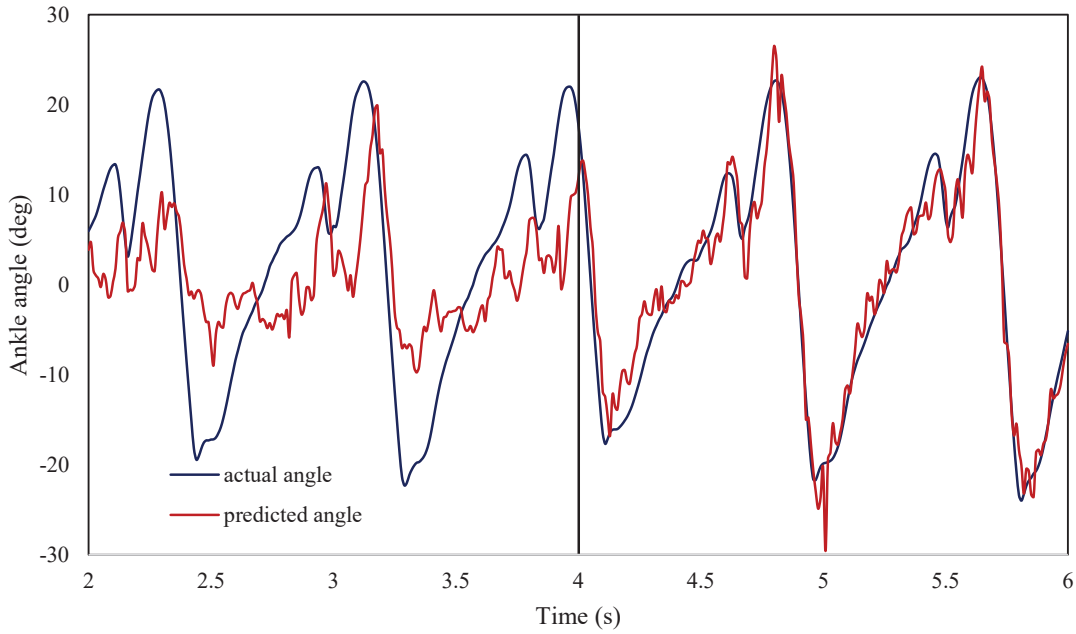


(a) Static model

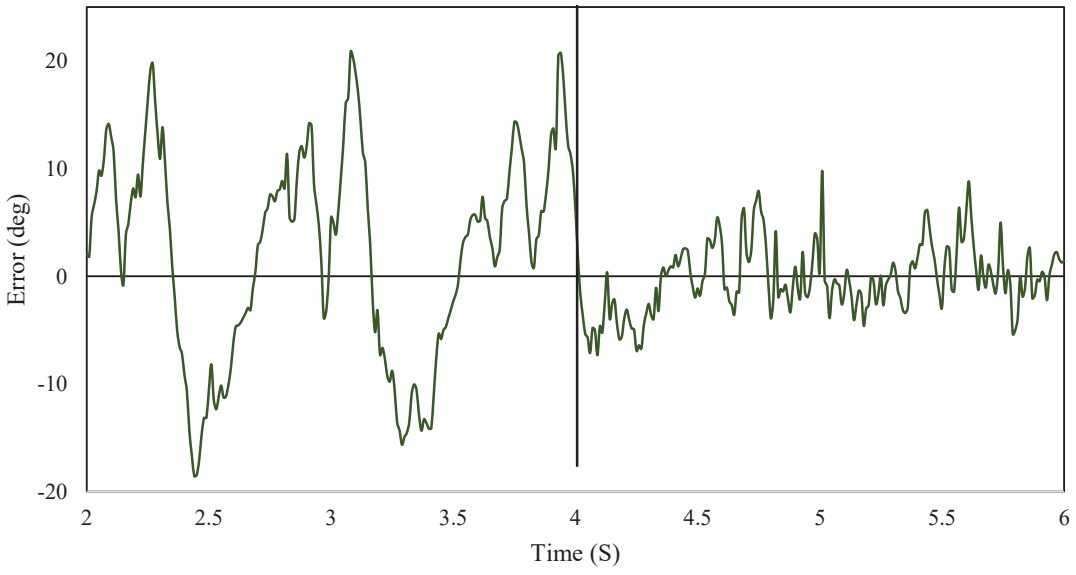


(b) Adaptive model

Figure 7.4 *Model performance for ankle angle prediction while accelerating the walking speed from 0.5 m/s to 2.5 m/s at a constant rate. Model output for ankle angle prediction is represented for the (a) static model and the (b) dynamic model. The model outcome was tested on data from a randomly selected participant (ID 6) whose data was not used for static model training.*



(a) Ankle angle prediction and actual value



(b) Error between predicted and actual ankle angle

Figure 7.5 *a) A snapshot of model output with four seconds duration. The updating phase engaged in the process at second 4. b) The difference between a predicted angle and actual angle value represented in (a).*

A visual comparison between the adaptive and static models in ankle angle prediction during level walking at 1.5 m/s is represented in Figure 7.6. Comparison of model output to the inverse kinematics results (ground truth) showed how accurately the adaptive model could predict ankle angle regardless of variation in EMG patterns because of speed changes (Figure 7.6).

A more detailed comparison between the static and adaptive models was conducted to evaluate their performance in ankle angle prediction during walking at different constant speeds and acceleration during level walking (Table 7-4). As illustrated in Table 7-4, the static model only worked accurately at this speed for this participant and could not adapt to other walking speeds for this individual or others. However, the adaptive model accurately predicted the ankle angle in all conditions ($0.86<R^2<0.96$).

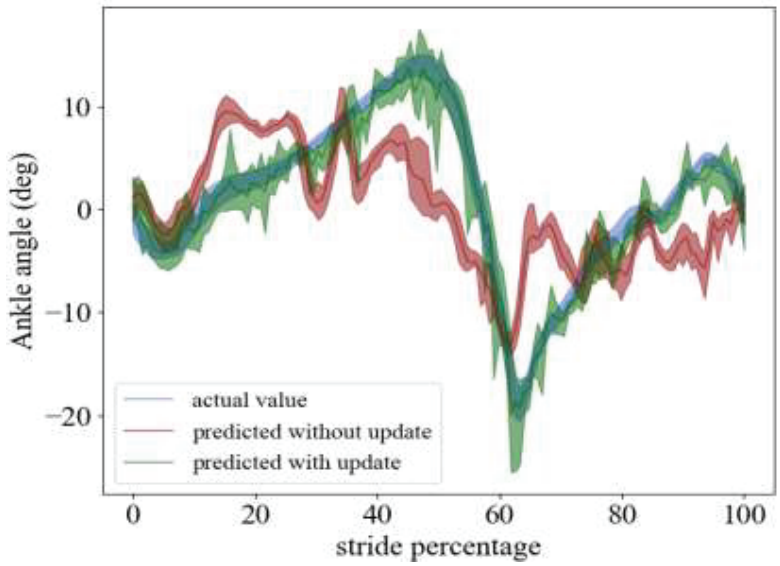


Figure 7.6 *Model output comparison when the updating phase was active (green line) and when it was not (red line) versus the measured ankle angle (blue). The variation represents the minimum and maximum ankle angle at each data point across 30 seconds of walking data at 1.5 m/s.*

Based on our experiment results, the model that trained on level ground walking could not predict ankle angles after ±10% change in ground incline.

Table 7-5 and Table 7-6 show the lack of accuracy in ankle angle prediction of the static model even for the individual and speed used for model training (subject ID 4, 1 m/s) after ±10% change in incline. However, the adaptive model was compatible with incline variation and could accurately predict ankle angle for multiple subjects walking at different speeds.

Table 7-4 *Model performance comparison between the (a) static model and (b) adaptive model during level-ground walking at multiple speeds. The model performance was reported in terms of RMSE (numbers) and R^2 (colour) for ten individuals. (red = R^2 of 0.4 and green = R^2 of 0.96). The column "acc" refers to data during acceleration from 0.5 m/s to 2.5 m/s.*

(a) Static model

ID	Test Speeds				
	1 m/s	1.5 m/s	2 m/s	2.5 m/s	acc
1	6.53	7.36	9.24	9.88	7.57
2	5.52	7.72	7.61	9.79	7.99
3	5.33	7.56	6.76	10.02	7.70
4	**1.98**	**6.63**	**6.41**	**8.06**	**6.42**
5	6.09	8.99	10.09	10.67	9.54
6	6.20	7.57	10.09	8.95	8.94
7	6.61	7.21	8.11	9.05	7.83
8	5.72	7.71	9.30	9.75	9.22
9	5.43	7.37	8.91	8.88	8.30
10	6.17	8.01	9.85	10.16	8.87

(b) Adaptive model

ID	Test Speeds				
	1 m/s	1.5 m/s	2 m/s	2.5 m/s	acc
1	1.85	2.16	2.94	2.86	1.51
2	1.94	3.82	2.74	3.43	2.50
3	2.57	3.08	2.59	4.13	2.32
4	**1.66**	**2.98**	**2.27**	**2.51**	**2.03**
5	3.56	4.47	3.88	3.88	2.90
6	2.42	3.16	3.34	3.51	2.64
7	1.82	2.56	2.39	2.96	1.67
8	2.30	3.14	2.87	3.37	2.31
9	2.94	3.51	3.22	3.99	2.80
10	1.98	2.95	3.16	2.87	2.62

Table 7-5 *Model performance comparison between the static and adaptive models during inclined walking (+10% slope) at multiple speeds. The model performance was reported in terms of RMSE (numbers) and R^2 (colour) for ten individuals. (red = R^2 of 0.32 and green = R^2 of 0.98).*

ID	adaptive model			static model		
	1 m/s	1.5 m/s	2 m/s	1 m/s	1.5 m/s	2 m/s
1	1.99	5.70	3.18	9.29	10.94	13.52
2	5.70	3.84	1.68	9.51	9.69	9.85
3	4.54	3.98	2.24	13.75	14.67	14.16
4	5.07	4.54	4.25	11.45	11.61	11.26
5	4.98	5.51	5.70	17.96	18.25	19.08
6	3.77	4.21	4.33	15.50	12.90	14.50
7	2.93	3.77	2.51	18.11	13.10	16.25
8	4.71	4.34	3.27	14.33	17.20	18.52
9	3.71	5.10	3.22	11.80	16.80	12.90
10	5.06	4.63	3.91	13.90	14.33	17.34

Table 7-6 *Model performance comparison between the static and adaptive models during declined walking (-10% slope) at multiple speeds. The model performance was reported in terms of RMSE (numbers) and R^2 (colour) for ten individuals. (red = R^2 of 0.21 and green = R^2 of 0.94).*

ID	adaptive model			static model		
	1 m/s	1.5 m/s	2 m/s	1 m/s	1.5 m/s	2 m/s
1	4.16	4.79	4.36	6.01	6.04	8.89
2	2.90	2.21	1.74	7.75	9.84	9.20
3	2.71	1.91	1.54	6.95	7.63	6.47
4	3.54	2.22	2.08	6.27	6.66	6.43
5	3.90	3.86	3.49	9.89	10.19	10.56
6	2.95	2.55	3.66	10.11	9.23	8.85
7	3.75	3.65	2.85	8.56	8.58	9.90
8	3.96	3.94	2.96	9.23	7.77	7.35
9	4.28	4.11	3.64	8.90	10.60	6.90
10	3.63	3.77	4.01	7.12	9.12	8.90

Table 7-7 compared the average of RMSE and R^2 (over nine unseen individuals) between the static and adaptive model for each test speed separately.

