FALL RISK ASSESSMENT USING GAIT PARAMETERS

by

Nezihe Pehlivan

B.S., Computer Engineering, Boğaziçi University, 2015

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Computer Engineering
Boğaziçi University
2017

ACKNOWLEDGEMENTS

I would love to give my endless gratitude to my supervisor Prof. Cem Ersoy. It has been pleasure to work with him. I always appreciate his effort and guidance in my academic studies; however, it is not the all. I am grateful for his wisdom and perspective that lighten my life.

I would like to give my thanks to my jury members Bert Arnrich and Özlem Durmaz İncel for their invaluable comments and feedback.

I deeply thank to Gülüstü Salur for her guidance throughout my research. Without her help, I would not been able to conduct this research. I would like to give my endless thanks to Nagme Ak for her patience with me and all the laughter we had together. I appreciate the guidance and valuable feedbacks of Bert Arnrich throughout my research. I thank Bilgin Koşucu and Şeyma Erdem for the joyful moments we had together. I enjoyed being a part of Healthware Research group, and I am so grateful for all we shared together.

I would like to thank my colleagues who have shared their knowledge and help; and provided me with their lovely friendships. Especially, I would like to thank Can Tunca for his help and support whenever I needed, without his help I would not succeed to write my thesis; Serhan Daniş for being a great mentor and a dear friend to me; Ahmet Cihat Baktır for all the lovely tea breaks and being my buddy for taking fresh air; Orhan Ermiş for his advices and conversations we had in the breaks. Without the guidance, support and friendship of you all, I would not have such a great experience during my graduate years. Thank you!

I would like to thank Hande Alemdar and Suzan Üsküdarlı for their mentorship, continuous encouragement and support. They always inspired me with their hard work and personalities. I admire them and they are role models to me.

I would like to thank to my lovely friends Burçin Camcı and Selin Erkan for their endless support and encouragement. They always provided me with motivation and love. I am so happy to have you in my life.

I would like to thank my dearest friends Mustafa Kaba, Taha Ceritli and Berkant Kepez for being great friends for me and being next to me whenever I needed. Their friendship, love and support always helped me to become a better person.

I would like to thank all CmpE people for creating such a lovely place. I always appreciate their hard work and attitudes. I enjoyed every moment I shared with them and learned a lot from them.

Last but not least, I would like to thank my beloved sister Aslı for her unconditional love and support. I would like to thank my brother-in-law Mehmet for his support during my thesis. Finally, I would like to thank my parents for always believing in me and providing me with their love and encouragement. I dedicate this thesis to my father.

ABSTRACT

FALL RISK ASSESSMENT USING GAIT PARAMETERS

Falls lead to severe public health problems resulting in devastating psychological and physical consequences for older people and their families. Therefore, fall risk assessment has been a popular field of research in the last decades to understand the underlying reasons of the fall and eventually to identify people at high fall risk so that effective preventive strategies for falls can be taken at an early stage. Clinical and functional fall risk assessment tools are not objective and reliable for the assessment of fall risk as they require human observation and judgement. Related studies aim to assess the fall risk based on gait parameters in a more objective way, however, they either propose obtrusive systems composed of numerous sensors or do not utilize machine learning methodology. In this study, we employ machine learning techniques and aim to develop an accurate, unobtrusive, objective and continuous fall risk assessment system based on gait parameters. For this purpose, we studied two gait analysis techniques: a Kinect-based gait analysis system, and a foot-mounted inertial sensor-based gait analysis system developed by our research group. Using the latter, experiments are conducted to extract gait parameters of 37 subjects of whom 21 have neurological conditions with gait implications. Gait features are computed from these gait parameters. Feature selection techniques are used to determine the most significant gait features. Based on these features, different machine learning algorithms are employed to classify people into high and low fall risk groups. This system enables us to identify people who are likely to experience a fall in the near future and inform their caregivers to take preventive interventions. The predictions are evaluated based on the accuracy, sensitivity, specificity and F-measure.

ÖZET

YÜRÜME PARAMETRELERİNİ KULLANARAK DÜŞME RİSKİ DEĞERLENDİRMESİ

Düşmeler, yaşlılar ve aileleri için ciddi sonuçlar içeren önemli sağlık sorunlarına neden olur. Düşmeleri önlemek için etkili önlemler erken bir aşamada alınmalıdır. Bu nedenle, düşme riskinin değerlendirilmesi, düşme riski seviyesinin saptanması ve bu seviyeye uygun bir önlem alınması için önemlidir. Kullanılan düşme riski değerlendirme ölçekleri, kişiden kişiye değişen gözlem ve yargıya bağlı olduğundan dolayı objektif ve güvenilir değildir. Yürüme parametrelerini kullanarak düşme riskini daha objektif ve doğru bir şekilde değerlendirmeyi amaçlayan sistemler vardır. Ancak bu sistemler ya çok sayıda sensörden oluşur ve kullanımı rahatsızlık vericidir ya da makine öğrenme metodolojisinden faydalanmamaktadır. Bu çalışmada, makine öğrenme tekniklerinden faydalanarak, yürüyüş parametrelerine bağlı, doğruluk oranı yüksek, rahatsızlık verici olmayan, objektif ve sürekli ölçümleme yapmayı hedefleyen bir sistemin geliştirilmesi amaçlandı. Bu amaçla, iki yürüme analizi tekniği incelendi: Kinect tabanlı bir sistem ve araştırma grubumuz tarafından geliştirilen ayağa monte edilen hareket sensörü tabanlı bir sistem. İkinci sistemi kullanarak, deneyler yaptık ve deneylerimize 21 tanesi nörolojik sorunlara sahip olmak üzere toplam 37 denek katıldı. Bu deneylerin amacı yürüme parametreleri üzerinden, yürüme özniteliklerini hesaplamaktı. En önemli yürüme öznitelikleri, öznitelik seçme yöntemleri kullanılarak belirlendi. Bu öznitelikleri kullanarak, çeşitli makine öğrenme yöntemleri ile denekler iki farklı gruba sınıflandırıldı: yüksek ve düşük düşme riskli grup. Bu sistem, düşme riski yüksek olan kişileri tespit etmemizi sağladı ve sistem sayesinde kişinin yakınları önlem almaları için uyarıldı. Sınıflandırma sonuçları doğruluk, duyarlılık, özgüllük ve F-ölçüsü temel alınarak değerlendirildi.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS/ABBREVIATIONS

3D	Three Dimensional
AIC	Akaike Information Criterion
ANOVA	Analysis of Variance
AUC	Area Under Curve
BBS	Berg Balance Scale
CV	Coefficient of Variation
DGI	Dynamic Gait Index
DLB	Dementia Lewy Body
DT	Decision Tree
F	Female
FO	Foot Off
FPR	False Positive Rate
FR	Functional Reach
FRA	Fall Risk Assessment
IC	Initial Contact
IMU	Inertial Measurement Unit
KNN	K Nearest Neighbor
NB	Naive Bayesian
L	Left
LR	Logistic Regression
L/R	Left/Right
M	Male
MCI	Mild Cognitive Impairment
NB	Naive Bayesian
NN	Neural Network
NPH	Normal Pressure Hydrocephalus
OLB	One Leg Balance
PD	Parkinson's Disease

POMA	Tinetti Performance Oriented Mobility Assessment
R	Right
RF	Random Forest
ROC	Reciever Operating Characteristics
SDK	Software Development Kit
SI	Symmetry Index
SVM	Support Vector Machines
TUG	Timed Up and Go
TPR	True Positive Rate
VaMCI	Vascular Mild Cognitive Impairment
VaD	Vascular Dementia
VaP	Vascular Parkinsonism

1. INTRODUCTION

Falls cause severe public health problems with potentially devastating consequences and significant financial implications for older people and their families. Falls mostly lead to serious psychological and physical consequences for older adults such as fear of falling, post fall anxiety, loss of independence and confidence, decreased physical ability, injury, hospitalization and even mortality [1]. These consequences result in a further restriction in daily activities, which causes further damage in the psychological and physical wellbeing of people. Moreover, falls incur both direct and indirect costs to the family and society. Direct costs to the family are the cost of hospitalization, treatment and care. Falls also cause societal and indirect costs that are severe for the family and the society such as loss of productivity of family caregivers which would have been used more effectively if the family member had not experienced fall [2, 3].

Approximately one third of people aged 65 and over fall each year, and for those who aged over 70, the rates are even higher. Between 20% and 30% of the falls result in mild to severe injuries, and approximately 15% of all emergency department visits result from falls [4]. This requires taking effective intervention strategies to prevent falls and finally to eliminate damaging consequences and fall incurred financial costs. Moreover, a fall may be the first indicator of an undetected illness. Therefore, it is a vital issue to identify people at higher risk of fall so that the most appropriate interventions can be implemented and ultimately the risk of fall can be reduced. Thus, elderly's quality of life is improved. In this sense, fall risk assessment plays a significant role in determining the fall risk level of individuals and eventually in helping taking steps to prevent the possible harm and cost to potential fallers and the relevant stakeholders [5].

Falls are caused by the complex interaction of multiple risk factors [3]. Therefore, it is of great significance to investigate these factors and identify the most crucial ones in order to take effective interventions accordingly and eventually to lower the risk and consequences of falling. There has been extensive research to identify which risk factors have the greatest association with the falls [6–9]. These studies focus on different

aspects of the causes and research the association between the potential causes of falls and fall occurrences among both healthy people and people with different pathological conditions in order to make appropriate decisions on which factors need intervention.

The main risk factors of the fall are the combination of various health determinants that directly or indirectly affect the wellbeing of people. Mainly fall risk factors can be divided into two main categories: intrinsic (personal) and extrinsic (environmental) risk factors. Intrinsic risk factors include individual characteristics such as age, previous falls, muscle weakness, gait and balance problems, vision deficit, postural hypotension, chronic conditions including arthritis, diabetes, stroke, movement disorders such as Parkinson's disorder, incontinence, medications, impaired cognition such as dementia, and psychological status such as fear of falling and anxiety. Extrinsic factors refer to hazards found around the environment of the house such as slippery floors, poor lighting or improper use of walking aids and assistive devices [3]. It should be noted that falls often result from dynamic interactions of the factors in all of the categories.

Fall risk of the elderly can be identified by healthcare professionals using fall risk assessment tools. There are a number of fall risk assessment tools in the literature [10–12]. These tools can be divided into two categories, namely clinical (nursing) assessment tools and functional assessment tools. The clinical fall risk assessment of the patients are performed in hospital and nursing home settings, and typically specific screening instruments or forms are employed for the assessment. These instruments consist of questions about intrinsic characteristics of the patient including psychological status, cognitive impairment, mobility dysfunction, fall history, disorders and sensory deficits. The scoring of these questions determines the fall risk level of the patients. Healthcare professionals use these instruments on a regular basis depending on the severity of the patients. Because of the regular use of the tools, they tend to be short and do not require detailed assessment of the patient. However, these assessments depend on the observation and judgement of the healthcare professionals or the self-report from the patient. Therefore, the objectivity and reliability of these tools are questionable [10].

Another assessment type performed to evaluate fall risk is functional mobility assessment. The functional fall risk assessment is mostly performed for community-dwelling elderly or outpatients. The functional assessment tools consist of tests that measure the functional capabilities and limitations in gait, balance and strength such as Tinetti Performance Oriented Mobility Assessment (POMA) [13], Berg Balance Scale (BBS) [14], Functional Reach (FR) [15], Dynamic Gait Index (DGI) [16], Timed Up and Go (TUG) [17], One Leg Balance (OLB) [18]. They provide standardized measures of disability and functional limitations and assess various motion characteristics in a more objective way than the clinical fall risk assessment tools. The scoring of these tests determines the fall risk level of the patients. However, they do not provide a detailed understanding of the theoretical relationships between the test result and the fall risk; therefore, the functional assessment tools are not totally reliable [12].

Among methods for identifying fall risk, fall risk assessment tools are the most commonly used technique; however, these tools require healthcare professionals to company the elderly during the test and this repeats on a regular basis, which is costly, overwhelming and time consuming. Also, no single fall risk assessment tool can be used for the elderly living in different settings because of lack of validation of the tools with a large group of people living in different settings and with different conditions [19]. Thus, these instruments lack standardization which causes lack of reliability and validity. Moreover, they lack sensitivity and objectivity as they require human observation and judgement, which often causes inaccurate results. While modern motion capture laboratories provide precise quantitative data, this method is obtrusive and cumbersome [20]. Therefore, a new, unobtrusive, objective, inexpensive, portable and user friendly fall risk assessment method is needed. Recently, researchers and physicians rely on assessments of a patient's gait, balance and the daily activities of the people in order to assess fall risk. Systems with wearable sensors have been developed to assess functional ability and mobility of people during their daily lives [21]. Different from clinical or functional assessment tools performed by healthcare professionals, these systems provide assessment of the fall risk in an objective way without requiring any expert knowledge. Also, most of these systems can be used outside the clinical environment, which gives an early alert of high fall risk and, therefore helps caregivers

implement preventive strategies at an early stage. However, current body-worn inertial sensor based systems are obtrusive, imprecise, and even inaccurate.

In this thesis, we study an unobtrusive foot-mounted inertial sensor-based fall risk assessment system. We assess fall risk based on gait characteristics that are significant fall risk factors. First of all, gait analysis techniques are studied. A Kinect-based gait analysis system is developed for unobtrusive and continuous monitoring of the gait. Later, an inertial sensor-based gait analysis system developed by our research group is utilized. Using this system, experiments are conducted to extract gait parameters of the subjects. Gait features are extracted out of these gait parameters to utilize for the assessment of the fall. Feature selection techniques are used to determine the most significant gait features having impact on fall risk. Different machine learning algorithms are employed to identify people at high and low fall risk based on selected gait features. In this way, people who are most likely to fall are detected. Thus, this fall risk assessment system enables us to identify people at high fall risk and inform their caregivers to take preventive interventions.

We tested our system with 37 subjects. Among 37 subjects, 11 of them are labeled as high fall risk, and 26 of them are labeled as low fall risk based on their history of fall and use of walking aid. Twenty one of the subjects are identified with different neurological conditions with gait implications. A questionnaire including factors that affect fall risk is prepared in collaboration with the neurologist to evaluate the characteristics of the subjects. The questionnaire is filled for every subject and the results of fall risk level identification methods are discussed based on these factors.

When compared to related studies in the literature, the proposed system has contributions to fall risk assessment as follows.

- Objective assessment: Our proposed system yields an objective assessment of the fall risk when compared with the functional fall risk assessment tools. These tools evaluate the mobility and gait of the subjects depending on the observation and the judgement of the healthcare professionals, so it does not provide an objective

assessment of the fall. Also, they do not provide the detailed explanation of the relationship between the test score and the fall risk, which does not help to take effective preventive strategies. Moreover, these tools require a great amount of time and effort for the healthcare professionals to apply on a regular basis. However, our system overcomes these difficulties by providing an objective and automated solution to the fall risk assessment.

- Unobtrusive design: Our proposed system provides an unobtrusive solution to fall risk assessment, which does not intervene the daily lives of the users. Any change in the elderly's physical, psychological or health condition might result in a change in their fall risk. Also, some medications or combination of medications might put the elderly at high fall risk. Considering that these are the common scenarios for the elderly, daily fall risk assessment systems are of great significance. Many studies assessing fall risk based on gait parameters utilize obtrusive sensor networks or motion capture systems and ground reaction platforms, which are expensive and require dedicated space to be set up, therefore, they can not be used for daily fall risk assessment purposes. However, our proposed system overcomes this limitation and provides an unobtrusive design.
- High accuracy: Our proposed system outperforms the fall risk assessment systems identifying people at high fall risk using gait parameters. We achieve a high accuracy of 94.59% for the identification of fall risk levels. This is a very promising result when the similar studies [22, 23] detailed in Chapter 3 are considered.

The rest of the thesis is organized as follows. In Chapter 2, we provide background information about inertial sensors, Microsoft Kinect and gait analysis. In Chapter 3, we present related works in the context of fall risk assessment and provide a comparative list of the related studies in the literature. In Chapter 4, gait analysis technique developed using Kinect is detailed. Inertial sensor based gait analysis system developed by our research group is explained in Chapter 5. In Chapter 6, we present the methodology for fall risk assessment using gait features. We provide our conclusions and directions for the future research in Chapter 7.

2. BACKGROUND

In this chapter, we present introductory information about the background concepts and tools which are used throughout the thesis. Particularly, we provide the overview of the tools used in our study, namely IMUs and Microsoft Kinect, and also we provide the overview of gait analysis, human gait characteristics and gait parameters that are key concepts in our study.

2.1. Inertial Measurement Units

Inertial measurement unit sensors (IMUs) are electronic devices that typically include an accelerometer and a gyroscope. The accelerometer measures changes in velocity, namely acceleration; while the gyroscope captures angular velocity changes. In theory, accelerometers can be used to estimate position changes by double integration of the acceleration, and changes in orientation can be estimated by integration of angular velocity. However, practically this is not the case as measurement errors cause drift in the signal. Therefore, it is worth noting that this must be taken into consideration while developing systems with IMUs.

Recently, due to reduced cost, miniaturization and improvement of precision of measurement, IMUs have been widely used for quantitative gait analyses as they enable gait analysis outside the laboratory. Therefore, IMU based gait analysis systems enable unobtrusive and continuous assessment of the gait.

In this study, we employ EXEL ExLs3 IMUs [24] as shown in Figure 2.1. The raw data is collected at a sampling rate of 100 Hz. ExLs3 IMUs can send the sensor readings via Bluetooth or can store the data in a local storage. Therefore, the data can be either used in real-time or can be stored in the device for a later use. The IMUs are strapped on both feet via adjustable tailor-made velcro straps as depicted in Figure 2.2.

Figure 2.1. ExL IMUs utilized in this study.

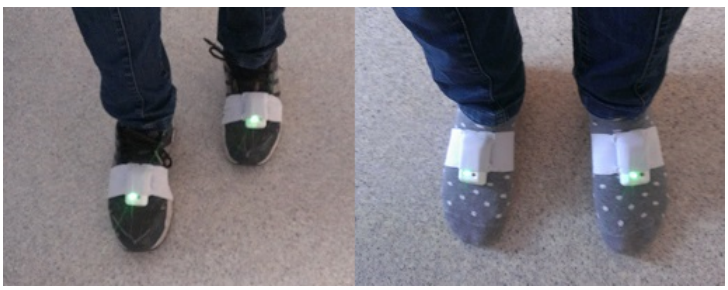

Figure 2.2. IMUs strapped on the feet.

2.2. Microsoft Kinect v2 Sensor

Kinect [25] is a low-cost, portable device that includes a camera and a depth sensor. In addition to the raw depth image, Kinect extracts a 3D virtual skeleton of the body with a rate of 30 frames per second. Each acquired frame data includes information of the skeleton that consists of twenty five joints as shown in Figure 2.3. Each joint information contains the 3D position of that joint in the Kinect coordinate system. The best operational range of Kinect to be able to obtain skeleton data is between 1.5 m and 4 m.

In this study, we use the skeleton data which enables us to track human walking using 3D positions of the joints. We have developed a data acquisition software for the Microsoft Kinect v2 sensor. The software was developed in the Visual Studio 2015 development environment using the C# programming language.

