CLASSIFICATION OF BODY MOVEMENTS USING A MATTRESS-BASED SENSOR ARRAY

by

Yi Jui Lee

B.Sc., University at Buffalo, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in
THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Biomedical Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2019

© Yi Jui Lee, 2019
The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

**Classification of Body Movements Using a Mattress-Based Sensor Array**

submitted by Yi Jui Lee in partial fulfillment of the requirements for the degree of Master of Applied Science in Biomedical Engineering.

**Examining Committee:**

Dr. H.F. Machiel Van der Loos, Mechanical Engineering  
*Supervisor*

Dr. Osman Ipsiroglu, Pediatrics  
*Additional Examiner*

Dr. Lyndia Wu, Mechanical Engineering  
*Additional Examiner*
Abstract

Movement events during sleep could be used to infer underlying sleep physiologies and disorders based on their motor presentations. Periodic Limb Movement Disorder (PLMD), for instance, mostly occurs in the lower extremities and usually involves the dorsiflexion of the ankle. Evaluation of sleep disorders is typically done through clinical polysomnography (PSG). While PSG remains the most reliable and comprehensive tool for such assessments, the studies are intensive in terms of time, cost and labor. Certain motor indices might be underestimated due to the nature of PSG instrumentation, and for some populations, these studies could be considered intrusive and uncomfortable. In this work, SleepSmart, a mattress-based sensing system composing of an 8x6 array of 3-D accelerometer sensors, was developed to provide data for machine learning algorithms to classify body movements into different levels of granularity (coarse/fine-grained labels).

A study with 10 subjects was conducted. A movement protocol was adapted to simulate movements during sleep. Three classification domains were defined for the movements: a) Domain A – 3 classes inferring general movement characteristics, b) Domain B – 8 classes indicating movements at various body locations, and c) Domain C – 22 classes, where each class corresponds to a specific movement descriptor. Four learning algorithms were tested and compared. Random Forest (RF), Support Vector Machines (SVM), Naïve- Bayes (NB), and the k-Nearest Neighbor (k-NN) algorithms were used.

The classification accuracies averaged across all domains were 96.91%, 94.10%, 88.91%, 83.88% for subject-dependent models, and 89.87%, 89.45%, 73.95%, 69.21% for subject-independent models for the RF, SVM, NB and k-NN algorithms, respectively. In RF models,
averaged recall and precision measures were 96.29% and 96.74% for subject-dependent models, and 89.23% and 89.91% for subject-independent models. The investigation of the effect of different training sizes revealed small sample requirements for training (as low as 3 training samples per class) to attain accuracies higher or comparable to the baseline value (84%) for each domain.

In this work, we have proposed a non-invasive sensor system and demonstrated the generalizability and the effectiveness of the system in classifying movements at different label granularities under subject-dependent and subject-independent considerations.
Lay Summary

Movement during sleep can have different meanings based on their motor patterns. Periodic limb movement disorder, for instance, is characterized by the bending of the ankle and movements at the lower limbs. To diagnose sleep-related movement disorders, a clinical study called polysomnography (PSG) is usually performed, but these studies are often expensive, time-consuming, and sometimes underestimates movement severity. In this thesis, we proposed a non-invasive mattress system (SleepSmart) where acceleration sensors were placed on the bed for measurements. Using machine learning techniques, we showed that the system was able to classify movements into different levels of descriptions effectively. As reproducible data might be hard to collect in the sleep setting, we demonstrated that large training datasets are not necessarily needed to achieve optimal classification performance when using this mattress-array approach.
Preface

The original design of the SleepSmart system was proposed by Dr. Mike Van der Loos in 1997. Subsequent prototypes of the SleepSmart system: SleepSmart 1.0, and SleepSmart 2.0 were developed. Iterations of circuit design and scale prototypes for SleepSmart 2.0 were developed by Yuta Dobashi, Candice Ip, and the 2016 ECE Capstone team (Andrea Addo, Fan Jiang, Riley Marsh, and Clint Zhang). The serial communication interface was improved by Simon Zheng, and the color bands were prepared by Sherry Wang. Developments for SleepSmart 1.0 were made by Samantha Sterling, and Abenezer Teklemariam.

The author performed the background literature review and was responsible for the identification of the research objectives and experimental designs in this work. The author recruited subjects, conducted the study, revised, assembled, and tested the final build of the SleepSmart 2.0 system. The author developed, modified, and integrated the software and hardware for the studies, performed data analysis and written the associated manuscripts. Dr. Mike Van der Loos and Maram Sakr provided guidance and review of the study protocols, analysis, and manuscripts. Dr. Osman Ipsiroglu contributed ideas to the research direction in this work.

The human-subject experiment conducted was approved by the UBC Clinical Research Ethics Board (Certificate number: H15-01090). The associated work in this thesis have been published in:

# Table of Contents

Abstract ................................................................................................................................. iii
Lay Summary ........................................................................................................................... v
Preface ................................................................................................................................. vi
Table of Contents ............................................................................................................... vii
List of Tables ....................................................................................................................... x
List of Figures ...................................................................................................................... xii
List of Abbreviations .......................................................................................................... xv
Acknowledgements ............................................................................................................ xvi
Dedication ............................................................................................................................. xvii

Chapter 1: Introduction ...................................................................................................... 1
  1.1 Purpose and Overview of This Thesis ........................................................................... 3

Chapter 2: Background and Literature Review ................................................................ 5
  2.1 Overview of Sleep ........................................................................................................... 5
  2.2 Movement during Sleep ................................................................................................. 6
    2.2.1 Normal Movements during Sleep .......................................................................... 7
    2.2.2 Abnormal Movements during Sleep ..................................................................... 8
  2.3 Sleep Assessment Devices ......................................................................................... 10
    2.3.1 Clinical Devices .................................................................................................... 10
    2.3.2 Research Devices and Related Work .................................................................... 11
    2.3.3 Summary ............................................................................................................... 14