Table 7-7 *The average of RMSE for ankle angle prediction at different walking speeds for nine unseen individuals.*

		Test speeds			
		1 m/s	1.5 m/s	2 m/s	2.5 m/s
Static model	RMSE	5.93	7.84	8.76	9.86
	R^2	0.49	0.33	0.35	0.44
Adaptive model	RMSE	2.47	3.34	3.10	3.56
	R^2	0.91	0.89	0.92	0.93

One item that had a significant impact on the model performance was the combination of input time series. Using a mixture of EMG and kinematics (IK) as input improved the model accuracy compared to only using EMGs as input (Figure 7.7).

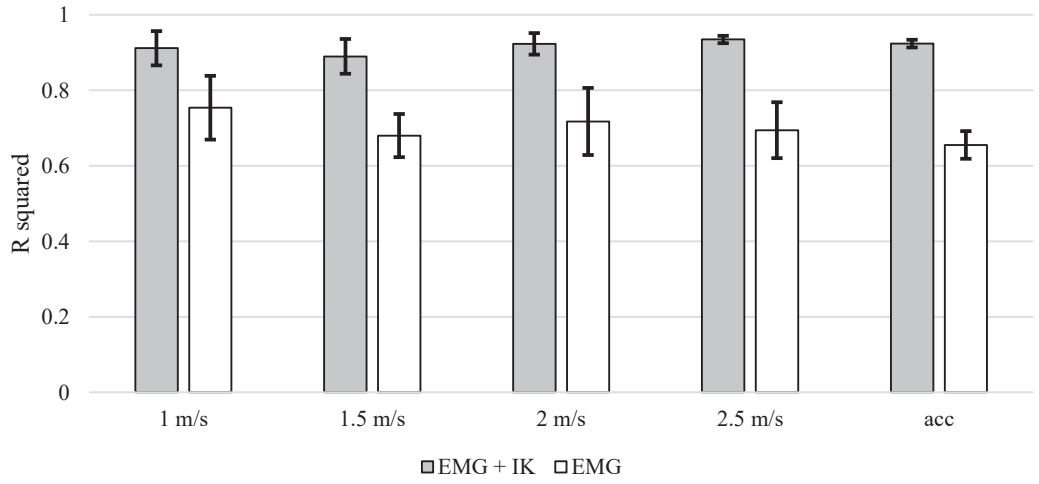


Figure 7.7 *Evaluate the impact of input time series on model accuracy. Muscle activity (EMG) combined with kinematics (IK) improved model accuracy compared to EMG alone.*

7.4 Discussion

The aim of this paper was to develop an adaptive model to predict the ankle joint angle as an indicator of motion intention based on EMG signals. The main characteristics of this model are the capability of continuous ankle angle forecasting 30 ms ahead of time and compatibility with different walking speeds and slopes across people without a requirement for training for each individual or walking condition separately.

The ankle angle is considered the desired output for this chapter's developed model. Ankle angle was selected instead of the moment because the model requires feedback of the actual output to update the model parameters on-the-fly. Ankle moment calculation in real-time is not feasible for

an assistive robot controller due to the need for solving inverse dynamic equations and the need for continuous ground reaction force measurement. However, ankle angle measurement is feasible by using a potentiometer at the ankle joint.

Testing the model on data from the acceleration trial illustrated the model's potential in terms of compatibility with variable speed conditions (Figure 7.4b). The variation in speed cause difference in EMG patterns, and a data-driven model trained on data related to constant speed walking will not predict joint angle (as shown by the static model in Figure 7.4a). Testing across 50 seconds of acceleration data showed that the value of R^2 for the adaptive model was 71% higher than the static model (Figure 7.4). The key factors behind the compatibility of the model with variation in walking speed, even for unseen individuals, are the simplicity of the model structure and the fast updating phase, which made it possible to update the model parameters according to the error between the predicted ankle angle and actual ankle angle under 10 ms (Figure 7.5).

EMG-based data-driven models are mainly used for motion classification and gait phase detection in the literature [144, 217, 244, 245]. EMG-based classifiers would not provide enough information to control an assistive robot smoothly and accurately as they don't offer a high-resolution estimation of joint angle or moment. So a continuous estimation of kinematics is needed, which is possible to provide by regression models including linear and nonlinear (neural network-based) models [21, 155, 246, 247]. However, the majority of the regression models introduced in the literature are mainly evaluated on a limited dataset and exposed to the test dataset during the training process. The adaptive model presented in this study predicted ankle angle during walking at speeds other than what the model trained on initially, even for unseen individuals (Figure 7.6). The value of R^2 under all of the testing speeds over ten individuals was over 0.85. The difference between the model output and actual ankle angle was constantly less than 5 deg, in line with state-of-the-art models (Table 7-4). In Table 7-4, the average R^2 value for the static model was 0.42 (for all speeds and all individuals), while the adaptive model was 0.92.

Moreover, the average RMSE for the static and adaptive models was 7.9 deg and 2.8 deg, respectively (Table 7-4). These findings illustrate the performance gains by using an adaptive model in comparison to a static model. Walking on different terrains is a condition that causes variation in EMG patterns, which is not investigated comprehensively in the literature due to limitations in data collection. Most of the developed models used instrumented treadmill for data collection, so the applied data for model training and evaluation was limited to walking at different speeds during level-ground walking. In this study, walking on different terrains was simulated by changing the slope of the treadmill during the experiment. A data-driven model trained on level ground walking cannot predict ankle angle based on EMG signals during inclined or declined ground walking

(Table 7-5 & Table 7-6). 10% slope is enough to change the pattern of EMG signals to the point that disabled the trained model on the level ground walking to predict the ankle angle during walking at any speed, even for the same individual whose data were used for training. However, the adaptive model was compatible with the variation of EMG signals due to terrain changes. The adaptive model estimated the ankle angle based on EMG signals for unseen individuals during ±10% slope walking with less than 5 deg error (Table 7-5 & Table 7-6).

The accuracy of the adaptive model in ankle angle prediction was not only subjective to the structure of the predictive model and efficiency of the updating phase. The number of selected features during the feature engineering process was another important factor that changed the model accuracy. Moreover, the list of input time series was another critical factor. It showed that adding the history of kinematics to EMG signals enhanced the model accuracy (4 deg improvement on RMSE). Figure 7.7 shows a 21% improvement in the value of R^2 after adding the kinematics to input EMG signals over testing speeds.

In this study, the ankle angle 30 ms ahead of the current time step was the predictive model's target. The reason for selecting the ankle angle as an indicator of motion intention instead of ankle moment was the requirement of real-time feedback of the actual value of the model output for updating the adaptive model parameters. It is practical to measure the ankle angle via a linear or rotary potentiometer in an actual application. However, having real-time feedback of other kinematics-related properties, such as moment, is not feasible. Because of this, the joint angle was predicted instead of moment as an indicator of motion intention.

The reported results in this study suggest that the implemented model is robust to the variation of EMG signals because of walking speed differences, the ground incline variation, and anatomical and biological differences across people. Having robustness to this level is important for an EMG-based model to control assistive robots, making the model adaptive to different working conditions. However, testing the model performance under a wider range of activities (such as stair climbing and sudden change in speed or walking direction) is required to support the idea of adaptability with different daily basis movements. Moreover, testing the model on the dataset from ten individuals works as proof of concept and evaluating the model on a larger dataset from a greater number of people with more age diversity is required.

7.5 Conclusion

This study introduced an adaptive model for ankle angle prediction 30 ms ahead of time, based on measured muscle activity and ankle angle at a current time step. The model can predict ankle flexion/extension angle during various walking conditions, including walking at different speeds

and inclines. The model also provides accurate predictions across different individuals, showing excellent generalizability. The optimized list of input features and the simple structure of the model made it possible to update the model's parameters whenever there was an error between prediction and actual ankle angle. This means the model doesn't need to be trained on an extensive dataset, and it adapts to variation in muscle activation as soon as the walking environment changes. Robustness to variation in muscle activation patterns caused by changes in speed or terrain makes the model ideal for controlling assistive robots in daily activities when there is no control over the working environment. Also, the compatibility of the model with anatomical diversity among people, which causes different muscle activation patterns, means there is no need for calibration for each individual separately. We plan to implement this model as part of a control loop for an ankle assistive robot to evaluate the real-world performance of this control strategy.

8 Conclusion

To develop an intuitive controller for assistive robots, the grand challenge is predicting the motion intention of the wearer forward in time. Data-driven models offer the potential to achieve this. The design and development of an adaptive data-driven model to predict ankle motion based on EMG signals have been explored in this thesis. The model was capable of predicting ankle angles 30 ms ahead of time, based on extracted features from input time series (including EMG signals from muscles crossing the ankle and ankle kinematics). Compatibility of the model with different tasks and individuals (with varying patterns in muscle activation) were two main characteristics of the ultimately designed model. On the way to developing this model, I had to find the answer to a set of critical research questions, including the required length of input time-series for motion prediction of the ankle at each time step, extracting the maximum information from the minimum number of input time series by applying feature engineering methods, and structure of the predictive model. The model's promising results suggest the potential applications in assistive robots. This model will be used to control an ankle assistive robot as part of future works.