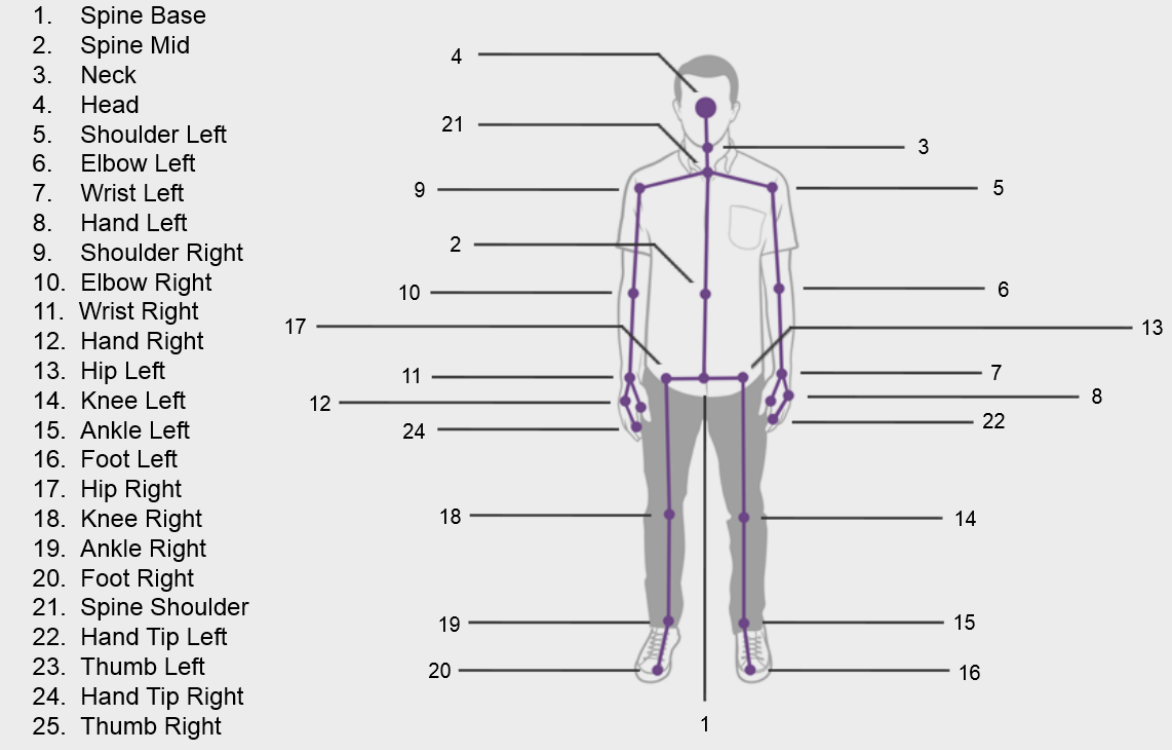

Figure 2.3. 25 Joints provided by the Microsoft Kinect v2 SDK.

2.3. Background to Gait

Human gait is a periodic movement of the body segments and includes repetitive motions. Gait analysis is the systematic study of the periodic movements and involves the measurement, description, and assessment of quantities that characterize human movement [26].

Gait analysis plays an essential role in identifying abnormalities or changes in human walking which are indicators of potential fall risk. Quantification of gait parameters helps people understand if they have any abnormalities in walking and helps assess the fall risk of people [27]. Especially, as people get older, their gait patterns change, and these changes might be indicators of different health problems and fall risk. Understanding these changes help them to maintain a stable walking pattern which reduces the risk of falling. Since the risk of falling in the elderly is among major problems that results in serious consequences and affects the quality of life negatively, monitor-

ing of the gait is of great importance for timely interventions aiming for preventing falls [28].

Monitoring of gait analysis is significant not only for fall risk assessment but also for medical inspection, follow-up and patient recovery. Neuro-degenerative diseases especially Parkinson's disease influence the gait and cause a decrease in mobility during the progression of the disease. Monitoring of the gait parameters is needed in order to obtain insight into progression of the disease for better treatment [29]. Furthermore, monitoring of the gait parameters is also important for assessment of the patient recovery. Clinicians are interested in measuring gait parameters for better assessment of the patient recovery [30].

2.3.1. Gait Cycle

A single gait cycle is defined as the period between two successive occurrences of one of the repetitive events of walking. Gait cycle starts with the initial contact of the reference foot with the ground and the cycle continues until this foot contacts the ground again. The main events and phases in the gait cycle are depicted in Figure 2.4.

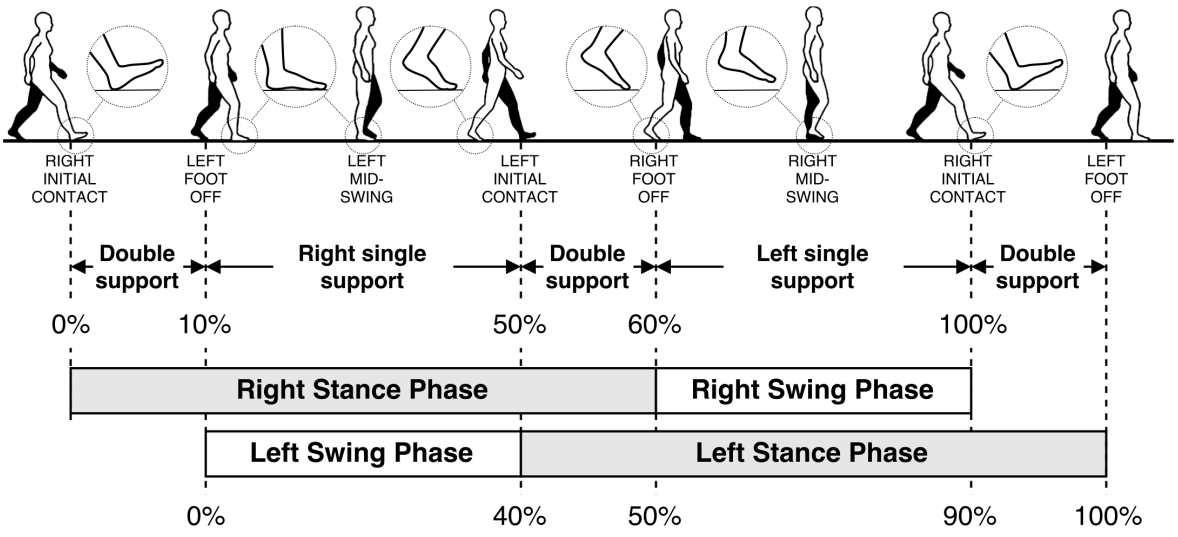

Figure 2.4. Human gait cycle.

The gait cycle consists of two main phases, namely stance phase and swing phase. Normally, the stance phase constitutes approximately 60% of the whole gait cycle, while the swing phase constitutes 40% [31]. The stance phase is the part of the gait cycle during which the foot remains in contact with the ground; hence the stance phase begins with the initial contact (IC) of the reference foot and ends with the foot-off (FO), which indicates the final contact of the reference foot with the ground. The swing phase is the part of the gait cycle that starts from the moment that the reference foot lifts off the ground until the reference foot contacts the ground again [31]. Gait phases and sub-phases are listed in Table 2.1 along with their definitions.

Table 2.1. Gait Phases.

Phase	**Sub-phase**	**Description**
Stance Phase	Loading Response	phase starting with initial contact and ending with opposite toe-off
	Mid-stance	phase starting with opposite toe-off and ending with heel strike
	Terminal stance	phase starting with heel strike ending with opposite initial contact
	Pre-swing	phase starting with opposite initial contact ending with toe-off
Swing Phase	Initial swing	phase starting with toe-off ending with feet adjacent
	Mid-swing	phase starting with feet adjacent ending with tibia vertical
	Terminal swing	phase starting with tibia ending with initial contact

2.3.2. Gait Parameters

Gait parameters are measurable characteristics of the gait. Gait parameters are divided into three categories, namely kinetic, kinematic and spatio-temporal gait parameters. Kinetic parameters are measurement of the force made during walking without any detailed knowledge of the position and orientation of the foot. Force platforms are used for extraction of kinetic parameters. Kinematic parameters describe the movement of the body through the space, but without reference to the forces involved. Kinematic parameters include the position and orientation of the body segments, such as the angles of the joints. Kinematic parameters can be extracted using camera systems or

inertial sensors [32]. Spatio-temporal parameters are based on temporal variables and provide fundamental timing and position information about the gait. Walking speed, stride length and cadence are examples of spatio-temporal gait parameters. In this study, we extract and use spatio-temporal gait parameters. Important spatio-temporal gait parameters are listed in Table 2.2 along with their definitions.

Table 2.2. Spatio-temporal Gait Parameters.

Gait Parameter	**Definition**
Stride Length (m)	Distance between two successive placement of the foot.
Step Length (m)	Distance one foot moves ahead of the other foot during the gait cycle.
Walking base (cm)	Side to side distance between the two feet.
Cadence (steps/mins)	Number of steps taken per minute.
Cycle Time (s)	Duration between two successive occurrences of the same gait event of the same foot.
Stance Time (s)	Duration of the stance phase, starting with IC and ending with FO of the same foot.
Swing Time (s)	Duration of the swing phase, starting with FO and ending with IC of the same foot.
Speed (m/s)	Stride Length / Cycle Time
Clearance (cm)	Elevation of the foot during swing phase.

3. RELATED WORK

Fall risk assessment has been a popular field of research in the last decades. Especially devastating psychological and physical consequences of the falls make this research field very significant to understand the underlying reasons of the fall and eventually to identify people at high fall risk so that preventive strategies for falls can be taken at an early stage. In this chapter, we present the overview of the related works in the field of fall risk assessment. In Table 3.1, the details of the related studies are summarized. These studies are presented in terms of features used to assess fall risk, the system setup such as the number of sensors and the sensor placement; design considerations such as daily life usability and unobtrusiveness; experiment design such as the number of subjects used in the experiments; and the methods applied for fall risk assessment.

Fall risk assessment studies in the literature focus on mainly three abilities of the people that are related to the fall risk, namely physical, social and cognitive abilities. In this chapter, related works are presented based on the studies that assess fall risk in terms of one of these three categories.

Walking patterns and physical functioning of people are among potential fall risk factors that need to be investigated. The gait and physical daily activities of people have been investigated in the literature to find the risk factors of the falls so that efficient fall risk assessment can be implemented based on the most important fall risk factors. Gait analysis, mobility and daily activities are the most studied physical aspects to analyse potential risk factors of the falls.

Gait analysis plays a crucial role in fall risk assessment. Gait parameters and gait variability features are extracted to investigate whether or not a specific gait feature is associated with falls. Walking speed, stride length, step width, cadence, stride time, swing time, stance time and variability of these parameters are the most commonly studied gait features [22, 23, 33–36]. Caby *et al.* [23] studied feature selection for fall

risk assessment. They use a rich set of gait characteristics extracted from 10 inertial sensors worn on different parts of the body. They validate their system with 20 hospitalized elderly. They find stride and step frequencies as the most important features distinguishing fallers from non-fallers in a cohort of hospitalized elderly. However, their setup is so obtrusive that it is not suitable for daily assessment of fall risk. Zhang *et al.* [22] studied identification of people at high and low fall risk based on gait features extracted from a motion capture system and a ground reaction force platform. Machine learning algorithms are employed to identify people at high and low risk of fall. Their system requires a laboratory setup and; therefore, it is not suitable for daily fall risk assessment. Other studies [33–36] focus on distinctive gait features between fallers and non-fallers. Statistical analysis is carried out to investigate the relationship between two groups of subjects and their extracted gait features. Therefore, distinguishing gait features between fallers and non-fallers are found.

Other gait characteristics that have association with falls are gait stability and symmetry. The gait is unsteady in the elderly who are prone to fall and; therefore, Paterson *et al.* [37] study the effect of gait stability and symmetry on fall risk. The raw acceleration signal is utilized to extract gait stability and symmetry features. According to the threshold of absolute values of these features, people are classified into three level of fall risk namely normal, attentive and dangerous.

One of the most important factors that characterize the different studies using gait features for fall risk assessment is the selection of sensors to extract gait parameters. Gait parameters are mostly extracted using body worn inertial sensors or instrumented walkways such as GAITRite [38] and ground reaction force platforms. While modern gait analysis laboratories or instrumented walkways provide precise quantitative analysis of the gait, these methods are costly and impractical, especially for the elderly and patients. However, systems with body-worn inertial sensors do not require such a laboratory environment and, therefore they are more suitable for daily life use. However, it should be noted that some of the gait analysis systems with inertial sensors are so obtrusive that they are not appropriate for daily use.

Besides gait characteristics, physical daily activities of people are related to fall risk. There are studies in the literature that investigate the risk factors of daily activities and mobility which indicate important clues about people's wellbeing in terms of their psychology and physical conditions [39, 40]. Schooten *et al.* [39] conduct a retrospective and prospective study to investigate the effect of daily life activities on fall occurrence. Daily activities of sitting, lying, standing, walking and walking characteristics are extracted from a trunk-mounted accelerometer during 8 consecutive days. Correlations between extracted physical mobility and fall occurrence are analysed in both retrospective and prospective manner in order to find the most significant factors correlated with falls. Weiss *et al.* [40] conduct a 3-day assessment of continuous gait and mobility, and extract features from PD patients. Not only mobility features but also disease related features are investigated to define the most important predictor for falls for PD patients. They conclude that body-worn sensors can continuously evaluate the fall risk in PD.

Raw data obtained from inertial sensors are analyzed in [41, 42] during functional fall risk assessment tests. Functional test scores such as TUG and BBS are not informative enough themselves in identifying people at high risk and low risk; therefore, these tests are instrumented with inertial sensors. In this way, these tests capture more information about the fall risk of people. Frequency and time domain features are extracted from the raw acceleration and gyroscope data collected during functional tests, and not only test scores but also extracted features are used in identifying distinctive fall risk factors between fallers and non-fallers; and eventually people at high risk and low risk of fall are identified. However, this study is not suitable for daily assessment of fall risk since it still requires physicians to company the elderly.

Postural stability and balance of people are among potential fall risk factors that need to be addressed [43–45]. Tests to measure postural stability and balance are performed in different conditions such as with open and closed eyes, on firm and compliant surfaces using force platforms. Muscular, strength and balance features are extracted. The relationship between these features and the history of fall is analysed to understand the effect of the features on fall occurrences. However, it should be noted

that these systems are costly and require a clinical environment because of the sensor setup and; therefore, they are not suitable for daily assessment of fall risk.

Additionally, there are camera-based systems that aim to assess the fall risk of the elderly. Rantz *et al.* [46] studied a continuous and unobtrusive in-home monitoring system that provides an accurate automated assessment of the fall risk. They track gait parameters of the people using Kinect. They investigate the correlation between gait parameters and fall risk assessment tool scores for fall risk assessment. In a further study [47], they improve their system and deploy the Kinect-based system in the private apartments of 19 older adults, and they track gait parameters of the people over two years. The fall risk is assessed and automated based on changes in the gait over a continuum of time.

Sociodemographic characteristics, cognitive and emotional conditions of people play a crucial role in fall risk assessment. These metrics have been studied in the literature to analyze risk factors of the falls. In [48, 49], risk factors of chronic diseases, medication use, life style factors, socio demographic characteristics and physical functioning are studied as potential predictors of the fall risk. In [50], the effectiveness of cognitive and motor tests are studied for prediction of falls with subjects who have different neurological diseases. Machine learning algorithms are employed based on these tests to identify patients who are likely to fall. They find that a single cognitive test is the most successful test in the prediction of fall for patients with neurological diseases.

Among the studies mentioned in this chapter, there are two studies that they have similarities to our study in terms of feature selection and applied fall risk identification techniques. Zhang *et al.* [22] aim to identify the fall risk level classifying people into two groups, high and low fall risk groups. They extract 32 gait features using a motion capture system and a ground reaction force platform. They collect data from 35 elderly people, in which 19 of them are with a fall history. They use dimensionality reduction techniques to prevent overfitting of the data and use KNN, NB, LR, NN, SVM classifiers for identification of the individuals with high fall risk.

The highest accuracy that they achieve with these techniques is 85.8%, whereas, in our study, we achieve a higher accuracy of 94.59%. Another point that our study differs from this study is the sensor setup. This study requires a laboratory environment for the extraction of gait features and does not enable unobtrusive daily assessment of fall risk. Moreover, our study differs from this study in terms of extracted gait features and techniques used in selecting features for classification algorithms. Caby *et al.* [23] studied feature selection for fall risk assessment using accelerometers. They use forward wrapper selection algorithm and apply four classification algorithms to select the features according to classification performances. Sixty seven features are extracted from the raw data obtained from 10 accelerometers mounted on different parts of the body. They collect data from 20 hospitalized elderly. Among the algorithms they applied, they report the highest accuracy as 95% with 80% specificity and the lowest accuracy as 80%. When our best performance metrics are compared with their best performing algorithm, we have almost the same accuracy (94.59%) with a higher specificity metric (96.15%). When the accuracies of all algorithms used for identification of the fall risk level are compared, they have average accuracy of 85% whereas we have higher average accuracy of 91.44%. Moreover, they extract a different set of gait-related features, and their proposed system is obtrusive and designed for hospitalized patients. Thus, our study differs from this study in terms of extracted features, and also we propose an unobtrusive inertial-sensor based fall risk assessment system that can be used in non-hospital settings.

Table 3.1. Related Work.

Ref.	Aim	Ground Truth / Methodology	Sensors	Features	Subjects	Strength / Weakness
[23]	Feature selection for fall risk assessment	Binary as high risk and low risk groups by self report and functional assessment tools; forward wrapper selection algorithm	Accelerometers (10 sensors on different parts of the limbs)	67 gait-related features (e.g. time and frequency features, the number of steps, variance of step period, arm and leg correlation, momentum and entropy analysis of the data, arm dynamics)	20 hospitalized elderly (F:5, NF:15)	Variety of gait features; Obtrusive, the number of sensors used, validated for hospital inpatients only
[22]	To classify people into two groups: fallers (high fall risk) and non-fallers (low fall risk)	Binary as high risk and low risk groups by self report; KNN, NB, LR, NN, SVM classifiers	Vicon motion capture system, ground reaction force platform	32 gait features (e.g. walking speed, stride length, step width, cadence, minimal toe clearance, maximal toe inclination, maximal dynamics trunk sway, maximal push-up force, maximal heel-contact force)	35 elderly (F:19, NF:16)	Variety of gait features; Obtrusive, requires laboratory setting
[33]	To improve TUG using body worn sensors for better fall risk assessment	Binary as high risk and low risk groups by self report; Logistic regression	Accelerometer and gyroscope (on the shank)	44 gait and balance parameters (e.g. stride, swing, stance, step times, cadence, turn time, return time, walk time ratio, TUG time, steps taken, turn angular velocity)	349 community-dwelling elderly (F:207, NF:142)	Variety of gait features Validated for community dwelling elderly only
[34]	To compare spatial temporal gait parameters and their variability between elderly fallers and non-fallers	Binary as high risk and low risk groups by self report; Kolmogorov-Smirnov test	Ground force plates (intsrumented walkway)	gait parameters (e.g. step time, walking speed, step width, distance between COP position)	125 healthy elderly (F:31, NF:94)	The number of subjects; Need for instrumented walkway, validated for healthy elderly only
[35]	To identify which gait parameters best predict the fall	Binary as high risk and low risk groups by self report; quasi-likelihood analysis	GAITRite	gait parameters (e.g. cadence, stride length, swing, double support, stride length variability and swing time variability)	597 adults aged 70 and over (F:226, NF:371)	The number of subjects; Need for instrumented walkway

Table 3.1. Related Work (cont.).

Ref.	Aim	Ground Truth / Methodology	Sensors	Features	Subjects	Strength / Weakness
[36]	To investigate if walking stability can predict future falls	Binary as high risk and low risk groups by self report; logistic regression, anova	GAITRite and accelerometers (on foot)	stride dynamics and gait variability	97 healthy women (F:54, NF:43)	Stability features; Need for instrumented walkway, validated for women only
[39]	To investigate potential risk factors among gait and daily activities	Binary as high risk and low risk groups by self report; logistic regression	Accelerometer (on the trunk)	gait and daily activity features (e.g. sitting, lying, standing, walking bouts, walking speed, stride time, stride length, gait intensity, gait variability, freq power, gait symmetry, harmonic ratio)	169 elderly (F:60, NF:109)	Retrospective and prospective study, 8-consecutive day of data collection; -
[40]	To find important gait features in fall risk assessment of PD patients	Binary as high risk and low risk groups by self report; regression analysis	Accelerometer and gyroscope (lower back)	3-day assessment of gait and mobility features (e.g. walking bouts, time spent walking, non-walking activity, the number of steps, cadence, frequency derived measures, step-to-step variabilities) + disease related features such as dosage of levodopa + fall risk measures such as BBS, DGI, TUG	107 PD Patients (F:40, NF:67)	not only variety of gait features but also daily activity and disease related features are considered. 3-day of continuous data collection; -

Table 3.1. Related Work (cont.).