Chapter 3: Methods .......................................................................................................... 16
3.1 Data Collection ............................................................................................................................ 16
  3.1.1 Study Design ............................................................................................................................ 16
  3.1.2 Study Participants ..................................................................................................................... 17
  3.1.3 Study Procedures ...................................................................................................................... 18
3.2 Materials and Sensors ................................................................................................................... 21
  3.2.1 SleepSmart Mattress Sensors .................................................................................................. 22
  3.2.2 Software and User Interfaces .................................................................................................. 30
  3.2.3 Video Device ............................................................................................................................ 32
3.3 Movement Classes Definition ....................................................................................................... 33
  3.3.1 Domain A .................................................................................................................................. 33
  3.3.2 Domain B .................................................................................................................................. 35
  3.3.3 Domain C .................................................................................................................................. 36
3.4 Data Processing and Analysis ....................................................................................................... 38
  3.4.1 Data Segmentation .................................................................................................................... 38
  3.4.2 Pre-Processing .......................................................................................................................... 41
  3.4.3 Feature Extraction .................................................................................................................... 42
  3.4.4 Classifier Algorithms ............................................................................................................... 43
3.5 Summary ......................................................................................................................................... 49

Chapter 4: Results ................................................................................................................................. 50
  4.1 Domain A Classification ............................................................................................................... 53
  4.2 Domain B Classification ............................................................................................................... 54
  4.3 Domain C Classification ............................................................................................................... 56
List of Tables

Table 3-1: The movement protocol consisting of 22 different types of movements (excluding initial postural states). The first movement in each group represents the initial postural state. .................. 19

Table 3-2: Classification frameworks and descriptions in previous movement assessment studies. Frameworks that included more than 4 classes were summarized. ........................................... 21

Table 3-3: Summary of the movement descriptions and their corresponding classes in Domain A .......................................................................................................................................................................................... 35

Table 3-4: Summary of the movement descriptions and their corresponding classes in Domain B .......................................................................................................................................................................................... 36

Table 3-5: Summary of the movement descriptions and their corresponding classes in Domain C .......................................................................................................................................................................................... 37

Table 3-6: Features computed from the time-series acceleration data. The features are defined as $x(i), ..., x(n)$, where $n$ represents the total number of features........................................ 43

Table 4-1: Number of movement samples in each class for the domains. ........................................ 51

Table 4-2: Classification accuracies for the learning algorithms in the domains. ...................... 57

Table 5-1: Performance measures for the RF model in all domains............................................. 70

Table 5-2: Minimum training samples per class to attain higher than baseline accuracy for all domains.......................................................................................................................................................................................... 78

Table 6-1: Overall performance measures in all classification domains and models.................. 114

Table 6-2: Domain A performance measures for subject-dependent models............................ 114

Table 6-3: Domain B performance measures for subject-dependent approach......................... 115

Table 6-4: Domain C performance measures for subject-dependent approach....................... 116
Table 6-5: Domain A performance measures for subject-independent approach.......................... 117
Table 6-6: Domain B performance measures for subject-independent approach....................... 117
Table 6-7: Domain C performance measures for subject-independent approach....................... 118
List of Figures

Figure 3-1: Experimental setup and the apparatuses used (highlighted) in the study. .................. 22
Figure 3-2: SleepSmart 1.0 physical sensors (top); SleepSmart 1.0 system diagram (bottom)...... 23
Figure 3-3: SleepSmart 2.0 system level diagram (left) and the physical components (right)..... 24
Figure 3-4: Pin diagram for the decoder component ............................................................... 25
Figure 3-5: Wiring diagram in one sensor row (top) and all 8 sensor rows (bottom). ............ 26
Figure 3-6: LIS3DH accelerometer sensor schematic ............................................................. 27
Figure 3-7: Zero-ohm resistor links for Decoder #1 and Sensor #1. The corresponding configuration varies for other decoders and sensors ................................................................. 28
Figure 3-8: Close-up view of a sensor-mounted PCB (left) and a decoder-mounted PCB, attached with a row of sensors (right) ................................................................. 29
Figure 3-9: Mattress topper for the SleepSmart sensor array (left), together with a waterproof-fitted layer (right) ............................................................................................................. 30
Figure 3-10: Graphical user interface (GUI) of the audio player and the video stream previewing the study setup .................................................................................................................... 31
Figure 3-11: Camera view of the subject with the color bands worn on specific body locations. 33
Figure 3-12: An example of a Class 1 movement (moving from right to left) ......................... 33
Figure 3-13: An example of a Class 2 movement (straightening left arm) ............................... 34
Figure 3-14: Two examples of Class 3 movements: flexing both ankles (top) and bending right leg (bottom) ..................................................................................................................... 34
Figure 3-15: Detected location candidates using the color thresholder function .................... 39
Figure 3-16: Tri-axial acceleration data from the mattress sensors displayed with Domain A labels. ........................................................................................................................................................................................................................................... 40

Figure 3-17: Tri-axial acceleration data from the mattress sensors displayed with Domain B labels. ........................................................................................................................................................................................................................................... 40

Figure 3-18: Tri-axial acceleration data from the mattress sensors displayed with Domain C labels. ........................................................................................................................................................................................................................................... 41

Figure 3-19: Frequency spectrum of the z-axis acceleration signal. ................................................................................................................................. 42

Figure 3-20: A simplified representation of the random forest method using majority voting to predict classes. ........................................................................................................................................................................................................................................... 44

Figure 3-21: Simplified representation of the k-NN classification using the Euclidean distance metric (left); illustration of the Cosine similarity metric. ........................................................................................................................................................................................................................................... 45

Figure 3-22: Simplified representation of a linear (left) and a non-linear decision boundary (right) using different kernel functions. ........................................................................................................................................................................................................................................... 47

Figure 4-1: Illustration of the cross-trial validation approach. Data from 1 trial were held out as a testing set while using the remaining data from 14 trials as the training set. ................................................................................................................................. 51

Figure 4-2: Illustration of the cross-subject validation approach. Data from 1 subject was held out as a testing set while data from the remaining 9 subjects were used as the training set. ................................................................................................................................. 52

Figure 4-3: Average classification accuracies and minimum-maximum accuracy ranges for subject-dependent and subject-independent models for Domain A. ........................................................................................................................................................................................................................................... 54

