Automated Foot Strike Identification and Fall Risk Classification for People with Lower Limb Amputations Using Smartphone Sensor Signals from 2 and 6-Minute Walk Tests

Pascale Juneau

Thesis submitted to the Faculty of Engineering
in partial fulfillment of requirements for the degree of

Master of Applied Science

in Biomedical Engineering
Faculty of Engineering

uOttawa

Ottawa Carleton Institute for Biomedical Engineering
University of Ottawa
Ottawa, Ontario

Abstract

Artificial intelligence (AI) algorithms for gait analysis rely on properly identified foot strikes for step-based feature calculation. Smartphone signals collected during movement assessments, such as the 6-minute walk test (6MWT), have been used to train AI models for foot strike identification and fall risk classification in able-bodied populations. However, there is limited research in populations with more asymmetrical gait. People with lower limb amputation can have high gait variability, adversely affecting automatic step detection algorithms. Hence, fall risk models for lower limb amputees have relied on manual foot strike labelling to calculate step-based features for model training, which is inefficient and impractical for clinical use.

In this thesis, decision tree and long-short term memory (LSTM) models were developed, optimized, and their performance compared for automated foot strike identification in an amputee population. Eighty people with lower limb amputations (27 fallers, 53 non-fallers) completed a 6MWT with a smartphone at the posterior pelvis. Automated and manually labelled foot strikes from the full 6MWT and from the first two minutes of data were used to calculate step-based features. A random forest model was used to classify fall risk. The best foot strike identification model was an LSTM with 100 hidden nodes in the LSTM layer, 50 hidden nodes in the dense layer, and batch size of 64 (99.0% accuracy, 86.4% sensitivity, 99.4% specificity, 82.7% precision). Automated foot strikes from the full 6MWT data correctly classified more fallers (55.6% versus 48.1%), whereas automated foot strikes from 2-minute data classified more non-fallers (90.6% versus 81.1%). Feature calculation using manually labelled foot strikes resulted in the best overall performance (80.0% accuracy, 55.6% sensitivity, 92.5% specificity).

This research created a novel method for automated foot strike identification in lower limb amputees that is equivalent to manual labelling and demonstrated that automated foot strikes can be used to calculate step-based features for fall risk classification. Integration of the foot strike identification model into a smartphone application could allow for immediate stride analysis after completing a 6MWT; however, fall risk classification model improvement is recommended to enhance clinical viability.

Table of Contents

List of Figures

List of Tables

Abbreviations and Definitions

2M-FS	2-Minute Walk Test Foot Strike
2MWT	2-Minute Walk Test
6M-FS	6-Minute Walk Test Foot Strike
6MWT	6-Minute Walk Test
AI	Artificial Intelligence
A-FS	Automated Foot Strike
ANN	Artificial Neural Network
AP	Anterior-Posterior
CFS	Correlation-based Feature Selection
CNN	Convolutional Neural Network
DT	Decision Tree
EMG	Electromyography
FFT	Fast Fourier Transform
FSR	Force Sensing Resistors
FQFFT	First Quartile of Fourier Transform
HMM	Hidden Markov Model
LSTM	Long Short-Term Memory
MDC	Minimal Detectable Change
M-FS	Manually Labelled Foot Strike
ML	Medio-lateral
REOH	Ratio of Even to Odd Harmonics
RMS	Root Mean Squared
RNN	Recurrent Neural Network
SVM	Support Vector Machine
TCN	Temporal Convolutional Neural Network
TOHRC	The Ottawa Hospital Rehabilitation Centre
TUG	Timed-Up and Go test

Acknowledgments

This thesis is dedicated to Madeleine Muriel Kort and Joni Adele Kort, and to their Nannie and Poppa, my Aunt Diane and Uncle Kees.

I would like to thank my supervisors, Dr. Edward D. Lemaire and Dr. Natalie Baddour for their advice and encouragement. I relied on their guidance when I needed to adjust my thesis due to the COVID-19 pandemic, and they were always available to support me through the transition. I'd also like to thank them for the opportunity to learn through hands on projects and training programs and providing their expertise.

I would also like to thank Dr. Helena Burger and Dr. Andrej Bavec for their assistance with data collection and their clinical expertise, and Natural Sciences and Engineering Research Council of Canada Discovery Grant and CREATE-READi grants for funding the research in this thesis.

I would like to thank my parents, Marc, and Dawn Juneau, for their love and support, I would not have been able to complete this thesis without them. My parents were always available when I needed them, whether it was suppling baked goods and home cooked meals, being there when I needed to talk through a problem, reminding me to take breaks to disconnect to relax and enjoy nature and to visit friends and family. I would like to thank my sisters, Alèque and Danielle Juneau. They are the best sisters a youngest sibling could ask for; they are my role models. The are intelligent, strong, and beautiful women who have taught me so much over the years, individually and through shared experience. I would like to thank my partner, Brandon Cardenas, for always being there when I needed a laugh.

Finally, thank you to all my friends and colleagues, some of whom I have not yet been able to meet in person, who I have worked on other projects with me these past years at The Ottawa Hospital Rehabilitation Centre, The Children's Hospital of Eastern Ontario, and the University of Ottawa.

1 Introduction

Foot strike identification is important when analyzing gait since foot strikes define the beginning and end of a gait cycle and are used to calculate stride parameters that relate to a person's mobility and functional level. Wearable sensors, such as accelerometers, gyroscopes, or magnetometers, are an increasingly common alternative to costlier and time-consuming techniques for gait analysis. These sensors can be combined with clinical movement assessments to provide more data to the clinician for decision-making. In general, gait event detection techniques use data from multiple sensors, requiring additional set-up time, with the most common placement on the shank or ankle/foot [1], which can possibly affect the person's walking pattern. A smartphone placed at the lower back can be used to collect data during walking that can be used for additional clinical analysis. For example, 3D signals collected during a 6-minute walk test (6MWT) from a single smartphone placed at the posterior pelvis were used with a rule-based algorithm to identify foot strikes in able-bodied participants [2]. These foot strikes could then be used to provide immediate results such as right and left step-time, stride time, and cadence.

While these techniques have been successful for able-bodied people, gait analysis is frequently required in populations that present with gait deviations, making automated step detection more difficult. For example, people with lower limb amputations have greater gait variability that can lead to instability during walking and increases their risk of falling compared to other rehabilitation patients. Steinberg et al. [3] reported ~32% fall incidence in lower limb amputees, with fall incidence increasing during in-patient rehabilitation compared to the post-surgery recovery period. Falls can cause a number of injuries, including fractures, amputated limb trauma, and can increase the amount of time a person spends in hospital or a rehabilitation facility. A fall can also affect a person's confidence while walking, increasing their fear of falling even if the fall did not result in injury [4]. People with lower limb amputations are at a greater risk of falls than age-matched able-bodied participants, despite the majority of fall risk analysis being focused on older adults [5, 6]. Early identification of people with elevated fall risk is necessary to appropriately implement fall prevention strategies, thereby minimizing potential future falls and injuries and reducing the burden on the healthcare system. Currently, clinicians use a combination of medical history and mobility assessments, such as the Timed-up and Go (TUG), to identify fall risk [7]. Often, multiple tests are required to identify those at risk. Predictive fall risk models in lower limb amputees have relied on manual foot strike labelling for feature calculation [8], so automated foot strike identification in this population would expedite clinical decision-making to prevent injury.

Artificial intelligence (AI) algorithms have been used for gait analysis in multiple disease populations. For example, a decision tree (machine learning approach) has been used in the past for clinical decision making and gait phase recognition [9]. A more complex subset of machine learning, called deep learning, has also been investigated for gait analysis. Deep learning approaches, such as recurrent neural networks (RNN), perform well on large data sets, requiring more training time but often result in higher accuracy [10]. A popular type of RNN is the long short-term memory (LSTM). LSTM contains feedforward and feedback components, like the RNN, but are also equipped with a forget gate that regulates information flow in dataset that have gaps between significant events, such as during gait. While both machine learning and deep learning approaches have been successful for gait analysis in other populations, their ability to classify foot strikes in lower limb amputees using smartphone signals has not yet been evaluated or compared.

This thesis developed a novel smartphone sensor-based algorithm for automated foot strike detection in lower limb amputees using data from a 6MWT. The clinical usefulness of automated foot strikes compared to manually labelled foot strikes was evaluated by comparing stride parameters. Automated foot strikes were subsequently used for step-based feature calculation to validate a previously developed fall risk classifier for lower limb amputees.

1.1 Rationale

Research has been lacking for foot strike identification and fall risk classification with lower limb amputees, despite their elevated risk of injury and falls. When algorithms are developed for use in lower limb amputee populations, for example gait-phase detection for microprocessor controlled prostheses, these algorithms are often trained on data from able-bodied participants [9]. The limited research for foot strike identification in lower limb amputees frequently combines signal data from wearable sensors and AI where the participant was required to wear multiple sensors [11, 12] or a single sensor located on the lower limb [13, 14]. These placements can be awkward or uncomfortable for the participant and may cause them to walk differently. A single sensor at the lower back is an acceptable placement for data collection and is more appropriate for clinical applications.

The 6MWT is a commonly used movement assessment that evaluates functional capacity. However, biomechanical data from a 6MWT can be used to extract additional clinical information. One study investigating foot strike identification in lower limb amputees applied a rule-based algorithm previously developed for able-bodied participants [2]; however, when the algorithm was applied to lower limb amputee gait, this accuracy decreased to ~87% [15]. More complex AI algorithms have been used for gait analysis in other populations who have increased gait variability [16, 17, 18, 19], but the literature is limited for lower limb amputee populations. Signal data from a 6MWT has also been used to calculate features for fall risk classification in both able-bodied and lower limb amputee participants [8, 20]. However, manual foot strike labelling was required for step-based features calculated for lower limb amputees, further demonstrating the need for automated foot strike detection. Additionally, the 2MWT is a viable option for those who cannot complete the full 6MWT, and is commonly completed by people with lower limb amputation. However, automated foot strike detection for the 2MWT in lower limb amputees has not yet been evaluated.

The main hypothesis for this thesis is that smartphone signals from a 6MWT can be used to train an AI model for automated foot strike detection and that automated foot strikes can be used to calculate step-based features for fall risk classification in lower limb amputees.

1.2 Objectives

1) Create and evaluate a viable AI model for automatically identifying foot strikes during walking in people with lower limb amputation using smartphone sensor data collected during a 6MWT, comparing machine learning and deep learning approaches.
    a. Hypothesis: A deep learning approach will classify foot strikes and non-foot strikes events with greater accuracy, sensitivity, and specificity than a machine learning approach [10] or previous rule-based approaches [15].
    b. Hypothesis: Stride parameters (step time, stride time, etc.) calculated from automated foot strikes will be equivalent to parameters calculated from manually labelled foot strikes.
2) Evaluate a random forest fall risk classification model for lower limb amputees using automated foot strikes and manually labelled foot strikes and smartphone signals collected

from a 6MWT for step-based feature calculation.

   a. Hypothesis: Step-based features calculated from automated foot strikes will result in fall risk classification with equivalent with accuracy, sensitivity, and specificity to manually labelled foot strikes [8].

3) Evaluate the viability of automated foot strike identification and fall risk classification models for lower limb amputees for the 2MWT.

   a. Hypothesis: An LSTM trained on smartphone signals from only 2 minutes of data will result in lower 2MWT foot strike classification accuracy, sensitivity, and specificity than an LSTM trained on smartphone signals from 6 minutes of data.
   b. Hypothesis: Stride parameters (step time, stride time, etc.) calculated from automated foot strikes will be equivalent to parameters calculated from manually labelled foot strikes
   c. Hypothesis: Step-based features calculated from automated foot strikes from 2 minutes of data will result in fall risk classification with equivalent with accuracy, sensitivity, and specificity to manually labelled foot strikes from 2 minutes of data.

1.3 Thesis Contributions

This thesis contributed to automated foot strike identification, fall risk classification, and smartphone applications for clinical movement evaluation in lower limb amputees. Specific contributions are:

- Developed a deep learning foot strike identification approach for lower limb amputees using smartphone signals collected during a 6MWT. For this population, previous rule-based algorithms were only able to detect 87% of foot strikes after error correction, which is insufficient for clinical applications [15]. The ability to automatically detect foot strikes from smartphone sensor data enables a range of capabilities, from stride parameter reporting to AI-based clinical-decision-making models.
- Developed a post processing method to correct for multiple positive classifications from the LSTM foot strike model, greatly improving prediction results. This post-processing method is essential for accurate foot strike timing.
- Demonstrated that automated foot strikes from a 6MWT can be used to calculate step-based features for fall risk identification with equal sensitivity to manually labelled, but lower

than clinical tests (e.g., TUG) for fall risk identification. This method should correctly identify at least half of the people at risk of falling after only completing a 6MWT, hence helping to mitigate falls for people who may not have been considered at risk. Additional fall-risk tests are recommended if the clinician believes a person is at risk of falling.

- Showed that 6MWT foot strike model can be applied to the first 2-minutes of data (surrogate for 2MWT) to calculate stride parameters, hence augmenting the 2MWT. However, fall detection results from the first 2 minutes of data were worse than the full 6MWT. Therefore, automated fall risk classification on lower limb amputees should be done with the 6MWT.

1.4 Thesis outline

This thesis manuscript is divided into 7 chapters. Chapter 2 is a literature review describing existing techniques for foot strike identification and previous research that investigated AI approaches for gait analysis in different populations. Chapter 2 also explains machine learning strategies, presenting feature extraction, feature selection, model optimization, and model evaluation methods.

Chapter 3 contains a manuscript published in Sensors, addressing objective 1 by comparing a machine learning and deep learning approach for foot strike identification. This chapter evaluates a decision tree and an LSTM for their ability to identify foot strikes in lower limb amputees.

Chapter 4 contains a manuscript submitted to Computers in Biology and Medicine that addresses objective 2 by investigating the ability of automated foot strikes to classify fall risk. This chapter evaluates fall risk classification of a random forest trained on step-based features calculated from automated and manually labelled foot strikes.

Chapter 5 contains a manuscript published in Sensors that addresses objective 3 by evaluating foot strike identification and fall risk classification using 2MWT data. This chapter compares LSTM models trained on a full 6MWT or the first two minutes of data for foot strike identification. This chapter also compares the ability of step-based features calculated using automated foot strikes from the first two minutes of data to classify fall risk in lower limb amputees.

Chapter 6 presents a thesis summary and presents suggestions for future work based on thesis conclusions.

2 Literature Review

1.1 Gait Biomechanics

Human gait is a complex, multi-stage movement that requires interaction between the cardiorespiratory system, nervous system, and musculoskeletal system, though most people walk unconsciously. Despite individual differences in gait characteristics, the gait cycle can be broken down into two main phases and 8 sub-phases, the order of which are consistent [21] (Figure 1). The main phases are the stance phase and the swing phase. Stance phase begins when the foot contacts the ground. The body's weight transfers to this limb and continues to support the body as it propels forward until the foot lifts, breaking the lower limb's contact with the ground, starting the swing phase. The leg swings forward until the foot contacts the ground again and the cycle repeats. These phases can be further broken down into sub-phases.

For stance phase (approximately 60% of the gait cycle):

- Initial contact, or foot strike: Begins when the foot, normally the heel, strikes the ground
- Loading response: The foot fully contacts the ground and the knee flexes to absorb the impact as weight begins shifting onto the stance leg.
- Mid-stance: The stance leg supports the full weight as the opposite foot lifts of the ground.
- Terminal stance: The body shifts from force absorption to forward propulsion.
- Pre-swing: The heel begins to lift off the ground and the body prepares to shift weight to the opposite foot as it contacts the ground.

For swing phase (approximately 40% of the gait cycle):

- Initial swing: The knee flexes and the toe lifts off the ground, breaking the limb's contact with the ground.
- Mid-swing: Additional knee flexion and hip flexion lift the leg to clear the ground and the leg swings forward as the knee begins to extend.
- Terminal swing: The body prepares to re-distribute the weight as the leg completes its forward swing and the heel begins to lower to contact the ground.

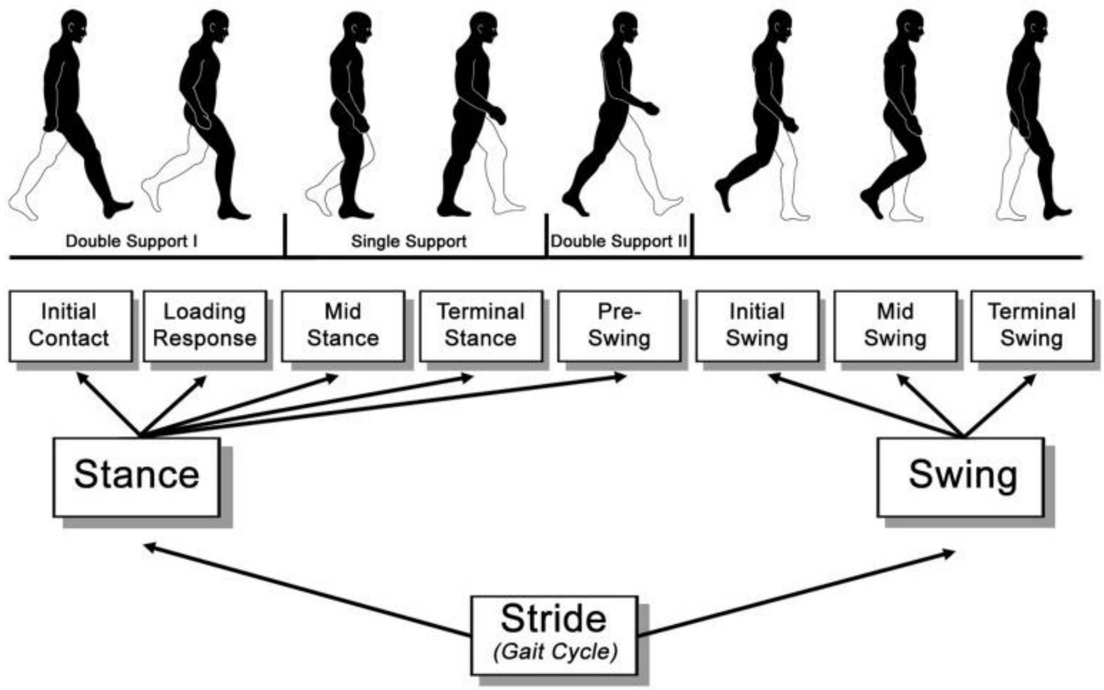

Figure 1 Gait cycle for the right leg [22].

By analyzing their gait, clinicians can gain insight into their patient's mobility and track changes in health over time. A person's mood can even be gleaned from observing body posture and walking speed [23]. Moreover, changes in gait characteristics could be used to identify neurodegenerative disorders or musculoskeletal injury. Gait analysis is also frequently used in rehabilitation programs to assess recovery after injury or surgery. While gait analysis can have subjective components, visual analysis depends on the clinician's experience and does not provide extensive detail. Quantitative gait parameters can be calculated to provide a more objective measure. These parameters can be divided into three different categories, temporal, spatial, and spatiotemporal.

Temporal parameters are calculated by observing sub-phases timing and includes step time, stride time, and cadence.

- Step time is the time between consecutive foot strikes (e.g., left and then right foot).
- Stance time is the time from foot strike until foot off for the same leg.
- Stride time is the time between consecutive foot strikes for one leg.
- Cadence is the number of steps per minute.

Spatial parameters are measures of distance between the feet during walking and include step length, step width, and stride length (Figure 2).

- Step length is the anterior-posterior distance from posterior foot to leading foot at initial contact.
- Step width is the lateral distance between the midpoint of each heel to the midline.
- Stride length is the anterior-posterior distance travelled between consecutive foot strikes for one foot.

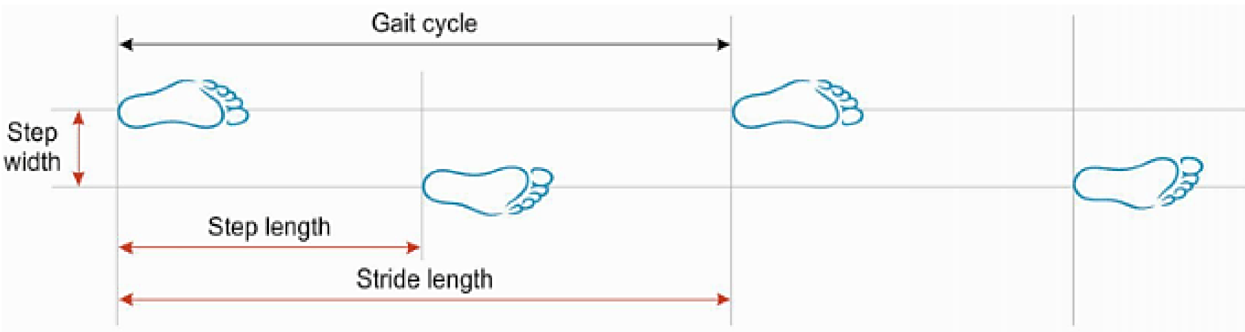

Figure 2 Spatial gait parameters [24].

Spatiotemporal parameters are a measure of distance travelled over time. For example, gait speed is usually measured in meters per second. Gait parameters can be calculated as an average for both legs, or the average for each leg can be calculated separately and compared. Another important quantitative measure is the variability of each parameter, which can be expressed as the standard deviation or coefficient of variability.

2.1 Lower limb Amputee Gait

The gait characteristics of people with lower limb amputation are distinct from the able-bodied population. Lower limb amputees present with gait deviations that include increased variability and asymmetrical movements patterns. These deviations can be subjectively observed; however, comparing gait parameters to normative values developed from able-bodied and healthy older adults can allow clinicians to develop a better picture of the person's mobility.

Table 1 displays average and standard deviation values for commonly assessed stride parameters, displaying values for healthy younger adults [25, 26], healthy older adults [27, 28], and transtibial and transfemoral amputees [29].