8.1 Thesis summary

Understanding the motion intention is the main requirement to providing an intuitive interaction between the assistive robot and the wearer. This thesis investigated an adaptive model to predict the motion intention in the ankle joint based on the recorded EMG signals from muscles involved in the motion generation in this joint. To make the process of mode training fast and efficient, the minimum length of information from input time series and the optimum number of features extracted from input time series were the questions that have been addressed in this thesis. The most efficient training datasets were investigated to make the model compatible with a wide range of activities. In addition, different types of data-driven models were analysed to find the most compatible model with various activities and people. In the end, a polynomial regression model was developed, adapting to a new working situation on the fly. The adaptive model performance was evaluated on collected data from ten healthy individuals walking at different speeds on an instrumented treadmill at three different inclines. The results showed the model potential for ankle angle prediction accurately during walking in different circumstances. Also, the model can get adapted to any new working situation out of the training dataset. The promising results in this thesis showed the potential for the model to be implemented as part of the control method for assistive robots, which will inform the controller about the motion intention of the wearer accurately.

8.2 Achievements and novel contributions

In this chapter, the novel contributions from each chapter are summarised. On the way toward developing a data-driven model for estimation of ankle dynamics based on EMG signals during different walking speeds and activities, a set of research questions needed to be answered. Each of these research questions was covered in one chapter in this thesis. This chapter covers the significant findings and outcomes of each chapter. The limitation of the developed model is also discussed in this chapter, and a pathway toward a more practical model in a real-life application for assistive robots control is proposed as future work.

8.2.1 Finding the best configuration of input time series for a data-driven model to predict ankle moment (Chapters 3 and 4)

To predict motion intention at each time step, it is required to feed the prediction model with a batch of information about the past. The length of this information is important to the accuracy of the model. It was reported in chapter 3 that a window of one second of the EMGs and kinematics provides enough information to predict ankle moment. Considering the 100 Hz sampling rate, one second of data was equivalent to 100 data points. Also, it has been shown that

EMG alone is not enough for moment prediction in ankle joint and additional information from kinematics (ankle angle and angular velocity) enhances the model accuracy. Further information from kinematics reduced the average of RMSE between 4% and 11% for ankle moment prediction at different time steps into the future (chapter 4).

8.2.2 Extract the most valuable information from input time series for a motion prediction data-driven model (Chapter 5)

To minimise the required physical equipment for data collection, it is necessary to find a way to extract the maximum information from minimum input channels. Chapter 5 showed that the mathematical combination of input time series gives valuable information to the model to predict ankle moment. Adding the subtraction of EMG channels and the difference between the right and left leg kinematics to the original input time series improved model accuracy from 0.945 of R^2 to 0.985.

One limitation of data-driven models is the requirement to train before testing on data from new individuals. In chapter 5, a feature engineering pipeline (including feature extraction and selection) is introduced to address this limitation. It discovered that instead of using all of the extracted features from the input time series, the top 50 most informative features were enough to achieve over 0.95 accuracy (R^2 value). Secondly, a further investigation showed among the top 50 features, sixteen 'generic' features were repeated in the list of top features for ten participants. Using the generic feature set (16 features) instead of a personalised feature set (50 features) only reduced the accuracy by 1%. However, a generic feature set offers tremendous advantages to model development. Firstly, because of using a small number of features as input, the process of decision making for the predictive model is faster in comparison to using all of the personalised features as input, while the model accuracy doesn't change significantly. Secondly, there is no need to go through the feature engineering process for each new individual as the generic feature set works perfectly for unseen individuals. In this case, the model training process for unseen individuals only includes the determination of model parameters. In other words, the generic feature set makes the data-driven models for motion prediction more compatible with the variation of EMG patterns among different individuals and makes the training process magnificently faster.

8.2.3 Make a data-driven model compatible with a wide range of activities. (Chapter 6)

In chapter 6, a random forest model was developed to predict ankle moment during walking at different speeds. The results showed that this model could accurately predict ankle moment only when the train and test were on data from the same speed. The variation in walking speed changes

the muscle activation pattern to the point that a data-driven model would not be able to predict the moment based on EMGs. This example shows that data-driven models are task-specific. Training the model on data from all of the walking speeds makes the training process significantly time-consuming. In chapter 6, a minimum required training dataset to make the model compatible with different walking speeds was investigated. It was observed that training the random forest model on acceleration data from 0.5 m/s to 2.5 m/s makes the model compatible with ankle moment prediction during walking at any speed in this range. The trained model on acceleration data predicted ankle moment during walking at 1 m/s, 1.5 m/s, 2 m/s, and 2.5 m/s accurately with 0.86 to 0.96 value of R^2.

8.2.4 Develop an adaptive data-driven model for motion intention prediction (adaptable with different individuals and tasks) (Chapter 7)

Data-driven models for motion prediction based on EMG signals are generally task-specific and subject-specific. Chapter 7 explored a solution to make a data-driven model adaptive. A linear regression model was developed as the predictive model, which took features from EMG and kinematics time series as input and predicted ankle angle 30 ms ahead of the current time step. This predictive model coupled with an updating block, adjusting the model parameters whenever the difference between the predicted ankle angle and actual ankle angle was more than 3 deg.

An adaptive model doesn't need to be trained on data related to every task and each new individual. The developed model in chapter 7 was initially trained on data from one individual walking at 1 m/s constant speed. A broad range of walking speeds from slow walking to jogging, level to inclined and declined conditions were used for testing to determine the model compatibility to different activities. Also, the test was repeated for unseen individuals.

The results showed the adaptive model was able to follow the pattern of ankle angle variation during walking at all test speeds for unseen individuals with at least 0.87 of R^2. Also, testing the model on acceleration data was successful as the model predicted ankle angle with 0.90 R^2. The results in chapter 7 proved the adaptability of the model with a wide range of activities and different muscle activation in different individuals. Moreover, the quick updating procedure (10 ms) suggests the application of this model in the real-time control of assistive robots.

The developed model in chapter 7 is useful for position control of assistive robots, while the models developed in chapters 5 and 6 are practical in moment control. In general, position control is useful for repetitive movements such as walking, but it has limitations when perturbations are applied to the system. For occasions involving a rapid change in movement, moment control will be necessary.

Implemented model for ankle movement prediction will be used to implement a controller for the ankle assistive robot developed at the University of Auckland. The ankle angle is required for the controller of this robot as it adjusts the engagement and disengagement of the spring based on the ankle angle and the gait phase. As a proof of concept, to show that the implemented model in this thesis has the potential to predict ankle angle into the future, the model output changed from moment to angle in chapter 7.

8.3 Limitations and Suggestions

This thesis focused on overcoming the limitations of data-driven models for the prediction of motion intention based on EMG signals, namely, task specificity and subject specificity. These limitations make the training process time consuming and restrict the application of these models in control of assistive robots. The proposed methodologies addressed these limitations with promising results. However, the methods presented have their own limitations and assumptions, which should be considered.

8.3.1 Constrained participant population

The data collected in this experiment were from ten participants aged between 24 and 30 years old. This homogenous population provides limited variation in muscle activation patterns caused by potential age differences. It was not possible to evaluate the robustness of the model to these changes. Moreover, due to limited time and access to the laboratory, the data collection was constrained to ten individuals. This was enough to show clinical proof-of-concept, but a larger cohort across a wider age range would enable further evaluation of model robustness.

In addition, the participants in the data collection process were healthy individuals without any lower limb pathology or disorders that might alter their muscle or walking patterns. Considering the final application of the model as an ankle assistive robot for stroke patients, it will be necessary to evaluate the model on stroke patients. The dataset in this thesis does not address questions regarding stroke patients, such as the influence of muscle deficits, quality of EMG signals and the inconsistency of walking patterns on model prediction accuracy.

8.3.2 A limited number of tasks

The dataset presented here provides a comprehensive assessment for EMG-based data-driven models for predicting lower limb motion with varying walking speeds and inclines. Compared to the literature, this is a considerable improvement beyond testing level ground walking at constant speeds. However, to approach real-life application, it is necessary to evaluate the model performance on other activities of daily living, such as stair climbing and walking across uneven

terrain or cambers, which introduce different challenges for stability and movement. This is considered as future work of this thesis.

The ankle moment and angle during walking at different speeds were two parameters of joint dynamics which were explored in this thesis. System identification techniques are needed to investigate other aspects of ankle joint dynamics, such as joint stiffness and the forces that must be generated to perform a voluntary movement. A more comprehensive dataset is required, including perturbation of ankle joint at different frequencies and external force amplitudes during postural control, to understand the physiological characteristics of the ankle and dynamics of this joint.

8.3.3 Data were collected on a treadmill in a motion capture laboratory.

Using a motion capture system with 11 cameras provides accurate measurement of body movement, but the constraints of performing these experiments in a laboratory limit the potential variability that is inherent in the real world. Also, using an instrumented treadmill to collect ground reaction forces during walking experiments limits the model's application to overground walking in the real world. Ideally, wearable sensors, such as inertial measurement units (IMUs), would be explored in future iterations of the model to provide these 'real world' data.

8.3.4 Signal processing assumptions

The signal processing and data preparation steps necessary for training and testing the model were performed 'offline' following data collection in this thesis. A real-world application of the model will require these signal processing steps (including filtering and windowing) to be performed 'online' or in real time. The process of filtering for the input signals will introduce a delay in the data preparation process depending on the type of the applied filter. In real-time applications the model needs to deal with this delay. However, as the model is predicting the output into the future, there will be enough time for signal processing and filtering.

8.3.5 A real-world assistive robot scenario

The final implemented model in this thesis will be used to control an ankle assistive robot as a real-world scenario. The robot has an active clutch system that engages and disengages a spring (Figure 8.1).

Figure 8.1 *The semi-active ankle assistive wearable robot. Designed and manufactured at the University of Auckland, New Zealand [248].*

The spring compresses during the early stance phase and releases the energy during the push-off phase to help the wearer complete the gait cycle [248]. It is necessary to activate and trigger the clutch at the right stride phase to provide the right amount of support. The developed model in this thesis can forecast the ankle angle and provide information for the robot's controller about the gait phase. A potentiometer is embedded in the robot's structure, which gives real-time feedback about the actual ankle angle to update the model parameters whenever the estimation is not accurate enough. In this scenario, the motion caption system will be replaced with a linear potentiometer, making it possible to implement this model outside of the laboratory. As future work, the robot controller performance will initially be assessed on healthy individuals, and then its performance will be evaluated on stroke patients.

The developed model to predict ankle moment and ankle angle in this thesis followed a generic approach. In other words, the feature engineering pipeline which was used for model development is applicable to other prediction problems. So it is possible to apply this model to predict other musculoskeletal model parameters, including the muscle and tendon characteristics. In this case, the data-driven model will make the process of calibration and parameter adjustment for a musculoskeletal model faster. For instance, the predictive model developed in this study can be used to define muscle and tendon parameters for a new individual based on the information from the previous calibration processes.