Ref.	Aim	Ground Truth / Methodology	Sensors	Features	Subjects	Strength / Weakness
[41]	To compare the performance of functional tests scores and features obtained from inertial sensors and pressure platforms to discriminate between low and high risk of fall	Binary as high fall risk and low fall risk group based on the history of falls questionnaire and usage of walking; Feature selection and then classification using DT, NB.	Accelerometer, Gyroscope (lower back) and pressure platform	frequency and time domain features from sensors during TUG, STS and 4-stage tests	296 community-dwelling older adults (F:51, NF:245)	- Obtrusive system, need for laboratory environment, tested only during functional tests
[42]	To instrument Timed Up and GO Test with inertial sensors for better assessment	Binary as fallers (high risk) and non-fallers (low risk) by POMA score; t-test	Smartphone, accelerometer and gyroscope (in the pocket or fixed in the tight)	Statistical measures and frequency based metrics obtained using th FFT of the acc.	18 community-dwelling older adults (F:5, NF:13)	Unobtrusive Tested only during functional tests
[43]	To identify functional parameters for the classification of older female fallers and prediction of 'first-time' fallers	Binary as high risk and low risk groups by self report; Logistic regression	dynamo-meter (knee and ankle); force platform; motion capture system	92 features of muscular, strength and gait analysis (e.g. torque measurements; CoP time series; cadence, double support time, foot clearence, step width)	84 elderly women (F:42, NF:42)	Prospective study; Need for laboratory environment
[44]	To determine the ability of postural stability to identify fallers among the elderly	Binary as high risk and low risk groups by self report; Anova	Forceplate	COP (Center of pressure) data and Dorsiflexion strength, plantar flexion, knee extensors, knee flexors strength. Six stability tests when (wide stance, eyes open/closed, eyes open standing on foam, and same as performed in narrow stance.)	143 elderly (F:19, NF:124)	- Obtrusive

Table 3.1. Related Work (cont.).

Ref.	Aim	Ground Truth / Methodology	Sensors	Features	Subjects	Strength / Weakness
[45]	To identify postural stability features that best distinguish between fallers and non-fallers	3 classess (fallers, non fallers, fallers) Between-group comparison tests	piezoelectric force plate	Static posturography tests under five conditions	130 cognitively-able elderly (F:45 , NF:67, RF:18)	- Need for laboratory environment
[46, 47]	To provide unobtrusive and continuous in-home fall risk assessment system	FRA Scores and GAITRite correlation analysis between FRA scores and gait parameters	Kinect	gait parameters	19 older adults Fallers	Unobtrusive and continuous; No mobility, specific to a certain location
[48]	To develop screening tool with five risk factors	3 classess (fallers, non-fallers, recurrent fallers) regression analysis	N/A	demographics, previous fall, medication and movement during One leg balance test (OLB)	1759 healthy subjects (F:563, NF:1196)	The number of subjects; -
[49]	To investigate potential predictors of fall-risk	2 classes (Falls ans reccurrent falls.) regression analysis	N/A	Physical, cognitive, emotional, and social functioning preceding the registration of falls.	1285 community dwelling	The number of subjects; -
[50]	To demonstrate that cognitive and motor tests can be used to create a robust predictive tool for falls	Binary as fallers (high risk) and non-fallers (low risk) by self report; Machine learning methods	N/A	Direct measurement of cognitive and physical function. Attention and executive function tests, a measure of physical function (Walk-12 test), a series of questions (concerning recent falls, surgery and physical function) and demographic information	323 patients with neurological condition (F:54, NF:269)	- Validated for hospital inpatients only.

4. EXTRACTION OF GAIT PARAMETERS USING KINECT

We propose a Kinect-based gait analysis system, which includes a single Kinect placed in a corridor in a home setting or clinic, so that as people walk through the corridor in their daily lives or during medical inspection, their gait parameters are estimated and reported. For this reason, we conducted a preliminary work in the laboratory and the clinical setting.

Kinect provides skeleton data for every frame. Skeleton data consists of 3D positions of 25 joints. We use 3D position of the left and right foot joints provided by Kinect in its own coordinate system. To be able to work with the data better, it is needed to convert the skeleton data in Kinect coordinate system to the real world coordinate system as shown in Figure 4.1, in which the origin corresponds to projection of the camera position on the floor: The X axis is to the sides of the camera, the Y axis is straight forward from the camera and the Z axis is straight up from the ground toward to the Kinect. For this purpose, the tilt angle of the Kinect is estimated using the floor clipping plane vector given together with every frame by the Kinect Software Development Kit (SDK). After the tilt angle is estimated, the Kinect coordinate system is rotated around its origin so that its z-axis is parallel to the floor. Then, the rotated data is translated into the real world coordinates. Described algorithm is given in Figure 4.2.

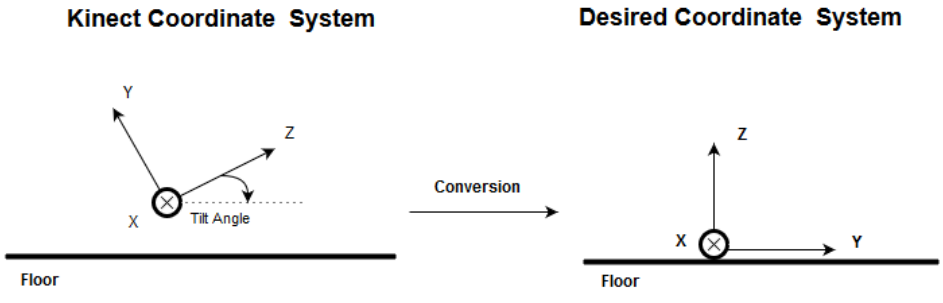

Figure 4.1. Kinect coordinate system to real world coordinate system.

Require: $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ floor-clipping-plane vector coefficients provided by Kinect SDK
where $a \times x + b \times x + c \times x + d = 0$ **and** $\sqrt{a^2 + b^2 + c^2} = 1$
Require: $m \times 3$ joint 3D position vector **jointPosition** provided by Kinect SDK
kinectFloorVector $\leftarrow \begin{bmatrix} a & b & c \end{bmatrix}$
worldFloorVector $\leftarrow \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$
$\theta \leftarrow \arccos(\mathbf{kinectFloorVector} \cdot \mathbf{worldFloorVector})$

$$\mathbf{Rotation} \leftarrow \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

NewJointPosition $\leftarrow$ **Rotation** $\times$ **jointPosition**$^{\mathbf{T}}$
NewJointPosition $\leftarrow$ **NewJointPosition**$^{\mathbf{T}}$
for $i = 1$ to m **do**
NewJointPosition$[\mathbf{i}, \mathbf{3}] \leftarrow$ **NewJointPosition**$[\mathbf{i}, \mathbf{3}] + d$
end for
return **NewJointPosition**

Figure 4.2. Algorithm to convert from Kinect coordinate system to the real world coordinate system.

The gait cycle begins when the foot initiates a contact with the ground and ends when the same foot initiates the subsequent contact with the ground. The stride length is computed measuring the distance between these two subsequent contacts. The trajectory of the feet while walking is shown in Figure 4.3.

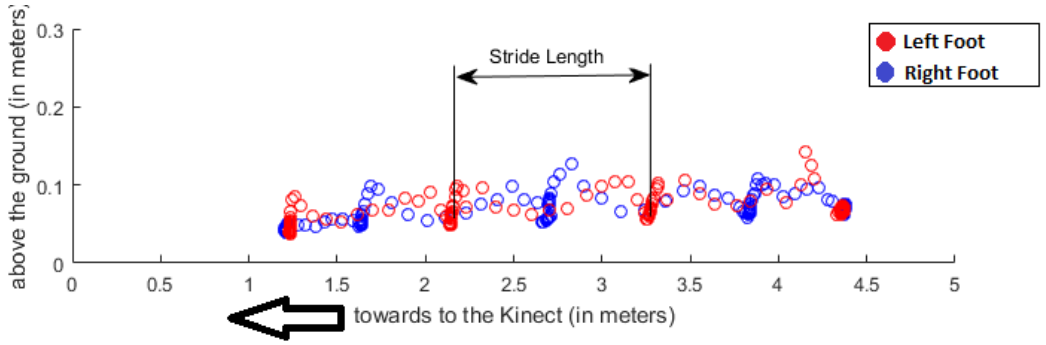

Figure 4.3. Trajectory of the right and left foot during walking through the corridor.

We employ a heuristic algorithm as explained in Figure 4.4 and 4.5 to estimate the gait parameters of stride length and cadence. This implemented algorithm to measure gait parameters makes use of the fact that Kinect measurement is more accurate when 3D position of the foot has the lowest elevation in z axis, where the foot is stationary. Therefore, the displacement between two consecutive lowest-z periods of the specific foot gives very accurate stride length estimations as shown in Figure 4.3. After extracting individual stride lengths, the average stride length and the standard deviation of the stride lengths are estimated. Moreover, the cadence is computed as the number of steps (i.e. 2 steps equal to 1 stride) per minute.

The accuracies of the Kinect based gait analysis systems are investigated in the literature [51, 52]. When compared with the gold standard instrumented walkways and camera motion capture systems, kinect-based gait analysis systems have reported highly accurate results. In [35] and [38], the mean absolute displacement errors of 2.5 cm and 2.9 cm per stride are reported, respectively. Also, the accuracy of Kinect based systems for analyzing pathological gait is studied in the literature [53, 54].

The validation of the system has been conducted with 5 healthy subjects on a designed walking path such that every 0.5 and 0.75 meter (m) of the path is marked. In this validation study, data collection is performed in two different sessions, with subjects being asked to walk on the designed path such that each step corresponds to every 0.5 m and 0.75 m marks respectively. Estimated stride lengths and predefined lengths are compared. The error rate is computed as 2.8% when walked with the step length of 0.5 m, and as 2.9% with the step length of 0.75 m.

```
Require: m × 3 position matrix M where m is number of data points;
  Define window size p;
  Define threshold z;
  Define StanceCount ← 0;
  Define Stance ← m × 1;
  {Find potential zero-velocity periods}
  for i = 1 to m do
    if M[i][2] < z then
      Stance[i] ← true;
    else
      Stance[i] ← false;
    end if
  end for
  for i = 1 to m do
    if Stance[i] == true then
      if StanceCount < p then
        Stance[i] ← true;
      end if
      StanceCount + +;
    else
      if StanceCount >= p then
        for j = 1 to p do
          Stance[i − j] ← false;
        end for
      end if
      StanceCount ← 0;
    end if
  end for
  return Stance
```

Figure 4.4. The algorithm for finding the stance periods (lowest-z periods).

Require: $m \times 1$ stance matrix **Stance** using Figure 4.4;
  StrideLengths $\leftarrow$ []
  for $i = 2$ to $m - 1$ **do**
    if Stance[i] == true and Stance[i − 1] == false then
      StanceStart $\leftarrow i$;
    else if Stance[i] == true and Stance[i + 1] == false then
      StanceEnd $\leftarrow i$;
      TimeDiff $\leftarrow$ *findTimeBetweenPoints*(**StanceStart**, **StanceEnd**);
      DistanceDiff $\leftarrow$ *findDistanceBetweenPoints*(**StanceStart**, **StanceEnd**);
      StrideLengths.*append*(**DistanceDiff**);
      TimePassed $\leftarrow$ **TimePassed** + **TimeDiff**;
    else
      Continue;
    end if
  end for
  AverageStrideLength $\leftarrow$ *getAverage*(**StrideLengths**);
  StdStrideLength $\leftarrow$ *getStdDeviation*(**StrideLengths**);
  Cadence $\leftarrow (2 *$ **TotalStrideCount**/**TimePassed**$) * 60$;
  return **AverageStrideLength**, **StdStrideLength**, **Cadence**

Figure 4.5. The Algorithm for Extracting Gait Parameters.

4.1. Experiments

In this section, experiments conducted to test the proposed Kinect-based gait analysis system is presented. Experimental setup, data collection procedure and experiment results are presented and discussed.

4.1.1. Experimental Setup and Data Collection

The best operational range of Kinect for skeleton data is between 1.5 m and 4 m. Subjects are asked to walk through a corridor in this range towards Kinect. The subjects are asked to repeat the walking through the corridor three times. Kinect is mounted on a tripod at a height of 0.7 m as shown in Figure 4.6.

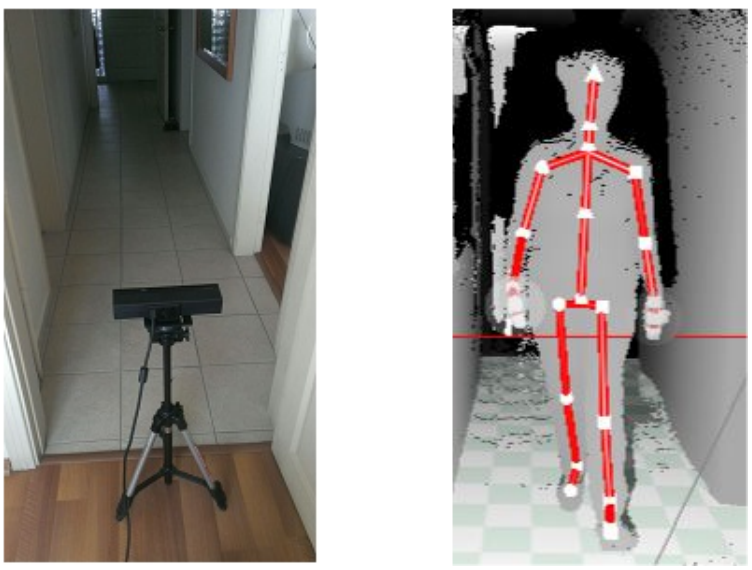

Figure 4.6. Experiment Setup and Data Collection.

21 subjects participated in the study. Fifteen of these subjects (8 males, 7 females) are healthy and six of these (3 males, 3 females) have different neurological conditions with gait implications; particularly, one with dementia, one with vascular dementia, one with progressive dysarthria, one with normal pressure hydrocephalus and one with Parkinson's disease and one with frontal gait apraxia. The healthy subjects' average age is 44.9±16.1 (min:25, max:76) and the average height is 1.68±0.08 m. The subjects

with neurological conditions have an average age of 68.5±12.2 (min:52, max:82) and an average height of 1.67±0.08 m. As Kinect operates in a limited range, every subject is asked to walk through the corridor three times at their normal pace.

4.1.2. Results

The average stride length, the standard deviation of stride length and the cadence of all subjects are extracted. The results are presented in Table 4.1. The gender, age, height and health status of the subjects are also presented. The subject ID is assigned to each subject for easy reference.

Considering the results, it is observed that there are differences between people having healthy gait and people having pathological gait. For example, one of the observations is that standard deviation of the stride length is higher in people with gait implications (avg: 0.122) than people with healthy gait (avg: 0.078). This difference is especially observed in Subjects 18 and 21. Another observation is that people with pathological gait have smaller stride lengths compared to stride lengths of people with healthy gait when also their heights are taken into consideration. Stride lengths and the standard deviation of stride lengths for both healthy and pathological gait are shown in Figure 4.7.

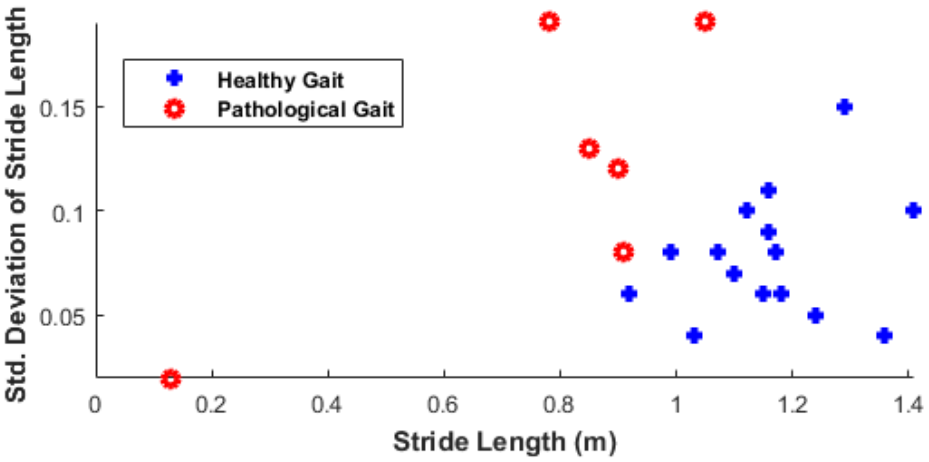

Figure 4.7. Standard deviation of stride length and stride length.

Table 4.1. Experiment Results.

Subject No	**Gender**	**Age**	**Height**	**Health Status**	**Avg. Stride Len(m)**	**Std. Stride Len**	**Cadence**
1	F	49	1.55	Healthy	1.07	0.08	94.59
2	M	26	1.72	Healthy	1.10	0.07	97.35
3	M	27	1.82	Healthy	1.36	0.04	93.18
4	F	25	1.68	Healthy	1.17	0.08	86.25
5	M	34	1.76	Healthy	1.18	0.06	90.04
6	M	29	1.73	Healthy	1.24	0.05	89.24
7	M	39	1.65	Healthy	1.16	0.09	96.68
8	M	36	1.8	Healthy	1.16	0.11	80.28
9	M	53	1.74	Healthy	1.41	0.10	89.88
10	F	76	1.55	Healthy	0.92	0.06	90.66
11	F	63	1.64	Healthy	0.99	0.08	82.29
12	F	67	1.59	Healthy	1.15	0.06	100.94
13	M	48	1.73	Healthy	1.29	0.15	90.98
14	F	42	1.63	Healthy	1.03	0.04	85.94
15	F	60	1.6	Healthy	1.12	0.10	109.83
16	M	82	1.74	Vascular dementia	0.91	0.08	86.06
17	F	56	1.58	Dementia	0.90	0.12	95.05
18	M	68	1.73	Progressive dysarthria	1.05	0.19	106.97
19	M	74	1.7	Frontal gait apraxia	0.85	0.13	98.47
20	F	79	1.57	Normal pressure hydrocephalia	0.13	0.02	67
21	F	52	1.56	Parkinson	0.78	0.19	102.61

The stride length and cadence comparison is depicted in Figure 4.8. The cadence is another important gait parameter that gives us information about gait abnormalities. Normally, when people intend to increase their pace, the average stride length and cadence increase together. For example, in Subject 21 it is observed that the average stride length of the subject is observed to be small, and the cadence is observed to be high compared to the rest of the subjects. Therefore, small stride length with high cadence makes staying in balance difficult, which might be an indicator of abnormality and an increase in the fall risk. Furthermore, when Figure 4.8 is examined, it can be observed that people with healthy gait and pathological gait form two different clusters, which can be used in further studies for identification of the pathological gait.

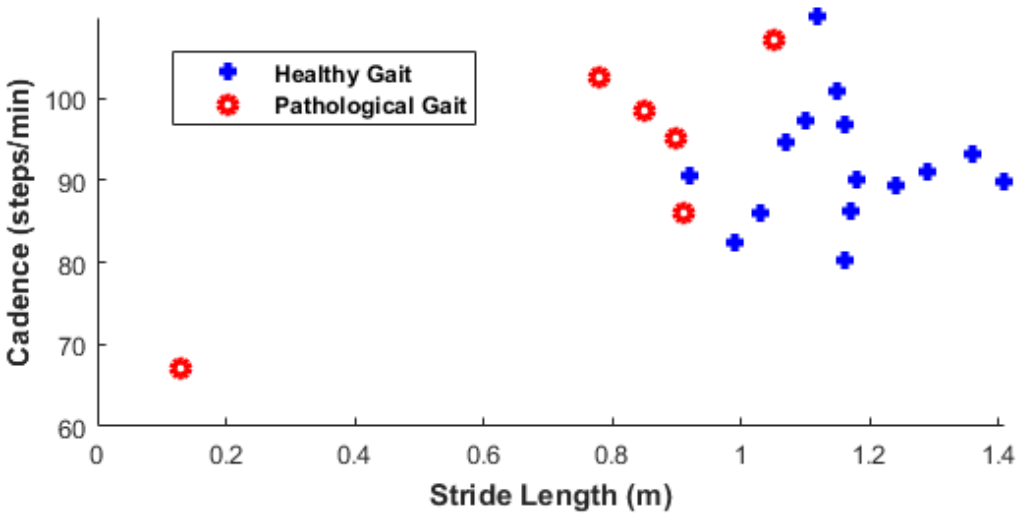

Figure 4.8. Stride length and cadence.