Table 1 Average and standard deviation stride parameters for able-bodied, healthy older adults, and lower limb amputees

	Healthy young adults	**Healthy older adults (age 60+)**	**Transtibial amputee**		**Transfemoral amputee**	
			Prosthetic	**Intact side**	**Prosthetic**	**Intact side**
Speed (m/s)	1.30 (0.11)	1.22 (0.23)	1.10 (0.16)	1.20 (0.21)	1.00 (0.13)	1.19 (0.08)
Cadence (steps/min)	115.65 (6.63)	107.38 (7.64)	99.15 (11.82)	104.40 (10.50)	79.53 (8.06)	
Stride length (m)	1.36 (0.08)	1.34 (0.19)	1.20 (0.12)	1.00 (0.15)	1.22 (0.14)	1.05 (0.18)
Step length (m)	0.68 (0.04)	0.57 (0.04)	0.65 (0.07)	0.65 (0.07)	0.71 (0.06)	0.73 (0.06)
Step width (m)	0.02 (0.005)	0.09 (0.03)	0.13 (0.03)	0.16 (0.03)	0.15 (0.04)	
Stride time (s)	1.05 (0.05)	1.12 (0.12)	1.22 (0.11)	1.11 (0.09)	1.40 (0.13)	1.30 (0.20)
Stance time (s)	0.69 (0.04)	0.71 (0.09)	0.63 (0.06)	0.65 (0.06)	0.80 (0.06)	0.86 (0.08)
Swing time (s)	0.36 (0.02)	0.41 (0.04)	0.43 (0.03)	0.38 (0.03)	0.57 (0.04)	0.44 (0.04)

Compared to healthy older adults, people with lower limb amputation have longer average swing and stride times, slower gait speed, and reduced cadence on their prosthetic side. Lower limb amputees also have greater step widths and step lengths for both legs. Intact leg stance time is also greater on average than the prosthetic side, with a greater difference for transfemoral amputees. Transfemoral amputees also present with larger step lengths and increased timing for all temporal gait parameters than transtibial amputees.

Increased step width, slower walking, and decreased stance timing for the prosthetic limb are compensatory movements to improve balance during walking by increasing the base of support and spending more time in stance on the more stable limb. People with lower limb amputations have greater instability during walking, putting them at elevated risk of injury or falls during the post-operative and post-rehabilitation periods [30, 31], therefore mobility and balance are frequently assessed.

2.2 Wearable Inertial Sensors

Advancements in data collection techniques have enabled more precise measurement and analysis of human movement. Increasingly, wearable sensors are used for human movement analysis, not only in research or laboratory-based clinical settings, but also for more practical applications. Sensors, such as accelerometers or gyroscopes, can be attached to a person and worn during movement assessments to extract richer data that can be used for clinical decision making. For example, initial contact marks the start of each gait cycle and is used to calculate a number of gait parameters. Automated detection of initial contacts during gait would allow for immediate calculation of stride parameters after completing a short walk test. Compared to fixed systems, such as image processing or in-floor sensors, wearable inertial sensors are more cost-effective, require less setup and data collection time than marker-based video systems. In some cases, instrumented walkways (depending on the number of wearable sensors used), are portable, and can be used in any movement scenario, making them a viable tool for foot strike identification [32]. Wearable sensors are commonly placed on the lower limb, inside the participant's shoes, on the pelvis, or a combination of locations for a multi-sensor system. Inertial sensor signals can then be extracted and used for automated foot strike detection and gait parameter calculations.

2.2.1 **Single Sensor Systems**

Compared to multi-sensor systems, a single sensor is preferred for clinical applications to minimize set-up and signal processing time. Recent research has shown that a single sensor placement can be used to detect foot strikes in both able-bodied and disease populations. The type and placement of the sensor varies throughout the literature, although the most popular device is an inertial measurement unit (IMU), and the most frequently used placements are the shank or foot [1]. However, sensors on the lower limb can interfere with the participant's natural movement, especially if the sensor is not properly fixed to the limb and moves distally during walking. Recently, the pelvis has emerged as an alternative sensor placement [18, 33, 34, 35].

2.2.2 **Smartphones for Data Collection**

Smartphones are emerging as a new tool for data collection. A ubiquitous piece of technology, smartphones contain built-in sensors, including accelerometers and gyroscopes. These sensors are used to determine smartphone location, orientation, and acceleration and provide information to numerous applications that are used every day. Additionally, smartphone applications can collect data from the integrated sensors and export the data for analysis. Algorithms for foot strike detection could benefit from these sensor signals. In the current literature, smartphone accelerometer and gyroscope signals were used for human activity recognition, detecting changes in movement behaviour to distinguish walking, ascending, or descending stairs, and running [36, 37]. Smartphones have also been used to collect data during short bouts of walking

[38], or during clinical movement assessments, such as the 6MWT [2], to detect foot strikes and automatically calculate gait parameters in able-bodied individuals. Smartphones have also been proposed as a potential analysis tool for numerous mHealth applications, including frailty assessments, fall risk, and promotion of active living [39].

2.3 Foot strike identification in lower limb amputees

A number of techniques have been used for foot strike identification in lower limb amputees. Previously, image processing was used for foot strike detection. These systems (e.g., Vicon [40]) allowed for foot strike identification during walking, running, sit-stand transitions and walking up or down stairs [41, 42, 43]. Foot strike detection is accomplished by training an algorithm to track the coordinates and/or velocity of markers placed on the lower extremity, specifically the heel and toe. A popular method is the algorithm described by Zeni et al. [44]. A biomechanical marker set typically needs to be placed on the person and the system requires one or more specialized cameras and video processing software. Recently, a video-based system was proposed that does not require marker placement, rather deep learning methods were applied to automatically identify anatomical landmarks in able-bodied participants [45]. Foot strikes were detected with 86.1% accuracy, using a combination of the Zeni et al. [44] algorithm and an additional algorithm to recover gait initiation and gait termination. However, this method was limited by the view of the camera; step length, for example, which relies on measurement for both ipsilateral and contralateral foot were less accurate.

Force plates have also been used to identify foot strikes from ground reaction forces and are often used as the ground truth comparator when evaluating foot strike identification approaches [46]. However, this necessitates specialized equipment (e.g., instrumented plates) and multiple force plates. Instrumented walkways are longer versions of force plates that enable stride parameter analysis for multiple strides, but these systems are also expensive and require a long, dedicated data collection area [47].

2.3.1 Wearable sensors for foot strike detection in lower limb amputees

Wearable sensors have provided a more accessible tool for foot strike identification in populations with variable gait, including lower limb amputees. A recent review [48] investigated the field of gait event detection algorithms for controlling lower limb prosthetic devices. Often, foot strike detection alone is not the goal of the research, but is necessary to segment the gait cycle into phases to extract other clinical outcome measures, like stride parameters, walking speed, or functional mobility. The review highlighted that, as with other populations, when a single sensor is used, the shank or the foot are frequent placements for lower limb amputees. IMUs are a common sensor selection. The types of algorithms used to detect gait events varied, including both rule-based and machine learning methods. For example, Maqbool et al. [13] presented a rule-based gait event detection system for foot strike detection in lower limb amputees using an IMU mounted on the shank using the angular velocity and linear acceleration of the shank. However, the system used in this analysis was only validated in one transfemoral amputee. Other research has demonstrated that a thresholding algorithm could perform well for foot strike identification in slightly larger data sets of transfemoral amputees, using data collected from a shank-mounted IMU [46, 49].

Insole and shoe-mounted sensors have also been proposed for foot strike detection in lower limb amputees with some success. IMUs can be mounted to the shoe or to the person's foot/ankle. Insole or in-shoe pressure sensors allow for a more direct measure of initial contact by analyzing ground reaction forces, similarly to a floor force place or instrumented mat. Pressure sensors are also lightweight and can be inserted into most footwear. Data collected from insole sensors and shoe-mounted sensors have been used in combination with machine learning models and rule-based approaches to identify initial contact and toe-off phases with some success, but these algorithms are often trained on data collected from able-bodied participants, or participants with hemiparesis [50, 51, 52].

Despite their popularity in a research context, sensors located on the lower limb can be impractical for clinical use in lower limb amputees since the location my alter how the person walks. As described above, the posterior pelvis has emerged as a viable sensor location, and the use of smartphones for signal collection have become increasingly popular. However, smartphone signals have had mixed success for foot strike identification in lower limb amputees.

2.4 Artificial intelligence for gait analysis

Increased gait variability of lower limb amputees and the use of gait aids make it more difficult for rule-based approaches to detect foot strikes. Rule-based algorithms, such as thresholding or peak detection, can be successful for foot strike detection when shank-mounted sensors are used. Distinct negative peaks can be identified in the shank angular velocity signal and used to classify both foot strike and foot-off events. Heuristic rule-based algorithms that have been successful in other populations have had limited success for lower limb amputees [14, 46, 53]. Consequently, higher complexity AI algorithms are proposed, and have been used with inertial sensor signals for foot strike detection in populations with pathological gait [18, 19, 54].

2.4.1 Machine learning approaches

Machine learning is a branch of AI, where algorithms are trained to sort structured data into discrete categories, or classes. Machine learning approaches can be supervised, where the algorithm is trained on labelled data (i.e., class is known), or unsupervised, where the training data is unlabelled, and the algorithm discovers difference in the underlying structure of the data [55]. Supervised machine learning is more popular, and includes algorithms such as linear regression, hidden Markov model (HMM), support vector machines (SVM), and decision trees. In general, machine learning algorithms are easy to build and do not require long training times, so they are a popular choice for classification problems and are becoming a popular tool for clinical analysis.

As discussed in the previous section, machine learning approaches have been used for gait phase detection for prosthetic device control. While rule-based approaches are popular for gait event detection, machine learning models can improve detection in populations with variable gait. Sánchez Manchola et al. [51] compared a threshold-based approach and a machine learning approach (HMM) for gait phase recognition to control robotic lower limb exoskeletons. An equal number of healthy participants and participants who had hemiparesis were included in the experimental protocol. Sensor signals were collected from an IMU placed on their foot instep. The machine learning approach outperformed the rule-based approach, and could be implemented to evaluate patient variability in addition to controlling lower limb exoskeletons.

Another supervised machine learning algorithm for gait phase recognition is SVM. SVMs are a robust prediction method typically used for binary linear classifications, but can be used for non-linear classification as well. SVM classifies data in a way that maximizes the distance, or margin, between the classes. The greater the separation between classes, the more confidently future data can be classified into either category. The decision boundary placement is based on the placement of different support vectors, data points of either class that are closest to the decision boundary (Figure 3). Boschmann et al. [56] used an SVM for gait phase detection in lower limb amputees using signals collected from lower limb electromyography (EMG), to identify muscular activity changes, and insole-embedded force sensing resistors (FSR) placed under the heel and toes.

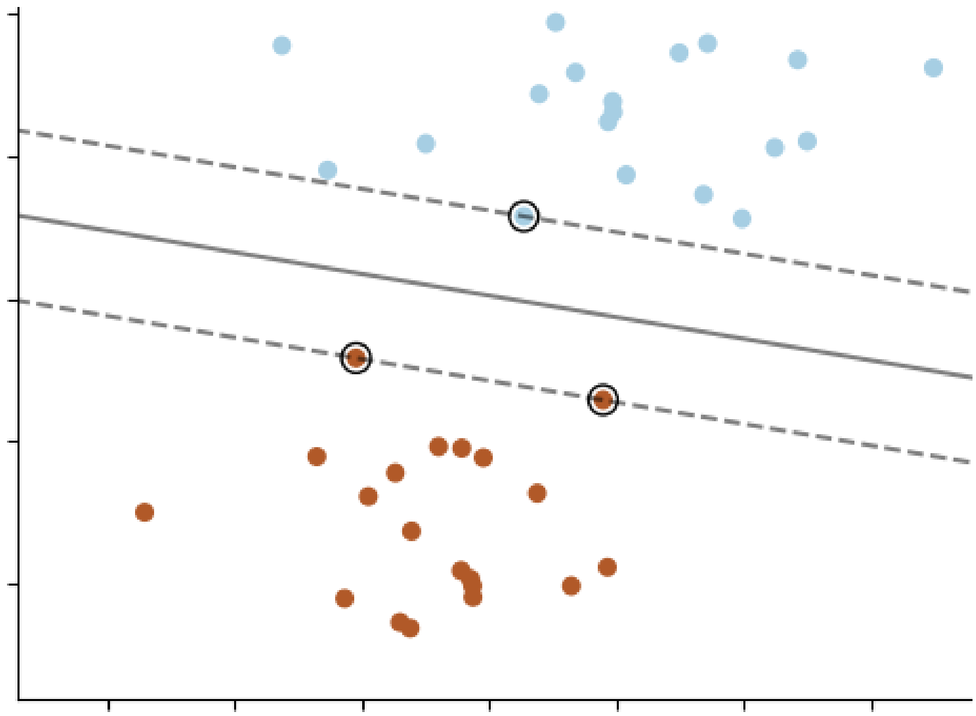

Figure 3 Support vector machine. The support vectors used to determine the decision boundary are circled. These samples demonstrate the widest margin between the nearest two points of different classes [57].

Decision trees are machine learning approaches that can be used to classify discrete data or regression of continuous data by forming flowchart-like decision boundaries to split the dataset into subsets or "nodes" based on features [58]. The subsets can be further divided, forming a tree structure, with the original dataset labelled the "root node" and the nodes connected by the decision "branches", until a final classification is made at the terminal node, or "leaf" (Figure 4).

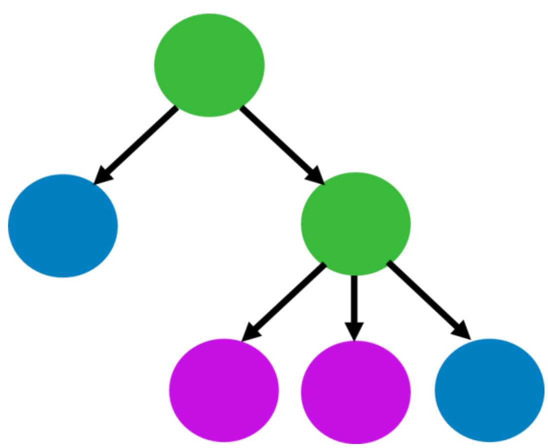

Figure 4 Example decision tree framework.

Decision trees have been used for numerous medical classification applications, such as breast cancer research to classify tissue as benign or malignant [59, 60]. Additionally, decision trees have been used to classify people with Parkinson's disease based on gait characteristics to study personalized medication protocols [61]. In this study, a decision tree outperformed three other classifiers, k-nearest neighbour (KNN), SVM, and naïve bayes. Decision trees have also been used in combination with signal data from local sensors for gait event detection and foot strike identification for disease populations, where irregular gait and the use of assistive devices preclude the application of more basic algorithms. For example, a decision tree was used to perform linear regression for gait phase detection to improve powered prosthetic devices for lower limb amputees, though the algorithm was trained on able-bodied participants [9].

2.4.2 Deep learning approaches for gait analysis

Deep learning is a more complex subset of machine learning. Neural networks, or artificial neural networks (ANN) are the basic framework for deep learning approaches. The neural network structure is similar to the human brain, often represented as a network of interconnecting nodes, similar to neurons, with information flowing in a single direction and structured in three main layers: input layer, hidden layers, and output layer (Figure 5**Error! Reference source not found.**). ANN architecture has varying levels of complexity depending on the selected approach. Deep learning algorithms perform well on large datasets, performing feature extraction along with classification, often achieving higher accuracy than machine learning approaches [10]. Due to this, different deep learning approaches have been used for a number of applications, including, speech recognition and natural language processing, image recognition and restoration, and medical image analysis [62, 63, 64, 65, 66]. Deep learning techniques can be applied for classification problems and, as with machine learning, include both supervised and unsupervised methods.

Convolutional neural networks (CNN) are a subset of ANN. CNNs are commonly used for image analysis (recognition, classification, etc.), but have also been employed for clinical applications. For image analysis, the neurons are organised to identify spatial dimensionality. CNNs consist of a network of three types of layers, with each layer extracting additional information: convolutional layers, pooling layers, and fully-connected layers (Figure 6). The convolutional layer determines the output of the neurons connected to the input layer, the pooling layer down samples the given input, and the fully-connected layers produce class scores from the activations [67]. While commonly used for image processing, a variation of CNN, temporal CNN (TCN) was developed to process sequential data, which is a limitation of traditional CNN. Filtjens et al. [19] employed TCN approach to identify initial contact to segment gait cycles in people with Parkinson's disease. Their results demonstrate that the TCN model could robustly identify initial contact, even for people experiencing freezing of gait, which can be greatly debilitating.

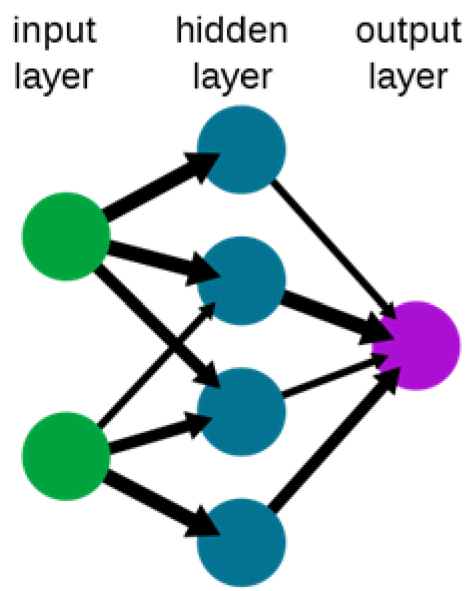

Figure 5 An example of a basic neural network (NN) framework [68].

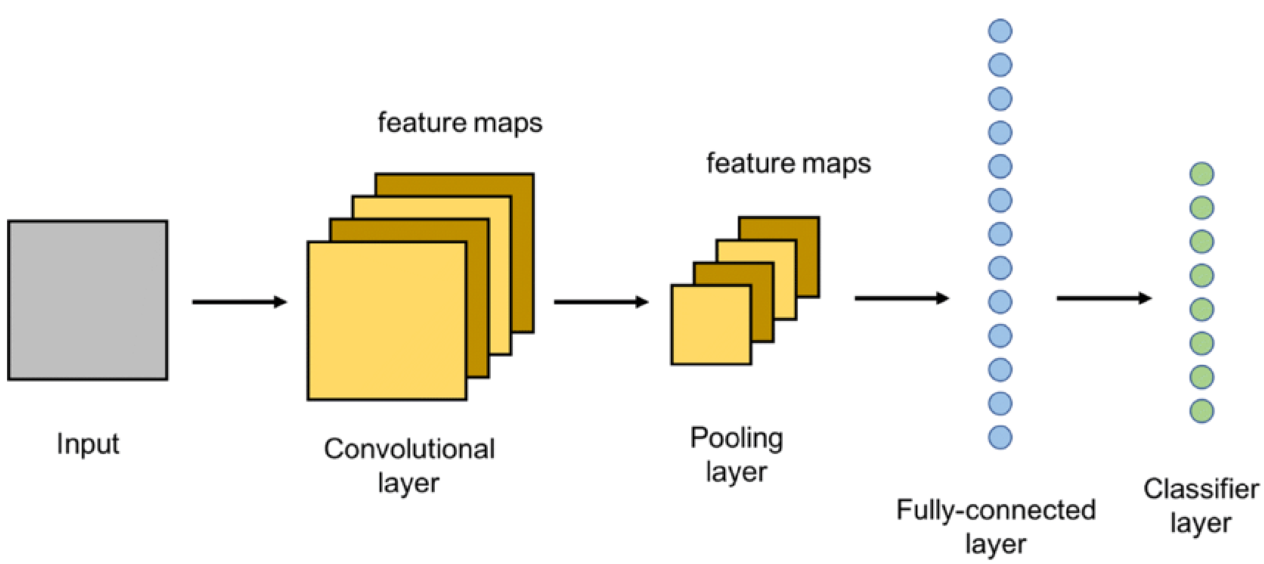

Figure 6 Basic convolutional neural network (CNN) framework [69].

RNN are another subset of ANN that perform well on temporal or sequential data. Unlike traditional feed-forward networks, RNN contain an internal ‘memory’ unit. This memory unit has the ability to store information about the previous layer in the network and use its output to inform the input of the next layer, updating weights via back propagation (Figure 7) [70]. For gait analysis, an RNN could be trained to identify a foot strike using information directly preceding as well as immediately after the initial contact. In previous research, RNNs were trained to segment walking data using instrumented footwear [54]. Gait data collected from in-shoe sensors were used as the training data for the model, which identified heel strikes and toe-offs within -5.9 ± 37.1 and 11.4 ± 47.4 ms.

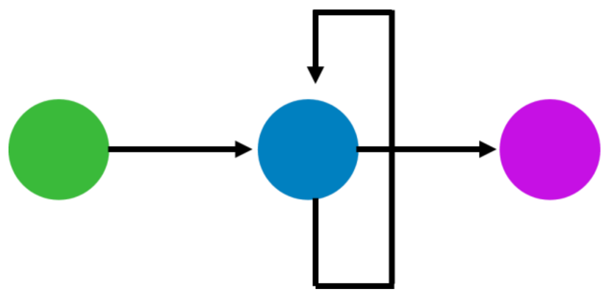

Figure 7 An example framework of a Recurrent Neural Network (RNN).

Long Short-Term Memory (LSTM), an RNN architecture, builds on the traditional framework (Figure 8). In traditional RNN trained using back-propagation, the gradients used to update weighting can become unstable, exploding to very large numbers or tending to zero, also called vanishing. LSTM can partially overcome this problem by including feedforward and feedback components, and a “forget gate” to assist with information regulation [48]. This forget unit allows for some information to be retained, while other information is dropped or forgotten. Like RNN, LSTM are well-suited to make predictions on sequential or time data and the addition of the forget unit means they perform well on large datasets with large gaps between potentially relevant information. Human gait is a sequential process broken down into a number of important sub-phases; however, initial contact is only a small percentage of that time. In practice, LSTM for gait analysis includes research on gait event detection in children with normal and pathological gait [18], in able-bodied participants for gait cycle detection [71], and smartphone signals have been used with LSTM for human activity recognition [37] in healthy adults.