Figure 4-4: Average classification accuracies and minimum-maximum accuracy ranges for subject-dependent and subject-independent models for Domain B. ........................................................................................................................................................................................................................................... 55
Figure 4-5: Average classification accuracies and minimum-maximum ranges for subject-dependent and subject-independent models for Domain A. ................................................................. 56

Figure 5-1: Domain C model accuracies illustrating variations in measures across different trials and subjects. .................................................................................................................................................. 60

Figure 5-2: The differences in performance between RF and SVM models under different training set sizes. .................................................................................................................................................................................. 61

Figure 5-3: Confusion matrices for Domain A (top) and Domain B (bottom) in cross-trial approaches .......................................................................................................................................................................................... 63

Figure 5-4: Confusion matrix for Domain C in the cross-trial approach ...................................................................................................................................................................................................... 65

Figure 5-5: Confusion matrices for Domain A (top) and Domain B (bottom) in cross-subject approaches ............................................................................................................................................................................. 67

Figure 5-6: Confusion matrices for Domain C for the cross-subject approach ...................................................................................................................................................................................................... 69

Figure 5-7: Iterative removal of trial data from the available training set in subject-dependent datasets .......................................................................................................................................................... 71

Figure 5-8: The corresponding model accuracies when the training set sizes were varied in each domain for subject-dependent datasets. ................................................................................................................................................................................................. 72

Figure 5-9: Iterative removal of subject data from the available training set in subject-independent datasets ............................................................................................................................................................................. 74

Figure 5-10: The corresponding model accuracies when the training set sizes were varied in each domain for subject-independent datasets. ...................................................................................................................................................................................................... 75

Figure 5-11: Remaining subject data further partitioned into 15 trials to train the classifier. ........................................................................................................................................................................................................... 77
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICSD</td>
<td>International Classification of Sleep Disorders</td>
</tr>
<tr>
<td>PLM</td>
<td>Periodic limb movement</td>
</tr>
<tr>
<td>PSG</td>
<td>Polysomnography</td>
</tr>
<tr>
<td>REM</td>
<td>Rapid eye movement</td>
</tr>
<tr>
<td>NREM</td>
<td>Non-rapid eye movement sleep</td>
</tr>
<tr>
<td>SRMD</td>
<td>Sleep-related movement disorder</td>
</tr>
<tr>
<td>PLMD</td>
<td>Periodic limb movement disorder</td>
</tr>
<tr>
<td>RMD</td>
<td>Rhythmic movement disorder</td>
</tr>
<tr>
<td>PLMD</td>
<td>Periodic leg movement during sleep</td>
</tr>
<tr>
<td>EMG</td>
<td>Electromyography</td>
</tr>
<tr>
<td>OSA</td>
<td>Obstructive sleep apnea</td>
</tr>
<tr>
<td>PCB</td>
<td>Printed circuit board</td>
</tr>
<tr>
<td>RF</td>
<td>Random forest</td>
</tr>
<tr>
<td>SVM</td>
<td>Support vector machine</td>
</tr>
<tr>
<td>NB</td>
<td>Naïve-Bayes classifier</td>
</tr>
<tr>
<td>k-NN</td>
<td>k-nearest neighbor</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to express my deepest gratitude to Dr. Mike Van der Loos and Dr. Osman Ipsioglu, who have been supportive and essential to my academic and research developments at UBC. I would like to thank the members of the Collaborative Advanced Robotics and Intelligent Systems (CARIS) lab for fostering a harmonious and collaborative environment for research work. I wish to acknowledge Maram Sakr for the valuable insights that she has provided in my thesis work, and I would also like to thank Mahsa Khalili, who have provided feedback and positive encouragements to the work challenges that I have faced. I gratefully acknowledge the funding received from Kids Brain Health Network (KBHN) that allowed me to explore this project. I would also like to thank the individuals who have participated in this study.
To my family.
Chapter 1: Introduction

Getting adequate and quality sleep is crucial for the healthy wellbeing and development of an individual, especially in growing children and adolescents. However, recent studies and surveys have reported that people are not getting enough sleep. According to a report published by the Centers for Disease Control and Prevention (CDC) in 2016, 35% of American adults are getting insufficient sleep [1]. Up to 70 million are affected by chronic sleep disorders [2], and in most cases, they remain undiagnosed and untreated [3][4], adversely influencing daily functions, health, and longevity. In Canada, surveys have also found that that over a third of Canadian adults and school-aged kids are sleep-deprived [5][6][7]. In younger individuals affected by neurodevelopmental, behavioral and emotional conditions, it is reported that up to 80% have sleep problems in this cohort [8]. Insufficient sleep, declared as a “public health problem” by the CDC, spans across other countries as well, such as Mexico, Japan, the United Kingdom and Germany [1].

Of particular interest in this thesis are the representations of nocturnal movements and the role they play in sleep disorders. The International Classification of Sleep Disorders (ICSD) has divided major sleep problems and disorders into several categories: parasomnias, sleep-related breathing disorders, circadian rhythm sleep disorders, sleep-related movement disorders, etc. [9]. Movements during sleep are commonly associated with several important sleep parameters: sleep stages and inter-stage transitions [10][11], disorder or disease indicators [12][13], and sleep quality [14]. For instance, abnormal sleep movements that are typically observed in sleep-related movement disorders are commonly presented in individuals with neurological disorders such as narcolepsy and Parkinson’s disease [15]. Sleep-related movement disorders, such as the periodic
limb movement (PLM) disorder, which affects 4-11% of adults [16], commonly produce the symptoms of discomfort and pain, worsening sleep quality and inducing excessive daytime sleepiness (EDS) [17][18], incurring detrimental effects to the individual and society.