Appendix A.1

Ethics approval for study

Research Office
Post–Award Support Services

The University of Auckland
Private Bag 92019
Auckland, New Zealand

Level 10, 49 Symonds Street
Telephone: 64 9 373 7599
Extension: 83711
Facsimile: 64 9 373 7432
ro-ethics@auckland.ac.nz

UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE (UAHPEC)

09-Nov-2018

MEMORANDUM TO:

Dr Thor Besier
Bioengineering Institute

Re: Application for Ethics Approval (Our Ref. 022112): Approved

The Committee considered your application for ethics approval for your study entitled **EMG-informed control strategies for assistive robotics**.

We are pleased to inform you that ethics approval has been granted for a period of three years.

The expiry date for this approval is 09-Nov-2021.

If the project changes significantly, you are required to submit a new application to UAHPEC for further consideration.

If you have obtained funding other than from UniServices, send a copy of this approval letter to the Activations team in the Research Office at ro-awards@auckland.ac.nz. For UniServices contracts, send a copy of the approval letter to the Contract Manager, UniServices.

The Chair and the members of UAHPEC would be happy to discuss general matters relating to ethics approvals. If you wish to do so, please contact the UAHPEC Ethics Administrators at ro-ethics@auckland.ac.nz in the first instance.

Please quote Protocol number **022112** on all communication with the UAHPEC regarding this application.

(This is a computer generated letter. No signature required.)

UAHPEC Administrators
University of Auckland Human Participants Ethics Committee

c.c. Head of Department / School, Bioengineering Institute

Dr Bryan Ruddy
Homayoon Zarshenas

Additional information:

1. Do not forget to fill in the 'approval wording' on the Participant Information Sheets, Consent Forms and/or advertisements, giving the dates of approval and the reference number. This needs to be completed, before you use them or send them out to your participants.

2. At the end of three years, or if the study is completed before the expiry date, please advise the Ethics Administrators of its completion.

3. Should you require an extension or need to make any changes to the project, please complete the online Amendment Request form associated with this approval number giving full details along with revised documentation. If requested before the current approval expires, an extension may be granted for a further three years, after which a new application must be submitted.

Appendix A.2

Experiment run sheet

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
AUCKLAND BIOENGINEERING INSTITUTE

Run Sheet for Experiment

Participant Code:	

Date:	

General data:

Gender	
Age	
Height	
Weight	

Participant preparation checklist:

equipment	check
Marker attachment (for Mo-Cap)	
IMU sensors attachment	
EMG sensors attachment	

Maximum voluntary contraction (MVC).

Muscle group	Done	Note	File name
Quadriceps muscles			MVC
Hamstring muscles			MVC
Dorsiflexion muscles			MVC
Plantarflexion muscles			MVC

Static pose: stand still on treadmill for two to three seconds

Attempts	Done	Note	File name
Attempt 1			stand

Jumping: (counter movement) - On the force plates

Attempts	Done	Note	File name
Attempt 1 (3 times)			jump

- The correct way of doing squat jump will be expressed to the participants beforehand.

Squat (body weight): one attempt which includes five trial. – On the force plates

Attempts	Done	Note	File name
Attempt 1 (5 times)			squad

- The correct way of doing squat will be expressed to the participants beforehand.

Sit to stand – standard chair height:

Attempts	Done	Note	File name
Attempt 1 (5 times)			sit_stand

- The height of chair adjusted in a manner that participants sit on chair with 90 degree knee.

1

Ascending and descending steps: tare the force plates

Attempts	Done	Note	File name
Attempt 1 (5 times)			steps

- In this experiment the participant will be asked to ascend four stairs in standard height.

Walk to run transition: (speed variation 0 to 2.5 m/s) – a =0.05

Attempts	Time (sec)	Done	Note	File name
normal	50			Normal_walk
Keep walking	50			Keep_walking
Keep running	50			Keep_running

- The participant asked to change speed from normal speed to run and back to normal speed.

Gait experiment: (0 degree incline) – a = 0.02

	Time (sec)	Done	Note	File name
0 to 1 (m/s)	50			Walk_00deg
1 (m/s)	30			Walk_00deg
1 to 1.5 (m/s)	25			Walk_00deg
1.5 (m/s)	30			Walk_00deg
1.5 to 2 (m/s)	25			Walk_00deg
2 (m/s)	30			Walk_00deg
2 to 2.5 (m/s)	25			Walk_00deg
2.5 (m/s)	30			Walk_00deg

Total time: 245 sec = 4.1 min

	Done	Speed	Duration	File name
Self-selected speed walking			1 min	Selfspeed_00walk
Faster speed walking		**+0.5**	1 min	Selfspeed_faster_00walk
Slower speed walking		**-0.5**	1 min	Selfspeed_slower_00walk

Gait experiment: (+10 degree incline) - a = 0.02

	Time (sec)	Done	Note	File name
0 to 1 (m/s)	50			Walk_15deg
1 (m/s)	30			Walk_15deg
1 to 1.5 (m/s)	25			Walk_15deg
1.5 (m/s)	30			Walk_15deg
1.5 to 2 (m/s)	25			Walk_15deg
2 (m/s)	30			Walk_15deg
2 to 2.5 (m/s)	25			Walk_15deg
2.5 (m/s)	30			Walk_15deg

Total time: 245 sec = 4.1 min

	Done	speed	Duration	File name
Self-selected speed walking			1 min	Selfspeed_15deg

Gait experiment: (-10 degree incline) - a = 0.02

	Time (sec)	Done	Note	File name
0 to 1 (m/s)	50			Walk_-15deg
1 (m/s)	30			Walk_-15deg
1 to 1.5 (m/s)	25			Walk_-15deg
1.5 (m/s)	30			Walk_-15deg
1.5 to 2 (m/s)	25			Walk_-15deg
2 (m/s)	30			Walk_-15deg
2 to 2.5 (m/s)	25			Walk_-15deg
2.5 (m/s)	30			Walk_-15deg

Total time: 245 sec = 4.1 min

	Done	speed	Duration	File name
Self-selected speed walking			1 min	Selfspeed_-15deg

References

1. Thrift, A.G., et al., *Global stroke statistics: An update of mortality data from countries using a broad code of "cerebrovascular diseases"*. International Journal of Stroke, 2017. 12(8): p. 796-801.
2. Mozaffarian, D., et al., *Heart disease and stroke statistics—2016 update: a report from the American Heart Association.* circulation, 2016. 133(4): p. e38-e360.
3. Ekker, M.S., et al., *Stroke incidence in young adults according to age, subtype, sex, and time trends.* Neurology, 2019. 92(21): p. e2444-e2454.
4. Benjamin, E.J., et al., *Heart disease and stroke statistics—2019 update: a report from the American Heart Association.* Circulation, 2019. 139(10): p. e56-e528.
5. Benjamin, E.J., et al., *Heart disease and stroke statistics—2017 update: a report from the American Heart Association.* circulation, 2017. 135(10): p. e146-e603.
6. Ranta, A., *Projected stroke volumes to provide a 10-year direction for New Zealand stroke services.* NZ Med J, 2018. 131(1477): p. 15-28.
7. Lo, E.H., T. Dalkara, and M.A. Moskowitz, *Mechanisms, challenges and opportunities in stroke.* Nature reviews neuroscience, 2003. 4(5): p. 399-414.
8. Kwakkel, G., et al., *Effects of intensity of rehabilitation after stroke: a research synthesis.* Stroke, 1997. 28(8): p. 1550-1556.
9. Arya, K.N., et al., *Movement therapy induced neural reorganization and motor recovery in stroke: a review.* Journal of bodywork and movement therapies, 2011. 15(4): p. 528-537.
10. Feng, Z.-Y., et al., *Angle measurement based on gyroscope and accelerometer signal fusion.* Journal of Southwest China Normal University, 2011. 4: p. 36.
11. Zhang, Y., K. Chen, and J. Yi, *Rider trunk and bicycle pose estimation with fusion of force/inertial sensors.* IEEE Transactions on Biomedical engineering, 2013. 60(9): p. 2541-2551.
12. Tariq, M., P.M. Trivailo, and M. Simic, *EEG-based BCI control schemes for lower-limb assistive-robots.* Frontiers in human neuroscience, 2018. 12: p. 312.
13. Al-Quraishi, M.S., et al., *EEG-based control for upper and lower limb exoskeletons and prostheses: A systematic review.* Sensors, 2018. 18(10): p. 3342.
14. Fernandes, P.N., et al. *EMG-based Motion Intention Recognition for Controlling a Powered Knee Orthosis*. in *2019 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. 2019. IEEE.
15. He, H. and K. Kiguchi. *A study on EMG-based control of exoskeleton robots for human lower-limb motion assist*. in *2007 6th International special topic conference on information technology applications in biomedicine*. 2007. IEEE.
16. Kiguchi, K. and Y. Imada. *EMG-based control for lower-limb power-assist exoskeletons*. in *2009 IEEE workshop on robotic intelligence in informationally structured space*. 2009. IEEE.
17. Cavanagh, P.R. and P.V. Komi, *Electromechanical delay in human skeletal muscle under concentric and eccentric contractions.* European journal of applied physiology and occupational physiology, 1979. 42(3): p. 159-163.
18. Ferreira, A., et al., *Human-machine interfaces based on EMG and EEG applied to robotic systems.* Journal of NeuroEngineering and Rehabilitation, 2008. 5(1): p. 1-15.
19. Lloyd, D.G. and T.F. Besier, *An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo.* Journal of biomechanics, 2003. 36(6): p. 765-776.