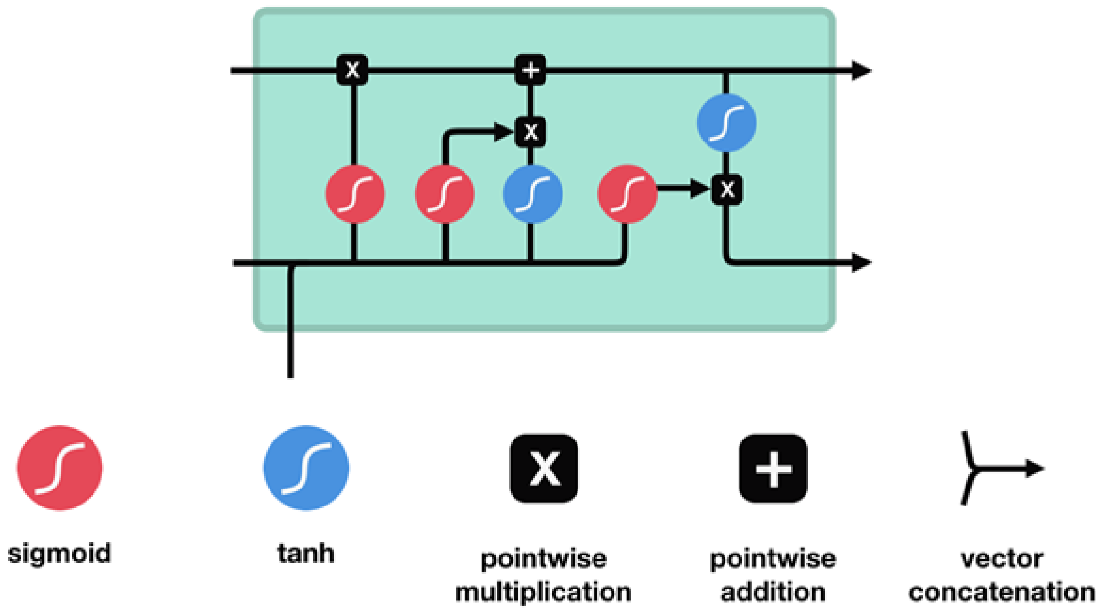

Figure 8 Long Short-Term Architecture [72].

2.5 Literature Review Summary

Through this literature review, the need for accurate foot strike identification algorithms for people with lower limb amputations was presented. The previous model for fall risk classification in amputees, using the 6MWT [8], relied on manual foot strike labelling to calculate step-based features, which is impractical for clinical use. Wearable sensors combined with AI algorithms could allow for the development of an automated foot strike identification model for lower limb amputees that uses smartphone signals collected from the posterior pelvis. Both machine learning and deep learning approaches have been used for human movement analysis and foot strike identification in other populations, however their ability to detect foot strikes from sensors at the pelvis has not yet been evaluated for amputee data. Decision trees and LSTM have been used for gait phase detection and are viable options for lower limb amputees. Integration of an automated foot strike detection model into a smartphone application could allow for data to be collected and processed for immediate clinical outcome reporting.

3 Comparison of decision tree and long short-term memory approaches for automated foot strike detection in lower extremity amputee populations

This chapter addresses Objective 1 by developing a novel foot strike identification model for lower limb amputees. Decision tree and LSTM models were developed for foot strike identification using signals from a 6MWT and foot strike identification accuracy for both models were compared. Stride parameter outcome measures calculated from automated foot strikes from both models were calculated and compared to manual foot strike labelling. The best model for foot strike identification was the LSTM. Stride parameters calculated from automated foot strikes were equivalent to manually labelled.

The content of this chapter was submitted and published in Sensors.

Pascale Juneau, Natalie Baddour, Helena Burger, Andrej Bavec, and Edward D. Lemaire. Comparison of decision tree and long short-term memory approaches for automated foot strike detection in lower extremity amputee populations. Sensors. 21 (2021), 6974. https://doi.org/10.3390/s21216974

3.1 Abstract

Foot strike detection is important when evaluating a person's gait characteristics. Accelerometer and gyroscope signals from smartphones have been used to train artificial intelligence (AI) models for automated foot strike detection in able-bodied and elderly populations. However, there is limited research on foot strike detection in lower limb amputees, who have a more variable and asymmetric gait. A novel method for automated foot strike detection in lower limb amputees was developed using raw accelerometer and gyroscope signals collected from a smartphone positioned at the posterior pelvis. Raw signals were used to train a decision tree model and long short-term memory (LSTM) model for automated foot strike detection. These models were developed using retrospective data (n = 72) collected with the TOHRC Walk Test app during a 6-min walk test (6MWT). An Android smartphone was placed on a posterior belt for each participant during the 6MWT to collect accelerometer and gyroscope signals at 50 Hz. The best

model for foot strike identification was the LSTM with 100 hidden nodes in the LSTM layer, 50 hidden nodes in the dense layer, and a batch size of 64 (99.0% accuracy, 86.4% sensitivity, 99.4% specificity, and 83.7% precision). This research created a novel method for automated foot strike identification in lower extremity amputee populations that is equivalent to manual labelling and accessible for clinical use. Automated foot strike detection is required for stride analysis and to enable other AI applications, such as fall detection.

3.2 Introduction

Foot strike identification is necessary for human gait evaluation, providing insight into a person's activity levels, mobility, and gait pattern. For example, foot strikes identify the start and end of a gait cycle and can be used to calculate the step time, stride time, and double support time for each leg. Previously, foot strike identification was completed by visual analysis with video-tracking systems (Vicon, etc.), ground reaction force analysis with force plates, or 3D accelerometer and gyroscope signal analysis from sensors placed at the foot/ankle or shank. While these methods have been successful for both able-bodied and disease populations [13, 73, 74], they can be expensive, difficult, and timely to set-up. More recently, 3D signals collected from a smartphone located at the pelvis have provided a more accessible analysis of the movement status [2, 38, 75]. While these models can identify foot strikes with a high accuracy, the foot strike identification models were typically based on able-bodied participant data.

Lower limb amputee populations can present with a high variability and inconsistent walking patterns that put them at a high risk of injury and falls [3]. Instability, an asymmetrical gait, or using a walking aid can make it difficult to automatically detect steps from sensor data. Algorithms for gait phase detection for microprocessor controlled prostheses are frequently trained on input data from able-bodied individuals wearing a single sensor on the thigh, shank, or foot, or on signals from multiple sensors on the body [76, 77]. While this may be effective for research purposes, using a single sensor location would facilitate use in clinical environments, where the time is not available to configure multisensory systems on a patient.

Recently, a rule-based foot strike identification algorithm for lower limb amputees was developed using anterior–posterior (AP) linear acceleration collected from a smartphone affixed to the posterior pelvis during a six-minute walk test (6MWT) [15]. This approach achieved 87% foot strike detection accuracy, with error correction. When analyzing lower limb amputee gait data for clinical use, manual foot strike labelling would be needed to ensure appropriate stride timing. For example, recent research by Daines et al. [8] used smartphone data during a 6MWT and manually labelled foot strikes to predict the fall risk for people with lower limb amputations with 81.3% accuracy. While these results are promising, manual labelling is time-consuming and impractical for clinical use, where immediate results reporting is desirable to support decision making at the point of patient contact.

Artificial intelligence (AI) algorithms have been proposed as an alternative method for gait analysis in populations that have a more variable gait, such as cerebral palsy or Parkinson's disease [16, 17, 18, 19]. Machine learning is a subset of artificial intelligence where algorithms make predictions by evaluating structured data over time. Machine learning models are simple to build, easy to interpret, and require shorter training times than more complex models. A popular supervised machine learning algorithm is the decision tree.

Decision trees classify data based on a set of features for each input. The model splits the data based on feature values and their corresponding class labels by determining the most effective decision boundary. Decision trees have been used to diagnose coronary artery disease [78] and to distinguish healthy tissue from cancerous tissue [59]. These models have also been used to perform logistic regression analysis for gait phase recognition to improve dynamic knee–ankle–foot orthosis control [9].

Deep learning is a more complex subset of machine learning. Deep learning models require more training time, but often provide a higher accuracy [10] because they can perform automated feature extraction and classification concurrently, whereas a feature selection process is required prior to training a machine learning algorithm. Deep learning methods also require less time during testing than machine learning techniques if the data set is large.

Recurrent neural networks (RNN), a deep learning approach, are a class of artificial neural networks that contain both feed-forward and feedback loops, making it possible to loop relevant information back into the network. RNNs perform well on sequential data, such as handwriting recognition [62], and have been implemented for gait segmentation, recognizing heel-strikes and toe-offs by training on data from in-shoe sensors [54]. Long short-term memory (LSTM) is a popular RNN architecture that, like RNN, has both feedforward and feedback components, and has the addition of a forget gate that sorts data into short-term and long-term memory cells. This process helps to regulate information flow by determining what data should be remembered and what data can be forgotten, making them ideal for data sets that have gaps between relevant events. For example, there is constant movement during gait, but a foot strike only occurs once every 0.4–0.6 s, depending on the walking speed [37]. Recently, LSTM networks have been trained on smartphone sensor data for human activity recognition [28] and gait cycle detection [71].

The effectiveness of machine learning and deep learning algorithms to classify foot strikes in lower limb amputee populations using smartphone signals have not yet been evaluated or compared. This research developed a novel method for automated foot strike detection using filtered acceleration and gyroscope signals collected from a smartphone during a 6-min walk test, using both decision tree and LSTM approaches. A viable model will provide the basis for automated stride parameter calculation and stride segmentation, which is essential for using new fall risk and health status AI models within clinical environments.

In this paper, Section 2 details the methodology and experimental design. Section 3 details the results obtained from this research. Section 4 provides a discussion of the results and their implications. Section 5 provides a conclusion and details future research.

3.3 Methods

Section 2 is structured as follows. Section 2.1 details participant recruitment and characteristics. Section 2.2 describes the experimental setup and data collection process. Section 2.3 discusses pre-processing, with Section 2.3.1 covering signal filtering and processing and Section 2.3.2 detailing the manual labelling of the ground truth foot strikes. Section 2.4 describes the construction of the classification models, where Section 2.4.1 describes the decision tree classifier and Section 2.4.2 describes the LSTM classifier. Section 2.5 details the evaluation metrics for classification. Section 2.6 presents the post-processing error correction.

3.3.1 Recruitment and Participants

A convenience sample of 93 transtibial, transfemoral, and bilateral lower limb amputees were recruited from the University Rehabilitation Institute (Ljubljana, Slovenia). The inclusion criteria were: transtibial or higher amputation; ability to walk with single cane, 2 crutches, or without any walking aids; minimum of 6 months post-amputation; had a functional prosthesis; no wounds on the residual limb; and was willing to participate. Only participants who completed the full 6 min were included in this analysis. Excluded trials were due to incomplete trial (15), cell phone affixed to the side of the hip instead of lower back (5), and use of a non-rolling walker (1). Therefore, 72 participants (14 female, 58 male, age 62.3 12.7) were included in this study. Participants included 63 transtibial, 5 transfemoral, and 4 bilateral transtibial amputees. Ten participants (13.9%) completed the 6MWT with a single cane/crutch, 22 participants (30.6%) walked with double crutches, and 40 participants (55.5%) walked without gait aids. All participants provided informed consent.

3.3.2 **Data Collection**

An Android smartphone was placed on a belt at the lower back of each participant before completing a 6MWT along a 20 m hallway (Figure 9). Accelerometer, gyroscope, and smartphone orientation data were collected with the TOHRC Walk Test app at 50 Hz (Figure 10). A 50 Hz sampling rate is common for walking analysis and is compatible with most smartphone sensor data collection capabilities. Each participant was video recorded for the duration of their 6MWT.

Figure 9 Experimental set-up: smartphone on posterior pelvis.

3.3.3 **Pre-processing**

3.3.3.1 **Filtering and Signal Processing**

Once the test was complete, data were exported from the smartphone for pre-processing. Raw accelerometer data, gyroscope data, smartphone orientation, and timestamps for each recording were imported into MATLAB 2020b. Signals were filtered with a fourth-order zero-lag Butterworth low pass filter with a cut-off frequency of 4 Hz. Smartphone orientation, XYZ coordinates for raw and linear acceleration (m/s2), and angular velocity (rads/s) were the input data. Since smartphone signals are collected at a variable sampling rate, each signal was re-interpolated at 50 Hz for a total of 18049 data points per participant over the 6-min walk test.

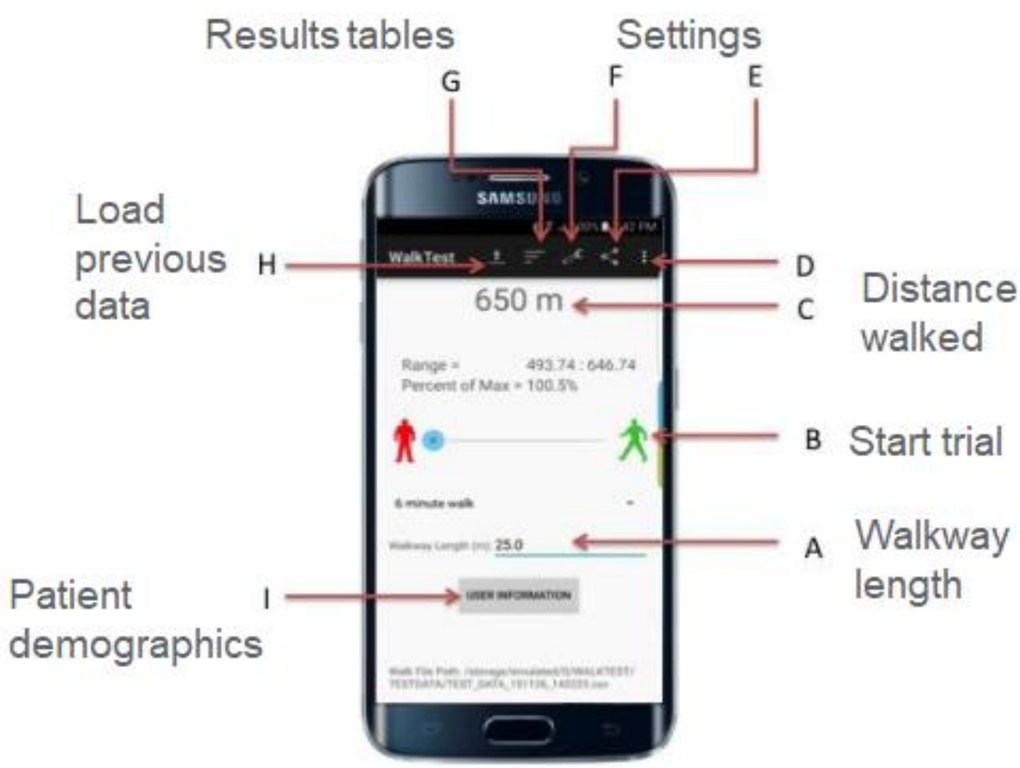

Figure 10 The Ottawa Hospital Rehabilitation Centre (TOHRC) Walk Test App [2].

3.3.3.2 **Manual Ground Truth Labelling**

Ground truth steps were manually identified and labelled by two assistants prior to model training as label 0 (no foot strike present) and label 1 (foot strike present) using the following procedure. Linear acceleration signals over time were graphed (Figure 11). In a typical gait cycle, AP acceleration peaks coincide with foot strike events, followed by a vertical acceleration peak.

Therefore, AP signal peaks immediately followed by a vertical signal peak were identified and the timestamp recorded as a foot strike event. Participant video was used to confirm timestamps. In cases where the AP peak was not well defined (e.g., gait irregularity, instability, etc.), a consensus of the two assistants was made and the most appropriate location was selected. All other timestamps were consequently labelled as "no foot strike present".

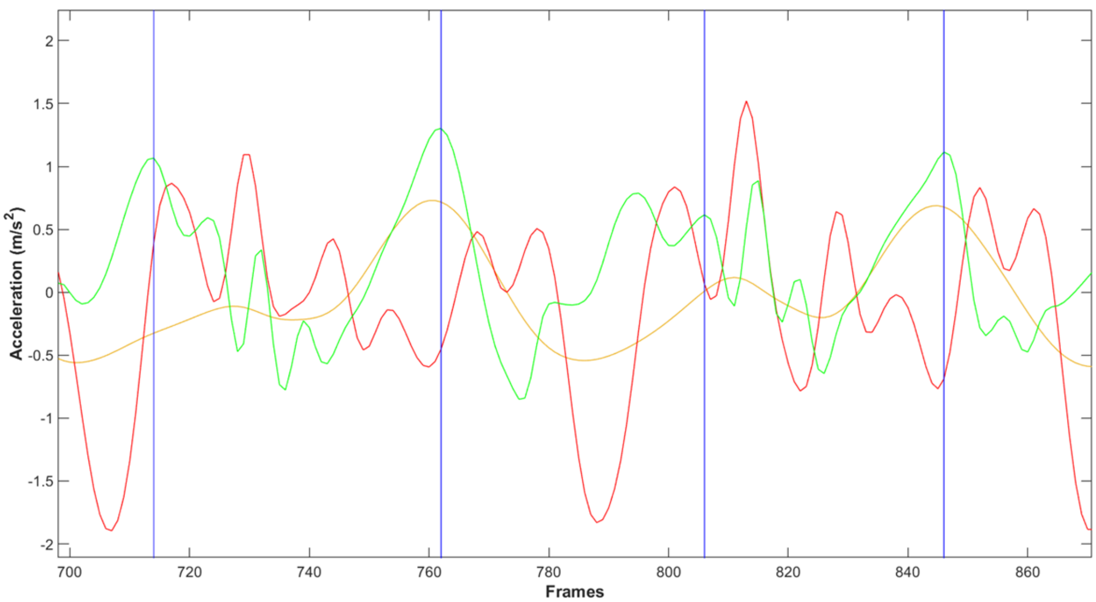

Figure 11: Ground truth foot strikes were manually identified prior to model training. Blue vertical lines indicate the timestamp of a manually identified ground truth label. All other timestamps were consequently labelled "no foot strike present".

3.3.4 Classification Models

3.3.4.1 Decision Tree

Models were written and evaluated in Python 4.1. The decision tree classifier and evaluation metrics were imported from scikit-learn library. Training data included 12 smartphone signals (3 orientation axes, XYZ coordinates for raw and linear acceleration. and angular velocity) and the corresponding ground truth labels for each data point. Hyperparameters evaluated included maximum tree depth and class weighting (1:2, 1:5, 1: 10, 1:20). The default options in the scikit-learn library were used for all other training parameters.

3.3.4.2 LSTM

The LSTM model was imported from Keras. Smartphone signals were formatted into data windows prior to model input. Each window spanned 15 frames (0.3 s) before the class label to 15 frames after the label (i.e., total window size of 31 frames). For each of the first 15 data points, the window size was 31 frames (i.e., current frame plus next 30 frames). Similarly, the previous 30 frames were used for each of the final 15 data points (i.e., previous 30 frames plus current frame). The 31-frame window size minimized the likelihood of more than one foot strike event occurring within the same window. Several hyperparameter combinations were evaluated, including batch size (32, 64, 128), number of hidden LSTM and hidden dense nodes (25, 50, 75, 100), dropout (0.3, 0.4, 0.5), and class weighting (1:2, 1:5, 1:10, 1:20). Since this was a binary classifier, binary cross-entropy was used as the loss function. Dense layer activation functions included ReLU in the input layer and sigmoid in the output layer. To evaluate the model, a confusion matrix module was imported from the sci-kit learn library.

3.3.5 **Classifier Evaluation**

Five-fold cross validation was used to evaluate performance of both AI models. A temporal tolerance of 2 frames (0.04 s) was used to match ground truth manually labelled foot strikes labels with predicted class labels. The results were evaluated based on sensitivity, specificity, accuracy, and precision. Stride parameters were calculated using both manually labelled ground truth and predicted foot strikes. The difference between these ground truths and predicted values for step time, stride time, and cadence were compared to the minimal detectable change (MDC) for each value. Since MDC was not available for lower limb amputee gait, MDC of stride parameters for healthy older adults was used [79, 80, 81].

3.3.6 **Post-Processing**

Foot strike predictions and linear acceleration signals over time were graphed to evaluate preliminary model performance. Typically, a single foot strike event would correspond with the AP signal peak. However, upon visual examination of the initial test results, periods of multiple consecutive predictions of the “foot strike present” class label corresponding to a single AP peak were observed. The predictions were often located prior to, at the foot strike instance, and immediately following the AP signal peak, causing a “banded” appearance on the graph. In order to correct for this “banding”, instances where two or more foot strike classifications occurred consecutively were identified in MATLAB. The timestamps of the start and end of each period of multiple foot strikes and the corresponding AP acceleration signal for this period were recorded. The peak AP acceleration within the period was identified. The foot strike label at this timestamp replaced all other foot strike labels in that period (Figure 12).

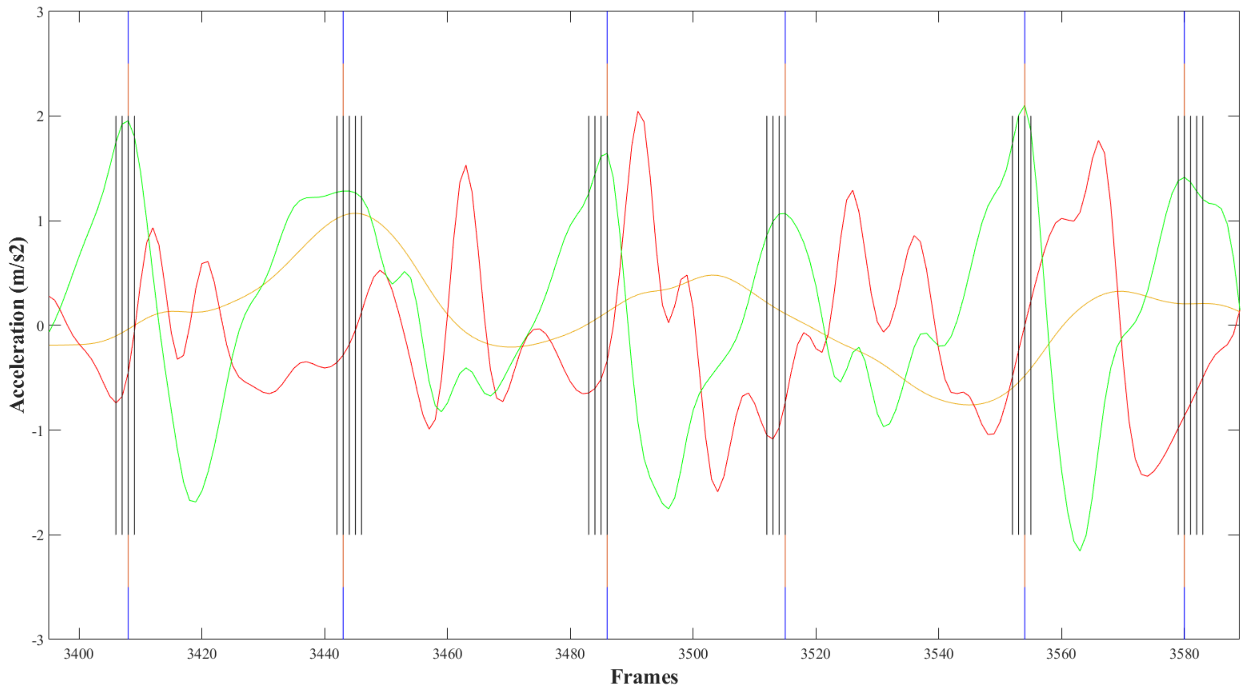

Figure 12 “Banded” groupings of foot strike predictions surrounding AP signal peak occurred in all participants. 6MWT accelerations are Anterior-posterior (green), vertical (red), and medio-lateral (yellow). Black vertical lines indicate model foot strike predictions prior to correction. Blue lines indicate ground truth foot strike labels. Orange lines indicate adjusted predictions corresponding with a peak in AP acceleration with banded periods.