Excessive daytime sleepiness due to poor sleep has been significantly associated with accidents involving transportation (planes, trains, automobiles) operators [19]. For instance, the U.S. National Highway Traffic Safety Administration estimates that sleep-deprived driving was responsible for 72,000 crashes, 44,000 injuries and 800 deaths in 2013. Even more so, these official statistics are regarded as substantial underestimates [20]. Major incidents and catastrophes have been linked with sleep deprivation and sleep disorders. Associated incidents include the Chernobyl nuclear accident, Exxon Valdez oil spill, and the Challenger Explosion [21]. Recent incidents that were linked with sleep disorders are the New Jersey/Long Island commuter train accidents in 2016 [22]. One study has suggested that the lower productivity levels and higher mortality risks associated with the lack of sleep cost the United States and Canada up to $411 billion and $21.4 billion in economic output, respectively, every year [1]. Furthermore, not getting enough restful sleep can cause significant developmental consequences, especially in younger individuals. Sleeping disorders are associated with a loss of quality of life [23], neurocognitive impairments and behavioral problems, growth and metabolic disorders, systemic inflammation, and adverse cardiovascular consequences [24][25]. Other unfavorable long-term outcomes such as delinquency, suicide, and substance abuse are also observed in many of the affected individuals [26][27].

Clinical polysomnography (PSG), occasionally accompanied by video recordings (video-polysomnography), is currently the gold standard and most commonly used technology for the
characterization of sleep states and sleep quality and the assessment of sleep problems [28][29].

While PSG remains the most reliable and comprehensive tool for sleep assessments, these studies are time, cost, and labor intensive [30], and potentially inaccessible depending on the region [31]. Manual scoring and annotation of polysomnograms is a time-consuming and tedious process, with varying differences between scorers [32][33]. Due to the invasive nature of polysomnography instrumentation, which requires the connection of multiple sensors and electrodes to the patient, most people find the exam uncomfortable [34]. Particularly in populations who have neurodevelopmental disorders, these individuals have difficulties tolerating intrusive monitoring in a new environment [35], and this could compromise the quality of the sleep assessment [36]. Moreover, in cases where detection of motor events is important to diagnose certain sleep-related movement disorders, PSG would underestimate motor-related parameters (such as the PLM-index) in the event where motor episodes are caused by other muscles, as they are not captured by the PSG [37]. Furthermore, the lack of sleep specialists and resources, especially in regions with pronounced geographical disparities, makes it difficult to access sleep studies and services [31]. The need for sleep assessments, especially for the pediatric population, is high, but there are currently insufficient resources to meet the demand for PSG sleep studies [31].

1.1 Purpose and Overview of This Thesis

The purpose of this thesis is to introduce a non-invasive, mattress-based sensor array system (O1) consisting of accelerometers and thermistors, in which we hypothesize that using this system would provide good performances in classifying voluntary body movements that are common during sleep into different domains of class descriptions (H1, H2, H3). The best performing
algorithm among 4 other tested algorithms was determined (O2), and the effect of training set size using this given algorithm was investigated (O3).

Chapter 2 of this thesis provides background and literature related to the thesis work, including an overview of clinical and other research devices used in the sleep domain. This chapter also outlines classification work that has been done in the literature. Chapter 3 outlines the study design and provide additional elaborations to the hypotheses and objectives of this work. The methods pertaining to the human-subjects study is elaborated. This chapter also introduces the mattress sensor system (O1) and provides the definition of the movement class domains. This chapter highlights the procedure pertaining to data segmentation, feature extraction and presents an overview of the learning algorithms used for the classification function. Chapter 4 details the cross-validation approaches and showed the accuracy measures for the subject-dependent and subject-independent learning models in the three classification domains (Domain A, Domain B, Domain C). Chapter 5 – Discussion – provides the interpretation of study results, comparing different classifier performances (O2) while also investigating the effect of the training set size in using subject-dependent and subject-independent data (O3). Chapter 6 concludes this thesis with a focus on summarizing the findings of this work, as well as covering recommendations for future work.
Chapter 2: Background and Literature Review

2.1 Overview of Sleep

Sleep can be defined as a prolonged duration of reduced activity. It results in a decreased responsiveness to external stimuli and it is a state that is easily reversible to wakefulness. Sleep is regulated by 2 internal biological mechanisms: the circadian rhythm and the homeostatic system. The circadian rhythm coordinates physiological and behavioral activities (such as when to sleep, wake, or eat) with daily environmental variations (such as sunlight or surrounding temperatures) in a 24-hour cycle [38], while the homeostatic system accumulates hypnogenic (sleep-inducing) compounds that generates the sleep drive [39].

There are two basic types of sleep: rapid eye movement (REM) sleep and non-rapid eye movement sleep (NREM), which has three (or four) different stages. Each type of sleep is associated with specific brain wave characteristics, body and neuronal activity [40]. Sleep generally progresses in a series of four or five sleep cycles each night and alternates between periods of NREM sleep stages and REM sleep. Each cycle typically follows the stages from NREM stage 1 to NREM stage 3/4 and finally into REM sleep, which is also known as deep sleep [40]. When sleepers progress from NREM sleep to REM sleep, they experience longer and deeper sleep. A stronger external stimulus is required to awake the sleeper at deeper states of sleep. They also experience more dream events during REM sleep [40]. There are also physiological changes associated with different stages of sleep. We can observe an increase in variability in the heart rate, respiration rate and eye movements [41], and a decrease of the body movement activity and body temperature when a sleeper progresses towards REM sleep [10]. Changes to these physiological signals and motor activities have been quantitatively measured and used in the sleep literature to
characterize sleep parameters. A few examples include distinguishing sleep stages [10], assessing sleep quality [42], and detecting sleep disorders [43], which will be detailed in the next section.