20. Ao, D., R. Song, and J. Gao, *Movement performance of human–robot cooperation control based on EMG-driven hill-type and proportional models for an ankle power-assist exoskeleton robot.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2016. 25(8): p. 1125-1134.
21. Chen, J., et al., *Surface EMG based continuous estimation of human lower limb joint angles by using deep belief networks.* Biomedical Signal Processing and Control, 2018. 40: p. 335-342.
22. Jørgensen, H.S., et al., *Recovery of walking function in stroke patients: the Copenhagen Stroke Study.* Archives of physical medicine and rehabilitation, 1995. 76(1): p. 27-32.
23. Neumann, D.A., *Kinesiology of the musculoskeletal system-e-book: foundations for rehabilitation*. 2016: Elsevier Health Sciences.
24. Huang, H.-C., et al., *The impact of timing and dose of rehabilitation delivery on functional recovery of stroke patients.* Journal of the Chinese Medical Association, 2009. 72(5): p. 257-264.
25. Shah, N., A. Basteris, and F. Amirabdollahian, *Design parameters in multimodal games for rehabilitation.* GAMES FOR HEALTH: Research, Development, and Clinical Applications, 2014. 3(1): p. 13-20.
26. Langhorne, P., J. Bernhardt, and G. Kwakkel, *Stroke rehabilitation.* The Lancet, 2011. 377(9778): p. 1693-1702.
27. Horn, S.D., et al., *Stroke rehabilitation patients, practice, and outcomes: is earlier and more aggressive therapy better?* Archives of physical medicine and rehabilitation, 2005. 86(12): p. 101-114.
28. Gillen, G., *Stroke rehabilitation: a function-based approach*. 2015: Elsevier Health Sciences.
29. Wang, C., et al., *Quantitative assessment of upper-limb motor function for post-stroke rehabilitation based on motor synergy analysis and multi-modality fusion.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2020. 28(4): p. 943-952.
30. Lindsay, P., et al., *World Stroke Organization global stroke services guidelines and action plan.* International Journal of Stroke, 2014. 9: p. 4-13.
31. Dickstein, R., *Rehabilitation of gait speed after stroke: a critical review of intervention approaches.* Neurorehabilitation and neural repair, 2008. 22(6): p. 649-660.
32. Brenner, I., *Effects of passive exercise training in hemiplegic stroke patients: A mini-review*. Sports Med Rehabil J. 2018; 3 (3), 2018. 1036.
33. Eng, J.J., *Strength training in individuals with stroke.* Physiotherapy Canada. Physiotherapie Canada, 2004. 56(4): p. 189.
34. Feng, W. and S.R. Belagaje. *Recent advances in stroke recovery and rehabilitation*. in *Seminars in neurology*. 2013. Thieme Medical Publishers.
35. Hosseini, Z.-S., H. Peyrovi, and M. Gohari, *The effect of early passive range of motion exercise on motor function of people with stroke: a randomized controlled trial.* Journal of caring sciences, 2019. 8(1): p. 39.
36. Nuyens, G.E., et al., *Reduction of spastic hypertonia during repeated passive knee movements in stroke patients.* Archives of physical medicine and rehabilitation, 2002. 83(7): p. 930-935.
37. French, B., et al., *Repetitive task training for improving functional ability after stroke.* Cochrane database of systematic reviews, 2016(11).
38. Wolf, S.L., et al., *Repetitive task practice: a critical review of constraint-induced movement therapy in stroke.* The neurologist, 2002. 8(6): p. 325.
39. Kwakkel, G., et al., *Constraint-induced movement therapy after stroke.* The Lancet Neurology, 2015. 14(2): p. 224-234.
40. Smith, M.A. and M.R. Tomita, *Combined effects of telehealth and modified constraint-induced movement therapy for individuals with chronic hemiparesis.* International Journal of Telerehabilitation, 2020. 12(1): p. 51.
41. Forster, A. and J. Young, *Incidence and consequences offalls due to stroke: a systematic inquiry.* Bmj, 1995. 311(6997): p. 83-86.
42. Buurke, J., et al., *The effect of walking aids on muscle activation patterns during walking in stroke patients.* Gait & posture, 2005. 22(2): p. 164-170.
43. Kim, S.-j., et al., *Effects of stationary cycling exercise on the balance and gait abilities of chronic stroke patients.* Journal of physical therapy science, 2015. 27(11): p. 3529-3531.

44. Prout, E.C., et al., *Patient characteristics that influence enrollment and attendance in aerobic exercise early after stroke.* Archives of physical medicine and rehabilitation, 2015. 96(5): p. 823-830.
45. Barbeau, H. and M. Visintin, *Optimal outcomes obtained with body-weight support combined with treadmill training in stroke subjects.* Archives of physical medicine and rehabilitation, 2003. 84(10): p. 1458-1465.
46. Selves, C., G. Stoquart, and T. Lejeune, *Gait rehabilitation after stroke: review of the evidence of predictors, clinical outcomes and timing for interventions.* Acta Neurologica Belgica, 2020. 120(4): p. 783-790.
47. Rossano, C. and P. Terrier, *Visually-guided gait training in paretic patients during the first rehabilitation phase: study protocol for a randomized controlled trial.* Trials, 2016. 17(1): p. 1-9.
48. Rugelj, D., *Low back pain and other work-related musculoskeletal problems among physiotherapists.* Applied ergonomics, 2003. 34(6): p. 635-639.
49. Rushton, D., *Functional electrical stimulation and rehabilitation—an hypothesis.* Medical engineering & physics, 2003. 25(1): p. 75-78.
50. Nightingale, E., et al., *Benefits of FES gait in a spinal cord injured population.* Spinal cord, 2007. 45(10): p. 646-657.
51. Rabelo, M., et al., *Overview of FES-assisted cycling approaches and their benefits on functional rehabilitation and muscle atrophy.* Muscle Atrophy, 2018: p. 561-583.
52. Freeman, C.T., et al., *A model of the upper extremity using FES for stroke rehabilitation.* 2009.
53. Alon, G., A.F. Levitt, and P.A. McCarthy, *Functional electrical stimulation enhancement of upper extremity functional recovery during stroke rehabilitation: a pilot study.* Neurorehabilitation and neural repair, 2007. 21(3): p. 207-215.
54. Avan, A., et al., *Socioeconomic status and stroke incidence, prevalence, mortality, and worldwide burden: an ecological analysis from the Global Burden of Disease Study 2017.* BMC medicine, 2019. 17(1): p. 1-30.
55. Bhattacharjya, S., *Assessing the Functionality of a Portable Rehabilitation System Used for Self-Management of Upper Extremity Function by Individuals With Stroke.* American Journal of Occupational Therapy, 2018. 72(4_Supplement_1): p. 7211505134p1-7211505134p1.
56. Bao, G., et al., *Academic review and perspectives on robotic exoskeletons.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2019. 27(11): p. 2294-2304.
57. Singh, R.M., S. Chatterji, and A. Kumar. *A review on surface EMG based control schemes of exoskeleton robot in stroke rehabilitation*. in *2013 International Conference on Machine Intelligence and Research Advancement*. 2013. IEEE.
58. Sivan, M., et al., *Home-based Computer Assisted Arm Rehabilitation (hCAAR) robotic device for upper limb exercise after stroke: results of a feasibility study in home setting.* Journal of neuroengineering and rehabilitation, 2014. 11(1): p. 1-17.
59. Butler, A., et al., *Effect of home-based rehabilitation on access to cost effective therapy for rural veteran stroke survivors.* Archives of Physical Medicine and Rehabilitation, 2017. 98(10): p. e58-e59.
60. McCabe, J., et al., *Comparison of robotics, functional electrical stimulation, and motor learning methods for treatment of persistent upper extremity dysfunction after stroke: a randomized controlled trial.* Archives of physical medicine and rehabilitation, 2015. 96(6): p. 981-990.
61. Pehlivan, A.U., D.P. Losey, and M.K. O'Malley, *Minimal assist-as-needed controller for upper limb robotic rehabilitation.* IEEE Transactions on Robotics, 2015. 32(1): p. 113-124.
62. Mohammed, S., Y. Amirat, and H. Rifai, *Lower-limb movement assistance through wearable robots: State of the art and challenges.* Advanced Robotics, 2012. 26(1-2): p. 1-22.
63. Marchal-Crespo, L. and D.J. Reinkensmeyer, *Review of control strategies for robotic movement training after neurologic injury.* Journal of neuroengineering and rehabilitation, 2009. 6(1): p. 1-15.
64. Huo, W., et al., *Lower limb wearable robots for assistance and rehabilitation: A state of the art.* IEEE systems Journal, 2014. 10(3): p. 1068-1081.
65. Valiente, A., *Design of a quasi-passive parallel leg exoskeleton to augment load carrying for walking*. 2005, MASSACHUSETTS INST OF TECH CAMBRIDGE MEDIA LAB.

66. Kawamoto, H. and Y. Sankai, *Power assist method based on phase sequence and muscle force condition for HAL*. Advanced Robotics, 2005. 19(7): p. 717-734.
67. De Rossi, S.M.M., et al., *Sensing pressure distribution on a lower-limb exoskeleton physical human-machine interface*. Sensors, 2011. 11(1): p. 207-227.
68. Suzuki, K., et al., *Intention-based walking support for paraplegia patients with Robot Suit HAL*. Advanced Robotics, 2007. 21(12): p. 1441-1469.
69. Zanotto, D., et al. *Improving transparency of powered exoskeletons using force/torque sensors on the supporting cuffs*. in *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*. 2013. IEEE.
70. Bogue, R., *Robotic exoskeletons: a review of recent progress*. Industrial Robot: An International Journal, 2015.
71. Herr, H., *Exoskeletons and orthoses: classification, design challenges and future directions*. Journal of neuroengineering and rehabilitation, 2009. 6(1): p. 1-9.
72. Banala, S.K., et al., *Novel gait adaptation and neuromotor training results using an active leg exoskeleton*. IEEE/ASME Transactions on mechatronics, 2010. 15(2): p. 216-225.
73. Argall, B.D. *Turning assistive machines into assistive robots*. in *Quantum Sensing and Nanophotonic Devices XII*. 2015. International Society for Optics and Photonics.
74. Kazerooni, H. and R. Steger, *The Berkeley lower extremity exoskeleton*. 2006.
75. Bi, L. and C. Guan, *A review on EMG-based motor intention prediction of continuous human upper limb motion for human-robot collaboration*. Biomedical Signal Processing and Control, 2019. 51: p. 113-127.
76. Ferris, D.P. and C.L. Lewis. *Robotic lower limb exoskeletons using proportional myoelectric control*. in *2009 Annual international conference of the Ieee engineering in medicine and biology society*. 2009. IEEE.
77. Kawamoto, H. and Y. Sankai. *Comfortable power assist control method for walking aid by HAL-3*. in *IEEE International Conference on Systems, Man and Cybernetics*. 2002. IEEE.
78. Song, R., et al., *Assistive control system using continuous myoelectric signal in robot-aided arm training for patients after stroke*. IEEE transactions on neural systems and rehabilitation engineering, 2008. 16(4): p. 371-379.
79. Karavas, N., et al., *Tele-impedance based assistive control for a compliant knee exoskeleton*. Robotics and Autonomous Systems, 2015. 73: p. 78-90.
80. Luo, J., C. Liu, and C. Yang, *Estimation of EMG-based force using a neural-network-based approach*. IEEE Access, 2019. 7: p. 64856-64865.
81. Hincapie, J.G. and R.F. Kirsch, *Feasibility of EMG-based neural network controller for an upper extremity neuroprosthesis*. IEEE transactions on neural systems and rehabilitation engineering, 2008. 17(1): p. 80-90.
82. Wang, L. and T.S. Buchanan, *Prediction of joint moments using a neural network model of muscle activations from EMG signals*. IEEE transactions on neural systems and rehabilitation engineering, 2002. 10(1): p. 30-37.
83. Bao, T., et al. *Surface-EMG based wrist kinematics estimation using convolutional neural network*. in *2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. 2019. IEEE.
84. ULKIR, O., G. Gokmen, and E. KAPLANOGLU, *EMG signal classification using fuzzy logic*. Balkan Journal of Electrical and Computer Engineering, 2017. 5(2): p. 97-101.
85. Bouteraa, Y., et al., *A fuzzy logic architecture for rehabilitation robotic systems*. International Journal of Computers Communications & Control, 2020. 15(4).
86. Adiputra, D., et al., *Fuzzy logic control for ankle foot orthoses equipped with magnetorheological brake*. Jurnal Teknologi, 2016. 78(11).
87. Baldacchino, T., et al., *Simultaneous force regression and movement classification of fingers via surface EMG within a unified Bayesian framework*. Frontiers in bioengineering and biotechnology, 2018. 6: p. 13.