To correct for missed steps, a method similar to Capela et al. [2] was employed. A locking period specific to each participant’s trial was defined from a 5-s sample of the filtered vertical acceleration signal from the beginning of the 6MWT trial. The time between positive zero-crossings for the vertical acceleration signal in the sample was used to calculate the locking period based on three procedures:

- The default locking period was half the maximum time between zero crossings.
- If the maximum time between zero crossings was greater than 0.6 s, the locking period was half the mean time between zero crossings.
- If the maximum time between zero crossings was less than 0.3 s, the maximum time between zero crossings was multiplied by 2.

To identify missed steps, periods where the duration between two consecutive steps was greater than 1.5 times the previous step were identified. The start of the period was increased by half the locking period, and end of the period was decreased by the same amount (i.e., so that the missed step was not inappropriately located at the start or end of the original selected period). A foot strike was inserted at the timestamp for the peak AP acceleration in this period (Figure 13).

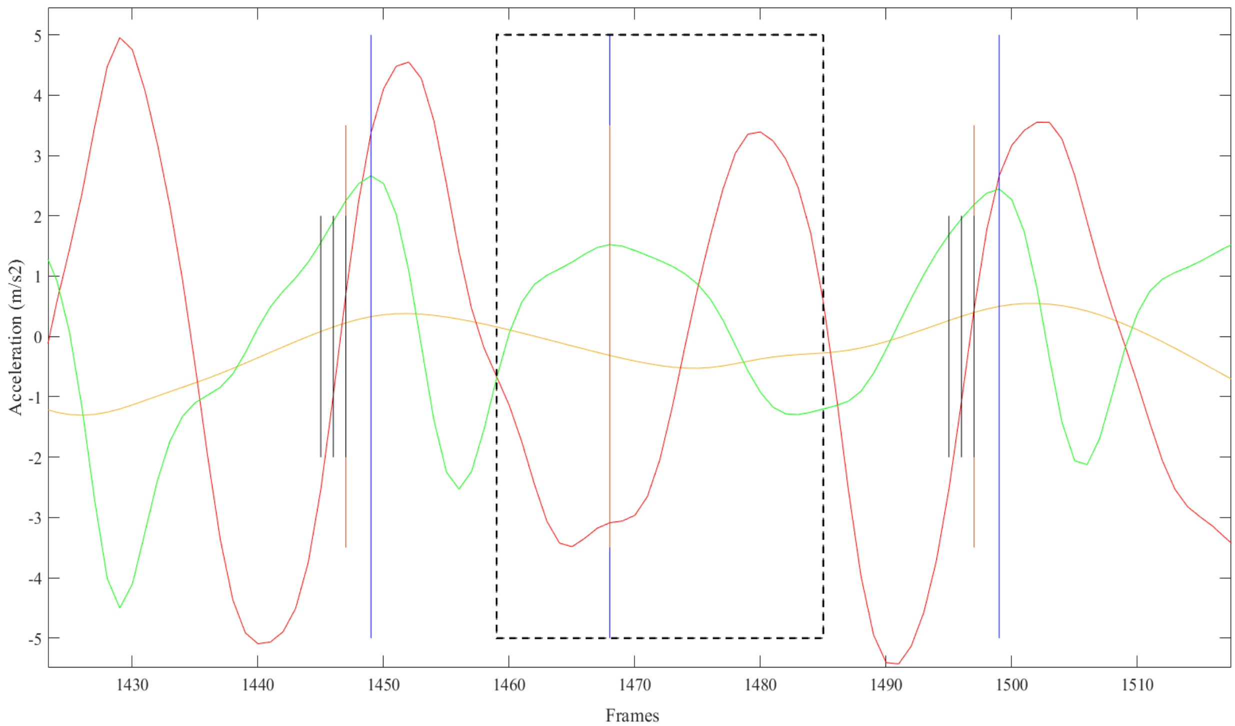

Figure 13 Missed step (vertical blue line) identified within adjusted search range (black dotted line). Foot strike inserted (vertical orange line) at timestamp for the peak AP acceleration (green line) in this period.

3.4 Results

A total of 39,561 foot strikes were identified and labelled in the ground truth data, accounting for 3.04% of total output labels (1,299,528). Table 2 displays confusion matrices for the decision tree and LSTM models. The best performing decision tree model had a maximum tree depth of 10 and class weighting of 1:20 (label 0: label 1). The decision tree classification accuracy was 98.7%, sensitivity was 83.0%, specificity was 99.2%, precision was 75.9% and F1-score was 0.79. The LSTM model with the best performance had a batch size of 64, dropout of 0.4, one LSTM layer with 100 hidden LSTM nodes, one dense layer with 50 hidden dense nodes, and a class weighting of 1:2 (label 0: label 1). The LSTM classification accuracy was 99.0%, sensitivity was 86.4%, specificity was 99.4%, precision was 82.7% and F1-score was 0.85.

Table 2 Confusion Matrices.

Decision Tree			LSTM		
	Foot Strike	**No Foot Strike**		**Foot Strike**	**No Foot Strike**
Foot strike	32,849	6,712	Foot strike	34,200	5,361
No foot strike	10,410	1,244,508	No foot strike	7,165	1,246,603

Differences in stride parameter outcome measures between manual and automated foot strikes for each model are displayed in Table 3. The step time and stride time differences were within the MDC for both models, whereas the differences in cadence were outside the MDC for both models. The LSTM model had smaller differences overall.

Table 3 Average and standard deviation (in brackets) difference between manual and automated foot strike stride parameter outcome measures for LSTM and decision tree (DT) models. MDC = minimum detectable change.

	LSTM	**DT**	**MDC**
Step time (s)	0.0010 (0.29)	−0.0139 (0.22)	0.042
Stride time (s)	−0.0006 (0.26)	−0.0149 (0.20)	0.772
Cadence (steps/min)	12.41 (-4.42)	15.64 (-6.22)	8.44

The automated band and missed step corrections were essential (Table 4). The LSTM results improved by 8.2% for sensitivity, 3.7% for specificity, 3.9% for accuracy, and, the most notable increase, 61.9% for precision.

Table 4 Evaluation metrics before and after automated corrections.

	Sensitivity	**Specificity**	**Accuracy**	**Precision**
After correction	86.4%	99.4%	99.0%	83.7%
Before correction	78.2%	95.7%	95.1%	21.8%

The foot strike identification error was 13.6%. Contributions to this error rate included automated foot strikes labelled within +/- five frames of manually labelled foot strikes (Figure 14), automated foot strikes greater than five frames from manually labelled foot strikes (Figure 15), steps missed by the AI not corrected for (Figure 16), and extra steps inserted by the AI model (Figure 17).

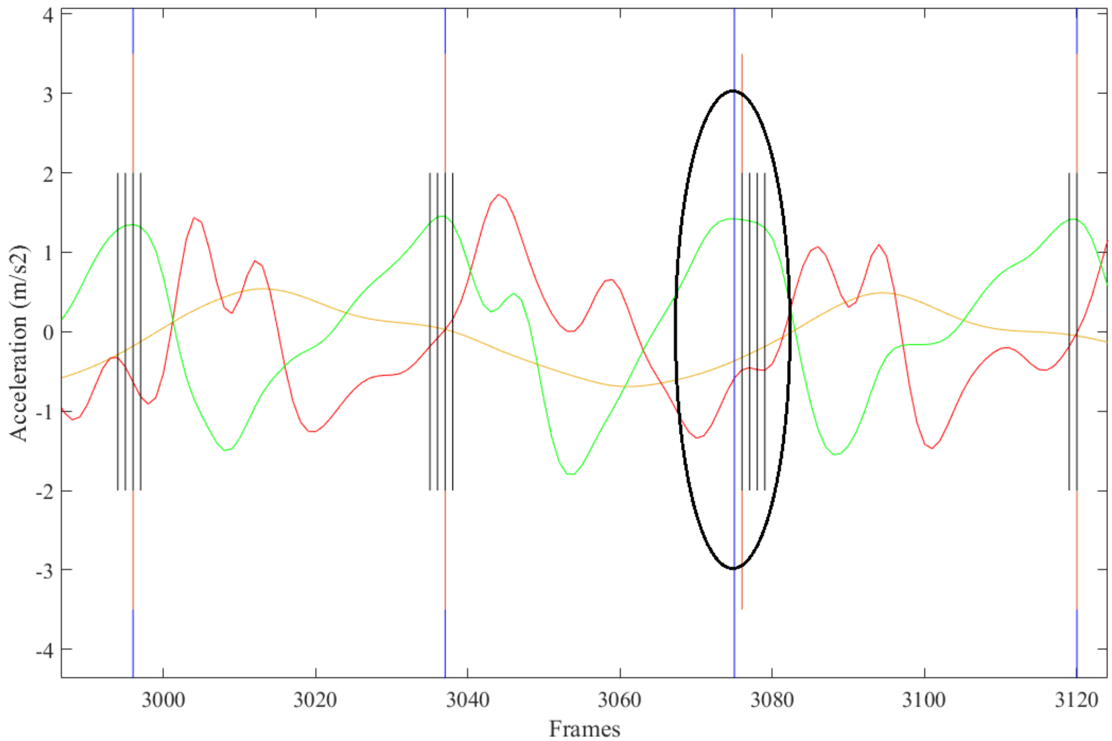

Figure 14 Automated foot strike (vertical orange line) inserted within five frames of manually labelled foot strike (vertical blue line).

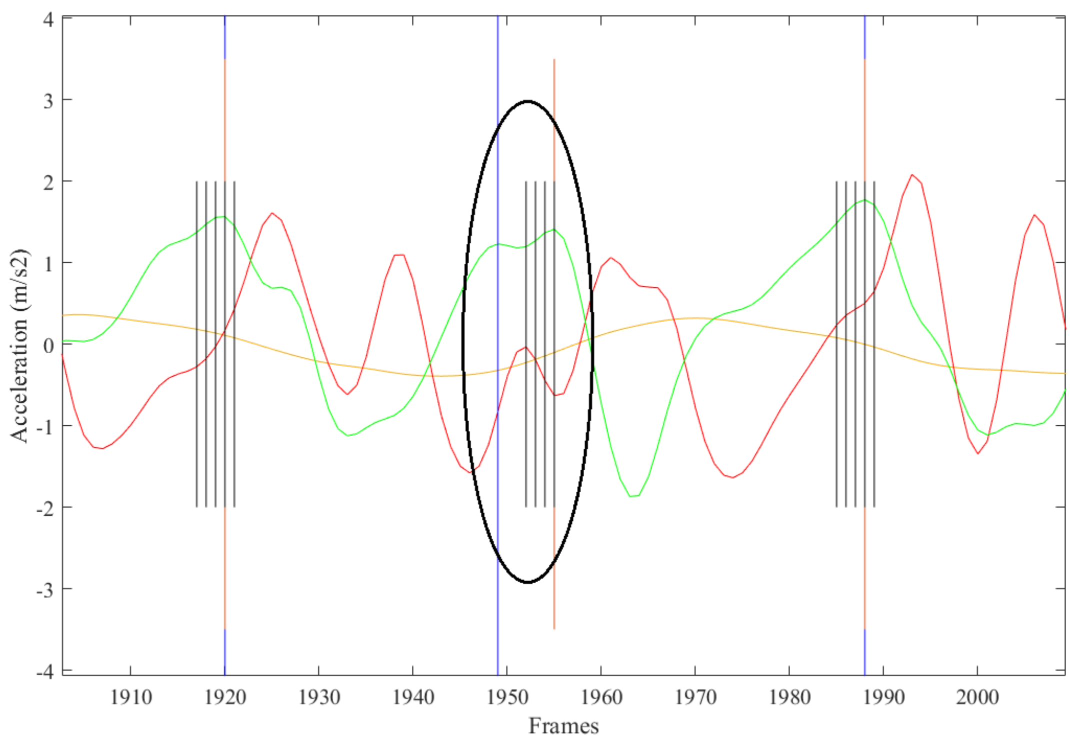

Figure 15 Automated foot strike inserted more than five frames from manually labelled foot strike.

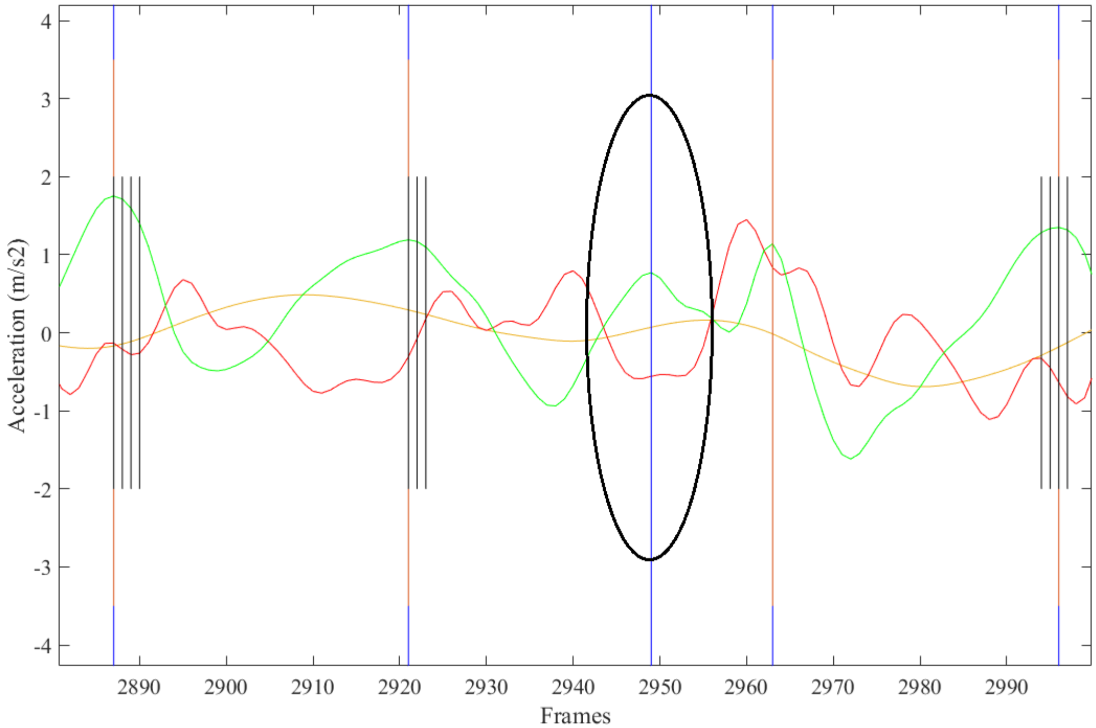

Figure 16 Manually labelled foot strike not identified by AI model.

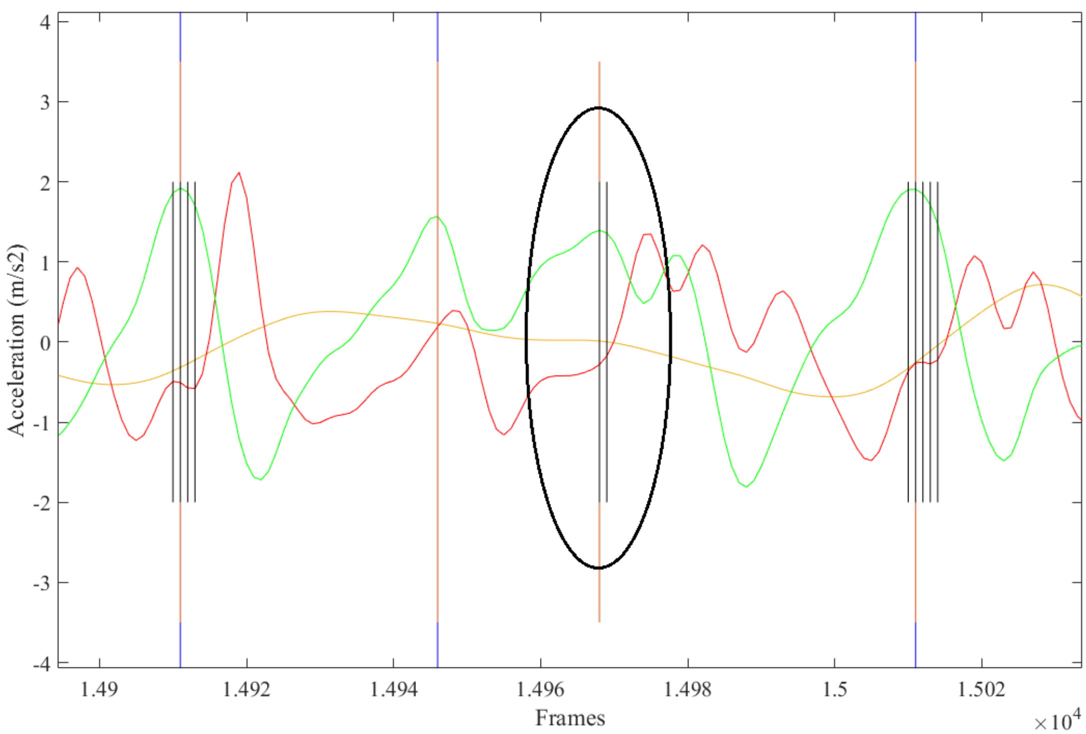

Figure 17 Extra foot strike inserted by AI model.

3.5 Discussion

This research successfully created an automated foot strike detection model that only requires smartphone acceleration, angular velocity, and orientation data from a posterior pelvis location. The LSTM model outperformed the decision tree in all areas of analysis and is the recommended model for future applications.

Compared to the previous rule-based lower limb amputee foot strike detection model in [15], the LSTM resulted in an improved foot strike classification. The algorithm described in [15] performed foot strike identification by using either AP acceleration or vertical acceleration; whichever provided a smoother signal. The LSTM model was trained on all 12 signals collected during the 6MWT. The inclusion of additional signals may have resulted in a more accurate foot strike identification by identifying patterns in all signals where a foot strike occurred. In addition, the increased complexity of the LSTM architecture may have been better suited to lower limb amputee gait variability.

Previous research [8] demonstrated that clinically relevant outcomes, such as fall risk, can be identified in amputees using 6MWT data and a random forest model. However, automated stride detection would be required to enable the system to automatically run the model, since data features are calculated for each stride. Implementation on a smartphone would allow any clinician to complete a 6MWT assessment and view the stride parameter outcome measures and fall risk status immediately after completing the trial (i.e., instant reporting).

The LSTM model had a foot strike prediction error rate of 13.6%. When the temporal tolerance was adjusted from ± two frames to ± five frames, as employed in Tan et al. [71], the foot strike prediction error decreased to 12.6%, showing that a small percentage of errors were within five frames (0.1 s) of the manually labelled foot strike, with 12% of errors occurring more than five frames away. Interestingly, in some cases where the automated foot strike was predicted to be more than five frames from the manually labelled foot strike, the AI model may have selected a more appropriate peak. For example, in Figure 15, a double peak is visible in the AP acceleration signal (green line). The manually labelled foot strike at frame 1949 corresponds with the first peak and the automatically labelled foot strike at frame 1955 corresponds with the second and greater peak. Looking at the previous step and the following step, the manually labelled foot strike corresponds with an AP acceleration peak immediately prior to a vertical acceleration peak (red). Given this pattern in acceleration signals prior to and after this timepoint, the predicted foot strike at frame 1955 is likely to be a more appropriate placement than the label at frame 1949 identified by a human observer.

Other errors included manually labelled steps that were not identified by the LSTM model that were not corrected for later, and extra foot strikes inserted in an inappropriate location. The use of walking aids, such as canes or crutches, can cause double peaks or abnormally-shaped curves in the acceleration signal, which can lead to these foot strike identification errors. In addition, steps that occur during a period of instability or if the person is walking asymmetrically (which is common for lower limb amputees) can cause similar errors. Another factor contributing to the error rate could be errors in manually identifying the ground truth foot strike events, as described above, where the visual identification of a foot strike could be off by multiple frames.

The error rate did not adversely affect the clinical outcome measures, where the difference between the automated and manually labelled foot strike step time and stride time was within the MDC. The difference in cadence was outside the MDC for both LSTM and decision tree models. However, the MDC for these parameters was not available for lower limb amputees; instead, values were compared with healthy older adult MDC. This suggests that, when extracting stride parameters from the 6MWT in lower limb amputees, clinical outcome measures from the automated foot strikes are equivalent to measures calculated from manually labelled events.

For both models, postprocessing to select one event within "banded" predictions was necessary to improve the model performance. Repetitive series of foot strike predictions surrounding the manually labelled ground truth foot strikes resulted in a greater number of false positives and fewer true negatives, affecting all classification results, and, in particular, there was a notable decrease in precision (Table 4). Band correction was very effective in automatically selecting the appropriate acceleration peak and post-processing could be integrated with the LSTM into a future smartphone application for foot strike identification.

This research had several limitations. Only those who completed the full 6MWT were included in this research. While stopping to rest during the 6MWT is permitted, the inability to complete the full 6 min is an indication of impairment that could be clinically relevant. As such, excluding them from the training data could limit model generalizability, reducing the accuracy for patients of decreased ability levels. A larger subset of people not completing the 6MWT would be required to improve the model. In addition, while participants with canes and crutches were included, those using non-rolling walkers were excluded from this analysis. Further subgroup analyses should be completed to investigate if the current model is also applicable to these groups. Since the study population sample only included five people with transfemoral amputation and four with bilateral amputation, further research could be performed to determine if the model would improve with more participants with these characteristics in the training set.

3.6 Conclusions

Foot strike identification is essential to define the gait cycle and calculate stride parameters. AI tools for clinical analysis (e.g., fall risk classification) rely on proper gait segmentation to calculate step-based features. In lower limb amputees, manual step identification was required due to the high gait variability and irregularity, limiting the clinical viability of such tools in this population [8]. This research developed a novel LSTM approach for automated foot strike detection in lower limb amputee populations using smartphone sensor signals at the posterior pelvis. An LSTM deep learning model was more effective for foot strike identification in lower limb amputees than a decision tree machine learning model. Post-processing further improved the classification results. Stride parameters calculated using predicted foot strikes were equivalent to those calculated from manually labelled foot strikes, demonstrating that the automated foot strikes with smartphone sensor data could be viable for clinical analysis. Future research could include a sub-group analysis of participants who did not complete the full 6MWT and those using mobility aids, such as wheeled walkers, since using these aids can provide signal characteristics that confuse AI classifiers. Additionally, this model could be validated for use in other disability groups, such as Parkinson's disease or cerebral palsy. The implementation of this foot strike detection model on a smartphone, as an improvement to the TOHRC Walk Test app, for example, would bring the advancements from this research to daily clinical use and improve clinical decision making for the lower limb amputee population.