2.2 Movement during Sleep

Body movements during sleep can be used to characterize sleep. For instance, motor activities could infer transition between sleep and wakefulness [44]. One study conducted with healthy elderly subjects showed that when body movements occur in elderly individuals, arousal from sleep is likely to happen in the next minute [45]. Other works have also found a strong relationship between the frequency of body movements and sleep stages [10][11]. A study has found that the rate of movement decreases according to the following sequence: Wake » NREM stage 1 » REM » NREM stage 2 » NREM stage 3/4 [10]. Another study has suggested a connection of motor activity to inter-stage transition during sleep, whereby the presence of body movements in Stage 2 sleep may delay the transition from NREM stage 2 to NREM stage 3 (slow wave sleep) [46]. The link between the types of body movements (major or minor movements) has been found to be related to specific sleep stages and certain sleep parameters [47]. For example, large postural movements may signal sleep stage changes (such as into or out from wake or REM sleep) [12]. Correlations have been made between body movements and sleep quality. According to other works, better sleep quality can be deduced from low motor activities, while an increase in motor activities would indicate a disturbed sleep [48]. Other parameters such as time in bed [49] and even sleep-related movement disorder (SRMD) indices such as the periodic limb movement index, have been derived from frequency, type and duration of the body movement activity in other works [50][51].
2.2.1 Normal Movements during Sleep

Sleep movements typically include gross movements (large body movements), movements from specific muscle groups (such as the lower limbs or the head), and myoclonus or jerks, which are phasic, flickering movements (such as twitches in the faciomandibular regions), typically lasting less than 0.5s [12][44]. These movements usually occur in the transitions between sleep/wake, and during light or NREM sleep [52]. The number of body movements during sleep is reduced in comparison to wakefulness and gradually decreases with the depth of sleep (from NREM stage 1 to NREM stage 4). During REM sleep, muscle tone is further reduced and goes into a state of atony, but myoclonic twitches or jerks may still occur [12].

In one night, 3 to 5 minutes of body motility and a total of 80 to 200 movements can typically be observed in good sleepers [12][53]. Good sleepers generally sleep for more than 7 hours and take no longer than 15 minutes to fall asleep [12]. During sleep, there are typically more smaller movements as compared to larger movements, and it takes approximately 5 to 10 seconds for the sleeper to shift to another position [12][54]. For the distribution of movements during the night, some have suggested that this is a unique characteristic of the sleeper [46], while others have indicated that greater mobility can be seen during the second half of the night, suggesting a gradual increase in movements through the night [12][54].

Sleep-related motor activities also vary in different population groups. From infancy, motor activities typically peak at the first three months of life and decreases at later ages [44][45]. A decrease in phasic motor movements is first observed, and this is followed by a decline in localized body movements [44]. Motor asymmetries during sleep were also reported according to different age groups. For example, in the early phases of sleep, higher activity in the non-dominant limb
was reported, whereas in infants, such occurrence is not yet present [44]. Only a limited literature exists on the body movement patterns in the elderly populations [12].

2.2.2 Abnormal Movements during Sleep

Abnormal movements during sleep typically happen as a result of a disrupted sleep physiology or mechanism [52]. These occurrences are usually more frequent in the younger populations, affecting up to 20% of adolescents and 4% of adults [55]–[57][58]. Most guidelines have characterized these motor behaviors according to the type or complexity of the event with their underlying sleep physiology, such as whether the movements are associated with transitions of sleep and wake, NREM, or REM sleep [52].

Abnormal movements can typically be divided into two categories: simple and complex movements. Complex movements typically include purposeful or even eccentric movements. These behaviors are typically seen in arousal disorders such as sleep walking, confusional arousals or during REM sleep, and they are considered parasomnias according to the ICSD [52]. Parasomnias, which are described by the American Academy of Sleep Medicine (AASM) as undesirable motor events or experiences during sleep, are generally considered benign phenomena and do not usually have a serious effect on sleep quality [59]. For simple movements, these could be represented as quick, discrete, or rhythmical movements and usually occur during NREM sleep stages and transitional sleep phases [52][60]. This type of movement commonly manifests in sleep-related movement disorders, such as rhythmic movement disorder (RMD), Restless Legs Syndrome (RLS), and Periodic Limb Movement Disorder (PLMD) [52][60].
Sleep-related movement disorders disturb sleep and can cause significant distress to patients and families. They present a unique challenge to practitioners due to the intrinsic movement patterns associated with each of the sleep disorders. RMD, for instance, is represented by stereotyped and repetitive movements that typically involves larger muscle groups. Head banging, body rocking, or body rolling are several common motor representations of this disorder. These movements can range in intensity and potentially cause injuries to the sleeper or their bed partner in the event of more violent movements [61]. The frequency of the movements varies, and the duration of these movements can last from several minutes to several hours [60].

Restless Legs Syndrome (RLS) is a disorder that creates an urge to move the legs and is commonly accompanied by an uncomfortable sensation in the lower limbs. Patients usually describe the sensations as itching or aching, while younger children may associate these symptoms as pain [60]. In severe cases, other parts of the body, including the arms, can also be affected. To relieve the discomfort, people with RLS often keep their limbs moving to reduce or relieve the sensations. They may constantly move their legs, roll in bed, or even pace around the floor [62][61]. These symptoms may last from a few minutes to several hours [60]. Most patients with RLS frequently have motor symptoms with similar characteristics to PLMS [63]. PLMS contributes to the diagnosis of RLS and occurs in approximately 80% to 90% RLS patients [64]. PLMD is represented by episodes of stereotyped and repetitive limb movements that occurs during sleep. In contrast to RLS, PLMD does not occur during wakefulness and only happens when the individual is asleep. Occurring every 5 to 90 seconds, these episodes of muscle contractions are usually accompanied by intermittent arousals or awakenings and they are more common during the first half of the night and during NREM sleep [60].
Aside from sleep disorders, abnormal sleep motor patterns could be a symptom to seizures or other physiological conditions (e.g. hypoglycemia). Therefore, physicians need to be wary that abnormal motor events that occur during nighttime are not necessarily symptoms unique only to sleep disorders, but they may also be an indicator to neurologic, medical or even psychiatric disorders [52].

2.3 Sleep Assessment Devices

2.3.1 Clinical Devices

In sleep medicine, objective methods that are most frequently used in sleep assessments are polysomnography (PSG) and actigraphy. Overnight, laboratory-based PSG is commonly used to screen for sleep disorders and is often considered the most reliable method for assessing sleep [61]. PSG is a clinical tool that monitors and quantifies physiological changes that occur during sleep. Throughout the study, a technician or technologist is typically present to perform the study. The technician will be responsible for setting up the equipment, placing electrodes and sensors on the patient’s body (scalp, temples, chest, legs), and monitoring the study over the course of the night. The sensors usually include nasal pressure sensors, respiratory inductance plethysmography (RIP) belts, microphones or even pulse oximeters. Basic PSG includes the recording of brain activity via electroencephalogram (EEG), muscle activity via electromyogram (EMG), and eye movements via electrooculogram (EOG) [65]. Physiologic variables such as respiratory effort, airflow, blood oxygen levels and heart rate are also recorded. Many portable PSG devices exist to assess sleep at home. These recordings are generally less expensive, and patients typically would prefer home-based monitoring as compared to clinical settings. However, due to the inability to directly monitor
behavior, remediate technical issues, and control recording and setup conditions, the diagnostic benefits of home-based portable monitoring is often diminished [61].