88. Liu, J., et al., *EMG-based real-time linear-nonlinear cascade regression decoding of shoulder, elbow, and wrist movements in able-bodied persons and stroke survivors*. IEEE Transactions on Biomedical Engineering, 2019. 67(5): p. 1272-1281.
89. Yokoyama, M., R. Koyama, and M. Yanagisawa, *An evaluation of hand-force prediction using artificial neural-network regression models of surface EMG signals for handwear devices*. Journal of Sensors, 2017. 2017.
90. Clarke, A.K., et al., *Deep learning for robust decomposition of high-density surface emg signals*. IEEE Transactions on Biomedical Engineering, 2020. 68(2): p. 526-534.
91. Koike, Y. and M. Kawato, *Estimation of dynamic joint torques and trajectory formation from surface electromyography signals using a neural network model*. Biological cybernetics, 1995. 73(4): p. 291-300.
92. Gui, K., H. Liu, and D. Zhang, *A practical and adaptive method to achieve EMG-based torque estimation for a robotic exoskeleton*. IEEE/ASME Transactions on Mechatronics, 2019. 24(2): p. 483-494.
93. Yu, H.-J., A. Lee, and Y. Choi, *Human elbow joint angle estimation using electromyogram signal processing*. IET signal processing, 2011. 5(8): p. 767-775.
94. Zhang, Q., R. Hosoda, and G. Venture. *Human joint motion estimation for electromyography (EMG)-based dynamic motion control*. in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2013. IEEE.
95. Chen, Y., X. Zhao, and J. Han, *Hierarchical projection regression for online estimation of elbow joint angle using EMG signals*. Neural Computing and Applications, 2013. 23(3): p. 1129-1138.
96. Muceli, S. and D. Farina, *Simultaneous and proportional estimation of hand kinematics from EMG during mirrored movements at multiple degrees-of-freedom*. IEEE transactions on neural systems and rehabilitation engineering, 2011. 20(3): p. 371-378.
97. Zhang, Q., et al., *Simultaneous and continuous estimation of shoulder and elbow kinematics from surface EMG signals*. Frontiers in neuroscience, 2017. 11: p. 280.
98. Akhtar, A., et al., *Estimation of distal arm joint angles from EMG and shoulder orientation for transhumeral prostheses*. Journal of Electromyography and Kinesiology, 2017. 35: p. 86-94.
99. Au, A.T. and R.F. Kirsch, *EMG-based prediction of shoulder and elbow kinematics in able-bodied and spinal cord injured individuals*. IEEE Transactions on rehabilitation engineering, 2000. 8(4): p. 471-480.
100. Ma, X., et al., *Continuous estimation of knee joint angle based on surface electromyography using a long short-term memory neural network and time-advanced feature*. Sensors, 2020. 20(17): p. 4966.
101. Al Kouzbary, H., et al., *Generating an Adaptive and Robust Walking Pattern for a Prosthetic Ankle-Foot by Utilizing a Nonlinear Autoregressive Network With Exogenous Inputs*. IEEE Transactions on Neural Networks and Learning Systems, 2021.
102. Jephil, P.B., et al., *Estimation of ankle joint torque and angle based on S-EMG signal for assistive rehabilitation robots*, in *Biomedical Signal Processing*. 2020, Springer. p. 31-47.
103. Xiong, D., et al., *Deep learning for EMG-based human-machine interaction: A review*. IEEE/CAA Journal of Automatica Sinica, 2021. 8(3): p. 512-533.
104. Wang, J., et al., *Surface electromyography based estimation of knee joint angle by using correlation dimension of wavelet coefficient*. IEEE Access, 2019. 7: p. 60522-60531.
105. Liu, Q., et al. *Knee joint angle prediction based on muscle synergy theory and generalized regression neural network*. in *2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. 2018. IEEE.
106. Tsinganos, P., et al. *Deep Learning in EMG-based Gesture Recognition*. in *PhyCS*. 2018.
107. Aung, Y.M. and A. Al-Jumaily, *Estimation of upper limb joint angle using surface EMG signal*. International Journal of Advanced Robotic Systems, 2013. 10(10): p. 369.
108. Anwar, T., Y.M. Aung, and A. Al Jumaily. *The estimation of knee joint angle based on generalized regression neural network (GRNN)*. in *2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)*. 2015. IEEE.

109. Hill, A.V., *The heat of shortening and the dynamic constants of muscle.* Proceedings of the Royal Society of London. Series B-Biological Sciences, 1938. 126(843): p. 136-195.
110. Hoy, M.G., F.E. Zajac, and M.E. Gordon, *A musculoskeletal model of the human lower extremity: the effect of muscle, tendon, and moment arm on the moment-angle relationship of musculotendon actuators at the hip, knee, and ankle.* Journal of biomechanics, 1990. 23(2): p. 157-169.
111. Durandau, G., D. Farina, and M. Sartori, *Robust real-time musculoskeletal modeling driven by electromyograms.* IEEE transactions on biomedical engineering, 2017. 65(3): p. 556-564.
112. Fleischer, C., C. Reinicke, and G. Hommel. *Predicting the intended motion with EMG signals for an exoskeleton orthosis controller.* in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2005. IEEE.
113. Sartori, M., et al. *A stiff tendon neuromusculoskeletal model of the knee*. in *2009 IEEE Workshop on Advanced Robotics and its Social Impacts*. 2009. IEEE.
114. Sartori, M., et al., *Estimation of musculotendon kinematics in large musculoskeletal models using multidimensional B-splines.* Journal of biomechanics, 2012. 45(3): p. 595-601.
115. Liu, L., et al. *EMG-driven model-based knee torque estimation on a variable impedance actuator orthosis*. in *2017 IEEE international conference on cyborg and bionic systems (CBS)*. 2017. IEEE.
116. Chandrapal, M., et al., *Investigating improvements to neural network based EMG to joint torque estimation.* Paladyn, Journal of Behavioral Robotics, 2011. 2(4): p. 185-192.
117. Hahn, M.E., *Feasibility of estimating isokinetic knee torque using a neural network model.* Journal of Biomechanics, 2007. 40(5): p. 1107-1114.
118. Morbidoni, C., et al., *A deep learning approach to EMG-based classification of gait phases during level ground walking.* Electronics, 2019. 8(8): p. 894.
119. Caesarendra, W., T. Tjahjowidodo, and D. Pamungkas. *EMG based classification of hand gestures using PCA and ANFIS*. in *2017 International Conference on Robotics, Biomimetics, and Intelligent Computational Systems (Robionetics)*. 2017. IEEE.
120. Ling-ling, C., et al. *Electromyographic movement pattern recognition based on random forest algorithm*. in *2015 34th Chinese Control Conference (CCC)*. 2015. IEEE.
121. Staporncharisit, S., et al., *Finger Angle estimation from Array EMG system using linear regression model with Independent Component Analysis.* Frontiers in neurorobotics, 2019. 13: p. 75.
122. Liu, J., et al., *EMG-based continuous and simultaneous estimation of arm kinematics in able-bodied individuals and stroke survivors.* Frontiers in neuroscience, 2017. 11: p. 480.
123. Gopura, R.A.R.C., K. Kiguchi, and Y. Li. *SUEFUL-7: A 7DOF upper-limb exoskeleton robot with muscle-model-oriented EMG-based control.* in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2009. IEEE.
124. Kiguchi, K., T. Tanaka, and T. Fukuda, *Neuro-fuzzy control of a robotic exoskeleton with EMG signals.* IEEE Transactions on fuzzy systems, 2004. 12(4): p. 481-490.
125. Rechy-Ramirez, E.J. and H. Hu, *Stages for Developing Control Systems using EMG and EEG signals: A survey.* School of Computer Science and Electronic Engineering, University of Essex, 2011: p. 1744-8050.
126. Huang, H.-P. and C.-Y. Chen. *Development of a myoelectric discrimination system for a multi-degree prosthetic hand*. in *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*. 1999. IEEE.
127. Oskoei, M.A. and H. Hu. *GA-based feature subset selection for myoelectric classification*. in *2006 IEEE international conference on robotics and biomimetics*. 2006. IEEE.
128. Englehart, K. and B. Hudgins, *A robust, real-time control scheme for multifunction myoelectric control.* IEEE transactions on biomedical engineering, 2003. 50(7): p. 848-854.
129. Phinyomark, A., C. Limsakul, and P. Phukpattaranont, *A novel feature extraction for robust EMG pattern recognition.* arXiv preprint arXiv:0912.3973, 2009.
130. Han, J.-S., et al. *New EMG pattern recognition based on soft computing techniques and its application to control of a rehabilitation robotic arm*. in *Proc. of 6th international conference on soft computing (IIZUKA2000)*. 2000.