4.1.3. Discussion

This study is a preliminary work to investigate whether or not Kinect is useful in gait analysis for both healthy and patients. The purpose of this work is to automate gait analysis and eventually fall risk assessment using environmentally placed sensors in the community dwelling facilities and at the homes of the elderly people so that fall risk can be measured in everyday living activities. Hence, the elderly and their caregivers can be alerted when fall risk increases and informed to take interventions to avoid devastating falls.

We show that the gait parameters of stride length and cadence can be computed using a Kinect-based system. Our system consists of a single Kinect and a notebook, which enables to measure gait parameters without requiring people to visit gait analysis laboratories or clinics. However, it still requires a simple infrastructural setup (a tripod to mount the device, a power source and a wired connection to a computer for data collection).

The Kinect system has a potential to be a low-cost, in-home sensor for extracting gait parameters. However, it should be noted that even though Kinect enables a non-intrusive and cheap solution for gait analysis, it has drawbacks such as the operational range limitation for skeleton frames and some technical limitations, such as its accuracy with different type of cloths and environmental lighting.

5. EXTRACTION OF GAIT PARAMETERS USING IMUs

Recently, our research group has proposed a foot-mounted inertial sensor-based gait analysis system [55]. In this thesis, we utilize the data provided by this gait analysis system. In this chapter, we present the details of this study and explain how gait parameters are extracted using this system.

Gait parameters are extracted using our foot-mounted inertial sensor-based gait analysis system that employs the zero-velocity update and Kalman filtering methodology. IMUs make it possible to analyse data outside the laboratory and capture information about the human gait during the person's everyday life activities, hence provide portable and unobtrusive solution to gait analysis. However, it is still challenging to work with them due to measurement errors. Theoretically, it is possible to obtain the change in the orientation of the IMU by integrating the angular velocity measured by the gyroscope, and the change in the position of the IMU can be computed by double integration of the acceleration signal of the accelerometer. As a result, when IMU mounted on the foot, gait parameters can be derived using the change in orientation and position. However, accelerometers and gyroscopes suffer from measurement errors, which causes the error to be accumulated when integrated. Therefore, estimation of the position and orientation is not accurate and reliable.

Implemented methodology solves the challenges of the IMUs using the zero velocity update method. We exploit the time period that feet remain stationary within the stance phase of the gait cycle, in order to correct errors in IMU measurements. When IMUs are placed on the feet, it is the situation that feet have near-zero velocity within the stance period of the gait cycle, thus correction can be applied on a step by step basis using the zero velocity update method and a Kalman filter to track and correct the correlations of the state errors [56, 57].

Corrected 3D position of the foot is obtained after applying zero velocity update with Kalman filter. Corrected 3D position signal is smoothed using Rauch-Tung-

Striebel smoother [58]. Corrected and smoothed 3D position is used for the detection of IC and FO events. Once IC and FO events are detected, gait parameters are estimated based on these events.

The estimation of gait spatio-temporal parameters requires the detection of the FO and IC events. Swing and stance phases are initiated and ended with gait events. Therefore, after identifying IC and FO events of the gait cycle, temporal gait parameters are derived. Gait parameters of the swing time is calculated as the time period between subsequent FO and IC events of the same foot and the stance time is calculated as the time period between subsequent IC and FO events of the same foot. Detection of IC events can be further used to segment strides, which enables the extraction of stride length, clearance, cycle time, cadence and speed. The stride length is calculated as the spatial distance between the positions of two consecutive IC events of the same foot. Cycle time is computed as the time difference between two consecutive IC events of the same foot. Clearance is computed as the maximum elevation between two consecutive FO and IC events. Cadence is computed as steps taken per minute. Speed is computed as the distance taken in a gait cycle time, which is the time taken between two consecutive IC or FO events. The parameters of stance and swing times, stride length and clearance are computed per stride and per foot, and hence their standard deviations are also calculated.

Using the system described above, we extract a variety of spatio-temporal gait parameters and use them for fall risk assessment. We are able to extract all spatio-temporal gait parameters listed in Table 2.2. However, gait parameters of step length and walking base can not be computed. By the definition, these two parameters require spatial synchronization between two feet. However, our proposed system tracks the spatial change of every single foot separately. Although there are methods in the literature that solves this issue by the introduction of an additional equipment, we prefer to keep our proposed system unobtrusive at the cost of inability to extract step length and walking base.

5.1. Experiments

In this section, experiments conducted to test the foot-mounted inertial sensor-based analysis system are presented. Experimental setup, data collection procedure and experiment results are presented and discussed.

5.1.1. Experimental Setup and Data Collection

Thirty seven subjects participated in the study. Sixteen of these subjects (8 males, 8 females) are healthy and their average age is 43.87±16.20 (min:25, max:76) and the average height is 1.67±0.08 m. 21 of the subjects (11 males, 10 females) have different neurological conditions with gait implications. Among these subjects, seven of them suffer from Parkinson's disease (PD), two have dementia with lewy bodies (DLB), one has mild cognitive impairment (MCI), one has vascular mild cognitive impairment (VaMCI), one has vascular dementia (VaD), two with normal pressure hydrocephalus (NPH), one has vascular parkinsonism (VaP), one has dementia, one has frontal gait apraxia, one has progressive dysarthria and the remaining three suffer from multiple conditions. They have an average age of 77.80±9.55 (min:52, max:90) and an average height of 1.65±0.08 m.

The IMUs are strapped on both feet of the subjects via adjustable tailor-made velcro straps. Subjects are asked to walk through a corridor. The subjects are asked to repeat the walking through the corridor three times.

5.1.2. Results

Average stride length, cadence, average cycle time, average L/R stance time, average L/R swing time, L/R stance ratio, speed and average L/R max clearance are extracted. Standard deviations of the parameters are also computed. The results of the experiments are presented in Table 5.1. The gender, age, height and health status of the subjects are also presented. The subject ID is assigned to each subject for easy reference.

Table 5.1. Experiment Results.

Subject ID	Gender	Age	Height (m)	Health Status	stride length (m)	cadence (steps/ min)	cycle time (s)	L stance time (s)	R stance time (s)	L swing time (s)	R swing time (s)	L stance ratio	R stance ratio	speed (m/s)	L max clearance (cm)	R max clearance (cm)
1	M	72	1.74	Mixed	1.2	88.9	1.35	0.84	0.84	0.5	0.52	0.62	0.62	0.88	8.6	9.1
2	M	85	1.65	DLB	0.99	109.3	1.1	0.66	0.71	0.44	0.39	0.6	0.65	0.9	7.3	6.9
3	F	73	1.62	PD	1.11	99.9	1.2	0.76	0.7	0.45	0.49	0.63	0.58	0.92	3.9	5.3
4	F	82	1.58	PD	0.5	72.4	1.66	1.18	1.18	0.47	0.49	0.72	0.71	0.3	4.5	4.4
5	M	80	1.76	Mixed	1.1	75	1.6	0.87	1.02	0.71	0.6	0.55	0.63	0.69	7.7	8.4
6	M	88	1.8	PD	0.36	87.1	1.38	0.99	1	0.37	0.4	0.73	0.72	0.26	2.1	3.7
7	F	85	1.65	NPH	0.81	88.3	1.36	0.86	0.85	0.49	0.52	0.64	0.62	0.6	5.3	6
8	F	81	1.6	PD	0.72	70.6	1.7	1.15	1.12	0.55	0.58	0.68	0.66	0.42	5.2	5.9
9	F	80	1.6	Mixed	0.73	82.6	1.45	0.94	0.96	0.53	0.48	0.64	0.67	0.5	4.5	5.2
10	M	82	1.65	VaD	1.16	88.4	1.36	0.78	0.82	0.57	0.56	0.58	0.59	0.86	9.4	10.7
11	F	90	1.6	PD	0.53	119.3	1.01	0.67	0.61	0.35	0.38	0.66	0.62	0.52	3.7	4
12	M	83	1.59	PD	0.93	95.6	1.26	0.77	0.76	0.48	0.5	0.62	0.6	0.74	6.9	6.1
13	M	76	1.82	MCI	1.31	92.2	1.3	0.76	0.76	0.56	0.53	0.58	0.59	1.01	9	9.3
14	M	84	1.69	VaP	0.45	90.9	1.32	0.89	0.87	0.43	0.45	0.67	0.66	0.34	3.7	3
15	F	82	1.5	DLB	0.83	83.8	1.43	0.96	0.84	0.47	0.59	0.67	0.59	0.58	5.2	4.8
16	M	82	1.74	VaMCI	0.87	84.3	1.42	0.91	0.91	0.51	0.51	0.64	0.64	0.61	6.6	6.6
17	F	56	1.58	Dementia	0.89	88.9	1.35	0.89	0.8	0.49	0.52	0.65	0.61	0.66	8.6	7.9
18	F	79	1.57	NPH	0.15	74.5	1.61	1.39	1.37	0.21	0.26	0.87	0.84	0.09	2.1	2.4
19	M	74	1.7	Frontal gait apraxia	0.93	107.5	1.12	0.69	0.67	0.43	0.45	0.62	0.6	0.83	8.3	7.3
20	M	68	1.73	Progressive dysarthria	1.23	102.1	1.18	0.8	0.69	0.45	0.42	0.64	0.62	1.04	8.2	8.1
21	F	52	1.56	PD	1.01	111.4	1.08	0.62	0.67	0.46	0.41	0.57	0.62	0.94	5.7	7.6
22	F	28	1.7	Healthy	1.13	92.9	1.29	0.84	0.74	0.5	0.51	0.63	0.59	0.87	7.3	7.2
23	F	49	1.55	Healthy	1.17	103.6	1.16	0.74	0.7	0.44	0.44	0.63	0.61	1.01	8.5	8.3
24	M	26	1.72	Healthy	1.11	96.8	1.24	0.73	0.76	0.51	0.49	0.59	0.61	0.89	8.2	8.1
25	M	27	1.82	Healthy	1.44	96.2	1.25	0.77	0.7	0.53	0.5	0.59	0.58	1.15	8.2	8.8
26	F	25	1.68	Healthy	1.16	92.1	1.3	0.82	0.8	0.47	0.52	0.64	0.61	0.89	8.5	8.5
27	M	34	1.76	Healthy	1.14	95.1	1.26	0.77	0.8	0.47	0.5	0.62	0.62	0.91	8.2	8.1
28	M	29	1.73	Healthy	1.22	92.1	1.3	0.82	0.84	0.5	0.45	0.62	0.65	0.94	9.3	8.9
29	M	39	1.65	Healthy	1.18	105	1.14	0.7	0.71	0.44	0.44	0.62	0.62	1.03	8.7	8.6
30	M	36	1.8	Healthy	1.04	83.2	1.44	0.91	0.88	0.55	0.55	0.63	0.62	0.72	9	8.7
31	M	53	1.74	Healthy	1.39	88.2	1.36	0.79	0.82	0.57	0.54	0.58	0.6	1.03	9	8.9
32	F	76	1.55	Healthy	0.93	87.1	1.38	0.86	0.88	0.5	0.52	0.63	0.63	0.68	7	7.2
33	F	63	1.64	Healthy	1.02	91.7	1.31	0.78	0.69	0.58	0.58	0.58	0.55	0.78	7.6	7.4
34	F	67	1.59	Healthy	1.22	96.8	1.24	0.74	0.7	0.51	0.53	0.59	0.57	0.99	7.7	7.7
35	M	48	1.73	Healthy	1.37	95.2	1.26	0.76	0.76	0.49	0.51	0.61	0.6	1.09	9.1	9
36	F	42	1.63	Healthy	1.01	81.2	1.48	0.91	0.92	0.58	0.55	0.61	0.63	0.68	6.9	6.8
37	F	60	1.6	Healthy	1.01	100	1.2	0.78	0.78	0.42	0.43	0.65	0.65	0.85	6.7	6.5

The gait parameters of the healthy subjects can be treated as a reference for assessing the gait abnormalities of the subjects with neurological conditions. The summary of the findings of the healthy gait is as follows:

- Stride length is dependent on age, height, gender and the intended walking speed of the subject at the moment. Therefore, not only stride length but also characteristic features of the subject should be taken into consideration to assess the gait abnormality. However, extreme low values might be an indicator of gait abnormality. The average stride length among the healthy subjects is 1.16±0.14 m, with the lowest stride length of 0.93 m. It is observed that the standard deviation is low and extreme low values are not common. Hence, rather than small inconsistencies, extreme values should be considered while assessing the gait.
- Cadence is also dependent on age, height, gender and the subject's intended pace. Moreover, cadence should be assessed considering stride length because these two metrics affect each other. For example, when a person intends to increase his/her pace, the average stride length and cadence increase together. Therefore, a small stride length value together with high cadence might be an indicator of the gait abnormality, since the person tries to have a certain speed at the cost of the small steps and a faster rate.
- One of the most important metrics in determining a gait abnormality is gait asymmetry, which reflects the differences for the same metric on the left and right sides of the feet. Asymmetries are mostly observed in the swing and stance times, stance ratios and maximum clearance metrics. It is observed that there are no considerable asymmetries for the healthy subjects.
- Another observation in the subjects with healthy gait is that they have stance ratios very close to the normal stance ratio of 60%. Therefore, the healthy subjects have not only the similarity of the left and right stance ratios which indicates symmetry and normal gait, but also normal stance ratios with average 61% for the left side and 61% for the right side, with very low standard deviations of 2.1% and 2.6% respectively. Any significant deviation from the nominal value of stance ratio might be an indicator of gait abnormality, hence it must be considered while assessing the gait.

Clinical observations of the subjects are compared with the extracted gait parameters. The clinical observations and the abnormalities observed in the gait of the

subjects with neurological conditions can be summarized as follows:

- Subject 1 has DLB and VaMCI with no history of fall. Extracted gait parameters do not suggest any gait abnormality. Also, there is no clinical findings of gait abnormality.
- Subject 2 has DLB with no history of fall. It is observed that there is a slight temporal asymmetry with stance ratios. Right and left feet stance ratios are measured to be 60% and 65% respectively. There is no significant gait abnormality verified by the clinical observation.
- Subject 3 has PD pronounced on one side and has a history of fall. The left-right asymmetry is observed by the neurologist. This observation is reflected in the extracted gait parameters, as there is a significant difference in clearance for the left and right feet, and a difference between the left and right stance ratios.
- Subject 4 has PD with classical Parkinsonian gait. The rigidity on the feet results in difficulty in taking off the feet from the ground. This difficulty causes deviations from the normal stance ratio of 60% with high stance ratios for the left and right feet, 72% and 71% respectively. It is also observed that the stride lengths and feet clearance are low, which indicates Parkinsonian gait.
- Subject 5 has VaD and DLB, without any extrapyramidal symptoms, but with overall slow movement. There is no significant asymmetry reflected in gait parameters. Also this is verified by the clinical observation. The slowness is reflected in the cadence and speed.
- Subject 6 has Parkinsonian gait. Small stride lengths, high stance ratios and an asymmetry in clearance are observed. These findings are verified by the clinical observations.
- Subject 7 has NPH and has problems in maintaining balance. This is not observed in any gait parameters, and the gait parameters of this subject do not indicate any significant abnormalities.
- Subject 8 has Parkinsonian gait with slow movement. This is reflected in small stride length, cadence and speed values.

- Subject 9 has VaD, DLB and Parkinsonism. It is observed that the stride lengths are small and the stance ratios are slightly higher than nominal value. Another observation is that right stance times are slightly higher than the left, which indicates an asymmetry.
- Subject 10 has VaD and exhibits no gait-related symptoms. The gait parameters also do not indicate any abnormality.
- Subject 11 has PD. The stride lengths are observed to be small. Also, a slight stance ratio asymmetry is reflected in the gait parameters, which is verified by the clinical observation of slightly higher left-sided rigidity.
- Subject 12 has PD with no significant gait-related symptoms. This is consistent with the gait parameters which do not indicate any significant abnormalities.
- Subject 13 has MCI and does not show any gait-related symptoms as expected from MCI. The gait parameters support this clinical observation.
- Subject 14 has VaP with Parkinsonism symptoms. High stance ratios, low feet clearance values and small stride lengths are observed, which is verified by the clinical observation.
- Subject 15 has DLB, causing left-right gait asymmetry. This is reflected in the higher left stance ratio. No other gait abnormalities are observed.
- Subject 16 has VaMCI. No significant abnormality is reflected in the gait parameters, which is verified by the clinical observation.
- Subject 17 has Dementia. Axial rigidity and slowness are observed by the clinical observation. This is reflected in slight asymmetry on stance ratios and clearance, and also low speed.
- Subject 18 has NPH. Small stride lengths, low cadence and speed are observed as a result of the experiments with this subject. Also, very high stance ratios on both feet are measured. These observations are in agreement with the clinical observation.
- Subject 19 has frontal gait apraxia. A slight asymmetry on clearance is observed which is verified with the clinical observation of balance problem.
- Subject 20 has progressive dysarthria. A slight asymmetry on stance times is observed. Clinical treatment verifies this with the observed balance problem.

- Subject 21 has PD. Asymmetry on stance ratio and clearance are observed. Clinical observation verifies this with the identified dyskinesia on the right foot.

5.1.3. Discussion

A mobile, unobtrusive and cheap inertial sensor-based gait analysis system is proposed. This system includes two small IMUs mounted on both feet and provides an unobtrusive solution to gait analysis without requiring any infrastructural setup, which can be used during daily activities of the people. Unlike the Kinect-based system, this system yields more mobility to the users, since it can be used outside the hospital or their homes. However, the battery life of the sensors remains to be a drawback in this system.

There are advantages of our system over existing inertial sensor-based systems in terms of sensor placement, the number of sensors in the system, sensor orientation and sensor noise and drift. The proposed system requires to mount two inertial sensors only on the feet without requiring any extra sensor placed on different parts of the body, which enables daily and friendly use of the system. Also, it does not require to place the sensor in a certain orientation, and IMUs can be placed in any orientation, which facilitates the daily use of the system. Moreover, the applied techniques overcome the challenges of inertial sensors such as errors in measurements and eventually a rich set of gait parameters are extracted.

The proposed system is tested with the healthy subjects and the subjects with neurological conditions. Healthy subjects are used as a reference in assessing gait parameters and in understanding gait abnormalities in the patient subjects. When extracted gait parameters are compared with the clinical observations, it is found that gait parameters reflect the abnormalities in the gait. Therefore, we conclude that the proposed system enables to monitor the gait parameters of both healthy people and people with neurological conditions.

6. METHODOLOGY FOR FALL RISK ASSESSMENT USING GAIT PARAMETERS

The purpose of this study is to investigate the effect of gait features on fall risk and then to identify people who are at high risk of fall based on the most significant gait characteristics. Objective assessment of gait analysis is valuable to identify the fall risk level since gait is an important intrinsic cause in fall risk. Previous studies showed that gait quality characteristics are related to fall risk [59,60]. We further study whether gait parameters that we extract from our foot-mounted inertial sensor-based gait analysis system can be utilized in the assessment of fall risk.

Using estimated gait parameters by utilizing the foot-mounted inertial sensor-based system described in Chapter 5, we extract significant gait features in terms of gait variability and symmetry, and determine their predictive effects in fall risk. We use feature selection algorithms and determine the most significant gait features which are related to the fall risk. Selected features are used in the classification of people into two groups, namely high and low fall risk.