4 Automated step detection with 6-minute walk test smartphone sensors signals for fall risk classification in lower limb amputees

This chapter addresses Objective 2 by evaluating a random forest model for fall risk classification in lower limb amputees. Step-based features were calculated using automated and manually labelled foot strikes and their ability to classify fall risk was compared. Sensitivity for automated and manually labelled foot strikes was equal, but specificity was lower for automated foot strikes.

The content of this chapter was submitted for publication to PLOS Digital Health.

Pascale Juneau, Edward D. Lemaire, Helena Burger, Andrej Bavec, and Natalie Baddour. Artificial intelligence-based fall risk classification and automated foot strike detection in lower limb amputees from smartphone sensors.

4.1 Abstract

Predictive models for fall risk classification are valuable for early identification and intervention for people at risk of falling. However, lower limb amputees are often neglected in fall risk research despite having increased fall risk compared to age-matched able-bodied individuals. Previous research developed a random forest fall risk classification model for lower limb amputees using smartphone signals collected during a six-minute walk test (6MWT), but manual labelling of foot strikes was required for feature calculation. Compared to manual foot strike labelling, how effective are features calculated from automated foot strike identification for fall risk classification in lower limb amputees? 80 participants (27 fallers, 53 non-fallers) with lower limb amputations completed a 6MWT with a smartphone attached to their posterior pelvis. Smartphone orientation, accelerometer, and gyroscope signals were collected with The Ottawa Hospital Rehabilitation Centre (TOHRC) Walk Test app. Automated foot strike detection was completed using a new Long Short-Term Memory (LSTM) deep learning approach. Step-based features were calculated using manually labelled foot strikes or automated foot strikes from smartphone signals. A random forest model was used to classify fall risk. The LSTM classified foot strike and non-foot strike events with 99.2% accuracy. Manually-labelled foot strikes correctly classified fall risk for 64 of 80 participants (accuracy 80%, sensitivity 55.6%, specificity 92.5%). Automated foot strikes correctly

classified 58 of 80 participants (accuracy 72.5%, sensitivity 55.6%, specificity 81.1%). Both approaches had the same fall risk classification results, but automated foot strikes had 6 more false positives. This research showed that automated detected foot strikes can be used to calculate step-based features for fall risk classification in lower limb amputees. Automated foot strike detection and fall risk classification could be integrated into a smartphone app to provide fall risk assessment immediately after completing a 6MWT.

4.2 Introduction

Falls are the leading cause of death by unintentional injury in Canada [82] and the second highest cause worldwide [83], with adults aged 65 and older at the highest risk [84]. While most falls are non-fatal, injury and permanent disability are common. On average, 20% of falls experienced by older adults in the U.S. resulted in an injury [85]. Early identification and intervention for those at an elevated risk is critical to preventing falls and prolonging well-being. Screening tests such as the Timed Up and Go (TUG) can be used to identify those at elevated risk of falling. While single screening tests can be easy to administer and completed in minutes, they are often done in tandem with a battery of other movement assessments that can be time-consuming and draining for both patient and clinician.

Artificial intelligence (AI) has been proposed as a method for fall risk prediction and classification in the elderly by using wearable sensors to collect data during different movement assessments [86, 87, 88]. For example, waist-mounted triaxial accelerometer data collected during the TUG were used to train machine learning algorithms, such as logistic regression, to estimate postural stability and classify individuals as fall risk or non-fall risk [89]. Another commonly used movement assessment is the six-minute walk test (6MWT), which evaluates functional capacity and can be completed in most clinical settings. While the 6MWT is not clinically used for fall risk predictions, Drover et al. [20] used accelerometer data from the pelvis and shank during a 6MWT to train a random forest model to classify fall risk in an elderly population (73.4% accuracy). This demonstrated that richer knowledge can be extracted from a simple movement assessment instead of requiring multiple assessments from the patient.

While seniors account for a large proportion of fallers, people with lower limb amputations are also at elevated risk of falling at all stages of rehabilitation and also post-rehabilitation. People with lower limb amputations have highly variable gait patterns, even if they are very active [3].

This variability can result in greater instability, leading to a higher likelihood of falling than age-matched able-bodied populations [5]. Until recently, despite the high prevalence of falls in the lower limb amputee population, there was limited research available for classifying fall risk using artificial intelligence. Daines et al. [8] proposed a method for fall risk classification in a lower limb amputee population using sensor data from a smartphone located at the posterior pelvis during a 6MWT. Most smartphones have integrated sensors (e.g., accelerometers, gyroscopes, magnetometers), are widely available, and are an accessible alternative to specialized dedicated equipment. Smartphone accelerometer and gyroscope signals were used to train a random forest machine learning model. Manually labelled foot strikes from turns during the 6MWT were used to calculate features from the raw smartphone signals. The random forest model achieved 81.3% fall risk classification accuracy. However, manual foot strike labelling is time-consuming (i.e., project assistant inspected each acceleration signal and video data frame to identify each foot strike) and not viable for clinical use. Automated foot strike detection would improve the feasibility of implementing a fall risk classification model in a clinical setting. Rule-based algorithms can identify steps in elderly individuals with very high accuracy (99.95%) [2]. However, the unstable and asymmetrical gait of lower limb amputees makes foot strike detection challenging. There is limited research on heuristic models for foot strike detection in lower limb amputees, though the dataset is small, and often use sensors located at the lower limb [14, 53]. Thibault et al. [15] explored a custom rule-based algorithm for foot strike identification in the same retrospective lower limb amputee 6MWT dataset as [8] using only the anterior-posterior (AP) linear acceleration signal. The rule-based approach resulted in 87% foot strike identification accuracy, noting that steps often needed to be removed, added, or relocated, suggesting that a more complex algorithm may be required to analyze irregular gait patterns in lower limb amputees.

Recently, a novel method for automated foot strike detection in lower limb amputees using a deep learning approach was developed [90]. A long-short term memory (LSTM) deep learning approach was trained on smartphone signal data collected during a 6MWT completed by lower limb amputees. Smartphone orientation, XYZ coordinates for raw and linear acceleration, and angular velocity collected from the posterior pelvis were used as input. The approach in [90] achieved 99.0% foot strike identification accuracy and stride parameters calculated from these automated foot strikes were equivalent to those of manually labelled foot strikes for most participants. These results suggested that automated foot strike methods can be used to calculate temporal features such as step time, stride time, and cadence for clinically decision-making. Therefore, this study sought to further validate the clinical usefulness of automatically-labelled foot strikes for fall risk classification.

To develop a fall risk classification tool that would be accessible to clinicians, automated foot strike detection is essential since the fall risk model's features are based on stride analysis, and foot strikes are needed to define each stride. This research compared fall risk classification accuracy of the random forest fall risk classifier in [8] using manually labelled foot strikes to automated foot strike detection. Successful fall risk classification with automated foot strike detection could lead to an accessible smartphone-based tool to enhance 6MWT utility by identifying people who may be at risk of falling without using another specific fall risk test.

4.3 Methods

4.3.1 Participants

A convenience sample of 93 transtibial, transfemoral, and bilateral lower limb amputees were recruited from the University Rehabilitation Institute (Ljubljana, Slovenia) (Table 5). Clinical records provided self-reported number of falls, with falling at least once in the past six months prior to testing considered fall risk. The inclusion criteria were: transtibial or higher amputation; ability to walk with single cane, two crutches, or without any walking aids; minimum of six months post-amputation; had a functional prosthesis; no wounds on the residual limb; and was willing to participate. Participants who could not complete the full 6MWT test were excluded from analysis. Excluded trials were due to unknown fall risk status (8) and cell phone affixed to the side of the hip instead of lower back (5).

All participants provided informed consent. This research was approved by the Ethic Committee of the University Rehabilitation Institute, Slovenia (# 46/2018) and re-approved for an additional 30 participants (# 27/2019). Each participant's self-reported fall history was used for classifying participants as "no fall risk" or "fall risk".

Table 5 Participant characteristics. Data presented as mean ± SD (range) or number (percentage).

Characteristic	**Value**
Age (years)	64.2 ± 12.2 (19-90)
Male	63 (78.8%)
Female	17 (21.2%)
Fall risk	27 (33.8%)
No fall risk	53 (66.2%)
Transtibial	72 (90.0%)
Transfemoral	3 (3.8%)
Bilateral (Transtibial)	5 (6.2%)
Time since amputation (years)	15.7 ± 18.0 (<1-65)
No aids	42 (52.5%)
Double crutches	25 (31.3%)
Single cane/crutch	12 (15.0%)
Rolling walker	1 (1.2%)

4.3.2 **Data Collection**

An Android smartphone was placed on a belt at the lower back of each participant before completing a 6-minute walk test (6MWT) along a 20m hallway (Figure 18). Each participant completed one trial. Participants were video recorded during their assessment. Accelerometer, gyroscope, and smartphone orientation data were collected with The Ottawa Hospital Rehabilitation Centre (TOHRC) Walk Test app at 50 Hz [2]. Raw accelerometer data, gyroscope data, smartphone orientation, and timestamps for each recording were imported into MATLAB 2020b. Smartphone signals were re-interpolated to 50Hz using linear interpolation, then a fourth-order zero-lag Butterworth low pass filter with a cut-off frequency of 4 Hz was applied [91].

Figure 18 Experimental set-up.

4.3.3 Step Identification and Fall Risk Classification

Ground truth steps were manually identified and labelled by two assistants prior to model training as label 0 (no foot strike present) and label 1 (foot strike present) using the following procedure. Smartphone linear acceleration over time was graphed. Foot strike events typically correspond to AP acceleration peaks followed by a vertical acceleration peak. Therefore, AP signal peaks immediately followed by a vertical signal peak were identified and the timestamp recorded as a foot strike event. Participant video was used to confirm foot strike identification. In cases where the foot strike event was not easily determined due to poor AP peak definition, double-peak, or irregular signal shape, the most appropriate location was determined by consensus. All other timestamps were consequently labelled as “no foot strike present”.

The automated step-detection approach is described in detail in [90]. Smartphone 3D orientation, acceleration, and angular velocity signals from 6MWT trials were used as input data for an LSTM deep-learning approach for foot strike identification. Predicted foot strike labels were post-processed in MATLAB 2020b to correct for model prediction errors. This included extra foot strike predictions and missed steps. To identify missed steps, a method similar to that employed by Capela et [2] was applied. An adaptive locking period specific to each participant's trial was defined from a 5 second sample of the filtered vertical acceleration signal from the beginning of the 6MWT trial. Periods where the duration between two consecutive steps was greater than 1.5 times the previous step were identified and searched for potential missed steps. The start of the period was increased by half the locking period, and end of the period was decreased by the same amount to prevent an inappropriately inserted step at the start or end of the original selected period. A foot strike was inserted at the timestamp for the peak AP acceleration within the adaptive locking period (Figure 19). Extra predictions were removed by identifying instances where two or more consecutive foot strike classifications occurred. The start and end of periods of consecutive predictions were located and the peak AP acceleration within the band was identified. The foot strike event corresponding to the AP peak was selected and all other predictions in this period were removed. Final cleaned predictions were used for feature calculation.

A random forest machine learning model developed by Daines et al. [8] was used for amputee fall risk classification. To evaluate the fall risk model using input based on manually labelled foot strikes or automated foot strike detection, two sets of features were calculated. Smartphone acceleration and angular velocity signals were used to calculate step-based features, using manually labelled foot strikes (M-FS) or automated foot strikes (A-FS). 62 features were extracted for each feature set (Table 6). Once features were extracted for each step, the minimum, maximum, mean, and standard deviation were calculated over all included steps for a total of 248 features (62 features multiplied by 4 statistics) per data set (i.e., 248 features each for M-FS and A-FS groups). The random forest fall risk model was trained using M-FS features and A-FS features to classify fall risk for each person, with results compared to the participant's ground truth fall risk. Correlation-based feature selection (CFS) was used to reduce dimensionality, based on research completed in [8]. Leave-one-out cross validation was used to evaluate the performance of the model.

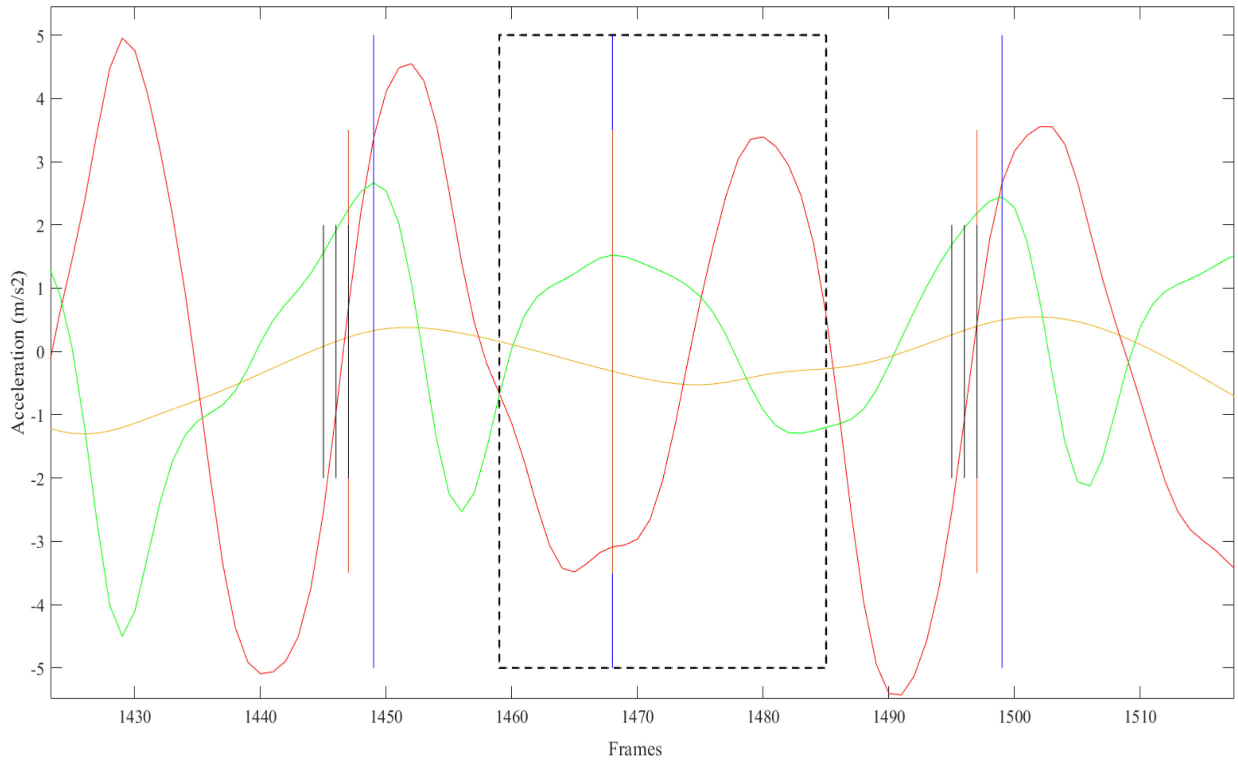

Figure 19 Example frame demonstrating locking period. A missed step (vertical blue line, frame 1468) identified within adjusted search range (black dotted line). A foot strike was inserted (vertical orange line, frame 1468) at timestamp corresponding to the peak AP acceleration (green curve) in this period.

4.4 Results

For this study, 80 participants were suitable for fall risk classification, 27 fall risk and 53 no fall risk. Table 7 displays the fall risk classifier confusion matrices for A-FS and M-FS. The fall risk classifier trained on features calculated from M-FS correctly classified 64 of 80 participants and achieved 80% accuracy, 55.6% sensitivity, 92.5% specificity, 78.9% precision and F1-score of 0.65. The A-FS approach resulted in 58 of 80 correctly classified participants and achieved 72.5% accuracy, 55.6% sensitivity,81.1% specificity, 60.0% precision and F1-score of 0.58. Classification of fall risk was the same for both foot strike identification groups.

Manual labelling was used as ground truth comparator for LSTM foot strike identification. For the participants included in this analysis, the LSTM foot strike identification model achieved 99.2% accuracy, 81.8% sensitivity, 99.7% specificity, 90.2% precision, F1-score of 0.86.

Table 6 Feature list for fall risk classification. AP = anterior-posterior; ML= medio-lateral; RMS= root-mean square; FFT= fast Fourier transform; REOH= ratio of even/odd harmonic frequencies

Temporal	**Descriptive Statistics**	**Frequency Domain Features**
Cadence	Minimum ML	Quartile FFT ML
Step time right	Minimum AP	Quartile FFT AP
Step time left	Minimum Vert	Quartile FFT Vert
Stride time	Maximum ML	Quartile FFT Tilt
Symmetry index	Maximum AP	Quartile FFT Rotation
	Maximum Vert	Quartile FFT Obliquity
	Mean ML	Maximum FFT ML
	Mean AP	Maximum FFT AP
	Mean Vert	Maximum FFT Vert
	Mean Tilt	Maximum FFT Tilt
	Mean Rotation	Maximum FFT Rotation
	Mean Obliquity	Maximum FFT Obliquity
	Range Tilt	Standard Deviation FFT ML
	Range Rotation	Standard Deviation FFT AP
	Range Obliquity	Standard Deviation FFT Vert
	Standard Deviation ML	Standard Deviation FFT Tilt
	Standard Deviation AP	Standard Deviation FFT Rotation
	Standard Deviation Vert	Standard Deviation FFT Obliquity
	Standard Deviation Tilt	Peak Distinction FFT ML
	Standard Deviation Rotation	Peak Distinction FFT AP
	Standard Deviation Obliquity	Peak Distinction FFT Vert
	RMS ML	Peak Distinction FFT Tilt
	RMS AP	Peak Distinction FFT Rotation
	RMS Vert	Peak Distinction FFT Obliquity
	RMS Tilt	REOH ML
	RMS Rotation	REOH AP
	RMS Obliquity	REOH Vert
		REOH Tilt
		REOH Rotation
		REOH Obliquity
Symmetry index: symmetry in right and left limb step times [92]		

Table 7 Confusion matrices

A-FS		
	No fall risk	Fall risk
No fall risk	43	10
Fall risk	12	15

M-FS		
	No fall risk	Fall risk
No fall risk	49	4
Fall risk	12	15

4.5 Discussion

This research demonstrated that automated foot strike identification can be used to calculate step-based features for the purpose of classifying fall risk in lower limb amputees. Fall risk classification was similar for both automated and manually labelled foot strike approaches, but the automated foot strike fall risk approach had more false positives (i.e., non-fall risk classified as fall risk). Automated foot strike detection is necessary in a clinical environment, where timely manual labelling is not feasible. A smartphone-based fall risk classification model from a 6MWT can benefit the patient and clinician since one assessment with a single sensor placement can provide functional capacity, stride parameters, and preliminary fall risk information.

Over 50% of fall risk participants and >70% of non-fall risk participants were correctly classified from signals collected during a 6MWT. The 6MWT is used to measure a person's functional capacity, and fall risk information is not typically available from this assessment. A smartphone application integrating the automated foot strike detection with fall risk classification could provide fall risk information to a clinician immediately after completing a 6MWT. While some people at risk of falling may not be identified, this should not negate the benefit of appropriately classifying more than half of the people, and thereby enabling interventions that can help maintain their safe mobility. However, to ensure that those who are at risk of future falls, but were misclassified as no fall risk, are not overlooked, patients should complete an additional assessment with higher fall risk sensitivity (e.g., TUG) to determine if an intervention strategy is necessary.

Six additional people who had not fallen in the 6 months previous to the study were misclassified as fall risk when features were calculated from automated foot strikes. A sub-group analysis of these participants did not identify any notable similarities that might explain their misclassification. They had varying ages, time since amputation, gait aid use, etc., mirroring the diversity of the people included in this study. Fall mechanisms are diverse and individual differences from person to person can contribute to the cause of a fall. Additional participant information other than fall history could be used to improve fall risk model performance. For example, an individual's score on a balance test could be informative since lower limb amputees who demonstrate better balance, and greater balance confidence, are at a greater risk of falling [3]. More research is needed to better understand these mechanisms and determine what, if any, gait characteristics can be used to predict fall risk when using AI for fall risk classification.

Foot strike identification accuracy of the 6 participants misclassified as fall risk was also investigated. For these 6 participants, the LSTM foot strike identification model had 99.1% accuracy, 79.8% sensitivity, and 99.7% specificity. While the accuracy and specificity were similar to the classification results of the full dataset, sensitivity was 2% lower, representing a greater number of missed foot strikes. Errors in foot strike identification could have contributed to the misclassification of these participants. Increasing the training set size for foot strike identification and may help to improve foot strike detection and increase specificity.

As described in [15], amputee gait differs greatly from the able-bodied population. People with lower limb amputations have more irregular and asymmetrical gait and the use of canes or crutches during walking can result in abnormally-shaped AP acceleration curves. Due to this, rule-based methods for foot strike detection were not as effective. Incorrectly identified foot strikes could adversely affect gait cycle segmentation and stride outcome measures for models that use step and stride-based features. Indeed, manual step identification can be affected similarly if the acceleration curves do not have distinct peaks or if they are abnormally shaped due to unexpected movements or gait aids. The deep learning foot strike classification approach described in [90] was able to detect foot strikes in this population from smartphone sensor signals without requiring feature extraction, demonstrating an improvement over the rule-based method developed by Capela et al. [2, 15].

The LSTM, with post-processing to correct errors such as extra predictions and missed steps, classified foot strike and non-foot strike events with 99.2% accuracy for the participants in this analysis. Other errors in foot strike identification included foot strike predictions that were within ±2 frames (±0.04 seconds). Some errors could not be corrected during post-processing, such as manually labelled steps that were not identified by the LSTM and extra foot strikes inserted in an inappropriate location. Despite these errors, clinical outcome measures, including stride parameters, were equivalent to manually labelled foot strikes for most participants.

A limitation of this research was the small number of people with transfemoral and bilateral amputations included in the dataset. While transtibial amputees are at a higher risk of falling in the post-operative period, people with transfemoral and bilateral amputations are at an elevated risk of falling post-rehabilitation [30, 31]. A greater number of transfemoral and bilateral participants included in training and testing sets could improve fall risk model generalizability. Future research in this area would benefit from a greater number of transfemoral and bilateral amputees, with sub-group analysis.