An actigraph is a wearable, wrist-size instrument that captures movement activity, and it is commonly worn on the non-dominant hand or ankle. Certain models have an event button which could be pressed to mark unique occurrence such as lights off/on or waking up from bed. It can be worn at home and it is capable of recording data for a long period of time [65]. Sleep and wakefulness are typically derived from the measurements, but it is also capable of inferring sleep parameters such as total sleep duration, sleep onset latency, and sleep efficiency. It has also been used to characterize circadian rhythms and detect sleep disturbances in children [65]. One of the main advantages of actigraphy over PSG is that the actigraph could be used to record sleep continuously and for a longer duration.

2.3.2 Research Devices and Related Work

Work related to sleep-sensing technologies can be divided into three categories: wearable, non-contact, and mattress-based devices. In the domain of wearable sensors, several studies have employed wrist-worn devices to detect sleep disorders, such as obstructive sleep apnea (OSA) [66], and estimate sleep stages from movements [67]. Similar to an actigraph, these wrist-worn devices are battery-powered and can include sensors such as accelerometers or an optical pulse photoplethysmograph [66][67]. Recent works have also utilized polyvinylidene fluoride (PVDF) sensors attached at various body regions (chest, abdomen, and upper lip) for sleep assessments [68][69]. Commonly used in other medical applications, PVDF is a piezopolymer that is capable of measuring changes in temperature, pressure or strain [68]. Another study has used a biometric
shirt, embedded with accelerometer sensors, RIP bands, and electrocardiogram (ECG) leads for the classification of sleep / wake states using information from body movements and heart rate variability [70]. One other work has used stretchable socks embedded with fabric electrodes and adapters to detect periodic leg movements (PLMs) based on surface electromyogram measurements [71].

In the domain of non-contact devices, others have [72] developed a video monitoring system to detect the frequency of leg movements. The videography system consists of a commercial off-the-shelf camera capable of infrared sensing, depth sensing, and color video recording, and it is mounted and positioned over top of the sleeper. Other authors have also utilized similar technologies to detect sleep postures [73][74]. Another work has leveraged acceleration signals acquired from a smartphone placed near the sleeper’s pillow and compared common sleep parameters such as sleep efficiency and total sleep time to an actigraph. Numerous other works have also employed Doppler radar technologies to detect limb movements [75], estimate physiological signals and patterns [76], and determine sleep-wake states [77]. Doppler radar is typically used to determine the velocity of an object, and it works by bouncing pulses of radio waves off the target region and measuring the frequency changes in the radio waves [78]. One work have utilized near-infrared and thermal imaging techniques to estimate sleep biosignals [79]. Another study has used a tent-type Clean Unit System Platform (CUSP), a fan-filter unit that is capable of monitoring air particles to record air particle fluctuations corresponding to various body movements [80].

Many mattress-based solutions with different sensing modalities and setups can be found in the literature. Several works have utilized a static charged sensitive bed (SCSB) to detect frequency
of movements [53] and periodic movement activity [51]. In the SCSB system, static charges are formed between the mattress and the clothing when the sleeper moves. These charges induce potential differences and are measured by the system [53][12]. One work has utilized a mechanoelectrical transducer in conjunction with PSG to detect and classify movement activities [10]. PVDFs have also been implemented as pad-based solutions for sleep monitoring. Another study have utilized a setup composed of a 4 x 1 PVDF sensor configuration and encapsulated these sensors within a silicone pad to prevent damage [81]. Another work has proposed an on-mattress temperature monitoring system consisting of 16 thermistors to assess body movements based on changes in bed temperature. These sensors are arranged 6 cm apart on a cable, situated across the trunk or the lateral parts of the hips of each subject [49].

Represented in different design forms, numerous studies have also utilized bed-based pressure-sensing methods to characterize sleep parameters. Pressure sensors emit a signal that is correlated to the weight or force applied to the sensor [82]. Van der Loos et al. [83] have proposed a system called SleepSmart, which is composed of a multi-sensor mattress pad that is placed on the top of a mattress to detect breathing rate, temperature changes, and body postures. This system consists of 54 Force Sensing Resistors (FSRs) and 54 thermistors that are sampled at a 100 Hz rate. These sensors are more densely placed under the torso than the lower extremities. Several other works have implemented different configurations of pressure sensors, some under the mattress [50] or packaged in textiles [84][85], to detect and classify movements [86][87], characterize sleep states [88][85], estimate respiratory events [82][84][87], and detect poses [89]. A few work have also installed force sensors at the corners of a bed frame [12], or under bedposts [90] to assess sleep movements. These studies have also reported that the size (twin or full bed)
and type of mattress (spring or foam rubber) or frame did not have a significant effect on the performance of the classifier [90][12].

2.3.3 Summary

As patients are still required to wear devices on their body, wearables are not ideal for sleep monitoring, especially in syndromic or neuroatypical children. For instance, children with autism spectrum disorder (ASD) often have sensory challenges [91], including tactile defensiveness [92][93][94], exhibiting inaccurate behavioral information as they are frequently in a high state of alertness [92]. Tactile defensiveness is a type of sensory processing disorder where the affected individual feels overwhelmed by the touch stimuli and often results in behavioral and emotional responses that are undesirable, exaggerating, and aversive [92]. Parents have also raised concerns about children wearing these devices [91][95].

Non-contact devices have several drawbacks. Lights emitted by the camera or sensors can be distracting or intimidating for younger children at bedtime [95][96]. In some cases, the sleeper may sleep under bed coverings or blankets to avoid the camera or to look for comfort, making it challenging for the observer to accurately record behaviors and movement events in such conditions [97][95][72]. Proper setup and control of the environment and recording conditions (such as the removal of reflective, or light-absorbing artifacts for video-infrared setups) are also required to maintain the data quality and integrity of these systems.