Subject characteristics, including fall history, age and height are obtained from medical records and self reports. A questionnaire including subject characteristics that play important role in fall risk is prepared in collaboration with the neurologist. The classification results are evaluated based on these factors. We label our subjects who have a history of fall or who use a walking aid or cane as the high fall risk group and the subjects who have no history of fall in the past and use no walking aid are labeled as the low risk group. We examine gait features of these two groups and employ different machine learning methods for the classification of these two groups.

This chapter presents the details about the methodology employed in fall risk assessment. In Section 6.1, we present how we extract gait features out of estimated gait parameters. An insight into utilized dataset is given and, the visualization and

interpretation of extracted features are presented in Section 6.2. In Section 6.3, we provide the details of feature selection from a rich set of gait features. In Section 6.4, we employ different machine learning methods using selected features for fall risk classification. We conclude the chapter with the results and our discussion in Section 6.5.

6.1. Feature Extraction

Gait parameters are extracted using our inertial sensor-based gait analysis system. In addition to these parameters, symmetry and variability of gait parameters are computed. In this section, how we compute these features are detailed. The full set of features utilized in the fall risk assessment is presented in Table 6.1.

Gait symmetry provides valuable information concerning the fall risk in the elderly. For this reason, we choose to investigate the effect of gait symmetry in fall risk. We use Equation 6.1 to compute the symmetry indexes of the gait parameters.

$$SI = \frac{|\mathbf{X}_L - \mathbf{X}_R|}{0.5 \times (\mathbf{X}_L + \mathbf{X}_R)} \times 100\% \tag{6.1}$$

The Symmetry Index (SI) is the most commonly used method in the measurement of gait symmetry. SI factor is a method of percentage assessment of the differences between the gait parameters for both sides of the lower limbs during walking. The value of SI which is equal to zero indicates full symmetry, whereas SI equal to or greater than 100% indicates its asymmetry [61].

Another gait feature extracted in this study is the variability of gait parameters. We quantify the variability in stride length, max clearance, stance time and swing time per foot for every gait cycle. We use the coefficient of variation (CV) to reflect the variability for each of the parameters, which is computed as in Equation 6.2.

$$CV = \frac{\mathbf{X}_\sigma}{\mathbf{X}_\mu} \times 100\% \tag{6.2}$$

The reason why we compute the variability with coefficient of variation is that the standard deviation is not informative enough itself without considering the average values, therefore we choose to use the coefficient of variation to reflect variations of the gait.

Table 6.1. Full Set of Gait Features.

Extracted Gait Variability Features	CV Stride Len	CV Cycle Time
	CV Swing Time L	CV Swing Time R
	CV Stance Time L	CV Stance Time R
	CV Clearance L	CV Clearance R
Extracted Gait Symmetry Features	SI Stance Time	SI Swing Time
	SI Stance Ratio	SI Clearence
Extracted Gait Parameters	Avg Stride Length	Avg Cycle Time
	Speed	Cadence
	Avg Stance Time (L)	Avg Stance Time (R)
	Avg Swing Time (L)	Avg Swing Time (R)
	Stance Ratio (L)	Stance Ratio (R)
	Avg Clearance (L)	Avg Clearance (R)

6.2. Data Visualization and Interpretation

In this section, we present a detailed insight into the dataset that is used in this study. We provide the details of the subjects participated the conducted experiments and visualize important gait features of the subjects so that it gives us a better understanding of the problem.

Thirty seven subjects participated the study. Among 37 subjects, 11 of them are labeled as high fall risk and 26 of them are labeled as low fall risk based on their fall history by their self reports and use of walking aid or cane as these factors are related to fall risk [62].

- High fall risk group: People who experienced at least a fall in the past; or use either walking aid or cane.
- Low fall risk group: People who have no history of fall in the past; and do not use any walking aid or cane.

Additionally, 21 of the subjects are identified with different neurological conditions with gait implications. We collaborate with the neurologist and identify the most important factors having impact on the fall risk. These factors are listed in Table 6.2. Fall risk level identification of the subjects are interpreted based on the listed items.

Table 6.2. Questionnaire Topics.

Health	Hearth Rhythm Blood pressure Orthostatic hypotension Significant Polyneuropathy Significant Vision Defect Sleep disorder	**Medication**	Sedatives Anti depressant Dopaminergic Antipshotic ACHEI Neuroleptics Antiarhythmic Antihypertensives Alcohol
Behaviour	Anxiety Depression Agitation Psychosis		
Parkinsonizm	Tremor Gait Disorder Rigidity Bradykinesia	**Cognition**	Pure movement disorder Mild cognitive impairment (MCI) MCI with movement disorder Dementia pure cognitive Dementia with movement disorder
Mobility	History of fall Use of walker/cane	**Orthopedic**	Prothesis Orthopedic surgery

Participant characteristics are compared between groups using health status (patient or healthy) and fall risk (high fall risk group and low fall risk group) as two categorical factors in an analysis of variance (ANOVA). Subject characteristics and their main effects on their health status and fall risk are presented in Table 6.3. Significant p-values are starred. From the results, there is a statistically significant difference between the groups in terms of age. So, it can be said that age plays a crucial role in fall risk level and health status of the participants.

Table 6.3. Participant Characteristics Between Groups.

	High Risk	Low Risk	**Health Status**		**Risk Level**	
	Mean±SD	Mean±SD	F-ratio	p-value	F-ratio	p-value
Female/Male	8/3	10/16				
Age (year)	82.2±4.6	55.07±20.3	63.112	2.96e-09***	18.452	0.000138***
Height (m)	1.63±0.08	1.68±0.08	0.793	0.379	0.258	0.614
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$						

Additionally, the effect of gait characteristics on health status and fall risk are compared using ANOVA between high and low fall risk groups, and also between people with neurological diseases and healthy people. Gait characteristics and their main effects on health status and fall risk are presented in Table 6.4. Significant p-values are starred. From the results, it can be inferred that gait variability features are significantly different between people with healthy gait and people with pathological gait. Therefore, it can be said that the gait variability is an important indicator of the health status. Also, some of these gait variability features are significantly different between people with high fall risk and low fall risk such as the variability of stride length, swing time and clearance for the left foot. Symmetry indexes are other significant features between groups. The symmetry index of stance ratio and clearance are the most significant features between groups according to the health status, while the symmetry indexes of stance and swing time are the most significant features between fall risk groups. Speed and clearance of the left foot have significant

effect on the health status and fall risk. Moreover, it is observed that the average cycle and swing time for the right foot are significant on the health status of the people while cadence, the average stance time for the right foot are significant on fall risk.

Table 6.4. Gait Characteristics Between Groups.

	Health Status		**Risk Level**	
	F-ratio	p-value	F-ratio	p-value
CV Stride Length	1630.264	2.61e-13***	48.847	2.3e-05***
CV Cycle Time	8.835	0.01269**	0.000	0.997771
CV Stance Time (L)	385.967	6.47e-10***	1.083	0.320446
CV Stance Time (R)	5.113	0.04499**	0.843	0.378184
CV Swing (L)	411.025	4.62e-10***	32.205	0.000143***
CV Swing (R)	450.906	2.81e-10***	1.920	0.193322
CV Clearence (L)	4834.682	6.76e-16***	11.134	0.006631***
CV Clearence (R)	7.523	0.01913**	0.021	0.886583
SIAvgStanceTime	1.069	0.32327	4.584	0.055508*
SIAvgSwingTime	1.700	0.21895	3.429	0.091083*
SIStanceRatio	18.155	0.00134**	1.724	0.215941
SIAvgClearance	14.673	0.00279**	1.674	0.222292
Cadence	0.130	0.72504	16.394	0.001919**
Speed	6.254	0.02947**	4.100	0.067855*
Avg Stride Length	1.519	0.24343	0.013	0.910299
Avg Cycle Time	4.932	0.04830**	0.364	0.558638
Avg Stance Time (L)	1.093	0.31823	0.020	0.889434
Avg Stance Time (R)	1.222	0.29262	4.030	0.069909*
Avg Swing Time (L)	2.634	0.13291	0.238	0.635376
Avg Swing Time (R)	3.925	0.07314*	0.931	0.355439
Stance Ratio (L)	0.990	0.34105	3.184	0.101959
Stance Ratio (R)	1.498	0.24649	2.634	0.132866
Avg Clearance (L)	6.712	0.02511**	15.697	0.002227***
Avg Clearance (R)	0.016	0.90077	0.430	0.525374

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Significant gait characteristics between fall risk groups are visualized in Figures 6.1-6.7. Significant gait characteristics in figures can be summarized as follows.

- In Figure 6.1, left and right feet clearance of the participants is shown. From the figure, it can be observed that people who have experienced at least a fall in the past have lower clearance compared to people who do not have a fall history. This might suggest us that a low foot clearance might cause falls. However, as participants' gait characteristics are examined in a retrospective manner with respect to the history of fall, it might be also said that a low foot clearance is a result of the fall as people develop fear of falling after falls, they hesitate to take their foot off the ground afterwards.
- In Figure 6.2, the stance time of the left and right feet is shown. It can be observed that people at high fall risk in general have higher stance time than people at low risk. This might be due to the situation that they find it difficult to move on and so, they prefer to stay in the stance period longer, which helps them to balance themselves better. Another reason for a longer stance period might be that if people have fear or anxiety of falling, they might intentionally choose to stay in the stance period longer.
- In Figure 6.3, cadence and speed comparison is depicted. This figure shows that people at high fall risk in general have lower speed compared to people at lower fall risk. Also, the former group has slightly lower cadence than the latter group. It can be inferred that people who are likely to fall walk slowly than the others. Another reason for slow speed might be the preventive behaviour they have against the thought of a possible fall.
- In Figure 6.4, the swing time variability for both feet is shown. It can be observed that people at high fall risk in general have higher swing variability than people at lower fall risk. This might be due to irregularities and abnormalities in their walking.
- In Figure 6.5, the stance time variability for both feet is shown. It can be observed that people at high fall risk in general have higher stance variability than people at lower fall risk. There are a few outliers in the low risk group, this might be

due to the fact that they have not experienced a fall in the past but they have a some kind of gait problem that might cause to fall in the future.

- In Figure 6.6, the average swing time for both feet is shown. When looked at the figure, no general pattern is observed. However, it can be said that people at high risk in general have smaller swing time than people at low risk. This might be because they have difficulty in staying in balance, and they hurry themselves to do the transition to the stance period.
- In Figure 6.7, left and right stance ratio is shown. It is observed that people at higher fall risk have higher stance ratio than low fall risk group. This observation is in the agreement with the previous observations that they stay in the stance period longer and they stay in the swing period less than the fall risk group. Furthermore, normally, the stance phase constitutes approximately 60% of the whole gait cycle, while the swing phase constitutes 40% [31]. Therefore, it is expected that the stance ratio in a gait cycle will be around 60% of the whole gait cycle. From the figure, it is observed that people in the low fall risk group have a stance ratio between 55%-65% whereas, people in the high fall risk group usually have a stance ratio higher than 65%.
- In Figure 6.8, cadence and stride length comparison is shown. From the figure, it can be observed that some of the people in the high fall risk group have higher cadence with smaller stride lengths compared to the low risk group; or they have smaller cadence with higher stride lengths than the low risk group. This might be due to abnormalities in their walking, which cause falls.

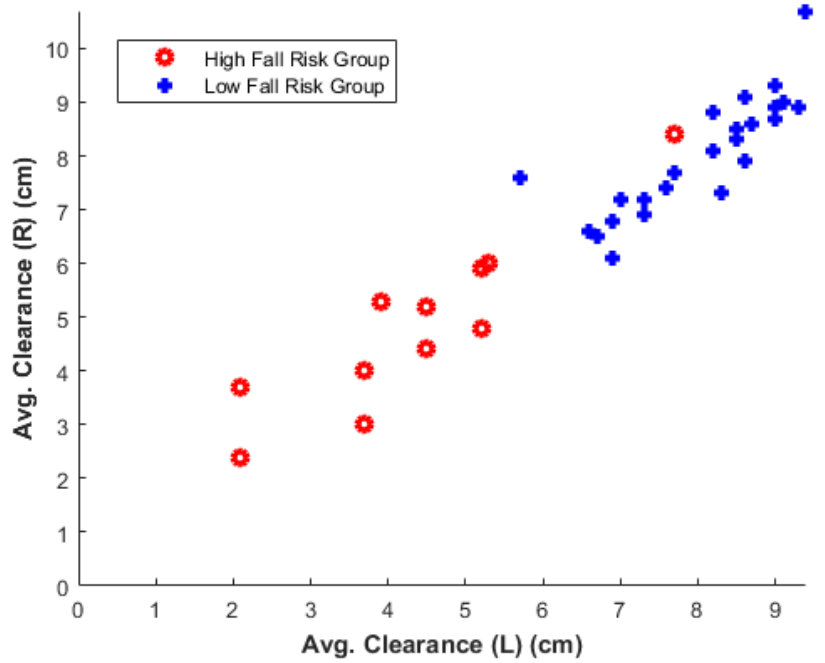

Figure 6.1. Clearance for both feet.

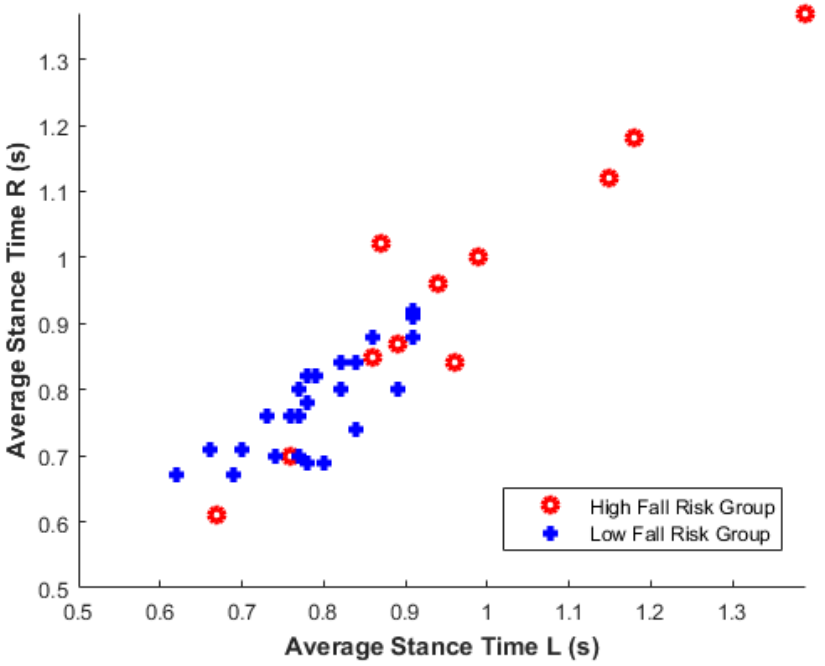

Figure 6.2. Average Stance Time.

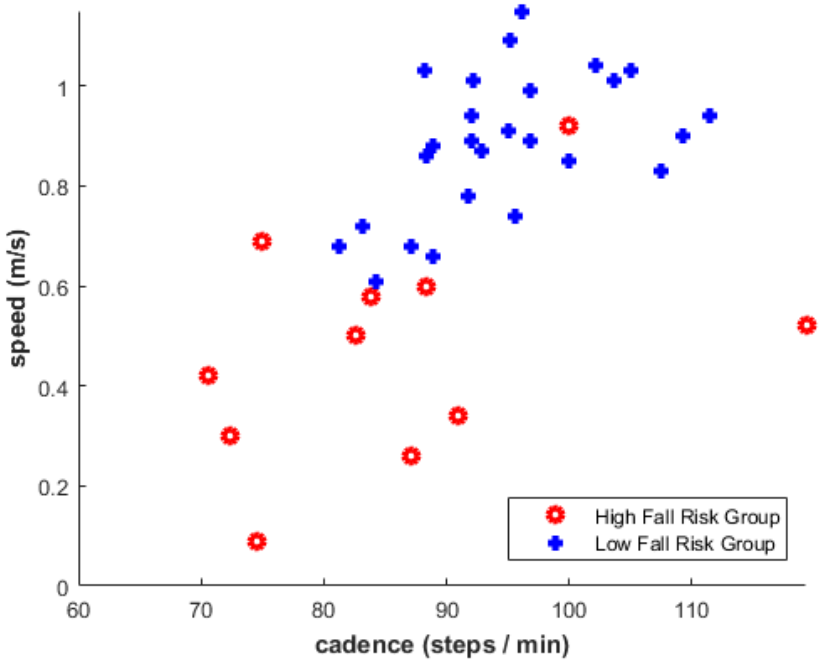

Figure 6.3. Cadence and Speed.

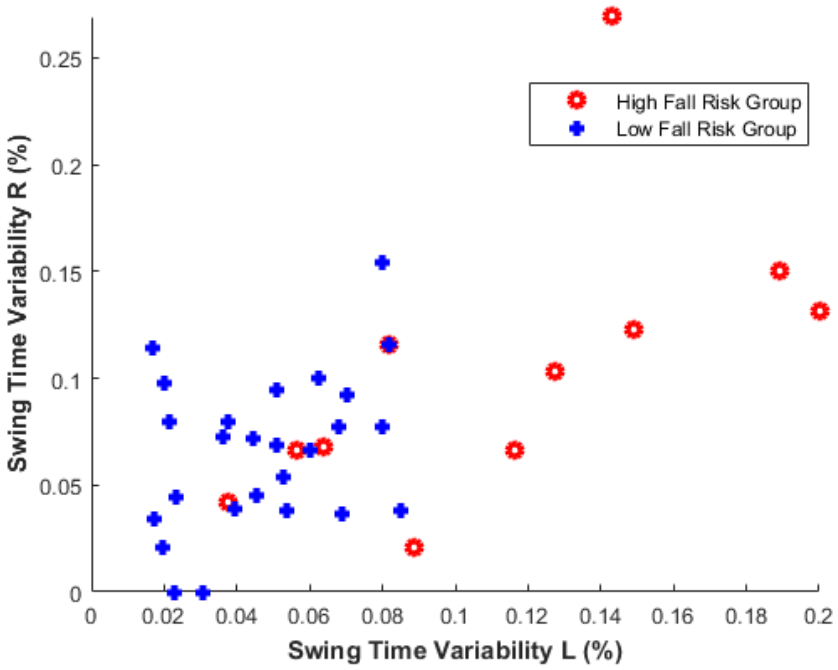

Figure 6.4. Swing Time Variability.

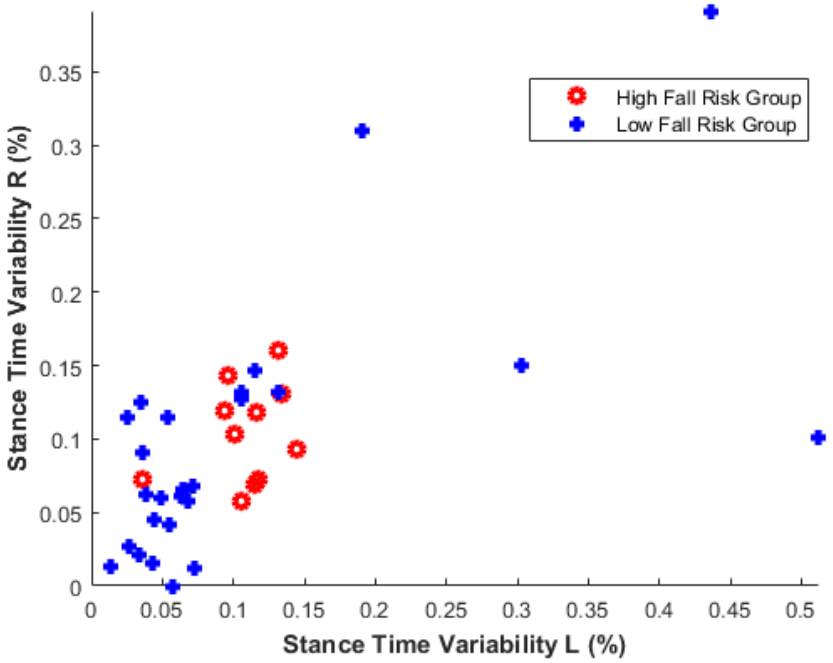

Figure 6.5. Stance Time Variability.