4.6 Conclusions

This study demonstrated that automatically detected foot strikes from a single smartphone sensor location on the body can be used to calculate step-based features for lower limb amputees after completing a 6MWT, leading to preliminary fall risk classification, an outcome that is not typically available for the 6MWT. This AI-enhanced 6MWT could be used to screen for people at risk of falls and then proceed with further assessments. Integration of this fall risk model into a smartphone application would improve the immediacy of the results, providing instant decision-making information. Future model development should include a greater number of people with transfemoral and bilateral transtibial amputations. This could improve the ability of the fall risk model to appropriately identify fall risk in these populations for further clinical assessment.

5 Amputee fall risk classification using machine learning and smartphone sensor data from 2-minute and 6-minute walk test

This chapter addresses Objective 3 by evaluating the ability of an LSTM trained on smartphone signals from two minutes or six minutes of data to identify foot strikes in the first two minutes of data. An LSTM was trained on 6MWT data and then re-trained on data from the first two minutes of the test. Foot strike identification accuracy was compared. Two minutes of data was not sufficient to detect foot strikes with the LSTM, so clinical outcome measures should be calculated using automated foot strikes identified by the LSTM trained on 6-minute data.

This chapter also addresses Objective 3 by evaluating a random forest model for fall risk classification in lower limb amputees using features calculated from 2 minutes of data. Features calculated using automated and manually labelled foot strikes identified in the first two minutes of 6MWT data were used to classify fall risk in amputees and their performance compared. Less than 50% of fall risk participants were correctly classified using automated foot strikes and 2 minutes of data.

The content of this chapter was published in Sensors.

Pascale Juneau, Natalie Baddour, Helena Burger, Andrej Bavec, and Edward D. Lemaire. Amputee fall risk classification using machine learning and smartphone sensor data from 2-minute and 6-minute walk test. 2022, 22(5):1749. https://doi.org/10.3390/s22051749

5.1 Abstract

The 6-min walk test (6MWT) is commonly used to assess a person's physical mobility and aerobic capacity. However, richer knowledge can be extracted from movement assessments using artificial intelligence (AI) models, such as fall risk status. The 2-min walk test (2MWT) is an alternate assessment for people with reduced mobility who cannot complete the full 6MWT, including some people with lower limb amputations; therefore, this research investigated automated foot strike detection and fall risk classification using data from a 2MWT. A long short-term memory (LSTM) model was used for automated foot strike detection using retrospective data (n = 80) collected with the Ottawa Hospital Rehabilitation Centre (TOHRC) Walk Test app during a 6-min walk test (6MWT). To identify foot strikes, an LSTM was trained on the entire six minutes of data, then re-trained on the first two minutes of data. The validation set for both models was ground truth foot strike labels from the first two minutes of data. Foot strike identification with the 6-min model had 99.2% accuracy, 91.7% sensitivity, 99.4% specificity, and 82.7% precision. The 2-min model achieved 98.0% accuracy, 65.0% sensitivity, 99.1% specificity, and 68.6% precision. To classify fall risk, a random forest model was trained on step-based features calculated using manually labeled foot strikes and automated foot strikes identified from the first two minutes of data. Automated foot strikes from the first two minutes of data correctly classified fall risk for 61 of 80 (76.3%) participants; however, <50% of participants who fell within the past six months were correctly classified. This research evaluated a novel method for automated foot strike identification in lower limb amputee populations that can be applied to both 6MWT and 2MWT data to calculate stride parameters. Features calculated using automated foot strikes from two minutes of data could not sufficiently classify fall risk in lower limb amputees.

5.2 Introduction

The six-minute walk test (6MWT) is a sub-maximal movement assessment used to evaluate aerobic capacity and mobility [93]. The 6MWT was originally developed for those with chronic respiratory or cardiovascular disease [94], but is now used to assess a number of populations, commonly older adults, people who have suffered a stroke, people with Parkinson’s disease, and lower limb amputees [95, 96, 97, 98]. The test is brief and requires minimal set up and space (minimum 12m) [99]. Participants are allowed to walk with mobility aids, stopping is allowed if needed, and distance walked moderately correlates with more complex aerobic capacity tests, like VO2 max, minimizing the burden for patient and clinician [100, 101, 102, 103, 104]. However, some people are unable or unwilling to complete a 6MWT due to limited mobility. The two-minute walk test (2MWT) is a similar assessment to the 6MWT but only requires two minutes of walking. Distances walked during a 2MWT correlate well with distance walked during a 6MWT, so the 2MWT is a viable alternative [105, 106, 107].

Recent research has sought to extract richer information from 2MWT and 6MWT using artificial intelligence (AI). The Ottawa Hospital Rehabilitation Centre (TOHRC) Walk Test app collects acceleration, angular velocity, and orientation data from a smartphone during walking.

These smartphone signals can then be used to automatically detect foot strikes. Foot strike detection is a necessary step in gait analysis because a foot strike defines the start and end of a gait cycle and can be used to calculate stride parameters such as step time and cadence. Capela et al. [108] demonstrated that smartphone signals collected using the TOHRC Walk Test app could be used as input data in a rule-based algorithm to identify foot strikes and calculate stride parameters for a 2MWT or 6MWT in able-bodied participants. Capela et al. [2] further examined that, when applied to 6MWT smartphone data of healthy older adults, the rule-based algorithm identified foot strikes with 99.9% accuracy. With the TOHRC Walk Test App, stride parameters are calculated immediately after completion of the walk test to provide real-time reporting to a clinician. However, when a similar rule-based algorithm was applied to lower limb amputee gait data, accuracy decreased to 87.0% and offline error correction was required [15]. Amputee gait differs markedly from healthy adults, which can make it difficult for AI algorithms to automatically detect steps.

Juneau et al. [90] developed a novel long-short term memory (LSTM) deep learning approach for automated foot strike detection in lower limb amputees. The LSTM was trained on filtered smartphone signals collected from the TOHRC Walk Test app during a 6MWT. Foot strike and non-foot strike events were classified with 99.0% accuracy, using offline error correction. Stride parameters calculated from the automated foot strikes were equivalent to manually labelled results for most participants. Stride parameters are not an outcome measure that is typically available from a 6MWT, demonstrating that clinical outcome measures can be calculated using automated foot strikes from 6MWT data.

Lower limb amputees typically have greater instability and higher variability during walking than healthy older adults, leading to elevated risk of falling [5]. Due to this, people may prefer to complete a 2MWT instead of a 6MWT. However, the deep learning automated foot strike approach has not yet been validated with 2MWT data. The LSTM foot strike identification model described in [90] was originally trained on data from a 6MWT; however, it is possible that a new algorithm specific to 2MWT data is required. Therefore, this study evaluated LTSM foot strike identification accuracy on 2MWT data after the model was trained on signals from a 6MWT or signals from a 2MWT, stride parameters were compared between automated foot strikes and manually labelled foot strikes and fall risk classification was assessed using step-based features from automated and manually labelled foot strikes. A successful 2MWT model would enhance the range of applications for this smartphone-based assessment approach, thereby enhancing access and immediacy of movement-based analyses for people with physical disabilities.

5.3 Methods

5.3.1 Recruitment and Participants

A convenience sample of 93 transtibial, transfemoral, and bilateral lower limb amputees were recruited from the University Rehabilitation Institute (Ljubljana, Slovenia) and gave informed consent for this study (Table 8). This research was approved by the Ethic Committee of the University Rehabilitation Institute, Slovenia (# 46/2018) and re-approved for an additional 30 participants (# 27/2019).

Table 8 Participant Characteristics. Data presented as mean ± SD (range) or number (percentage).

Characteristic	**Value**
Age (years)	64.2 ± 12.2 (19-90)
Male	63 (78.8%)
Female	17 (21.2%)
Fall risk	27 (33.8%)
No fall risk	53 (66.2%)
Transtibial	72 (90.0%)
Transfemoral	3 (3.8%)
Bilateral (Transtibial)	5 (6.2%)
Time since amputation (years)	15.7 ± 18.0 (<1-65)
No aids	42 (52.5%)
Double crutches	25 (31.3%)
Single cane/crutch	12 (15.0%)
Rolling walker	1 (1.2%)

Each participant’s self-reported fall history was used to classify participants as no fall risk or fall risk. Participants were considered a fall risk if they reported falling at least once in the past six months prior to testing. The inclusion criteria were: transtibial or higher amputation; ability to walk with single cane, two crutches, or without any walking aids; minimum of six months post-amputation; had a functional prosthesis; no wounds on the residual limb; and was willing to participate. Excluded trials were due unknown fall risk status (8 participants) and cell phone affixed to the side of the hip instead of lower back (5 participants).

5.3.2 **Data collection**

Each participant completed a 6MWT along a 20 m hallway with an Android smartphone affixed to the posterior pelvis (Figure 20). The 6MWT was video recorded using a second Android smartphone for each participant. The TOHRC Walk Test app collected smartphone acceleration (m/s), angular velocity (rads/s), and smartphone orientation at an average of 50 Hz [2].

5.3.3 **Pre-Processing**

5.3.3.1 **Filtering and Signal Processing**

Raw accelerometer data, gyroscope data, smartphone orientation, and timestamps were exported for pre-processing and were imported into MATLAB 2020b. Signals were filtered with a fourth-order zero-lag Butterworth low pass filter with a 4Hz cut-off frequency. Since smartphone signals have a variable sampling rate, each signal was re-interpolated at 50 Hz for a total of 18049 data points per signal per participant over the 6MWT [109, 110].

Figure 20 Experimental set-up: smartphone on posterior pelvis.

To determine if the automated foot strike detection model from [90] can predict foot strikes in 2MWT data, a simulated 2MWT dataset was created. The first two minutes of the 6MWT trial was determined to replicate most closely that of a 2MWT, so the first two minutes of each participant's trial was exported for a total of 6000 data points per participant.

5.3.3.2 **Manual Ground Truth Labelling**

Two assistants manually identified and labelled ground truth steps prior to model training. The two class labels were label 0 (no foot strike) and label 1 (foot strike) and were identified using the following procedure. Linear acceleration signals over time were graphed and anterior-posterior (AP) acceleration signal peaks were identified (Figure 21). Usually, foot strike events correspond with AP acceleration peaks that are followed by a peak in vertical acceleration. Acceleration signals matching this pattern were visually identified and a foot strike event was recorded at the timestamp of AP signal peaks immediately followed by a vertical signal peak. Timestamps were confirmed using participant video. An agreement of the two assistants was required in cases where the AP peak was not well defined or in cases of multiple peaks to select the most appropriate location for the foot strike event. All other timestamps were therefore labelled as "no foot strike".

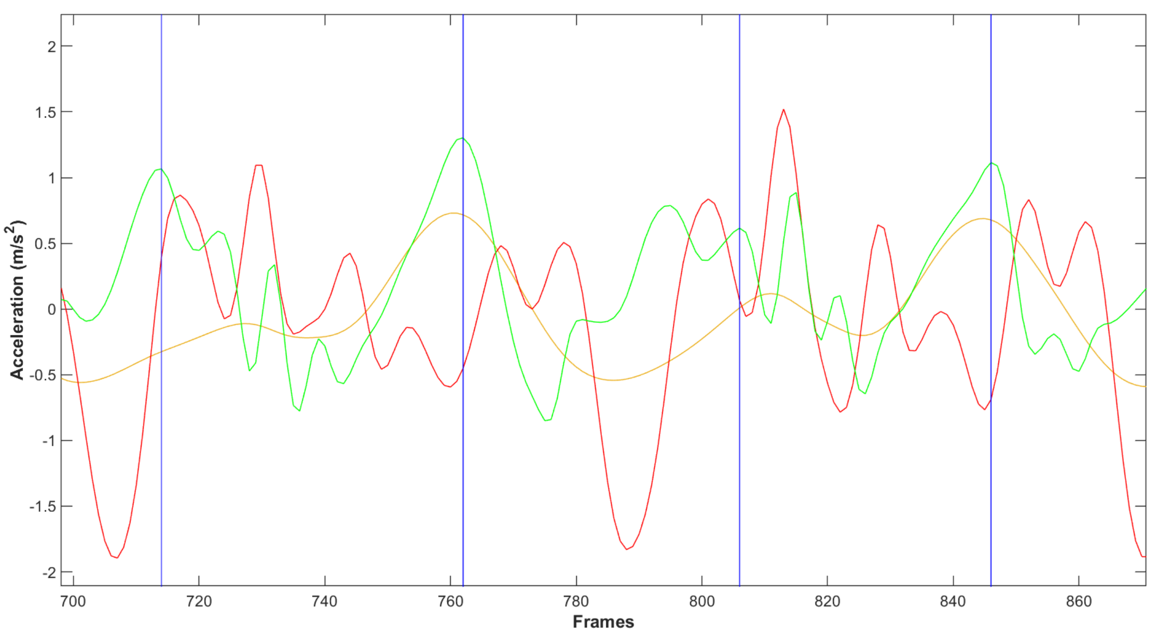

Figure 21 Filtered smartphone signals over time. Medio-lateral acceleration (yellow curve), vertical acceleration (red curve), and anterior-posterior (AP) acceleration (green curve) were used to identify ground truth foot strikes. Typically, foot strikes correspond with a AP acceleration peak followed by a vertical acceleration peak. Video recording of the trial was used to confirm the timestamp of foot strikes. Vertical blue lines indicate frames manually identified as ground truth labels.

5.3.4 Foot Strike Classification Models

An LSTM deep learning approach was developed and evaluated for automated foot strike detection in [90]. The model was written and evaluated in Python 4.2. The LSTM layer was imported from Keras [111] and several hyperparameter combinations were evaluated. The LSTM from [90] with the best performance was subsequently used for this analysis, with 100 hidden nodes in the LSTM layer, 50 hidden nodes in the dense layer, a batch size of 64, and a dropout value of 0.4. Smartphone orientation, XYZ coordinates for raw and linear acceleration (m/s2), and angular velocity (rads/s) from the full 6MWT and the first two minutes of walk test data were the input.

Smartphone signals were formatted into data windows prior to model input. Each window spanned 15 frames (0.3 s) before the class label to 15 frames after the label. For the first 15 data points, 30 frames after the class label were used. Similarly, the previous 30 frames were used for the final 15 data points. The 31-frame window size (i.e., 15 before, labelled frame, 15 after) was selected to minimize the likelihood of multiple foot strike events occurring within the same window of signal data.

The LSTM foot strike identification model was trained and evaluated twice, initially trained on the full six minutes of data (6M-FS model) and then trained on the first 2 minutes from the 6MWT, to approximate a 2MWT trial (2M-FS). The validation set for both models was ground truth labels from the first two minutes of data (i.e., the same set to enable direct foot strike comparison between models).

5.3.5 Foot Strike Model Evaluation

Five-fold cross validation was used to evaluate performance of both foot strike models. A temporal tolerance of ±2 frames (±0.04 s) was used to match ground truth manually labelled foot strikes with predicted class labels. Evaluated metrics were sensitivity, specificity, accuracy, and precision.

Stride parameters were calculated using both manually labelled and automated foot strikes. The difference between step time, stride time, and cadence from each group was calculated. These differences were compared to the minimal detectable change (MDC) for each stride parameter. Since MDC was not available for lower limb amputee gait, stride parameter MDC for healthy older adults was used [79, 80, 81].

5.3.6 Post-Processing

In [90], periods of repeated foot strike predictions corresponding with a single AP acceleration peak were observed in the preliminary data. This was also observed for the model trained on 2MWT data. Predicted foot strike labels were post-processed in MATLAB 2020b to correct for model prediction errors, including extra foot strike predictions and missed steps. Extra predictions were removed by identifying instances where two or more consecutive foot strike classifications occurred. The start and end of periods of consecutive predictions were located and the peak AP acceleration within the band was identified. The foot strike event corresponding to the AP peak was selected and all other predictions in this period were removed. To identify missed steps, periods where the duration between two consecutive steps was greater than 1.5 times the previous step were identified. An adaptive locking period was applied and searched for potential missed steps. Within the adaptive locking period, the AP acceleration peak was identified, and a foot strike was inserted at this timestamp. Final cleaned predictions were used for feature calculations.

5.3.7 Feature Calculations and Fall Risk Classification

A random forest machine learning model developed by Daines et al. [8] was used for amputee fall risk classification. Smartphone acceleration and angular velocity signals were used to calculate step-based feature sets, one using automated foot strikes from the 6M-FS model, one using automated foot strikes from the 2M-FS model, and a comparator set using manually labelled foot strikes. 62 features were extracted for each feature set (Table 9). Once features were extracted for each step, the minimum, maximum, mean, and standard deviation were calculated over all included steps for a total of 248 features (62 features multiplied by 4 statistics) per data set.

5.4 Results

A total of 12,308 foot strikes were identified and labelled in the first 2 minutes of walk test data, accounting for 3.06% of total output labels (402,000). Table 10 displays foot strike classification confusion matrices for the 6M-FS and 2M-FS models. The 6M-FS accuracy was 99.2%, sensitivity was 91.7%, specificity was 99.4%, precision was 82.7%, and F1-score was 0.87. The 2M-FS accuracy was 98.0%, sensitivity was 65.0%, specificity was 99.1%, precision was 68.6%, and F1-score was 0.67. 2M-FS model foot strike classification was poor, with 35% of steps missed. Therefore, further analysis of stride parameters and fall risk classification were not possible using foot strikes from the 2M-FS model.

Automated foot strikes identified using the 6M-FS model and manually labelled FS from the first two minutes of data were used to calculate stride parameters. The average and standard deviation difference between manual and automated foot strike stride parameters were calculated and compared to MDC values for these outcomes (Table 11).

Table 9 Feature list. AP = anterior-posterior; ML= medio-lateral; RMS= root-mean square; FFT= fast Fourier transform; REOH= ratio of even/odd harmonic frequencies.

Temporal	**Descriptive Statistics**	**Frequency Domain Features**
Cadence	Minimum ML	Quartile FFT ML
Step time right	Minimum AP	Quartile FFT AP
Step time left	Minimum Vert	Quartile FFT Vert
Stride time	Maximum ML	Quartile FFT Tilt
Symmetry index	Maximum AP	Quartile FFT Rotation
	Maximum Vert	Quartile FFT Obliquity
	Mean ML	Maximum FFT ML
	Mean AP	Maximum FFT AP
	Mean Vert	Maximum FFT Vert
	Mean Tilt	Maximum FFT Tilt
	Mean Rotation	Maximum FFT Rotation
	Mean Obliquity	Maximum FFT Obliquity
	Range Tilt	Standard Deviation FFT ML
	Range Rotation	Standard Deviation FFT AP
	Range Obliquity	Standard Deviation FFT Vert
	Standard Deviation ML	Standard Deviation FFT Tilt
	Standard Deviation AP	Standard Deviation FFT Rotation
	Standard Deviation Vert	Standard Deviation FFT Obliquity
	Standard Deviation Tilt	Peak Distinction FFT ML
	Standard Deviation Rotation	Peak Distinction FFT AP
	Standard Deviation Obliquity	Peak Distinction FFT Vert
	RMS ML	Peak Distinction FFT Tilt
	RMS AP	Peak Distinction FFT Rotation
	RMS Vert	Peak Distinction FFT Obliquity
	RMS Tilt	REOH ML
	RMS Rotation	REOH AP
	RMS Obliquity	REOH Vert
		REOH Tilt
		REOH Rotation
		REOH Obliquity
Symmetry index: symmetry in right and left limb step times [92]		

Fall risk analysis was completed with foot strikes during the first 2 minutes of data, identified using the 6M-FS model, and the Daines et al. fall risk model [8] (Table 12). 61 of 80 participants were correctly classified (76.3% accuracy, 48.1% sensitivity, 90.6% specificity, 72.2% precision, F1-score 0.58). Classification using features from manually labelled foot strikes resulted in 63 out of 80 participants correctly classified (78.8% accuracy, 51.9% sensitivity, 92.5% specificity, 77.8% precision, F1-score 0.62).

Table 10 Foot strike classification

6M-FS			**2M-FS**		
	Foot Strike	No Foot Strike		Foot Strike	No Foot Strike
Foot strike	11,283	1,025	Foot strike	8,006	4,302
No foot strike	2,361	386,010	No foot strike	3,672	383,200

Table 11 Average and standard deviation (in brackets) difference between manual and automated foot strike stride parameter outcome measures for 6M-FS model. MDC = minimum detectable change.

	Automated foot strikes	**MDC**
Step time (s)	0.045 (0.11)	0.042
Stride time (s)	0.044 (0.09)	0.772
Cadence (steps/min)	7.36 (-0.26)	8.44

Table 12 Fall risk classification confusion matrices for automated and manual foot strike identification.

Automated FS			**Manual FS**		
	Fall Risk	No Fall Risk		Fall Risk	No Fall Risk
Fall risk	13	14	Fall risk	14	13
No fall risk	5	48	No fall risk	4	49

5.5 Discussion

This research had two outcomes, foot strike identification and fall risk classification. The automated LSTM foot strike identification approach from [90], when trained on smartphone signals from six minutes of data and applied to two minutes of data, identified foot strike and non-foot strike events with 99.2% accuracy. When an LSTM was trained on smartphone signals from only the first two minutes of data and applied to two minutes of data, foot strike identification was poor (35% of steps missed), even after offline post-processing for error correction. The LSTM trained on six minutes of data outperformed the LSTM trained on 2 minutes of data in all foot strike identification performance metrics.

Further investigation into the results of the LSTM trained on 2 minutes of data (i.e., 2M-FS) revealed that, for 23 participants (~34% of participants), fewer than 50% of the average number of foot strikes were detected. Furthermore, 7 of these 23 participants had fewer than 10 total steps detected by the LSTM. There were no identifiable similarities between these participants, and this sub-group included both people who were fall risk and people who were not at risk of falling. In this case, stride parameter analysis and step-based feature calculation for fall risk classification for these participants would not be feasible since most features are based on stride analysis. Therefore, to complete stride parameter analysis and fall risk classification for all participants, only automated foot strikes identified in the first two minutes of data by the 6M-FS model were used.

LSTM is a deep learning approach that is often trained on large datasets of sequential data. Deep learning approaches are more complex than decision trees and other machine learning approaches and can often perform with higher accuracy on large datasets. However deep learning often requires a greater amount of labelled data to prevent overfitting when training [10]. This could explain why the LSTM trained on the full six minutes of walk test data had better foot strike identification than the LSTM trained only on the first two minutes of walking; 2 minutes of walking for each participant was not enough data for an LSTM to isolate foot strike events.