For these reasons, we theorized that a key advantage of mattress-based implementations is their effectiveness in movement assessments. Mattress-based methods are not hindered by line of sight requirements. As they are direct or closer in proximity when measuring movements, spatial
and intensity inferences can easily be attained from sensor signals, especially when multi-sensor or fusion approaches were considered [98]. For non-contact devices, studies have mostly explored static assessments for movements (such as postures or number of movements) [74][73]. Some have reported suboptimal performances when parts of the body regions were concealed in bed coverings [99][74], while a few have managed to overcome this drawback by using different sensors, such as infrared cameras [73][74], and in various setup configurations (sensors installed above the feet of patient with an inclination of 30%) [73]. A few that have assessed dynamic movements (such as leg movements) operated using region of interest (ROI) segmentations and have only classified movements into few classes (typically leg movements) [37][100].

Similarly, previous mattress-based approaches have categorized movements into binary or 3-class frameworks [12][10]. Other works have utilized PSG to supplement classification workflows [10][101], and were limited in their ability to sufficiently describe movement events due to coarser class definitions [12][47][51][86]. Recent works have began to incorporate more granular labels into their classification approaches (6 to 9 classes) [102][103]. One other work have incorporated a total of 15 classes that consisted mostly of awake movement events from patients (such as lie down, sit up, exit bed) [104]. In some studies, the classification accuracies were lower (79.7%) [102], while others achieved better accuracies (90%) but binned movements into a framework with reduced number of classes [12][102][104][103]. Expanding on Van der Loos et al. [83] and previous classification work, the research described in this thesis seeks to provide finer-grained descriptions for movement events by performing classification operations into different label granularities (up to 22 classes) using subject-dependent and subject-independent approaches. The mattress-based system used in this research will be described in Section 3.2.1.
Chapter 3: Methods

3.1 Data Collection

3.1.1 Study Design

The main goal of this study is to evaluate the performance of the mattress-based sensor system in characterizing voluntary body movements that are common during sleep into three domains: Domain A, Domain B, and Domain C. The first domain (Domain A) consists of three separate classes: large postural movements, isolated movements from the head and the upper limbs, and leg movements. The second domain (Domain B) consists of 8 classes, labelling movements at specific body locations: head, torso and limbs, left arm, right arm, left leg, right leg, both legs, and both feet. The classes in the third domain (Domain C) fundamentally divides the movements into their discriminative descriptions, which correspond to a total of 22 unique labels. The definitions of the three domains are further detailed in Section 3.3. Originally introduced by Van der Loos et al. [83], the name of the device, SleepSmart, will be used throughout this thesis to refer to the mattress-based device used in this research.

The hypotheses investigated in this study were: using a mattress-based accelerometer array (O1), common in-bed movements can be classified into: (H1) Domain A – a tri-class system, (H2) Domain B – an eight-class system, and (H3) Domain C – a 22 class system. Due to the relatively high-dimensional feature space and a potentially non-linear relationship between the input and output datasets, the study seeks to compare performance between 4 machine learning algorithms (O2) in modelling this relationship between the input features and the target outcome (described in Section 3.4.4). This study also aims to investigate the effect of training set size (O3) with the best performing learning model in different classification domains.
The confirmation of hypothesis H1 would support the position that an accelerometer-based platform could be utilized to characterize body movements at a comparable performance to other sensing devices used in previous studies [12][37]. Subsequently, verification of hypothesis H2 and H3 would reveal the capability of the SleepSmart system in identifying movements into finer-grained labels (more descriptive) [103][104][102]. The automatic characterization, complemented with differentiating descriptions of motor events, could facilitate sleep studies, serve as better predictors in sleep assessments, and provide a deeper understanding of the connection between specific motor events and their underlying sleep physiologies or disorders. The investigation of the research objective, O2 would determine the best learning algorithm among 4 tested algorithms to model the classification problem of this mattress-human system, and the examination of the research objective O3 would inform ideal proportions of training set sizes necessary to attain good classification performance in different domains.

To test these hypotheses, the study consisted of one 15-trial session for each subject. In each trial, the subjects were instructed to perform pre-defined movements on the SleepSmart mattress. The experiment setup used in the study are detailed in Section 3.2, which includes the SleepSmart sensor system and a ceiling-mounted video camera.

### 3.1.2 Study Participants

Subjects were recruited through advertisements on social media, the author’s lab websites, as well as through recruitment flyers posted on public bulletin boards in the University of British Columbia (UBC) campus and surrounding communities. The inclusion criteria required the
subjects to be 14 years old or older, to be healthy with no heart or breathing problems, ambulatory and able to communicate in spoken and written English.

A total of 10 subjects (2 females and 8 males) with an age ranging from 17 to 21 were recruited for this study. Data collection took place in the Robotics for Rehabilitation Exercise and Assessment in Collaborative Healthcare (RREACH) Lab on UBC’s Point Grey campus. Each study lasted approximately 90 minutes. All participants provided written informed consent and received a compensation of $10 for their participation. The study protocol was approved by the UBC Clinical Research Ethics Board (H15-01090). The advertising materials and consent forms are presented in Appendix A.

3.1.3 Study Procedures

Before each study begins, the researcher introduced the study and asked the subject to sign a written consent form. After the subject provided written consent, the subject was asked to take off shoes and uncomfortable items. Then, the subject was instructed to wear 8 colored cloth bands at specific body locations (head, torso, arms, legs, feet) and lie down on the mattress.

Each subject performed 15 trials composed of predefined movements in each study session. Before the study begins, the subject was told to listen to the movement instructions played through an audio speaker, and to move after hearing the audio cue sound (“beep” sound). When the cue sound was played, the subject had approximately 6 seconds to perform the instructed movement and then to rest in a still position. Subsequently, audio instructions with the cue sounds were played until all the movement actions in the set had been performed by the subject. Each trial took
approximately 5 minutes to complete and the subjects were encouraged to move in a natural and comfortable manner.