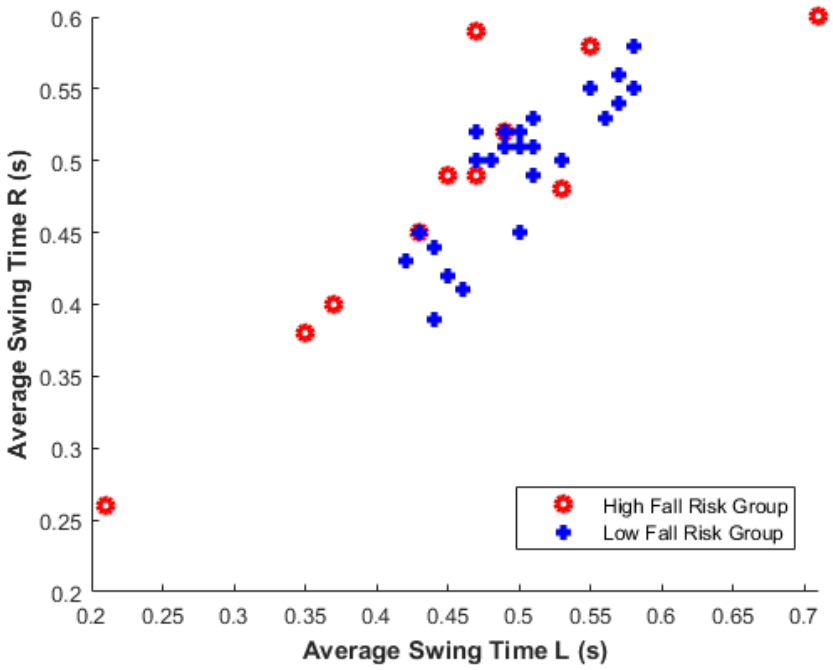

Figure 6.6. Average Swing Time.

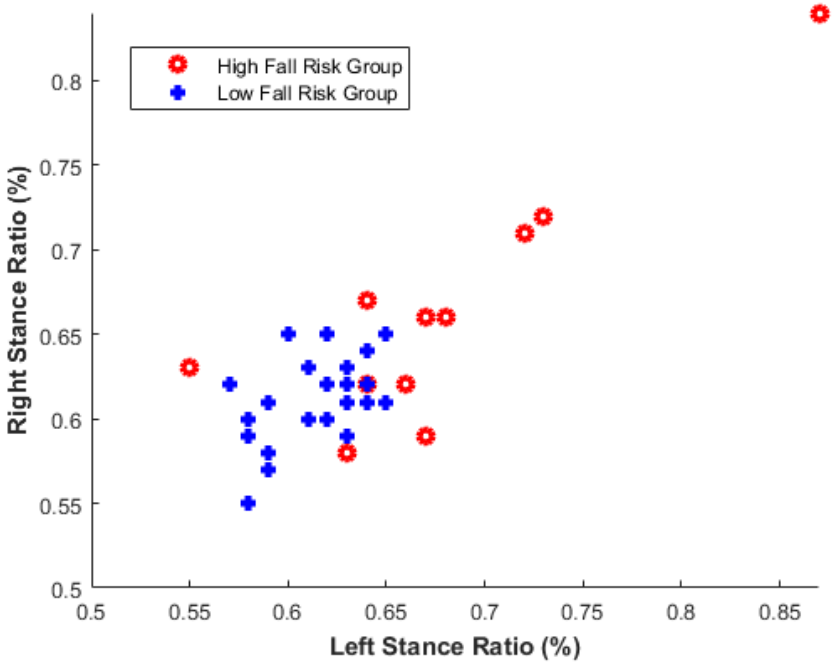

Figure 6.7. Left and Right Stance Ratio.

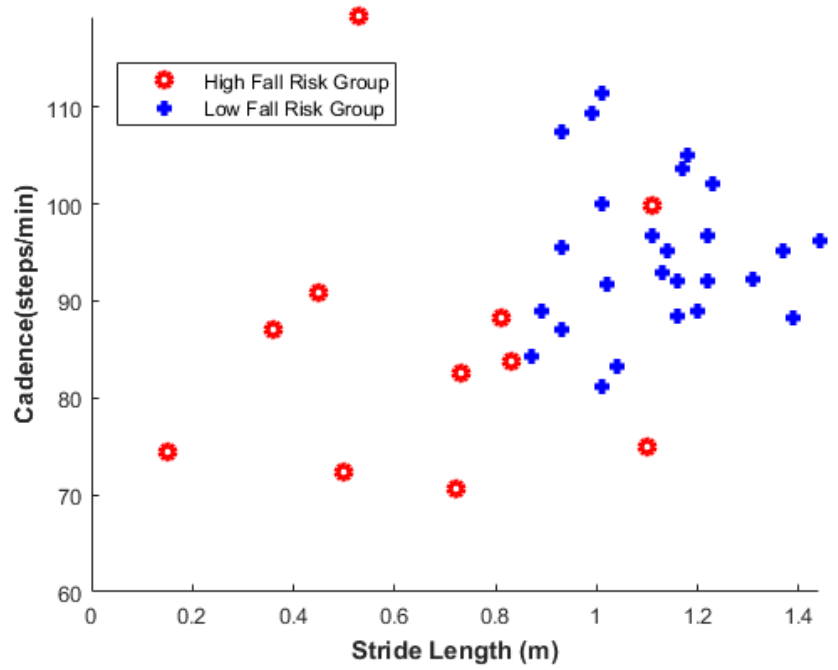

Figure 6.8. Cadence and Stride Length.

6.3. Feature Selection

Using more features does not always provide better results. Different features often mask distinctiveness of each other when they are used together and hinder better understanding of the underlying process that generated the data, which results in low performance. Also, it might be the case that some features include the same kind of information, which increases the complexity of the model. Especially, using more features with a small dataset creates the overfitting problem and reduces the generalization of the classifier. Therefore, an ideal subset of relevant features must be selected for model construction in order to obtain better performance in classification. In addition to the better performance, there are a number of advantages of selecting and working with subset of features. It, for instance, facilitates data visualization and data understanding; which is very significant for the solution of the problem. Also, it reduces the storage requirements and training time; which, in turn, decreases the computational cost of the system [63].

In this section, feature selection algorithms that are employed to obtain the ideal subset of features to classify high and low fall risk groups are detailed. We employ a pairwise correlation based feature selection method, a stepwise feature selection method, a decision tree for feature selection and a random forest for feature selection.

6.3.1. Pairwise Correlation based Feature Selection

One of the methods for understanding the relation of the feature to the response variable is the correlation analysis. The correlation is a measure of the strength of the association between the two variables. The most common measure of correlation is the Pearson Correlation [64], and it is used in this study. Pearson correlation analysis results in a coefficient that lies between -1 and +1, in which -1 represents perfect negative correlation whereas +1 means perfect positive correlation and 0 means no linear correlation between the two variables.

We apply a pairwise Pearson correlation technique to the full set of gait features listed in Table 6.1. It is aimed to investigate the relationship of features among themselves in a pairwise manner and the relationship of the features with the fall risk level.

The correlation coefficients which have absolute values higher than 0.3 are presented in Table 6.5. These are the most significant features which affect the fall risk level. However, pairwise correlations of all features must be considered while selecting the most appropriate features. Therefore, the relationship of the features are investigated in a pairwise manner and then features are selected.

Table 6.5. Correlated Features.

Feature	**Coefficient**	**Feature**	**Coefficient**
CV Stride Length	0.5416	Speed*	0.7475
CV Clearance (R)*	0.4744	Avg Stride Length	0.7174
CV Clearance (L)	0.5423	Avg Cycle Time	0.4552
CV Swing Time (L)	0.6725	Cadence	0.3833
CV Swing Time (R)	0.3626	Stance Ratio (R)	0.4892
SI Avg Swing Time*	0.5106	Stance Ratio (L)	0.5445
SI Stance Ratio*	0.3330	Average Stance Time (L)	0.5630
SI Average Clearance*	0.5149	Average Stance Time (R)	0.5451
*Selected Features			

For every feature in Table 6.1 and fall risk level, 25 $\times$ 25 correlation matrix is calculated, where 25 is the number of gait features including the fall risk level. The most correlated features with the fall risk level are found as listed in Table 6.5. The speed is found to be the most correlated feature with the fall risk level which has coefficient of 0.7475. It is also observed that other correlated features have high pairwise correlation with speed. Therefore, all correlated features can not be directly selected as a feature without considering pairwise correlations as it will create redundancy. Feature selection procedure is as follows. Initially, selected feature set is defined as an

empty set. The speed is included to the selected feature set. After including the speed feature to the selected feature set, other features are included to this set considering pairwise correlations with already selected features. Thus, the next most correlated feature is found and its pairwise correlations with the selected features are examined. If this feature has an absolute correlation coefficient higher than 0.5 with one of the selected features, this feature is not included to the selected feature set; otherwise it is included, as it is considered that it contains additional information which is not covered by already selected features. The process continues until pairwise coefficients of all correlated features are compared. As a result, selected features with the pairwise correlation method are found to be speed, CV of clearence (R), SI of average clearance, SI of average swing time and SI of stance ratio.

6.3.2. Stepwise Feature Selection

We apply logistic regression in a stepwise procedure to select the best subset of the features. Stepwise forward selection and stepwise backward elimination procedures for feature selection are utilized. Akaike information criterion (AIC) is a measure of goodness of the fit and the most common criterion used in evaluating the quality of models, so it is utilized for model selection [65]. At every step, both procedures' selection criterion is minimizing AIC information. In stepwise forward selection, we start with an empty feature set, and at each step we add the best feature based on the chosen criterion. The process stops when no more improvement in the result can be obtained with the remaining features in the set. Unlike the stepwise forward selection, in the stepwise backward elimination, we start with the full feature set and at each step we remove the feature that provides the least or no improvement based on the chosen criterion. This process terminates when it is not possible to improve the performance by removing any of the features from the selection. We also use a combination of the forward selection and backward elimination which is a bidirectional search selection. In this technique, at every step both stepwise forward selection and backward elimination are conducted, the one that yields the best result based on the chosen criterion is selected [66].

Considering the size of the dataset used in this study, three different models are constructed for different features. Stepwise feature selection methods are applied on gait symmetry, gait variability features and gait parameters separately. Full models are constructed as depicted in Figure 6.9. The first model is constructed to select significant gait variability features, whereas the second model is constructed to select significant gait symmetry features. The third model is constructed to investigate the significance of speed. From the pairwise correlation analysis, it is observed that the speed has high correlation with other gait parameters, therefore only the speed is included to the third model. As a result of the stepwise method, the final and the best model for the first model includes CV of stride length, cycle time and clearance. The final model for symmetry features consists of SI of swing and clearance. The best model for gait parameters includes the speed. We combine the selected features from three final models and feed them as input into the classification algorithms.

Require: CVStrideLength, CVCycleTime, CVClearence(R), CVClearence(L) CVSwingTime(R), CVSwingTime(L), CVStanceTime(L), CVStanceTime(R), gait variability features for all subjects

Require: SIStanceTime, SISwingTime, SIStanceRatio, SIClearance, gait symmetry index features for all subjects

Require: Speed , speed for all subjects

Require: FallRiskLevel, list of fall risk level

$full.model.variability \leftarrow FallRiskLevel \sim CVStrideLength + CVCycleTime + CVClearence(R) + CVClearence(L) + CVSwingTime(R) + CVSwingTime(L) + CVStanceTime(L) + CVStanceTime(R)$

$full.model.symmetry \leftarrow FallRiskLevel \sim SIStanceTime + SISwingTime + SIStanceRatio + SIClearance$

$full.model.gaitparameters \leftarrow FallRiskLevel \sim Speed$

final.model.variability $\leftarrow$ **stepwiseTechnique**(**full.model.variability**)

final.model.symmetry $\leftarrow$ **stepwiseTechnique**(**full.model.symmetry**)

final.model.gaitparameters $\leftarrow$ **stepwiseTechnique**(**full.model.gaitparams**)

Figure 6.9. Model Construction for Stepwise Selection Procedure.

6.3.3. Decision Tree for Feature Selection

Decision Trees (DT) are a non-parametric supervised learning methods used for classification and regression. A decision tree is constructed by learning simple decision rules inferred from the data. Each decision tree consists of decision nodes and terminal nodes. At each step, the feature that splits the input best is chosen as a decision node. While decision nodes represent the rules of the split, terminal nodes represent the decision results. Splitting continues recursively until there is no input to split where the terminal node is constructed [67].

The goodness of the split is defined by the impurity function that aims the best separation of the data. In our implementation, entropy function is used as an impurity measure. Splitting continues until all the leaves are pure based on impurity measure. However, this might cause overfitting. For instance, if splitting is continued to the point where each terminal node contains only one data point, then it means that the tree overfits to the training data, and performs poorly with testing data. Pruning methods are employed to overcome this problem [68]. We use pruning to prevent overfitting.

The decision tree selects features while constructing the tree. The decision node is chosen such that it has the feature that separates the data best. Therefore, features selected at each decision node represent the data best. The algorithm results in 2-node decision tree as depicted in Figure 6.10. The features used in the construction of rules in the decision tree are average clearance for the left foot and SI of average stance time. The most important feature utilized in decision tree is the clearance for the left foot. The other important feature is SI of stance time.

From Figure 6.10, it can be inferred that if a person has an average clearance for the left foot less than 5.5 cm, his or her fall risk level is classified as high risk. If a person has an average clearance higher than 5.5 cm and SI of stance time less than 12.4516%, his or her fall risk level is classified as low risk. Otherwise, the constructed terminal node is not pure and they are again classified as the low fall risk group because of the pruning.

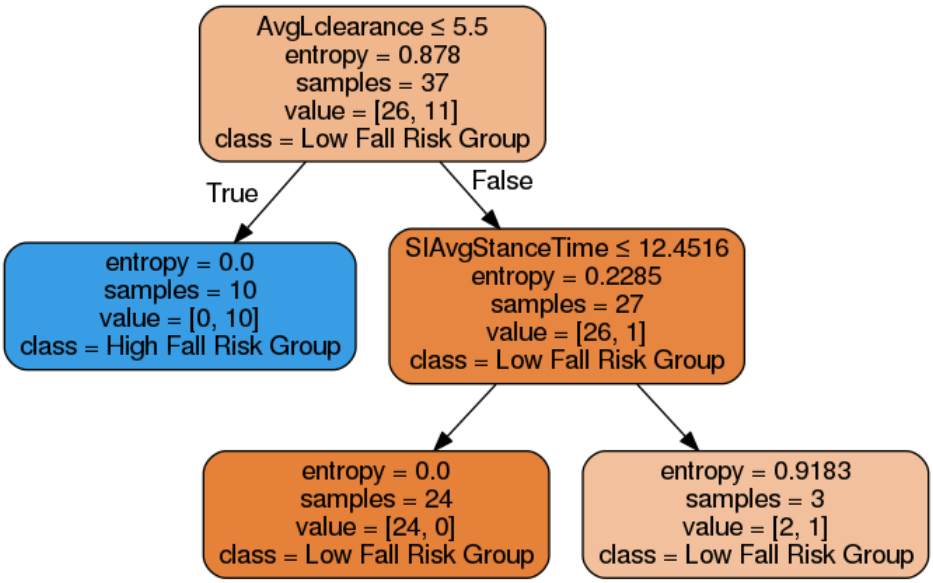

Figure 6.10. Decision Tree for Feature Selection.

6.3.4. Random Forest for Feature Selection

Random Forest (RF) is used as a classifier and also as a method for the feature selection. RF combines a number of decision trees built using different subsets of the training dataset. At each step of the tree construction, a different subset of features is randomly selected and a split is performed using a feature which separates the data best [69].

As a random forest includes many decision trees, different set of features are used by all of the decision trees. Feature importance of a single tree is computed based on the decrease in the model accuracy when the feature is removed. After computing feature importances for individual trees in RF, the importance of the features for RF is obtained by averaging importance measures for individual trees. Figure 6.11 shows the importance of features for the constructed random forest. We threshold the feature importances shown in the figure. We select the features which have importance more than 0.04. In the order of significance, average clearance for the left and right foot, speed, average stride length, CV of swing time for the left foot, stance ratio for the left

foot and SI of average clearance are selected features.

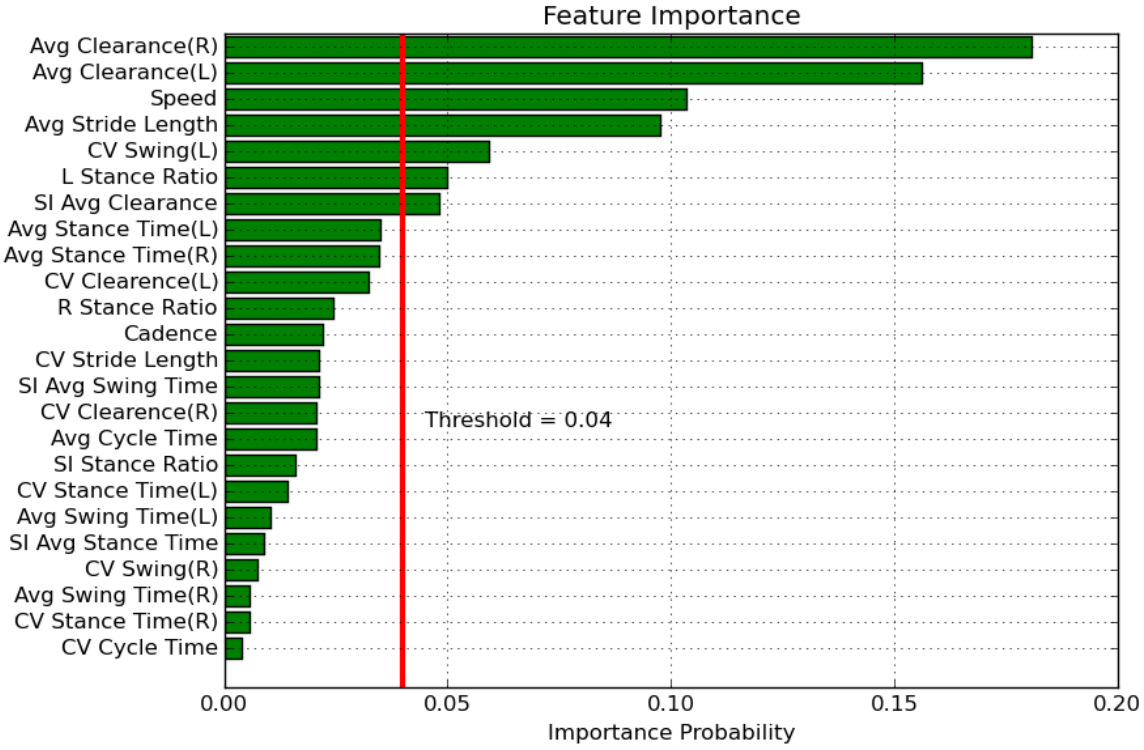

Figure 6.11. Random Forest for Feature Selection.

6.3.5. Discussion

Four different feature selection techniques are employed. Comparison of selected features by all of the methods is presented in Table 6.6. The speed is selected by all techniques. The symmetry index of swing time and clearance and stance ratio are selected by the pairwise correlation based feature selection method. The stepwise feature selection method selects only two of the symmetry indexes while the decision tree and random forest select only the symmetry index of stance time and stance ratio respectively. All four methods select different gait variability features.

Pairwise correlation based feature selection is used in this study to investigate fall risk level-feature and feature-feature correlations. This is the most commonly used

method in measuring linear correlations as it is efficient and reliable. However, it is not the most appropriate measure if correlations are not linear and if outliers exist in the dataset. For example, a single outlier especially in the small dataset can mask the strong correlation between two variables. Therefore, it is useful to study more robust measures of correlation for feature selection. On the other hand, tree based feature selection methods are not always able to exclude irrelevant features since an irrelevant feature also splits the data best [70]. Therefore, this must be taken into consideration while choosing the most appropriate feature selection method. Considering these aspects of pairwise correlation based and tree based feature selection methods, in Section 6.4 we employ features selected by the stepwise feature selection method.