Step time, stride time and cadence were calculated using automated foot strikes from the 6M-FS model and using manually labelled foot strikes from the first two minutes of data. The difference in stride time between automated foot strike and manually labelled foot strike calculations from the first two minutes of data was within the MDC for healthy older adults, while step time and cadence were outside of the MDC for healthy older adults. MDC for healthy older adults was used as a comparator since these values do not exist for lower limb amputees. Average step time and stride time for automated foot strikes were both within approximately 0.04s of the manually labelled foot strikes, suggesting step time and stride time calculated using automated foot strikes are comparable to manually labelled foot strikes for most participants.

Fall risk classification was performed using a random forest model trained on step-based features calculated from automated foot strikes identified in the first two minutes of data by the 6M-FS model. While the random forest correctly classified 76.3% of all participants and over 90% of non-fall risk participants, only 13 out of 27 (48.1%) fall risk participants were correctly classified. This means that ~52% of people who had fallen in the past six months were mistakenly classified as not being at risk of falling. In a clinical setting, this could translate to patients not being referred for further testing and a delay in the implementation of fall intervention strategies. These results indicated that step-based features from a 2MWT cannot be recommended for fall risk classification in lower limb amputees without further model development. Further refinement of the fall risk model using data from a 6MWT could improve classification results to a clinically usable standard.

It is important to consider that this study used the first 2 minutes of a 6MWT to approximate a 2MWT. When observed in a clinical setting, some people begin walking at a comfortable pace with a typical gait pattern, but more gait deviations and instability become apparent as they continue walking. People may also walk faster if they know they only have to walk for two minutes instead of six minutes. This may explain why fall risk classification was worse when features were calculated from two minutes of data than six minutes of data; the signals collected during the first two minutes of data were not sufficient to differentiate fall risk participants and non-fall risk participants.

It is possible that data from the full six minutes may not be required to differentiate fallers from non-fallers. In previous fall risk classification research, random forest models trained on signals from 6MWT trials focused on signal data during turns only. For example, Drover et al. [20] classified fall risk in older adults from 6MWT data collected from multiple sensors with 77.3% accuracy, 66.1% sensitivity, and 84.7% specificity and noted that turn data improved all classification metrics. In addition, Daines et al. [8] used turn data from 6MWT trials in lower limb amputees with 81.3% accuracy, 57.2% sensitivity, and 94.9% specificity, however manual foot strikes and turn identification was required. Future research could examine if features calculated from automated foot strikes during turns from 2MWT (and 6MWT) would improve fall risk classification. Additionally, differences in gait characteristics that distinguish fallers and non-fallers may become more apparent as the test progresses. Step-based features calculated from automated foot strikes during the final minutes of the walk test could be analyzed as a possible improvement on the fall risk classification model. Future studies could also investigate other gait-based features that may better differentiate fallers from non-fallers in a shorter walk test.

Fall risk information is not normally available from 2 or 6MWT, so this is promising for the future development of an AI-enhanced TOHRC Walk test app for use in lower limb amputees without requiring separate models integrated into the smartphone application for each walk test. Stride parameters can be automatically calculated from the automated foot strikes for either a 6MWT or 2MWT immediately after completing the walk test with reasonable confidence. A model for fall risk classification could be integrated into a future application and applied to 6MWT data, however, it is not recommended for 2MWT data.

5.6 Conclusions

Foot strike identification is essential to define the gait cycle and calculate stride parameters. AI tools for clinical analysis (e.g., fall risk classification) rely on proper gait segmentation to calculate step-based features. This research determined that a smartphone app can provide accelerometer and gyroscope signals during a 6MWT or 2MWT for AI-based analyses to automatically determine foot strikes. The foot strike identification model used a LSTM deep learning approach trained on six minutes of data, with this model being applicable for identifying foot strikes in both 6 minutes of data and 2 minutes of data with at least 99% accuracy. However, a model only trained on the first two 2 minutes of data had poor foot strike identification results, thereby not supporting use of this approach for outcome measurement.

Step and stride time calculated using automated foot strikes in the first two minutes of data identified by the 6M-FS model of smartphone data were equivalent to manually labelled foot strikes for most participants, indicating that the 2MWT stride outcomes measurements could be viable for clinical analysis. Integration of this foot strike detection model into the TOHRC Walk Test app could allow for immediate stride parameter analysis in lower limb amputees after completing a 6MWT or 2MWT. However, fall risk classification using step-based features calculated from automated foot strikes is not recommended for the 2MWT.

6 Thesis Conclusions and Future Work

Throughout this thesis a novel deep learning approach was developed and evaluated for foot strike identification in lower limb amputees and the clinical application of automated foot strikes was evaluated. Smartphone signals collected during a 6MWT were used to train a decision tree and an LSTM for foot strike identification. The LSTM outperformed the decision tree in all outcome measures and is the recommended model, between these two, for foot strike identification in lower limb amputees. Automated foot strikes were used to calculate stride parameters, such as step-time, and the results were equivalent to stride parameters calculated from manually labelled foot strikes. The LSTM trained on 6MWT data was used to identify foot strikes in 2 minutes of data, indicating that foot strikes could be detected in a 6MWT or a 2MWT without requiring a separate model. Step-based features were calculated using automated and manually labelled foot strikes from 6 and 2 minutes of data and were used to train a random forest for fall risk classification. Step-based features calculated using automated foot strikes from 6 minutes of data correctly classified more participants than features calculated using foot strikes from 2 minutes of data, though manual labelling improved fall risk classification for both 6 and 2-minute models.

This research used signal data from a single smartphone during a 6MWT to detect foot strikes and calculate stride parameters in lower limb amputees without feature calculation. The model proposed in this thesis could be integrated into a smartphone application (e.g., as an enhancement to the TOHRC Walk Test App) to allow for immediate outcome reporting after completing a single movement assessment (i.e., 6MWT or 2MWT). Furthermore, automated foot strikes from the 6MWT correctly classified ~56% of fall risk participants and >80% of non-fall risk participants, an outcome not typically available from a 6MWT. Therefore, manual foot strike labelling is not required for clinical analysis in lower limb amputees. However, <50% of fall risk participants were correctly classified using automated foot strikes from 2 minutes of data, so fall risk classification is not yet recommended for 2MWT data.

The conclusions for each thesis objective and the corresponding hypotheses are:

6.1 Objective 1

Create and evaluate a viable artificial intelligence model for automatically identifying foot strikes during walking in people with lower limb amputation using smartphone sensor data collected during a 6MWT, comparing machine learning and deep learning approaches.

- Hypothesis 1: A deep learning approach will classify foot strike and non-foot strike events with greater accuracy, sensitivity, and specificity than a machine learning approach [10] or previous rule-based approaches [15].

The best model for foot strike identification was the LSTM with 100 hidden nodes in the LSTM layer, 50 hidden nodes in the dense layer, and batch size of 64 (99.0% accuracy, 86.4% sensitivity, 99.4% specificity, precision was 82.7%, F1-score was 0.85). The best performing decision tree had an accuracy of 98.7%, sensitivity was 83.0%, specificity was 99.2%, precision was 75.9% and F1-score was 0.79. Initially the LSTM and decision tree were trained on 72 participants that completed the full 6MWT; however, after the LSTM demonstrated superior performance, the model was re-trained on the full dataset (n=80). The final model, discussed in Chapter 4, had 99.2% accuracy, 99.4% specificity, 81.8% sensitivity, 90.2% precision, and F1-score of 0.86.

LSTM is a deep learning approach, a subset of artificial intelligence algorithms that perform well on large, sequential datasets, often with higher accuracy than other machine learning approaches. Despite this, both the LSTM and the decision tree required post-processing for error correction. Nevertheless, both models developed in this thesis were an improvement over the previous rule-based method [15]. The increased complexity of the LSTM architecture could have identified patterns in amputee gait more effectively. Additionally, the LSTM was trained on 12 signals collected from the smartphone, where the previous rule-based method relied on only tri-axial acceleration data.

Post-processing was necessary for both models after numerous false positives were identified. The repeated patterns of false positives corresponded to AP signal peaks, and could therefore be easily identified and corrected. The post-processing method was developed specifically for this thesis and greatly improved model performance, notably increasing precision of the LSTM by 61.9%. In some cases, after post-processing was applied, a more appropriate peak was identified by the model as the location of the foot strike compared to the ground truth label, which was identified by a human observer. Post-processing also identified instances of missed steps by the model, improving sensitivity by 8.2%.

- Hypothesis 2: Stride parameters (step time, stride time, etc.) from automated foot strikes will be equivalent to parameters calculated from manually labelled foot strikes.

The difference between step time, stride time, and cadence when calculated using manually labelled foot strikes or automated foot strikes were compared between LSTM and decision tree approaches. For both models, step time and stride time differences between automated and manually labelled foot strikes were minimal, but cadence differences were outside the MDC for this parameter and cadence had larger variability than the other two parameters. There is no available MDC for lower limb amputees for these parameters, so the MDC for healthy older adults was used as a comparator. The LSTM had smaller differences overall compared to the decision tree, which was expected due to improved foot strike classification accuracy with the LSTM.

6.2 Objective 2

Evaluate a random forest fall risk classification model for lower limb amputees using automated foot strikes and manually labelled foot strikes and smartphone signals collected from a 6MWT for step-based feature calculation.

- Hypothesis: Step-based features calculated from automated foot strikes will result in fall risk classification with equivalent with accuracy, sensitivity, and specificity to manually labelled foot strikes.

The random forest fall risk model trained on features calculated using manually labelled foot strikes correctly classified fall risk for 64 of 80 participants. Automated foot strikes correctly classified 58 of 80 participants. While both approaches had the same sensitivity (55.6%), automated foot strike features resulted in 6 more false positives than manually labelled foot strikes, decreasing specificity and accuracy compared to manually labelled foot strikes. Therefore, the current foot strike model led to slightly poorer fall risk classification performance; however, automated foot strike detection will enable a clinically viable application that could identify at least half the people at risk of falling.

6.3 Objective 3

Evaluate viability of automated foot strike identification and fall risk classification models for lower limb amputees using smartphone signals collected from a 2MWT.

- Hypothesis: An LSTM trained on smartphone signals from only 2 minutes of data will result in lower foot strike classification accuracy, sensitivity, and specificity than an LSTM trained on smartphone signals from 6 minutes of data.

The LSTM trained on full trial data outperformed the LSTM trained on the first two minutes of data in all performance metrics. The 6-minute model had 99.2% accuracy, 91.7% sensitivity, 99.4% specificity, and 82.7% precision. The 2-minute model had 98.0% accuracy, 65.0% sensitivity, 99.1% specificity, and 68.6% precision. Similar to Chapter 3, both models required post-processing for error correction. Despite this, the 2-minute model still had a foot strike identification error rate of 35%. When these results were investigated more thoroughly, for approximately 34% of participants, fewer than half of the average number of foot strikes were identified. Due to the low sensitivity for 2-minute foot strike identification, clinical outcome measures were only analyzed from the first two minutes of data using foot strikes identified by the 6-minute model.

It is likely that two minutes of data is not sufficient to train a deep learning model for foot strike classification, especially due to the high variability in lower limb amputee gait. However, the LSTM high sensitivity and specificity when trained on full 6MWT trials suggests that this model can be applied to both 6MWT and 2MWT data for foot strike identification in lower limb amputees and that a secondary model is not required.

- Hypothesis: Stride parameters (step time, stride time, etc.) calculated from automated foot strikes will be equivalent to parameters calculated from manually labelled foot strikes.

Average differences between manual and automated foot strike stride parameter outcome measures calculated for the first two minutes of data were equivalent to manually labelled for most participants. The standard deviation for all three parameters analyzed was smaller than the standard deviation for parameters calculated in Chapter 3 using automated foot strikes from 6 minutes of data, indicating that gait deviations that might affect foot strike identification become more frequent as the 6MWT progressed.

- Hypothesis: Step-based features calculated from automated foot strikes from 2 minutes of data will result in fall risk classification with equivalent accuracy, sensitivity, and specificity to manually labelled foot strikes from 2 minutes of data.

Using features calculated from automated foot strikes identified in the first two minutes of data, the random forest correctly classified 61 of 80 participants (76.3% accuracy, 48.1% sensitivity, 90.6% specificity). Classification using features from manually labelled foot strikes resulted in 2 additional correctly classified participants, one fall risk and one non-fall risk, for a total of 63 out of 80 participants (78.8% accuracy, 51.9% sensitivity, 92.5% specificity). While manual foot strike labelling resulted in better overall classification, only ~52% of fall risk participants were correctly classified. Therefore, the model described in this thesis is not recommended for a 2MWT in its current form, even with manual foot strike labelling.

Interestingly, specificity for automated foot strike labelling from 2 minutes of data was greater than specificity from full trial data (90.6% compared to 81.1% for 6MWT). While full trial data resulted in higher fall risk sensitivity, it is important to note that the difference between the 2-minute and 6-minute fall risk models is only two misclassified participants. Fall risk classification could not be recommended for a 2MWT based on the research statistics; however, further 2MWT research with an appropriate sample size could show better results and allow for a change in the recommendation.

6.4 Future Work

This thesis is an important contribution to foot strike detection literature for lower limb amputees. However, there are areas of future development that can be pursued to further enhance this research, including pre-processing and post-processing methods, improvements to fall risk classification, participant diversity, and implementation for clinical application.

6.4.1 Pre-processing

The pre-processing method used in this thesis included linear integration of smartphone signals to 50Hz, followed by fourth-order zero-lag Butterworth low pass filter with a 4Hz cut-off frequency. Butterworth low pass filters provide the most robust performance for biomechanical signals however filter cut-offs differ in the literature [91]. Human movement is often filtered at 6Hz, but can vary between 3 and 10Hz [112]. A 4Hz filter was used in this thesis based on previous research that required a 4Hz cut off frequency to obtain a usable signal for the rule-based foot strike detection [2]. Future work could investigate different filtering frequencies for smartphone signals from the 6MWT to determine if there is a more appropriate cut off frequency. For example, keeping more of the signal with a higher cut-off frequency could provide useful data for the LSTM network.

The LSTM in this research was trained on 12 signals extracted from the smartphone; however, all 12 signals may not have been required for foot strike detection. Further investigation could determine which signals were important for foot strike detection, and which signals were extraneous for LSTM training in lower limb amputees. Removing unnecessary signals may improve the model performance, or perhaps improve efficiency when running the model.

6.4.2 **Post-processing**

In this thesis, post-processing was essential to the overall success of the foot strike identification model. However, there was still an error rate of 13.6% (i.e., foot strikes were unidentified or identified in an inappropriate location). This error rate could be improved with adjustments to the foot strike identification model; however, it is possible that improvements to the post-processing method may further improve classification results without having to retrain the LSTM. This could include changes or adjustments to how the data is searched for and inserts missed steps and improvements to the code that identifies false positive predictions.

6.4.3 **Fall risk classification**

The fall risk classification approaches in this thesis could be improved to achieve greater fall risk sensitivity and specificity. Confusion matrixes from Chapter 4 (Table 7) and Chapter 5 (Table 12) were similar; however, 6MWT analysis correctly classified 2 additional fall risk participants and first two minutes of data analysis correctly classified 5 additional non-fall risk participants. The full trial data was necessary for training the LSTM for foot strike classification; however, it is possible that full trial data is not needed for fall risk classification. One potential solution could be to only use data from automated foot strikes during turns. In previous research using manually labelled foot strikes, fall risk classification in lower limb amputees was improved by using turn data only [8]. Compared to automated foot strikes from full trial data, fall risk sensitivity using manually labelled turn data from [6] was similar to automated foot strikes in this research (57.2% vs. 55.6%), but fall risk specificity was much higher (94.9% vs 81.1%) using foot strikes from turn data, indicating that this could improve classification results of the automated foot strike approach.

Additionally, the number of gait deviations typically increases as the trial progresses. The increased presence of these deviations, which may exacerbate instability, could better discriminate those at risk of falls from participants who are not fall risks. Features calculated using foot strikes identified from the first two minutes of data resulted in greater specificity than full trial data. However, 2-minute analysis could not be recommended for fall risk classification due to the decreased sensitivity. Features calculated from automated foot strikes identified in the final minutes of the 6MWT could be researched to determine if this subset improves classification. Future work to improve fall risk classification in lower limb amputees using automated foot strikes could also benefit from investigating whether a combination of these two approaches optimizes performance.

Another area of future work for fall risk classification could be investigating different factors contributing to an individual's risk of falling and whether additional participant information could be a valuable feature for model training. In this research, participant's self-reported fall history was used for initial classification. However, there are numerous factors that could indicate if a person with a lower limb amputation is at risk of fall, including their age, other comorbidities, and time since amputation [3].

6.4.4 **Participant diversity**

The majority of the participants included in this thesis were transtibial amputees (90.0%) and mostly male (78.8%) limiting the generalizability of the results. While a small percentage of transfemoral and bilateral amputees were included, a greater number of these populations would benefit the fall risk classification model since walking patterns of these amputee groups typically differ. Additionally, this thesis did not include people with partial foot, ankle, or hip disarticulation amputations. Future work in this area should include a more diverse population to assess model performance on both foot strike identification and fall risk classification and, if required, perform additional model development including these cohorts.

6.4.5 **Clinical implementation**

The implementation of the foot strike identification model into a smartphone application would allow for automated evaluation of clinical outcomes for lower limb amputees. The TOHRC Walk Test App currently provides immediate outcome reporting for able-bodied participants, providing stride parameter analysis immediately after completing a 6MWT or 2MWT. Ideally, integration of this model into the TOHRC Walk Test App would provide better outcome analysis for clinicians treating people with lower limb amputation. However, the clinical applications demonstrated in this thesis (stride parameters, fall risk analysis) have not yet been validated in additional lower limb amputee participants. While this model was trained on data from, and developed for, people with lower limb amputation, future research should investigate if this model achieves acceptable results with other populations with variable gait.

7 References

[1] H. Prasanth, M. Caban, U. Keller, G. Courtine, A. Ijspeert, H. Vallery and J. von Zitzewitz, "Wearable Sensor-Based Real-Time Gait Detection: A Systematic Review.," *Sensors,* vol. 21, p. 2727, 2021.

[2] N. Capela, E. Lemaire and N. Baddour, "Novel algorithm for a smartphone-based 6-minute walk test application: Algorithm, application development, and evaluation.," *J. Neuroeng. Rehabilitation.* , vol. 12, no. 19, 2015.

[3] N. Steinberg, A. Gottlieb, I. Siev-Ner and M. Plotnik, "Fall incidence and associated risk factors among people with a lower limb amputation during various stages of recovery – a systematic review.," *Disability and Rehabilitation,* vol. 41, no. 15, pp. 1778-1787, 2019.

[4] W. C. Miller, M. Speechley and B. Deathe, "The prevalence and risk factors of falling and fear of falling among lower extremity amputees," *Archives of Physical Medicine and Rehabilitation,* vol. 82, no. 8, pp. 1031-1037.

[5] A. Olenšek, M. Zadravec, H. Burger and Z. Matjačić, "Dynamic balancing responses in unilateral transtibial amputees following outward-directed perturbations during slow treadmill walking differ considerably for amputated and non-amputated side," *J Neuroeng Rehabil,* vol. 18, no. 1, p. 123, 2021.

[6] C. K. Wong, S. T. Chihuri and G. Li, "Risk of fall-related injury in people with lower limb amputations: A prospective cohort study," *J Rehabil Med,* vol. 48, no. 1, pp. 80-85, 2016.

[7] S.-H. Park, "Tools for assessing fall risk in the elderly: a systematic review and meta-analysis," *Aging Clin. Exp. Res.,* vol. 30, no. 1, pp. 1-16, 2018.

[8] K. Daines, N. Baddour, H. Burger, A. Bavec and E. Lemaire, "Fall risk classification for people with lower extremity amputations using random forests and smartphone sensor features from a 6-minute walk test.," *PLoS ONE,* vol. 16, p. e0247574, 2021.

[9] J. Farah, N. Baddour and E. Lemaire, "Design, development, and evaluation of a local sensor-based gait phase recognition system using a logistic model decision tree for orthosis-control.," *J. Neuroeng. Rehabilitation,* vol. 16, p. 22, 2019.

[10] C. Kamath, S. Bukhari and A. Dengel, "Comparative Study Between Traditional Machine Learning and Deep Learning Approaches for Text Classification," in *Proceedings of the ACM Symposium on Document Engineering 2018 (DocEng '18); Association for Computing Machinery,* New York, NY, USA, 2018.

[11] E. Simonetti, C. Villa, J. Bascou, G. Vannozzi, E. Bergamini and H. Pillet, "Gait event detection using inertial measurement units in people with transfemoral amputation: a comparative study," *Medical & Biological Engineering & Computing,* vol. 58, no. 1, pp. 461-470, 2020.

[12] M. Sharifi Renani, C. Myers, R. Zandie, M. Mahoor, B. Davidson and C. Clary, "Deep Learning in Gait Parameter Prediction for OA and TKA Patients Wearing IMU Sensors," *Sensors,* vol. 20, no. 19, p. 5553, 2020.

[13] H. Maqbool, M. Husman, M. Awad, A. Abouhossein, P. Mehryar, N. Iqbal and A. Dehghani-Sanij, "Real-time gait event detection for lower limb amputees using a single wearable sensor.," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL., USA, 16-20 August 2016.

[14] H. F. Maqbool, M. A. B. Husman, M. I. Awad, A. Abouhossein, N. Iqbal, M. Tahir and A. A. Dehghani-Sanij, "Heuristic real-time detection of temporal gait events for lower limb amputees," *IEEE Sensors Journal,* vol. 19, no. 8, pp. 3138-3148, 2018.

[15] G. Thibault, "Amputee FS Identification Accuracy from 6 Minute Walk Test Raw Data.," Dept. Mechanical Engineering, Univ. of Ottawa, Ottawa, ON, Canada, 2019.

[16] A. Gupta, A. Jadhav, S. Jadhav and A. Thengade, "Human gait analysis based on decision tree, random forest and KNN algorithms.," in *Applied Computer Vision and Image Processing. Advances in Intelligent Systems and Computing,* Singapore, 2020.

[17] Y. Zhang and D. Gu, "A Deep Convolutional-Recurrent Neural Network for Freezing of Gait Detection in Patients with Parkinson's Disease," in *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI),* Suzhou, China, 19-21 October 2019.

[18] Ł. Kidziński, S. Delp and M. Schwartz, "Automatic real-time gait event detection in children using deep neural networks.," *PLoS ONE,* vol. 14, p. e0211466, 2019.