131. Khezri, M. and M. Jahed, *Real-time intelligent pattern recognition algorithm for surface EMG signals*. Biomedical engineering online, 2007. 6(1): p. 1-12.
132. Gabor, D., *Theory of communication. Part 1: The analysis of information*. Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering, 1946. 93(26): p. 429-441.
133. Weeks, M., *Digital Signal Processing Using MATLAB & Wavelets added for testing purpose*. 2010: Jones & Bartlett Publishers.
134. Zecca, M., et al., *Control of multifunctional prosthetic hands by processing the electromyographic signal*. Critical Reviews™ in Biomedical Engineering, 2002. 30(4-6).
135. Englehart, K., et al., *Classification of the myoelectric signal using time-frequency based representations*. Medical engineering & physics, 1999. 21(6-7): p. 431-438.
136. Phinyomark, A., P. Phukpattaranont, and C. Limsakul, *Feature reduction and selection for EMG signal classification*. Expert systems with applications, 2012. 39(8): p. 7420-7431.
137. Güler, N.F. and S. Koçer, *Classification of EMG signals using PCA and FFT*. Journal of medical systems, 2005. 29(3): p. 241-250.
138. Phinyomark, A., et al., *Application of linear discriminant analysis in dimensionality reduction for hand motion classification*. Measurement Science Review, 2012. 12(3): p. 82-89.
139. Naik, G.R., *Advances in Principal Component Analysis: Research and Development*. 2017: Springer.
140. Davis, B.L. and C.L. Vaughan, *Phasic behavior of EMG signals during gait: use of multivariate statistics*. Journal of Electromyography and Kinesiology, 1993. 3(1): p. 51-60.
141. Zebari, R., et al., *A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction*. Journal of Applied Science and Technology Trends, 2020. 1(2): p. 56-70.
142. Qi, J., et al., *Surface EMG hand gesture recognition system based on PCA and GRNN*. Neural Computing and Applications, 2020. 32(10): p. 6343-6351.
143. Zhang, D., et al. *A comparative study on PCA and LDA based EMG pattern recognition for anthropomorphic robotic hand*. in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. 2014. IEEE.
144. Rabin, N., et al., *Classification of human hand movements based on EMG signals using nonlinear dimensionality reduction and data fusion techniques*. Expert Systems with Applications, 2020. 149: p. 113281.
145. Anam, K. and A. Al-Jumaily. *A novel extreme learning machine for dimensionality reduction on finger movement classification using sEMG*. in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*. 2015. IEEE.
146. Junior, J.J.A.M., et al., *Feature selection and dimensionality reduction: An extensive comparison in hand gesture classification by sEMG in eight channels armband approach*. Biomedical Signal Processing and Control, 2020. 59: p. 101920.
147. Abbaspour, S., et al., *Evaluation of surface EMG-based recognition algorithms for decoding hand movements*. Medical & biological engineering & computing, 2020. 58(1): p. 83-100.
148. Wagner, J., J. Kim, and E. André. *From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification*. in *2005 IEEE international conference on multimedia and expo*. 2005. IEEE.
149. Liu, J., *Feature dimensionality reduction for myoelectric pattern recognition: A comparison study of feature selection and feature projection methods*. Medical engineering & physics, 2014. 36(12): p. 1716-1720.
150. Amin, P. and A.M. Khan, *A Study on the Effect of Dimensionality Reduction on Classification Accuracy of Myoelectric Control Systems*. Advances in VLSI, Signal Processing, Power Electronics, IoT, Communication and Embedded Systems: Select Proceedings of VSPICE 2020, 2021. 752: p. 25.
151. Christ, M., et al., *Time series feature extraction on basis of scalable hypothesis tests (tsfresh–a python package)*. Neurocomputing, 2018. 307: p. 72-77.
152. Christ, M., A.W. Kempa-Liehr, and M. Feindt, *Distributed and parallel time series feature extraction for industrial big data applications*. arXiv preprint arXiv:1610.07717, 2016.

153. Zhang, X., et al., *A regression-based framework for quantitative assessment of muscle spasticity using combined EMG and inertial data from wearable sensors.* Frontiers in neuroscience, 2019. 13: p. 398.
154. Bangaru, S.S., et al., *ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors.* Automation in Construction, 2021. 126: p. 103653.
155. Xia, P., J. Hu, and Y. Peng, *EMG-based estimation of limb movement using deep learning with recurrent convolutional neural networks.* Artificial organs, 2018. 42(5): p. E67-E77.
156. Besier, T.F., et al., *Repeatability of gait data using a functional hip joint centre and a mean helical knee axis.* Journal of biomechanics, 2003. 36(8): p. 1159-1168.
157. Stegeman, D. and H. Hermens, *Standards for surface electromyography: The European project Surface EMG for non-invasive assessment of muscles (SENIAM).* Enschede: Roessingh Research and Development, 2007. 10: p. 8-12.
158. Delp, S.L., et al., *OpenSim: open-source software to create and analyze dynamic simulations of movement.* IEEE transactions on biomedical engineering, 2007. 54(11): p. 1940-1950.
159. Falisse, A., et al., *OpenSim versus human body model: a comparison study for the lower limbs during gait.* Journal of Applied Biomechanics, 2018. 34(6): p. 496-502.
160. Bakke, D. and T. Besier, *Shape model constrained scaling improves repeatability of gait data.* Journal of Biomechanics, 2020. 107: p. 109838.
161. Bisseling, R.W. and A.L. Hof, *Handling of impact forces in inverse dynamics.* Journal of biomechanics, 2006. 39(13): p. 2438-2444.
162. Schmidt, A., *Movement pattern recognition in basketball free-throw shooting.* Human movement science, 2012. 31(2): p. 360-382.
163. Rupérez, M.J., et al., *Artificial neural networks for predicting dorsal pressures on the foot surface while walking.* Expert Systems with Applications, 2012. 39(5): p. 5349-5357.
164. He, W., Z. Li, and C.P. Chen, *A survey of human-centered intelligent robots: issues and challenges.* IEEE/CAA Journal of Automatica Sinica, 2017. 4(4): p. 602-609.
165. Teban, T.-A., et al. *Recurrent dynamic neural network model for myoelectric-based control of a prosthetic hand.* in *2016 Annual IEEE Systems Conference (SysCon)*. 2016. IEEE.
166. Nazmi, N., et al. *Fuzzy logic for walking patterns based on surface electromyography signals with different membership functions.* in *2016 6th International Workshop on Computer Science and Engineering, WCSE*. 2016.
167. Subasi, A., M. Yilmaz, and H.R. Ozcalik, *Classification of EMG signals using wavelet neural network.* Journal of neuroscience methods, 2006. 156(1-2): p. 360-367.
168. Stegeman, D. and H. Hermens, *Standards for surface electromyography: The European project Surface EMG for non-invasive assessment of muscles (SENIAM).* Enschede: Roessingh Research and Development, 2007: p. 108-12.
169. Sartori, M., D. Farina, and D.G. Lloyd, *Hybrid neuromusculoskeletal modeling to best track joint moments using a balance between muscle excitations derived from electromyograms and optimization.* Journal of biomechanics, 2014. 47(15): p. 3613-3621.
170. Moré, J.J., *The Levenberg-Marquardt algorithm: implementation and theory*, in *Numerical analysis*. 1978, Springer. p. 105-116.
171. Huang, R., et al., *Learning physical human–robot interaction with coupled cooperative primitives for a lower exoskeleton.* IEEE Transactions on Automation Science and Engineering, 2019. 16(4): p. 1566-1574.
172. Li, Q., et al., *Contact force/torque control based on viscoelastic model for stable bipedal walking on indefinite uneven terrain.* IEEE Transactions on Automation Science and Engineering, 2019. 16(4): p. 1627-1639.
173. Wu, X., et al., *Individualized gait pattern generation for sharing lower limb exoskeleton robot.* IEEE Transactions on Automation Science and Engineering, 2018. 15(4): p. 1459-1470.
174. Sartori, M., et al., *Toward modeling locomotion using electromyography-informed 3D models: application to cerebral palsy.* Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 2017. 9(2): p. e1368.

175. Pizzolato, C., et al., *CEINMS: A toolbox to investigate the influence of different neural control solutions on the prediction of muscle excitation and joint moments during dynamic motor tasks.* Journal of biomechanics, 2015. 48(14): p. 3929-3936.
176. Zhang, L., et al., *Ankle joint torque estimation using an emg-driven neuromusculoskeletal model and an artificial neural network model.* IEEE Transactions on Automation Science and Engineering, 2020. 18(2): p. 564-573.
177. Cui, C., et al. *sEMG-based prediction of human lower extremity movements by using a dynamic recurrent neural network*. in *2016 Chinese Control and Decision Conference (CCDC)*. 2016. IEEE.
178. Xiong, B., et al., *Intelligent prediction of human lower extremity joint moment: an artificial neural network approach.* IEEE Access, 2019. 7: p. 29973-29980.
179. Su, C., et al., *Ankle Joint Torque Prediction Based on Surface Electromyographic and Angular Velocity Signals.* IEEE Access, 2020. 8: p. 217681-217687.
180. Moreira, L., et al. *AI-based reference ankle joint torque trajectory generation for robotic gait assistance: First steps*. in *2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. 2020. IEEE.
181. Ullauri, J.B., et al. *On the EMG-based torque estimation for humans coupled with a force-controlled elbow exoskeleton*. in *2015 International Conference on Advanced Robotics (ICAR)*. 2015. IEEE.
182. Kempa-Liehr, A.W., et al. *Feature engineering workflow for activity recognition from synchronized inertial measurement units*. in *Asian Conference on Pattern Recognition*. 2019. Springer.
183. Chin, A., et al., *The off-break and "doosra": Kinematic variations of elite and sub-elite bowlers in creating ball spin in cricket bowling.* Sports Biomechanics, 2009. 8(3): p. 187-198.
184. Zhang, J., et al. *The MAP client: user-friendly musculoskeletal modelling workflows*. in *International Symposium on Biomedical Simulation*. 2014. Springer.
185. Zarshenas, H., et al. *Ankle torque forecasting using time-delayed neural networks*. in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2020. IEEE.
186. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the Royal statistical society: series B (Methodological), 1995. 57(1): p. 289-300.
187. Kennedy, A., et al., *Modelling the projected separation of microlensing events using systematic time-series feature engineering.* Astronomy and Computing, 2021. 35: p. 100460.
188. Schreiber, T. and A. Schmitz, *Discrimination power of measures for nonlinearity in a time series.* Physical Review E, 1997. 55(5): p. 5443.
189. Pedrinelli, A., L.E. Garcez-Leme, and R.d.S.A. Nobre, *The effect of physical training on locomotive apparatus in elderly people.* Revista brasileira de ortopedia, 2009. 44(2): p. 96-101.
190. Liu, Z., et al. *A new trajectory determination method for robot-assisted ankle ligament rehabilitation*. in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019. IEEE.
191. Mattacola, C.G. and M.K. Dwyer, *Rehabilitation of the ankle after acute sprain or chronic instability.* Journal of athletic training, 2002. 37(4): p. 413.
192. Dietz, V., G. Colombo, and L. Jensen, *Locomotor activity in spinal man.* The lancet, 1994. 344(8932): p. 1260-1263.
193. Zhang, M., et al., *Adaptive trajectory tracking control of a parallel ankle rehabilitation robot with joint-space force distribution.* IEEE Access, 2019. 7: p. 85812-85820.
194. Gassert, R. and V. Dietz, *Rehabilitation robots for the treatment of sensorimotor deficits: a neurophysiological perspective.* Journal of neuroengineering and rehabilitation, 2018. 15(1): p. 1-15.
195. Yao, S., et al., *Adaptive admittance control for an ankle exoskeleton using an EMG-driven musculoskeletal model.* Frontiers in neurorobotics, 2018. 12: p. 16.
196. Durandau, G., et al., *Voluntary control of wearable robotic exoskeletons by patients with paresis via neuromechanical modeling.* Journal of neuroengineering and rehabilitation, 2019. 16(1): p. 1-18.