Table 6.6. Selected features with different feature selection methods.

Full Set of Features	**Pairwise Correlation**	**Stepwise Selection**	**DT Selection**	**RF Selection**
CV Stride Length	×	✓	×	×
CV Cycle Time	×	✓	×	×
CV Stance Time (R)	×	×	×	×
CV Swing Time (L)	×	×	×	✓
CV Clearance (R)	✓	×	×	×
CV Clearance (L)	×	✓	×	×
SI Swing Time	✓	✓	×	×
SI Stance Time	×	×	✓	×
SI Stance Ratio	✓	×	×	×
SI Clearance	✓	✓	×	✓
Speed	✓	✓	✓	✓
Average Clearance (L)	×	×	×	✓
Average Clearance (R)	×	×	×	✓
Average Stride Length	×	×	×	✓
Stance Ratio (L)	×	×	×	✓

6.4. Classification

In this section, we employ different machine learning methods to build binary classifiers which can classify people into two groups, namely high and low fall risk groups. We use KNN, NB, LR, DT, RF and SVM classifiers and discuss their performances. We identify the outliers, and evaluate them based on the questionnaire.

Features selected by the stepwise feature selection method described in Section 6.3.2 are utilized in the training of the algorithms. Labeling of the subjects are determined based on their history of fall and use of walking aid as described in Section 6.2.

For the testing of the algorithms, leave-one out cross-validation technique is employed. The leave one out cross-validation procedure is performing the learning step with all but except one data point in the dataset and then testing the learned algorithm with the one left out at the beginning. This procedure is repeated for every data point in the dataset [67]. Average performances are reported. For evaluating the performances, accuracy, precision, sensitivity, specificity and F-measure metrics are used.

6.4.1. K Nearest Neighbours

K Nearest Neighbours (KNN) is a simple algorithm comparing a data point with its K-nearest neighbours and classifying this data point to the class that most frequently occurs among its K-neighbours [71]. Comparison of the data point with K-nearest neighbours is based on a defined similarity function. We use the Euclidean distance as the similarity function to identify the nearest neighbours. Thus, we normalize the dataset before applying KNN.

K is selected to be 3 considering the size and the nature of the dataset. Features selected by the stepwise method are used for the training of KNN. Otherwise, when full set of features feed into the KNN, it is observed that its performance decreases

as irrelevant features increase the complexity of the model. Moreover, unlike DT and RF, KNN does not perform feature selection; therefore, KNN takes advantage of the feature selection. KNN classifier performance is summarized in Table 6.7.

Table 6.7. KNN Confusion matrix for leave-one-out cross validation (K=3).

	Predicted Class		**Results (%)**				
True Class	Low Risk	High Risk	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low Risk	24	2	91.89	83.33	90.91	92.31	86.96
High Risk	1	10					

As shown in the Table 6.7, three of the subjects are misclassified. Subjects 2 and 21 are assigned to high fall risk. Based on the questionnaire filled for these subjects, Subject 2 uses antidepressants and exhibits depression and agitation. Even though this subject does not labeled as high fall risk, the effects of antidepressants and his or her psychological mood might have been reflected in the gait, and so he or she is classified into high fall risk group. Subject 21 has use of anti-dopamine, sedatives and has dyskinesia. Even though this subject has no fall history or use of walking aid, the effects of these characteristics might have been reflected in the gait and put his/her at high fall risk, that's why she or he is misclassified into the high fall risk group by the KNN. Subject 3 has a history of fall; however, this subject is misclassified into the low fall risk. It is observed in the clinical treatment that she or he has normal gait; however, experiences sudden changes in her or his postural tone, which might have caused a fall. Therefore, this misclassification suggests us continuous data collection and analysis for the identification of such cases.

6.4.2. Naive Bayesian

A Naive Bayesian (NB) is a simple probabilistic classifier based on the Bayes rule, so it probabilistically predicts the classes from the features based on Bayes' theorem. NB classifiers assume that the effect of a feature on a given class is independent of other

features. This assumption is called class conditional independence, which simplifies the computation involved and, therefore, is considered 'naive' [72].

NB classifier performance is depicted in Table 6.8. As shown in this table, three of the subjects are misclassified. Subject 3, 4 and 6 are misclassified to the low risk group. Initially they labeled as high fall risk because of history of fall and use of walking aid. However, it might be the case that they experienced a fall not because of gait characteristics but other reasons. For example, Subject 3 has sudden changes in postural tone and unpredictable changes in the gait. Therefore, this observation might cause misclassification of him to the low risk group. Subject 4 has orthopedic problems and also his or her family takes care of him/her very well. Therefore, the reason why she or he does not reflect any abnormality in the gait might be because of this protective care and eventually is classified to the low risk group. Subject 6 uses various medications and has slowness; however, she or he has protective behaviour concerning her/his walking. Therefore, this might cause misclassification of her/him to the low risk group.

Table 6.8. NB Confusion matrix for leave-one-out cross validation.

	Predicted Class		**Results (%)**				
True Class	Low Risk	High Risk	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low Risk	26	0	91.89	100.0	72.73	100.0	84.21
High Risk	3	8					

6.4.3. Logistic Regression

Binary Logistic Regression (LR) is another classification technique [66]. It computes the relationship between certain class and features. It estimates the probability that a data point belongs to certain class when the features are given. In our study, LR estimates the probability of being classified to certain fall risk group. The probabilities of being classified as high fall risk group is shown in Figure 6.12. LR classifier performance is depicted in Table 6.9.

Four of the subjects are misclassified. Subject 4, 7, 11 are misclassified to the low fall risk group, whereas Subject 21 is misclassified into the high fall risk group. Subject 4 has movement problems and her/his family takes care of her/him very well. Therefore, it might be the reason why she or he does not reflect any abnormality in the gait. Subject 7 has a problem in maintaining her/his balance. However, this is not reflected in her/his gait, which later causes misclassification. Subject 11 has significant cognitive skills and has awareness of fall. Even though she or he experienced a fall in the past, she or he is very careful with his/her walking. Therefore, any abnormality might not be reflected in the gait, which later causes misclassification. Subject 21 uses sedatives and also has dyskinesia and asymmetric movement disorder on the right foot. Even though this subject has neither fall history nor use of walking aid, these characteristics might be reflected in the gait and put his/her at high fall risk.

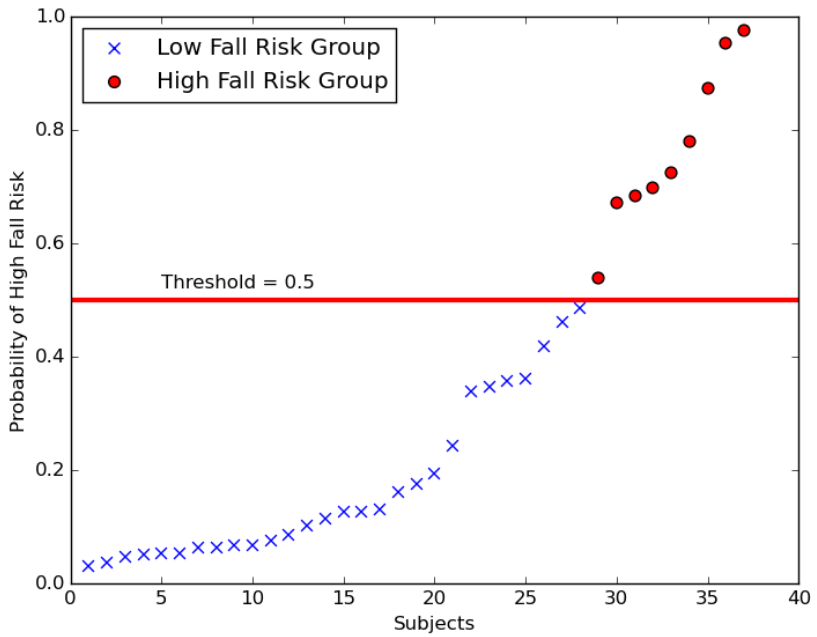

Figure 6.12. Probabilities of being classified as high fall risk group using LR.

Table 6.9. LR Confusion matrix for leave-one-out cross validation.

	Predicted Class		**Results (%)**				
True Class	Low Risk	High Risk	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low Risk	25	1	89.19	88.89	72.73	96.15	80
High Risk	3	8					

6.4.4. Decision Tree

Decision Trees (DT) form a well known supervised machine learning technique used for classification and regression. A decision tree is constructed by learning simple decision rules inferred from the data. DT is one of the most commonly used classification algorithms since they are simple to understand and to interpret [67]. However, decision tree learners can create over-complex trees that overfits to the training data and do not generalize the overall data. This is called overfitting and, methods such as pruning are necessary to avoid this problem. In our implementation, pruning methods are employed to prevent overfitting and also entropy function is used as an impurity measure.

Decision trees are also used for feature selection as described in Section 6.3.3. Therefore, we compare the DT performance with full feature set and with features selected by the stepwise feature selection algorithm. DT classifier performances with different feature sets are summarized in Table 6.10 and Table 6.11. It is observed that DT performs better with full set of features. This is because DT selects the most important features while constructing the tree and, these features might not be selected by the stepwise method.

Table 6.10. DT Confusion matrix for leave-one-out cross validation.

	Predicted Class		**Results (%)**				
True Class	Low Risk	High Risk	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low Risk	25	1	89.19	88.89	72.73	96.15	80.00
High Risk	3	8					

Table 6.11. DT Confusion matrix for leave-one-out cross validation with full set of features.

	Predicted Class		**Results (%)**				
True Class	Low Risk	High Risk	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low Risk	25	1	91.89	90.00	81.82	96.15	85.71
High Risk	2	9					

As shown in Table 6.10, four of the subjects are misclassified when the DT algorithm is applied with features selected by the stepwise procedure. Subjects 3, 5 and 7 are misclassified into the low fall risk group, whereas Subject 16 is misclassified into the high fall risk group. Moreover, when the DT algorithm is applied with the full feature set, three of the subjects are misclassified. Subjects 3 and 5 are misclassified into the low fall risk group, whereas Subject 21 is misclassified into the high fall risk group. Subject 5 is initially labeled as high fall risk because of the history of fall in the past. However, this subject has a hip prosthesis, which might suggest that this subject has experienced a fall because of the hip problem, which is not reflected in gait characteristics and; therefore, this causes misclassification. Based on the questionnaire, Subject 7 uses antidepressants and has a high hearth rhythm. Also in the clinical treatment, it is observed that she or he has difficulty in maintaining balance. These characteristics might not be reflected in the gait, and therefore, this subject is misclassified to the low fall risk group. Event though Subject 16 is not label as high risk, she or he is classified into the high fall risk group. She or he has MCI and uses antidepressants and sedatives, which might have impact on the gait and so might cause misclassification.

6.4.5. Random Forest

Random Forest (RF) is a supervised learning method for classification and regression. RF constructs a number of decision trees at training time using different samples of training set and feature set. RF outputs the class that is the most selected one among the individual trees [69]. In our implementation, we use pruning for individual decision trees in the forest to prevent overfitting. We construct 100 different decision trees in RF and use entropy as an impurity function.

Random Forests are also used for feature selection as described in Section 6.3.4. Therefore, we compare the performance of RF algorithm when both full feature set and features selected by the stepwise feature selection algorithm are given separately. RF classifier performances with different feature sets are summarized in Table 6.12 and Table 6.13. It is observed that RF performs better with full set of features. This is because individual trees in RF select the most important features while constructing themselves and, these features might not be selected by the stepwise method.

Table 6.12. RF Confusion matrix for leave-one-out cross validation.

	Predicted Class		**Results (%)**				
True Class	Low Risk	High Risk	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low Risk	25	1	91.89	90.00	81.82	96.15	85.71
High Risk	2	9					

Table 6.13. RF Confusion matrix for leave-one-out cross validation with full set of features.

	Predicted Class		**Results (%)**				
True Class	Low Risk	High Risk	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low Risk	26	0	94.59	100.00	81.82	100.0	90.00
High Risk	2	9					

As shown in Table 6.12, Subject 3, 5 and 17 are misclassified when the RF algorithm is applied with features selected by the stepwise procedure. Also, when the RF algorithm is applied with the full feature set, Subject 3 and 5 are misclassified into the low fall risk group. Event though Subject 17 is not label as high risk, she or he is classified into the high fall risk group. This subject has axial rigidity and uses antidepressants. These two factors might have effects on the gait that eventually classifier detects a risk of fall.

6.4.6. Support Vector Machines

Support Vector Machines (SVM) is a supervised machine learning technique. SVM projects the feature space onto a higher dimensional space in which the dataset becomes linearly separable. Kernel functions are used to construct a mapping into a high dimensional feature space. Then, a hyperplane is searched in the projected data space such that the margin between the hyperplane and both classes is maximized [73]. In our implementation, we use linear kernel function.

Features selected by the stepwise feature selection algorithm are used for the training of the SVM. Otherwise, when full set of features feed into the SVM, it is observed that its performance decreases, which is due to high complexity of the model that creates generalization problem. Also, unlike DT and RF, SVM does not perform feature selection. Therefore, SVM takes advantage of the feature selection. SVM classifier outperforms all of the classifier and its performance is depicted in Table 6.14.

Table 6.14. SVM Confusion matrix for leave-one-out cross validation.

	Predicted Class		**Results (%)**				
True Class	Low Risk	High Risk	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
Low Risk	25	1	94.59	90.91	90.91	96.15	90.91
High Risk	1	10					

Two of the subjects are misclassified. While Subject 4 is misclassified into the low fall risk group and Subject 21 is misclassified into the high fall risk group. These two subjects are also commonly misclassified by other methods. As mentioned, Subject 4 has orthopedic problems and so, his or her family takes care of him/her very well. Therefore, it might be the reason why she or he does not reflect any abnormality in the gait that might cause fall. Subject 21 has dyskinesia and uses sedatives. Even though this subject has no fall history or use of walking aid, these characteristics might be reflected in the gait and put his/her at high fall risk.

6.5. Results and Discussion

We study different feature selection techniques and employ different machine learning methods to identify people at high fall risk based on selected gait parameters. Feature selection and classification techniques are discussed in detail. The feature selection method and the classification technique should be chosen considering all the discussed aspects. The applied methodology yields promising results. However, there are limitations of the study. First of all, we work with a small dataset. It is very challenging to recruit people with a fall history so that their gait features are investigated. This suggests that generalization of the results should be considered carefully. Thus, even though our results are promising, more data must be collected and analysed.

Machine learning methods are employed to classify people into two fall risk groups, namely high and low fall risk groups. Performance of the machine learning methods are evaluated based on accuracy, precision, sensitivity, specificity and F-measure metrics. When people are classified to the high fall risk group; how accurate this classification is measured by precision. Sensitivity, also called true positive rate or recall, measures the accuracy of true classification when people have high fall risk, while specificity measures the accuracy of true classification when people have low fall risk. F-measure is harmonic mean of precision and recall values, and indicates the trade-off between precision and recall. Among these metrics, sensitivity or recall is particularly important as it can lead to unexpected consequences to classify people at high fall risk as people at low fall risk. All in all, the most appropriate classifier must

be chosen not only considering their accuracies but also taking into consideration their specificity, sensitivity, precision and F-measure.

Table 6.15 shows the comparison of accuracy, precision, specificity, sensitivity and F-measure of the applied different machine learning algorithms. As presented in Table 6.15, the SVM classifier and RF trained with the full feature set outperform the other classifiers. KNN, RF and NB classifiers and DT classifier trained with the full feature set have the same performance in terms of accuracy; however, other performance metrics differ. LR and DT classifiers have the same and the poorest performance among others.

Table 6.15. Performance comparison of different classification techniques.

	Result (%)				
Classification	**Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**F-measure**
SVM	94.59	90.91	90.91	96.15	90.91
RF (full feature set)	94.59	100.0	81.82	100.0	90.00
KNN	91.89	83.33	90.91	92.31	86.96
RF	91.89	90.00	81.82	96.15	85.71
DT (full feature set)	91.89	90.00	81.82	96.15	85.71
NB	91.89	100.0	72.73	100.0	84.21
LR	89.19	88.89	72.73	96.15	80.00
DT	89.19	88.89	72.73	96.15	80.00

DT and RF algorithms are applied with both features selected by the stepwise feature selection procedure and the full feature set. They perform better when the full feature set is given as they select the best features while constructing trees. Therefore, it should be taken into consideration while employing these algorithms.

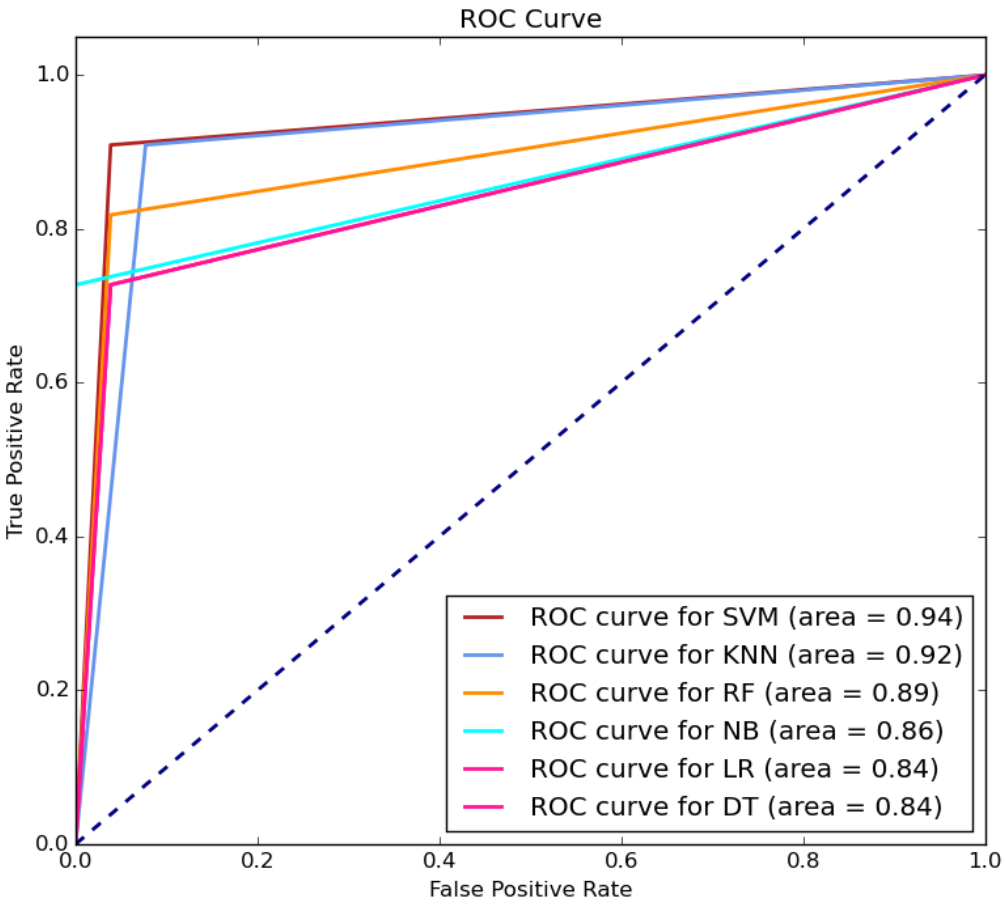

Figure 6.13. ROC curves of the classification algorithms.

Figure 6.13 shows the receiver operating characteristic (ROC) curves of the algorithms. The ROC curve is one of the methods for visualizing the classification quality, which shows the dependency between the true positive rate (TPR) and the false positive rate (FPR). TPR is equivalent to sensitivity and FPR is equivalent to (1 – specificity). All possible combinations of TPR and FPR compose a ROC space. One TPR and one FPR together determine a single point in the ROC space, and the position of a point in the ROC space shows the tradeoff between sensitivity and specificity. The steepness of the ROC curves is an important metric, as it is ideal to maximize the TPR while minimizing the FPR [74]. The quality of classification can be determined using the ROC curve by calculating the area under the ROC Curve (AUC) coefficient. AUC is

a measure of how well a parameter can distinguish between two groups. The higher the value of AUC coefficient, the better the classification accuracy is. AUC equal to 1 means a perfect classifier and, AUC equal to 0.5 is obtained for purely random classifiers. AUC less than 0.5 means the classifier performs worse than a random classifier. From Figure 6.13, it can be concluded that the SVM has the highest classification quality over other classification algorithms.