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

[20] D. Drover, J. Howcroft, J. Kofman and E. D. Lemaire, "Faller Classification in Older Adults Using Wearable Sensors Based on Turn and Straight-Walking Accelerometer-Based Features," *Sensors,* vol. 17, p. 1321, 2017.

[21] J. Perry and J. M. Burnfield, Gait Analysis: Normal and Pathological Function, vol. 9, Thorofare, New Jersey: SLACK Incorporated, 2010, p. 353.

[22] T. Stöckel, R. Jackstei, M. Behrens, R. Skripitz, R. Bader and A. Mau-Moeller, "The mental representation of the human gait in young and older adults," *Frontiers in Psychology,* vol. 6, p. 943, 2015.

[23] C. L. Roether, L. Omlor, A. Christensen and M. A. Giese, "Critical features for the perception of emotion from gait.," *Journal of Vision,* vol. 9, no. 6, p. 15, 2009.

[24] W. Pirker and R. Katzenschlager, "Gait disorders in adults and the elderly.," *Wien Klin Wochenschr,* vol. 129, pp. 81-95, 2017.

[25] A. G. Bernal, R. Becerro-de-Bengoa-Vallejo and M. E. Losa-Iglesias, "Reliability of the OptoGait portable photoelectric cell system for the quantification of spatial-temporal parameters of gait in young adults," *Gait & Posture,* vol. 50, pp. 196-200, 2016.

[26] T. M. Owings and M. D. Grabiner, "Variability of step kinematics in young and older adults," *Gait & Posture,* vol. 20, no. 1, pp. 26-29, 2004.

[27] O. Beauchet, G. Allali, H. Sekhon, J. Verghese, S. Guilain, J.-P. Steinmetz, R. W. Kressig, J. M. Barden, T. Szturm, L. C. P., S. Grenier, L. Bherer, T. Liu-Ambrose, V. L. Chester and M. L. Callisaya, "Guidelines for Assessment of Gait and Reference Values for Spatiotemporal Gait Parameters in Older Adults: The Biomathics and Canadian Gait Consortiums Initiative," *Frontiers in Human Neuroscience,* vol. 11, 2017.

[28] W. S. Kim and E. Y. Kim, "Comparing Self-Selected Speed Walking of the Elderly With Self-Selected Slow, Moderate, and Fast Speed Walking of Young Adults," *Ann Rehabil Med.,* vol. 38, no. 1, pp. 101-108, 2014.

[29] E. D. Lemaire, "International Society for Prosthetics and Orthotics Stride Parameter Database: Unilateral Transtibial and Transfemoral Amputees," ISPO Scientific Commitee, Brussels, 2014.

[30] S. W. Hunter, F. Batchelor, K. D. Hill, A. M. Hill, S. Mackintosh and M. Payne, "Risk Factors for Falls in People With a Lower Limb Amputation: A Systematic Review.," *PMR,* vol. 9, no. 2, pp. 170-180, 2017.

[31] C. K. Wong and S. T. Chihuri, "Impact of Vascular Disease, Amputation Level, and the Mismatch Between Balance Ability and Balance Confidence in a Cross-Sectional Study of the Likelihood of Falls Among People With Limb Loss," *American Journal of Physical Medicine & Rehabilitation.,* vol. 98, no. 2, pp. 130-135, 2019.

[32] A. Muro-de-la-Herran, B. Garcia-Zapirain and A. Mendez-Zorrilla, "Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications.," *Sensors,* vol. 14, no. 2, pp. 3362-3394, 2014.

[33] H. Lim, B. Kim and S. Park, "Prediction of Lower Limb Kinetics and Kinematics during Walking by a Single IMU on the Lower Back Using Machine Learning.," *Sensors,* vol. 20, no. 1, p. 130, 2020.

[34] I. Ghersi, M. H. Ferrando, C. G. Fliger, C. Castro Arenas, D. J. Edwards Molina and M. T. Miralles, "Gait-cycle segmentation method based on lower-trunk acceleration signals and dynamic time warping," *Medical Engineering & Physics,* vol. 82, no. 1, pp. 70-77, 2020.

[35] R. J. Mobbs, J. Perring, S. M. Raj, M. Maharaj, N. Yoong, L. W. Sy, R. D. Fonseka, P. Natarajan and W. J. Choy, "Gait metrics analysis utilizing single-point inertial measurement units: a systematic review," *mHealth,* vol. 8, no. 9, 2022.

[36] M. Straczkiewicz, P. James and J. P. Onnela, "A systematic review of smartphone-based human activity recognition methods for health research," *npj Digit. Med.,* vol. 4, p. 148, 2021.

[37] S. Mekruksavanich and A. Jitpattanakul, "LSTM Networks Using Smartphone Data for Sensor-Based Human Activity Recognition in Smart Homes.," *Sensors,* vol. 21, p. 1636, 2021.

[38] B. Manor, W. Yu, H. Zhu, R. Harrison, O.-Y. Lo, L. Lipsitz, T. Travison, A. Pascual-Leone and J. Zhou, "Smartphone App–Based Assessment of Gait During Normal and Dual-Task Walking: Demonstration of Validity and Reliability.," *JMIR MHealth UHealth ,* vol. 6, no. e36, 2018.

[39] P. Picerno, M. Iosa, C. D'Souza, M. G. Benedetti, S. Paolucci and G. Morone, "Wearable inertial sensors for human movement analysis: a five-year update," *Expert Review of Medical Devices*, vol. 18, no. sup1: Digital Health, pp. 79-94, 2021.

[40] Vicon Motion Systems Ltd, "Vicon," [Online]. Available: https://www.vicon.com/. [Accessed 28 February 2022].

[41] E. Reznick, K. Embry, R. Neuman, E. Bolívar-Nieto, N. P. Fey and R. D. Gregg, "Lower-limb kinematics and kinetics during continuously varying human locomotion," *Sci Data*, vol. 8, p. 282, 2020.

[42] M. Peterson, D. Ewins, A. Shaheen and P. Formento, "Evaluation of Methods Based on Conventional Videography for Detection of Gait Events.," in *VI Latin American Congress on Biomedical Engineering CLAIB 2014, Paraná, Argentina 29, 30 & 31 October 2014. IFMBE Proceedings*, Paraná, Argentia, 2015.

[43] Z. Aftab and R. Shad, "Estimation of gait parameters using leg velocity for amputee population," *PLoS ONE,* vol. 17, no. 5, p. e0266726, 2022.

[44] J. A. Zeni, J. G. Richards and J. S. Higginson, "Two simple methods for determining gait events during treadmill and overground walking using kinematic data," *Gait and Posture,* vol. 27, no. 4, pp. 710-714, 2008.

[45] C. McGuirk, "A multi-view video based deep learning approach for human movement analysis," M.A.Sc Thesis, University of Ottawa, Ottawa, October, 2021.

[46] E. D. Ledoux, "Inertial Sensing for Gait Event Detection and Transfemoral Prosthesis Control Strategy," *IEEE Transactions on Biomedical Engineering,* vol. 65, no. 12, pp. 2704-2712, December 2018.

[47] J. B. Dingwell, B. L. Davis and D. M. Frazder, "Use of an instrumented treadmill for real-time gait symmetry evaluation and feedback in normal and trans-tibial amputee subjects," *Prosthetics and Orthotics International ,* vol. 20, no. 2, pp. 101-110, 2009.

[48] Y. Yu, X. Si, C. Hu and J. Zhang, "A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures," *Neural Comput,* vol. 31, no. 7, pp. 1235-1270, 2020.

[49] J. Kim, N. Colabianchi, J. Wensman and D. H. Gates, "Wearable Sensors Quantify Mobility in People With Lower Limb Amputation During Daily Life," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 28, no. 6, pp. 1282-1291, June 2020.

[50] A. Rattanasak, P. Uthansakul, M. Uthansakul, T. Jumphoo, K. Phapatanaburi, B. Sindhupakorn and S. Rooppakhun, "Real-Time Gait Phase Detection Using Wearable Sensors for Transtibial Prosthesis Based on a kNN Algorith," *Sensors*, vol. 22, no. 11, p. 4242, 2022.

[51] M. Sánchez Manchola, M. Bernal, M. Munera and C. Cifuentes, "Gait Phase Detection for Lower-Limb Exoskeletons using Foot Motion Data from a Single Inertial Measurement Unit in Hemiparetic Individuals.," *Sensors* , vol. 19, no. 13, p. 2988, 2019.

[52] J. Rueterbories, E. G. Spaich and O. K. Anderson, "Gait event detection for use in FES rehabilitation by radial and tangential foot accelerations," *Medical Engineering & Physics*, vol. 36, no. 4, pp. 502-508, 2014.

[53] B. Mariani, H. Rouhani, X. Crevoisier and K. Aminian, "Quantitative estimation of foot-flat and stance phase of gait using foot-worn inertial sensors," *Gait & Posture*, vol. 37, no. 2, pp. 229-234, 2013.

[54] A. Prado, X. Cao, M. Robert, A. Gordon and S. Agrawal, "Gait Segmentation of Data Collected by Instrumented Shoes Using a Recurrent Neural Network Classifier.," *Phys. Med. Rehabilitation Clin. North Am.* , vol. 30, pp. 355-366, 2019.

[55] T. Jiang, J. L. Gradus and A. J. Rosellini, "Supervised Machine Learning: A Brief Primer," *Behavior Therapy*, vol. 51, no. 5, pp. 675-687, 2020.

[56] A. Boschmann, P. Kaufmann and M. Platzner, "Accurate Gait Phase Detection using Surface Electromyographic Signals and Support Vector Machines," in *IEEE Intl. Conf. Bioinformatics and Biomedical Technology*, 2011.

[57] Pedregosa et al., "Support Vector Machines," Scikit-learn: Machine Learning in Python, 2011. [Online]. Available: https://scikit-learn.org/stable/modules/svm.html#id15. [Accessed 14 June 2022].

[58] S. Shalev-Shwartz and S. Ben-David, "18. Decision Trees," in *Understanding Machine Learning*, Cambridge University Press, 2014.

[59] P. Sathiyanarayanan, S. Pavithra, M. Saranya and M. Makeswari, "Identification of breast cancer using the decision tree algorithm.," in *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, Pondicherry, India, 29-30 March 2019.

[60] R. Sumblay, N. Vishnusri and S. Jeyalatha, "Diagnosis of Breast Cancer using Decision Tree Data," *International Journal of Computer Applications*, vol. 98, no. 10, pp. 16-24, 2014.

[61] S. Aich, P. M. Pradhan, S. Chakraborty, H.-C. Kim, H.-T. Kim, H.-G. Lee, I. H. Kim, M. Joo, S. J. Seong and J. Park, "Design of a Machine Learning-Assisted Wearable Accelerometer-Based Automated System for Studying the Effect of Dopaminergic Medicine on Gait Characteristics of Parkinson's Patients," *Journal of Healthcare Engineering*, vol. 2020, pp. 1-11, 2020.

[62] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke and J. Schmidhuber, "A Novel Connectionist System for Unconstrained Handwriting Recognition.," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 31, pp. 855-868, 2009.

[63] U. Kamath, J. Liu and J. Whitaker, Deep Learning for NLP and Speech Recognition, Switzerland: Springer, Cham, 2019.

[64] L. Deng and Y. Liu, Deep Learning in Natural Language Processing, Singapore: Springer, 2018.

[65] R. Chen, L. Mihaylova and H. Zhu, "A Deep Learning Framework for Joint Image Restoration and Recognition," *Circuits Syst Signal Process,* vol. 39, pp. 1561-1580, 2020.

[66] J. Ker, L. Wang, J. Rao and T. Lim, "Deep Learning Applications in Medical Image Analysi," *IEEE Access,* vol. 6, pp. 9375-9389, 2018.

[67] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," *arXiv preprint arXiv:1511.08458.,* 2015.

[68] Wikipedia, "Neural Network," [Online]. Available: https://upload.wikimedia.org/wikipedia/commons/9/99/Neural_network_example.svg . [Accessed 28 February 2022].

[69] L. Zaniolo and O. Marques, "On the use of variable stride in convolutional neural networks," *Multimedia Tools and Applications,* vol. 79, no. 4, 2020.

[70] P. Dhruv and S. Naskar, "Image Classification Using Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN): A Review," in *Machine Learning and Information Processing. Advances in Intelligent Systems and Computing*, Singapore, Springer, 2020, pp. 367-381.

[71] H. Tan, N. Aung, J. Tian, M. Chua and Y. Yang, "Time series classification using a modified LSTM approach from accelerometer-based data: A comparative study for gait cycle detection.," *Gait Posture,* vol. 74, pp. 128-134, 2019.

[72] M. Phi, "Illustrated Guide to LSTM’s and GRU’s: A step by step explanation," Towards Data Science, 24 September 2018. [Online]. Available: https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21. [Accessed 28 February 2022].

[73] X. Meng, H. Yu and M. Tham, "ait phase detection in able-bodied subjects and dementia patients.," in *013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* Osaka, Japan, 3-7 July 2013.

[74] A. Paraschiv-Ionescu, C. Newman, L. Carcreff, C. Gerber, S. Armand and K. Aminian, "Locomotion and cadence detection using a single trunk-fixed accelerometer: Validity for children with cerebral palsy in daily life-like conditions.," *J. NeuroEngineering Rehabil.,* vol. 16, no. 24, 2019.

[75] L. Pepa, F. Verdini and L. Spalazzi, "Gait parameter and event estimation using smartphones.," *Gait Posture,* vol. 57, pp. 217-223, 2017.

[76] H. Vu, D. Dong, H.-L. Cao, T. Verstraten, D. Lefeber, B. VanderBorght and J. Geeroms, "A Review of Gait Phase Detection Algorithms for Lower Limb Prostheses.," *Sensors,* vol. 20, p. 3972, 2020.

[77] H. Vu, F. Gomez, P. Cherelle, D. Lefeber, A. Nowé and B. Vanderborght, " ED-FNN: A New Deep Learning Algorithm to Detect Percentage of the Gait Cycle for Powered Prostheses.," *Sensors,* vol. 18, no. 7, p. 2389, 2018.

[78] M. Ghiasi, S. Zendehboudi and A. Mohsenipour, "Decision tree-based diagnosis of coronary artery disease: CART model.," *Comput. Methods Programs Biomed.,* vol. 192, p. 105400, 2020.

[79] M. Almarwani, S. Perera, J. VanSwearingen, P. Sparto and J. Brach, "The test–retest reliability and minimal detectable change of spatial and temporal gait variability during usual over-ground walking for younger and older adults.," *Gait Posture,* vol. 44, pp. 94-99, 2016.

[80] C. Rábago, J. Dingwell and J. Wilken, "Reliability and minimum detectable change of temporal-spatial, kinematic, and dynamic stability measures during perturbed gait.," *PLoS ONE,* vol. 10, p. e0142083, 2015.

[81] J. Soulard, J. Vaillant, R. Balaguier and N. Vuillerme, "Spatio-temporal gait parameters obtained from foot-worn inertial sensors are reliable in healthy adults in single- and dual-task conditions.," *Sci. Rep.,* vol. 11, p. 10229, 2021.

[82] Parachute, "Cost of Injury in Canada," 2021. [Online]. Available: https://parachute.ca/en/professional-resource/cost-of-injury-in-canada/. [Accessed 27 October 2021].

[83] World Health Organization, "Falls," 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/falls/. [Accessed 27 October 2021].

[84] Centers for Disease Control and Prevention, "Injury Prevention & Control, Keep on Your Feet," 2021. [Online]. Available: https://www.cdc.gov/injury/features/older-adult-falls/index.html . [Accessed 27 October 2021].

[85] B. Moreland, R. Kakara and A. Henry, "Trends in Nonfatal Falls and Fall-Related Injuries Among Adults Aged ≥65 Years — United States, 2012–2018," *MMWR Morb Mortal Wkly Rep.,* vol. 69, pp. 875-881, 2020.

[86] B. R. Greene, A. O'Donovan, R. Romero-Ortuno, L. Cogan, C. N. Scanaill and R. A. Kenny, "Quantitative Falls Risk Assessment Using the Timed Up and Go Test," *IEEE Transactions on Biomedical Engineering,* vol. 57, no. 12, pp. 2918-2926, 2010.

[87] J. Howcroft, E. D. Lemaire, J. Kofman and W. E. McIlroy, "Dual-Task Elderly Gait of Prospective Fallers and Non-Fallers: A Wearable-Sensor Based Analysis.," *Sensors,* vol. 18, no. 4, p. 1275, 2018.

[88] J. Howcroft, J. Kofman and E. D. Lemaire, "Prospective Fall-Risk Prediction Models for Older Adults Based on Wearable Sensors.," *EEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 25, no. 10, pp. 1812-1820, 2017.

[89] C.-H. Lee, S.-H. Chen, B. C. Jiang and T.-L. Sun, "Estimating Postural Stability Using Improved Permutation Entropy via TUG Accelerometer Data for Community-Dwelling Elderly People," *Entropy,* vol. 22, p. 1097, 2020.

[90] P. Juneau, N. Baddour, H. Burger, A. Bavec and E. D. Lemaire, "Comparison of Decision Tree and Long Short-Term Memory Approaches for Automated Foot Strike Detection in Lower Extremity Amputee Populations.," *Sensors,* vol. 21, p. 6974, 2021.

[91] F. Crenna, G. B. Rossi and M. Berardengo, "Filtering Biomechanical Signals in Movement Analysis," *Sensors,* vol. 21, no. 13, p. 4580, 2021.

[92] H. Sadeghi, P. Allard, F. Prince and H. Labelle, "Symmetry and limb dominance in able-bodied gait: a review.," *Gait Posture,* vol. 12, no. 1, pp. 34-45, 2000.

[93] B. Balke, "A simple field test for the assessment of physical fitness," Rep Civ Aeromed Res Inst US, 1963.

[94] ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories, "TS statement: guidelines for the six-minute walk test.," *Am J Respir Crit Care Med,* vol. 166, pp. 111-117, 2002.

[95] S. Mudge and N. Stott, "Timed Walking Tests Correlate with Daily Step Activity in Persons with Stroke.," *Arch Phys Med Rehabil.,* vol. 90, pp. 296-301, 2009.

[96] T. Steffen and M. Seney, "Test-Retest Reliability and Minimal Detectable Change on Balance and Ambulation Tests, The 36-Item Short-Form Health Survey, and the Unified Parkinson Disease Rating Scale in People with Parkinsonism.," *Phys Ther.,* vol. 88, pp. 733-746, 2008.

[97] S.-J. Lin and B. N.H., "Six-Minute Walk Test in Persons with Transtibial Amputation.," *Archives of Physical Medicine and Rehabilitation.,* vol. 89, pp. 2354-2359, 2008.

[98] J. Bean, D. Kiely, S. Leveille, S. Herman, H. C., R. Fielding and W. Frontera, "The 6-Minute Walk Test in Mobility-Limited Elders: What Is Being Measured?," *The Journals of Gerontology: Series A.,* vol. 57, pp. M751-M756, 2002.

[99] "Core Measures: Six Minute Walk Test.," [Online]. Available: http://www.iupac.org/dhtml_home.html. [Accessed 10 December 2021].

[100] T. Gjøvaag, I.-M. Starholm, P. Mirtaheri, F. Hegge and K. Skjetne, "Assessment of Aerobic Capacity and Walking Economy of Unilateral Transfemoral Amputees.," *Prosthetics and orthotics international.,* vol. 38, 2013.

[101] J. F. Burr, B. S., M. Faktor and D. Warburton, "The 6-Minute Walk Test as A Predictor of Objectively Measured Aerobic Fitness in Healthy Working-Aged Adults.," *Phys Sportsmed.,* vol. 39, pp. 133-139, 2011.

[102] Q. Zhang, H. Lu, S. Pan, Y. Lin, K. Zhou and L. Wang, "6MWT Performance and its Correlations with VO2 and Handgrip Strength in Home-Dwelling Mid-Aged and Older Chinese.," *Int. J. Environ. Res. Public Health.*, vol. 14, p. 473, 2017.

[103] A. Mänttäri, J. Suni, H. Sievänen, P. Husu, H. Vähä-Ypyä, H. Valkeinen, K. Tokola and T. Vasankari, "Six-Minute Walk Test: A Tool for Predicting Maximal Aerobic Power (Vo2 Max) in Healthy Adults.," *Clinical Physiology and Functional Imaging*, vol. 38, pp. 1038-1045, 2018.

[104] E. F. Sperandio, "Reference Values for the 6-Min Walk Test in Healthy Middle-Aged and Older Adults: From the Total Distance Traveled to Physiological Responses.," *Fisioterapia em Movimento.*, vol. 32, p. e003231, 2019.

[105] N. Pratiwi, R. Satyawati, D. Tinduh, S. Melaniani and M. Mahmuddah, "Validity and Reliability of 2 Minutes Walking Test in Frailty Elderly.," in *Proceedings of the 11th National Congress and the 18th Annual Scientific Meeting of Indonesian Physical Medicine and Rehabilitation Association (KONAS XI and PIT XVIII PERDOSRI 2019)*, Jakarta, Indonesia, 2019.

[106] D. Scalzitti, K. Harwood, J. Maring, L. S.J., R. E.A. and C. E., "Validation of the 2-Minute Walk Test with the 6-Minute Walk Test and Other Functional Measures in Persons with Multiple Sclerosis.," *Int J MS Care.*, vol. 20, pp. 158-163, 2018.

[107] L. Reid, T. P., M. Besemann and N. Dudek, "Going places: Does the Two-Minute Walk Test Predict the Six-Minute Walk Test in Lower Extremity Amputees?," *J Rehabil Med*, vol. 47, pp. 256-261, 2015.

[108] N. Capela, E. Lemaire and N. Baddour, "A Smartphone Approach for the 2 and 6-Minute Walk Test.," in *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Chicago, Illinois, USA, 2014.

[109] D. G. E. Robertson, G. E. Caldwell, J. Hamill, G. Kamen and S. Whittlesey, Research methods in biomechanics, Champaign, Illinois: Human Kinetics, 2014.

[110] J. Challis, "Data processing and error estimation.," in *Biomechanical Evaluation of Movement in Sport and Exercise: The British Association of Sport and Exercise Sciences Guidelines*, London, Routledge, 2008.

[111] "LSTM layer," [Online]. Available: https://keras.io/api/layers/recurrent_layers/lstm/#lstm-class . [Accessed 17 January 2022].

[112] D. A. Winter, Biomechanics and Motor Control of Human Movement, New York, NY: John Wiley & Sons, 2009.