197. Sartori, M., et al., *Robust simultaneous myoelectric control of multiple degrees of freedom in wrist-hand prostheses by real-time neuromusculoskeletal modeling.* Journal of neural engineering, 2018. 15(6): p. 066026.
198. Zhang, M., et al., *A robot-driven computational model for estimating passive ankle torque with subject-specific adaptation.* IEEE Transactions on Biomedical Engineering, 2015. 63(4): p. 814-821.
199. Li, Z., et al., *Hybrid brain/muscle signals powered wearable walking exoskeleton enhancing motor ability in climbing stairs activity.* IEEE Transactions on Medical Robotics and Bionics, 2019. 1(4): p. 218-227.
200. Li, Z., et al., *Adaptive neural control of a kinematically redundant exoskeleton robot using brain–machine interfaces.* IEEE transactions on neural networks and learning systems, 2018. 30(12): p. 3558-3571.
201. Erdemir, A., et al., *Model-based estimation of muscle forces exerted during movements.* Clinical biomechanics, 2007. 22(2): p. 131-154.
202. Liu, H.-J. and K.-Y. Young, *An adaptive upper-arm EMG-based robot control system.* International Journal of Fuzzy Systems, 2010. 12(3): p. 181-189.
203. Liu, H.-J. and K.-Y. Young, *Upper-limb EMG-based robot motion governing using empirical mode decomposition and adaptive neural fuzzy inference system.* Journal of Intelligent & Robotic Systems, 2012. 68(3): p. 275-291.
204. Rahatabad, F.N., et al., *A fuzzy-genetic model for estimating forces from electromyographical activity of antagonistic muscles due to planar lower arm movements: the effect of nonlinear muscle properties.* Biosystems, 2012. 107(1): p. 56-63.
205. Zhai, X., et al., *Self-recalibrating surface EMG pattern recognition for neuroprosthesis control based on convolutional neural network.* Frontiers in neuroscience, 2017. 11: p. 379.
206. Montgomery, J.R. and A.M. Grabowski, *The contributions of ankle, knee and hip joint work to individual leg work change during uphill and downhill walking over a range of speeds.* Royal Society open science, 2018. 5(8): p. 180550.
207. Pan, L., D.L. Crouch, and H. Huang, *Comparing EMG-based human-machine interfaces for estimating continuous, coordinated movements.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2019. 27(10): p. 2145-2154.
208. Bagherzadeh-Khiabani, F., et al., *A tutorial on variable selection for clinical prediction models: feature selection methods in data mining could improve the results.* Journal of clinical epidemiology, 2016. 71: p. 76-85.
209. Zhang, Z. *Improved adam optimizer for deep neural networks*. in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*. 2018. IEEE.
210. Géron, A., *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. 2019: O'Reilly Media.
211. Pedregosa, F., et al., *Scikit-learn: Machine learning in Python.* the Journal of machine Learning research, 2011. 12: p. 2825-2830.
212. Liaw, A. and M. Wiener, *Classification and regression by randomForest.* R news, 2002. 2(3): p. 18-22.
213. Li, Z., et al., *Estimation of knee movement from surface emg using random forest with principal component analysis.* Electronics, 2020. 9(1): p. 43.
214. Gupta, A., et al. *3D pose from motion for cross-view action recognition via non-linear circulant temporal encoding.* in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014.
215. Kasap, Z. and N. Magnenat-Thalmann, *Intelligent virtual humans with autonomy and personality: State-of-the-art*, in *New Advances in Virtual Humans*. 2008, Springer. p. 43-84.
216. DiCicco, M., L. Lucas, and Y. Matsuoka. *Comparison of control strategies for an EMG controlled orthotic exoskeleton for the hand.* in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*. 2004. ieee.
217. Mukhopadhyay, A.K. and S. Samui, *An experimental study on upper limb position invariant EMG signal classification based on deep neural network.* Biomedical signal processing and control, 2020. 55: p. 101669.

218. Huang, H., et al., *Continuous locomotion-mode identification for prosthetic legs based on neuromuscular–mechanical fusion.* IEEE Transactions on Biomedical Engineering, 2011. 58(10): p. 2867-2875.
219. Pew, C. and G.K. Klute, *Turn intent detection for control of a lower limb prosthesis.* IEEE Transactions on Biomedical Engineering, 2017. 65(4): p. 789-796.
220. Young, A., T. Kuiken, and L. Hargrove, *Analysis of using EMG and mechanical sensors to enhance intent recognition in powered lower limb prostheses.* Journal of neural engineering, 2014. 11(5): p. 056021.
221. Zia ur Rehman, M., et al., *Stacked sparse autoencoders for EMG-based classification of hand motions: A comparative multi day analyses between surface and intramuscular EMG.* Applied Sciences, 2018. 8(7): p. 1126.
222. Park, S., et al., *Hierarchical motion segmentation through sEMG for continuous lower limb motions.* IEEE Robotics and Automation Letters, 2019. 4(4): p. 4402-4409.
223. Wang, J., L. Qi, and X. Wang. *Surface EMG signals based motion intent recognition using multi-layer ELM.* in *LIDAR Imaging Detection and Target Recognition 2017*. 2017. International Society for Optics and Photonics.
224. Clancy, E.A., et al., *Identification of constant-posture EMG–torque relationship about the elbow using nonlinear dynamic models.* IEEE Transactions on Biomedical Engineering, 2011. 59(1): p. 205-212.
225. Ding, Q., J. Han, and X. Zhao, *Continuous estimation of human multi-joint angles from sEMG using a state-space model.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2016. 25(9): p. 1518-1528.
226. Han, J., et al., *A state-space EMG model for the estimation of continuous joint movements.* IEEE Transactions on Industrial Electronics, 2015. 62(7): p. 4267-4275.
227. Meng, W., et al. *An EMG-based force prediction and control approach for robot-assisted lower limb rehabilitation*. in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2014. IEEE.
228. Zhao, Y., et al., *An EMG-driven musculoskeletal model for estimating continuous wrist motion.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2020. 28(12): p. 3113-3120.
229. Rajagopal, A., et al., *Full-body musculoskeletal model for muscle-driven simulation of human gait.* IEEE transactions on biomedical engineering, 2016. 63(10): p. 2068-2079.
230. Ibitoye, M.O., et al., *Estimation of electrically-evoked knee torque from mechanomyography using support vector regression.* Sensors, 2016. 16(7): p. 1115.
231. Youn, W. and J. Kim, *Feasibility of using an artificial neural network model to estimate the elbow flexion force from mechanomyography.* Journal of neuroscience methods, 2011. 194(2): p. 386-393.
232. Chéron, G., et al., *A dynamic recurrent neural network for multiple muscles electromyographic mapping to elevation angles of the lower limb in human locomotion.* Journal of neuroscience methods, 2003. 129(2): p. 95-104.
233. Huang, Y., et al., *Real-time intended knee joint motion prediction by deep-recurrent neural networks.* IEEE Sensors Journal, 2019. 19(23): p. 11503-11509.
234. Raj, R. and K. Sivanandan, *Elbow joint angle and elbow movement velocity estimation using NARX-multiple layer perceptron neural network model with surface EMG time domain parameters.* Journal of back and musculoskeletal rehabilitation, 2017. 30(3): p. 515-525.
235. Gupta, R., I.S. Dhindsa, and R. Agarwal, *Continuous angular position estimation of human ankle during unconstrained locomotion.* Biomedical Signal Processing and Control, 2020. 60: p. 101968.
236. Li, Q., et al. *sEMG based joint angle estimation of lower limbs using LS-SVM.* in *International Conference on Neural Information Processing*. 2013. Springer.
237. Meng, M., et al. *EMG signals based gait phases recognition using hidden Markov models*. in *The 2010 IEEE International Conference on Information and Automation*. 2010. IEEE.

238. Razin, Y.S., et al., *Predicting task intent from surface electromyography using layered hidden Markov models*. IEEE Robotics and Automation Letters, 2017. 2(2): p. 1180-1185.
239. Li, G., T. Liu, and J. Yi, *Wearable sensor system for detecting gait parameters of abnormal gaits: A feasibility study*. IEEE Sensors Journal, 2018. 18(10): p. 4234-4241.
240. Yi, J., et al., *Kinematic modeling and analysis of skid-steered mobile robots with applications to low-cost inertial-measurement-unit-based motion estimation*. IEEE transactions on robotics, 2009. 25(5): p. 1087-1097.
241. Zhang, L., et al., *Ankle Joint Torque Estimation Using an EMG-Driven Neuromusculoskeletal Model and an Artificial Neural Network Model*. IEEE Transactions on Automation Science and Engineering, 2021. 18(2): p. 564-573.
242. Kwon, S. and J. Kim, *Real-Time Upper Limb Motion Estimation From Surface Electromyography and Joint Angular Velocities Using an Artificial Neural Network for Human–Machine Cooperation*. IEEE Transactions on Information Technology in Biomedicine, 2011. 15(4): p. 522-530.
243. Dempsey, A.R., et al., *The effect of technique change on knee loads during sidestep cutting*. Medicine & Science in Sports & Exercise, 2007. 39(10): p. 1765-1773.
244. Ziegler, J., H. Gattringer, and A. Mueller. *Classification of gait phases based on bilateral emg data using support vector machines*. in *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*. 2018. IEEE.
245. Tuncer, T., S. Dogan, and A. Subasi, *Surface EMG signal classification using ternary pattern and discrete wavelet transform based feature extraction for hand movement recognition*. Biomedical Signal Processing and Control, 2020. 58: p. 101872.
246. Smith, L.H., T.A. Kuiken, and L.J. Hargrove, *Evaluation of linear regression simultaneous myoelectric control using intramuscular EMG*. IEEE Transactions on Biomedical Engineering, 2015. 63(4): p. 737-746.
247. Krasoulis, A., S. Vijayakumar, and K. Nazarpour. *Evaluation of regression methods for the continuous decoding of finger movement from surface EMG and accelerometry*. in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*. 2015. IEEE.
248. Mahsa, M., *A Semi-Passive Clutched Elastic Ankle Exoskeleton to Augment Human Locomotion*, in *Auckland Bioengineering Institute*. 2021, University of Auckland: Auckland, New Zealand.