Additionally, evaluation of the results are made carefully, considering the items in the questionnaire that are listed in Table 6.2. This helps us understand cause and effect relationships of the fall better. Misclassification of the each classifier is discussed. Subjects 3, 4, 5 and 21 are the most commonly misclassified subjects. These subjects are the outliers and needed to be investigated in detail. Subject 3 has hypertension and exhibits sudden changes in postural tone whereas Subject 4 has orthopedic problems. Subject 5 has a hip prosthesis and uses medication for hypertension. Subject 21 uses sedatives and has dyskinesia. The characteristics of the patients should be taken into consideration as they might have different relationships with the fall. Therefore, interventions for fall prevention should be taken carefully considering these characteristics and the future studies should be conducted considering the misclassified subjects.

All in all, we must be very careful in making conclusions regarding the importance of health issues in this problem; however, we can conclude that machine learning methods can be used in the prediction of the level of fall risk based on gait characteristics of the people.

7. CONCLUSION AND FUTURE WORK

In this thesis, we proposed a fall risk assessment system based on gait parameters. First of all, we studied a Kinect based gait analysis system and a foot-mounted inertial sensor-based gait analysis system, and discussed their advantages and drawbacks. We chose to utilize the foot-mounted inertial sensor-based gait analysis system for the fall risk assessment for two reasons: (1) it provides a portable and mobile solution to gait analysis, which does not limit the mobility of the people during their daily lives and, (2) it enables to extract a rich set of gait parameters, which provide a more comprehensive assessment of the fall.

The purpose of this study is two-fold: (1) to determine which gait features are the most related with the fall and (2) to identify people at high fall risk based on these significant gait features. For this reason, we employed four different feature selection algorithms to determine the most important gait features associated with the falls, and discussed the features selected by these different methods. Later, we employed different machine learning algorithms to identify people at high fall risk. Performance of the algorithms evaluated based on accuracy, precision, recall and F-measure. We discussed the performance of the algorithms and examined the misclassified subjects in order to understand the factors that cause them to be outliers so that more extensive metrics or methods are considered for further studies.

Compared to other studies in the literature, our study has contributions to fall risk assessment using gait parameters in three main points. First, our study provides a more objective assessment of fall risk when compared with the functional fall risk assessment tools which evaluate the mobility and gait of the subjects. These tools depend on the observation and the judgement of the healthcare professionals, so it does not yield objective assessment of the fall risk. Also, it requires a massive amount of time and effort for the healthcare professionals to implement these tools on a regular basis. However, our system overcomes these difficulties by providing an objective and automated solution to the fall risk assessment. Second, our study provides an unob-

trusive solution to fall risk assessment, which does not intervene the daily lives of the users. Many studies assessing fall risk based on gait parameters utilize motion capture systems, ground reaction platforms or obtrusive sensors, which are mostly expensive and require dedicated space to be set up, therefore, they can not be used for daily fall risk assessment purposes. Third, our system outperforms other fall risk assessment systems identifying people at high fall risk using gait parameters. We achieved a high accuracy of 94.59% at the identification of fall risk levels. This is a very promising result when similar studies that are mentioned in Chapter 3 are considered.

As a future work, prospective experiment design is suggested for the assessment of the fall. In this study, we designed a retrospective experiment to investigate the role of gait features in fall risk level. However, there might be cases where the history of the fall in the past might not be reflected by the current gait features. Therefore, a prospective study should also be studied. Furthermore, the algorithms applied in our methodology yield better results when there are enough data representing the particular group. Therefore, utilized dataset can be enriched for a more comprehensive assessment of the fall risk. Moreover, daily activities of the people play an important role in fall risk. It is suggested that not only gait parameters but also daily activities should be tracked and used in the assessment of the fall risk.

REFERENCES

1. Tinetti, M. E. and C. S. Williams, "Falls, injuries due to falls, and the risk of admission to a nursing home", *New England journal of medicine*, Vol. 337, No. 18, pp. 1279–1284, 1997.

2. Stevens, J. A., P. S. Corso, E. A. Finkelstein and T. R. Miller, "The costs of fatal and non-fatal falls among older adults", *Injury prevention*, Vol. 12, No. 5, pp. 290–295, 2006.

3. Organisation mondiale de la santé. Bureau régional de l'Europe, C. T., Dawn Skelton, *What are the main risk factors for falls amongst older people and what are the most effective interventions to prevent these falls?*, WHO Regional Office for Europe, 2004.

4. Ageing, W. H. O. and L. C. Unit, *WHO global report on falls prevention in older age*, World Health Organization, 2008.

5. Howcroft, J., J. Kofman and E. D. Lemaire, "Review of fall risk assessment in geriatric populations using inertial sensors", *Journal of neuroengineering and rehabilitation*, Vol. 10, No. 1, p. 91, 2013.

6. Thurman, D. J., J. A. Stevens and J. K. Rao, "Practice parameter: Assessing patients in a neurology practice for risk of falls (an evidence-based review) report of the quality standards subcommittee of the American Academy of Neurology", *Neurology*, Vol. 70, No. 6, pp. 473–479, 2008.

7. Ganz, D. A., Y. Bao, P. G. Shekelle and L. Z. Rubenstein, "Will my patient fall?", *Jama*, Vol. 297, No. 1, pp. 77–86, 2007.

8. Oliver, D., F. Daly, F. C. Martin and M. E. McMurdo, "Risk factors and risk assessment tools for falls in hospital in-patients: a systematic review", *Age and*

ageing, Vol. 33, No. 2, pp. 122–130, 2004.

9. Laessoe, U., H. C. Hoeck, O. Simonsen, T. Sinkjaer and M. Voigt, "Fall risk in an active elderly population–can it be assessed?", *Journal of negative results in biomedicine*, Vol. 6, No. 1, p. 2, 2007.

10. Perell, K. L., A. Nelson, R. L. Goldman, S. L. Luther, N. Prieto-Lewis and L. Z. Rubenstein, "Fall risk assessment measures an analytic review", *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, Vol. 56, No. 12, pp. M761–M766, 2001.

11. Close, J. and S. R. Lord, "Fall assessment in older people", *British Medical Journal*, Vol. 343, No. d5153, 2011.

12. Persad, C., S. Cook and B. Giordani, "Assessing falls in the elderly: should we use simple screening tests or a comprehensive fall risk evaluation?", *European journal of physical and rehabilitation medicine*, Vol. 46, No. 2, pp. 249–259, 2010.

13. Tinetti, M. E., T. F. Williams and R. Mayewski, "Fall risk index for elderly patients based on number of chronic disabilities", *The American journal of medicine*, Vol. 80, No. 3, pp. 429–434, 1986.

14. Berg, K., "Balance and its measure in the elderly: a review", *Physiotherapy Canada*, Vol. 41, No. 5, pp. 240–246, 1989.

15. Duncan, P. W., D. K. Weiner, J. Chandler and S. Studenski, "Functional reach: a new clinical measure of balance", *Journal of gerontology*, Vol. 45, No. 6, pp. M192–M197, 1990.

16. Shumway-Cook, A. and M. H. Woollacott, *Motor control: theory and practical applications*, Lippincott Williams & Wilkins, 1995.

17. Shumway-Cook, A., S. Brauer and M. Woollacott, "Predicting the probability for

falls in community-dwelling older adults using the Timed Up & Go Test", *Physical therapy*, Vol. 80, No. 9, p. 896, 2000.

18. Vellas, B. J., S. J. Wayne, L. Romero, R. N. Baumgartner, L. Z. Rubenstein and P. J. Garry, "One-leg balance is an important predictor of injurious falls in older persons", *Journal of the American Geriatrics Society*, Vol. 45, No. 6, pp. 735–738, 1997.

19. Scott, V., K. Votova, A. Scanlan and J. Close, "Multifactorial and functional mobility assessment tools for fall risk among older adults in community, home-support, long-term and acute care settings", *Age and ageing*, Vol. 36, No. 2, pp. 130–139, 2007.

20. Lockhart, T. E., A. T. Barth, X. Zhang, R. Songra, E. Abdel-Rahman and J. Lach, "Portable, non-invasive fall risk assessment in end stage renal disease patients on hemodialysis", *Wireless Health 2010*, pp. 84–93, ACM, 2010.

21. Scanaill, C. N., C. Garattini, B. R. Greene and M. J. McGrath, "Technology innovation enabling falls risk assessment in a community setting", *Ageing international*, Vol. 36, No. 2, pp. 217–231, 2011.

22. Zhang, L., O. Ma, J. M. Fabre, R. H. Wood, S. U. Garcia, K. M. Ivey and E. D. McCann, "Classification of older adults with/without a fall history using machine learning methods", *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pp. 6760–6763, IEEE, 2015.

23. Caby, B., S. Kieffer, M. de Saint Hubert, G. Cremer and B. Macq, "Feature extraction and selection for objective gait analysis and fall risk assessment by accelerometry", *Biomedical engineering online*, Vol. 10, No. 1, p. 1, 2011.

24. "Inertial Measurement Units", `http://www.excelmicroel.com`, *EXEL ExLs3*, accessed at April 2017.

25. "Microsoft Kinect", `https://developer.microsoft.com/windows/kinect`, accessed at April 2017.

26. Tao, W., T. Liu, R. Zheng and H. Feng, "Gait analysis using wearable sensors", *Sensors*, Vol. 12, No. 2, pp. 2255–2283, 2012.

27. Gabel, M., R. Gilad-Bachrach, E. Renshaw and A. Schuster, "Full body gait analysis with Kinect", *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 1964–1967, IEEE, 2012.

28. Simila, H., M. Immonen, J. Merilahti and T. Petakoski-Hult, "Gait analysis and estimation of changes in fall risk factors", *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pp. 6939–6942, IEEE, 2015.

29. Del Din, S., A. Godfrey, B. Galna, S. Lord and L. Rochester, "Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length", *Journal of neuroengineering and rehabilitation*, Vol. 13, No. 1, p. 1, 2016.

30. Pawar, J. S. and D. E. Upasani, "Gait Activity Monitoring using Distributed Wearable Sensors for Patient Recovery", *International Research Journal of Engineering and Technology (IRJET)*, Vol. 2, No. 7, 2015.

31. Whittle, M. W., *Gait analysis: an introduction*, Butterworth-Heinemann, 2014.

32. Hu, F., *Telehealthcare Computing and Engineering: Principles and Design*, CRC Press, 2013.

33. Greene, B. R., 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.

34. Svoboda, Z., L. Bizovska, M. Janura, E. Kubonova, K. Janurova and N. Vuillerme, "Variability of spatial temporal gait parameters and center of pressure displacements during gait in elderly fallers and nonfallers: A 6-month prospective study", *Public Library of Science (PloS) One*, Vol. 12, No. 2, 2017.

35. Verghese, J., R. Holtzer, R. B. Lipton and C. Wang, "Quantitative gait markers and incident fall risk in older adults", *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, Vol. 64, No. 8, pp. 896–901, 2009.

36. Paterson, K., K. Hill and N. Lythgo, "Stride dynamics, gait variability and prospective falls risk in active community dwelling older women", *Gait & posture*, Vol. 33, No. 2, pp. 251–255, 2011.

37. Jiang, S., B. Zhang and D. Wei, "The elderly fall risk assessment and prediction based on gait analysis", *Computer and Information Technology (CIT), 2011 IEEE 11th International Conference on*, pp. 176–180, IEEE, 2011.

38. "Gait Analysis Walkway System", `http://www.gaitrite.com/`, accessed at May 2017.

39. van Schooten, K. S., M. Pijnappels, S. M. Rispens, P. J. Elders, P. Lips and J. H. van Dieën, "Ambulatory fall-risk assessment: amount and quality of daily-life gait predict falls in older adults", *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, Vol. 70, No. 5, pp. 608–615, 2015.

40. Weiss, A., T. Herman, N. Giladi and J. M. Hausdorff, "Objective assessment of fall risk in Parkinson's disease using a body-fixed sensor worn for 3 days", *Public Library of Science (PloS) One*, Vol. 9, No. 5, 2014.

41. Silva, J., J. Madureira, C. Tonelo, D. Baltazar, C. Silva, A. Martins, C. Alcobia and I. Sousa, "Comparing Machine Learning Approaches for Fall Risk Assessment", *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 4: BIOSIGNALS, (BIOSTEC 2017)*, pp. 223–

230, 2017.

42. Silva, J. and I. Sousa, "Instrumented timed up and go: Fall risk assessment based on inertial wearable sensors", *Medical Measurements and Applications (MeMeA), 2016 IEEE International Symposium on*, pp. 1–6, IEEE, 2016.

43. König, N., W. Taylor, G. Armbrecht, R. Dietzel and N. Singh, "Identification of functional parameters for the classification of older female fallers and prediction of 'first-time'fallers", *Journal of The Royal Society Interface*, Vol. 11, No. 97, p. 20140353, 2014.

44. Melzer, I., N. Benjuya and J. Kaplanski, "Postural stability in the elderly: a comparison between fallers and non-fallers", *Age and ageing*, Vol. 33, No. 6, pp. 602–607, 2004.

45. Merlo, A., D. Zemp, E. Zanda, S. Rocchi, F. Meroni, M. Tettamanti, A. Recchia, U. Lucca and P. Quadri, "Postural stability and history of falls in cognitively able older adults: The Canton Ticino study", *Gait & posture*, Vol. 36, No. 4, pp. 662–666, 2012.

46. Rantz, M., M. Skubic, C. Abbott, Y. Pak, E. E. Stone and S. J. Miller, "Automated Fall Risk Assessment and Detection in the Home: A Preliminary Investigation.", *AAAI Fall Symposium: Artificial Intelligence for Gerontechnology*, 2012.

47. Rantz, M., M. Skubic, C. Abbott, C. Galambos, M. Popescu, J. Keller, E. Stone, J. Back, S. J. Miller and G. F. Petroski, "Automated in-home fall risk assessment and detection sensor system for elders", *The Gerontologist*, Vol. 55, No. 1, pp. 78–87, 2015.

48. Bongue, B., C. Dupré, O. Beauchet, A. Rossat, B. Fantino and A. Colvez, "A screening tool with five risk factors was developed for fall-risk prediction in community-dwelling elderly", *Journal of clinical epidemiology*, Vol. 64, No. 10, pp. 1152–1160, 2011.

49. Tromp, A., S. Pluijm, J. Smit, D. Deeg, L. Bouter and P. Lips, "Fall-risk screening test: a prospective study on predictors for falls in community-dwelling elderly", *Journal of clinical epidemiology*, Vol. 54, No. 8, pp. 837–844, 2001.

50. Mateen, B. A., M. Bussas, C. Doogan, D. Waller, A. Saverino, F. J. Király and E. D. Playford, "Machine Learning in Falls Prediction; A cognition-based predictor of falls for the acute neurological in-patient population", *arXiv preprint arXiv:1607.07751*, 2016.

51. Baldewijns, G., G. Verheyden, B. Vanrumste and T. Croonenborghs, "Validation of the Kinect for gait analysis using the GAITRite walkway", *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pp. 5920–5923, IEEE, 2014.

52. Stone, E. E. and M. Skubic, "Unobtrusive, continuous, in-home gait measurement using the Microsoft Kinect", *IEEE Transactions on Biomedical Engineering*, Vol. 60, No. 10, pp. 2925–2932, 2013.

53. Ťupa, O., A. Procházka, O. Vyšata, M. Schätz, J. Mareš, M. Vališ and V. Mařík, "Motion tracking and gait feature estimation for recognising Parkinson's disease using MS Kinect", *Biomedical engineering online*, Vol. 14, No. 1, p. 97, 2015.

54. Procházka, A., O. Vyšata, M. Vališ, O. Ťupa, M. Schätz and V. Mařík, "Use of the image and depth sensors of the Microsoft Kinect for the detection of gait disorders", *Neural Computing and Applications*, Vol. 26, No. 7, pp. 1621–1629, 2015.

55. Tunca, C., N. Pehlivan, N. Ak, B. Arnrich, G. Salur and C. Ersoy, "Inertial Sensor-Based Robust Gait Analysis in Non-Hospital Settings for Neurological Disorders", *Sensors*, Vol. 17, No. 4, 2017.

56. Hung, T. N. and Y. S. Suh, "Inertial sensor-based two feet motion tracking for gait analysis", *Sensors*, Vol. 13, No. 5, pp. 5614–5629, 2013.

57. Ferrari, A., P. Ginis, M. Hardegger, F. Casamassima, L. Rocchi and L. Chiari, "A Mobile Kalman-Filter Based Solution for the Real-Time Estimation of Spatio-Temporal Gait Parameters", *IEEE transactions on neural systems and rehabilitation engineering*, Vol. 24, No. 7, pp. 764–773, 2016.

58. Haykin, S. S. *et al.*, *Kalman filtering and neural networks*, Wiley Online Library, 2001.

59. Barak, Y., R. C. Wagenaar and K. G. Holt, "Gait characteristics of elderly people with a history of falls: a dynamic approach", *Physical therapy*, Vol. 86, No. 11, p. 1501, 2006.

60. Mortaza, N., N. Abu Osman and N. Mehdikhani, "Are the spatio-temporal parameters of gait capable of distinguishing a faller from a non-faller elderly", *European Journal of Physical and Rehabilitation Medicine*, Vol. 50, No. 6, pp. 677–691, 2014.

61. Błażkiewicz, M., I. Wiszomirska and A. Wit, "Comparison of four methods of calculating the symmetry of spatial-temporal parameters of gait", *Acta of Bioengineering and Biomechanics*, Vol. 16, No. 1, 2014.

62. Rispens, S. M., K. S. van Schooten, M. Pijnappels, A. Daffertshofer, P. J. Beek and J. H. van Dieën, "Identification of fall risk predictors in daily life measurements: gait characteristics' reliability and association with self-reported fall history", *Neurorehabilitation and neural repair*, Vol. 29, No. 1, pp. 54–61, 2015.

63. Guyon, I. and A. Elisseeff, "An introduction to variable and feature selection", *Journal of machine learning research*, Vol. 3, No. Mar, pp. 1157–1182, 2003.

64. De Silva, A. M. and P. H. Leong, *Grammar-based feature generation for time-series prediction*, Springer, 2015.

65. Bozdogan, H., "Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions", *Psychometrika*, Vol. 52, No. 3, pp.

345–370, 1987.

66. King, J. E., "Binary logistic regression", *Best practices in quantitative methods*, pp. 358–384, 2008.

67. Alpaydin, E., *Introduction to Machine Learning*, The MIT Press, 2nd edn., 2010.

68. Mehta, M., J. Rissanen, R. Agrawal *et al.*, "MDL-Based Decision Tree Pruning.", *KDD*, Vol. 21, pp. 216–221, 1995.

69. Genuer, R., J.-M. Poggi and C. Tuleau-Malot, "Variable selection using random forests", *Pattern Recognition Letters*, Vol. 31, No. 14, pp. 2225–2236, 2010.

70. "A decision rule-based method for feature selection in predictive data mining", *Expert Systems with Applications*, Vol. 37, No. 1, pp. 602 – 609, 2010.

71. Cover, T. and P. Hart, "Nearest neighbor pattern classification", *IEEE transactions on information theory*, Vol. 13, No. 1, pp. 21–27, 1967.

72. Leung, K. M., "Naive bayesian classifier", *Polytechnic University Department of Computer Science/Finance and Risk Engineering*, 2007.

73. Gunn, S. R. *et al.*, "Support vector machines for classification and regression", *ISIS technical report*, Vol. 14, pp. 85–86, 1998.

74. Zhu, W., N. Zeng, N. Wang *et al.*, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations", *NESUG proceedings: health care and life sciences, Baltimore, Maryland*, pp. 1–9, 2010.