Adapting from movement sets defined in other studies [12][102][103], the chosen movements consisted of 6 large movements of the torso, 6 movements of the head and arms, and 10 leg movements. Table 3.1 shows the movement set performed in the study protocol. The movement groups were randomized and began with an initial body posture state (e.g., lie down facing up), followed by a large movement then an assortment of isolated and leg movements. At the end of the study session, the experimenter removed the color bands and asked the subject to fill out a demographics form (listed in Appendix A.2).

Table 3-1: The movement protocol consisting of 22 different types of movements (excluding initial postural states). The first movement in each group represents the initial postural state.

<table>
<thead>
<tr>
<th>Movement Group</th>
<th>Movement Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lie down facing up</td>
</tr>
<tr>
<td></td>
<td>Move from back to right</td>
</tr>
<tr>
<td></td>
<td>Straighten both legs</td>
</tr>
<tr>
<td></td>
<td>Bend left leg</td>
</tr>
<tr>
<td></td>
<td>Straighten left leg</td>
</tr>
<tr>
<td></td>
<td>Bend both legs</td>
</tr>
<tr>
<td></td>
<td>Flex both ankles</td>
</tr>
<tr>
<td>2</td>
<td>Lie down facing up</td>
</tr>
<tr>
<td></td>
<td>Move from back to left</td>
</tr>
<tr>
<td></td>
<td>Straighten both legs</td>
</tr>
<tr>
<td></td>
<td>Bend right leg</td>
</tr>
<tr>
<td></td>
<td>Straighten right leg</td>
</tr>
<tr>
<td></td>
<td>Bend both legs</td>
</tr>
<tr>
<td></td>
<td>Flex both ankles</td>
</tr>
<tr>
<td>3</td>
<td>Lie down facing left</td>
</tr>
<tr>
<td></td>
<td>Move from left to right</td>
</tr>
<tr>
<td></td>
<td>Straighten left arm</td>
</tr>
<tr>
<td></td>
<td>Bend left arm</td>
</tr>
</tbody>
</table>
Table 3-1: The movement protocol consisting of 22 different types of movements (excluding initial postural states). The first movement in each group represents the initial postural state.

<table>
<thead>
<tr>
<th></th>
<th>Movement Protocol</th>
</tr>
</thead>
</table>
| 4 | Lie down facing right  
   | Move from right to left  
   | Straighten right arm  
   | Bend right arm          |
| 5 | Lie down facing right  
   | Move from right to back  
   | Turn head to the right |
| 6 | Lie down facing left  
   | Move from left to back  
   | Turn head to the left  |

The movement events performed in this study were chosen and supported by other studies in the domain of movement assessments [12][103][102]. These studies have incorporated a simulated movement protocol, in which a series of predefined movements in bed were defined to mimic those that are typically observed during sleep [12][102][103]. Table 3.2 expanded on a summary of classification frameworks that was defined in one other study [12].

One other movement event introduced in this study protocol (flexing both ankles) was based on the literature described in one other work [71]. The purpose of this inclusion is to incorporate movements of varying degrees of intensity (larger and smaller movements) in the protocol. From a biomechanical perspective, this intends to also simulate the common motor characteristic of periodic limb movements (PLMs) at the lower extremities (dorsiflexion of the ankles) [71].
Table 3-2: Classification frameworks and descriptions in previous movement assessment studies. Frameworks that included more than 4 classes were summarized.

<table>
<thead>
<tr>
<th>Study</th>
<th>Number of Classes</th>
<th>Framework Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zahradka et al. [102]</td>
<td>6</td>
<td>Major postural movements</td>
</tr>
<tr>
<td>Aaronson et al. [105]</td>
<td>2</td>
<td>Major postural movements</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Postural immobility</td>
</tr>
<tr>
<td>Alaziz et al. [103]</td>
<td>8</td>
<td>Movements at specific body regions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major postural movements</td>
</tr>
<tr>
<td>Gori et al. [45]</td>
<td>2</td>
<td>Duration of the body movements shorter than 15s</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Duration of the body movements longer than 15s</td>
</tr>
<tr>
<td>Wilde-Frenz [10]</td>
<td>3</td>
<td>Isolated movement of either the head, trunk or the limbs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Combined movements of distant parts of the body (such as the head or limbs)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major postural movements</td>
</tr>
<tr>
<td>Muzet et al. [46]</td>
<td>3</td>
<td>Movements without displacement that affect extremities of the body</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Movements that affect only one part of the body and modify its position</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major postural movements</td>
</tr>
<tr>
<td>Aaronoff et al. [104]</td>
<td>15</td>
<td>Combination of awake and body postural movements</td>
</tr>
<tr>
<td>Harada et al. [86]</td>
<td>2</td>
<td>Gross movements (large postural or articular movements)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slight movements (respiration and heart pulses)</td>
</tr>
<tr>
<td>Adami et al. [12]</td>
<td>3</td>
<td>Large postural movements</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Head and arm movements</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Leg movements</td>
</tr>
<tr>
<td>Eguchi et al. [71]</td>
<td>2</td>
<td>Dorsiflexion of the ankle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extension of the big toe</td>
</tr>
</tbody>
</table>

3.2 Materials and Sensors

This section introduces the experimental setup, which includes the mattress-based sensor array (SleepSmart), the software interface, and the video system. A personal computer (PC) running Windows 7 managed the audio player interface, video, and logging applications, while also
communicating with the SleepSmart sensors via USB connection. Figure 3.1 shows the experimental setup at the lab.

![Experimental setup and the apparatuses used (highlighted) in the study.](image)

### 3.2.1 SleepSmart Mattress Sensors

The original SleepSmart (v1.0) was first prototyped in 1999 at Stanford University. The patented system has undergone several improvements and modifications in its software and hardware, such as the employment of faster microchips and optimization of the serial communication interface, but fundamentally remains as a mattress-based, force-sensing device. Figure 3.2 illustrates the sensor, the system level diagram and the physical layout of SleepSmart 1.0.
The development of SleepSmart 2.0 integrated modern, digital sensors in place of the analog sensors in SleepSmart 1.0. Instead of force-sensitive resistors (FSRs), it now uses tri-axial accelerometers. An accelerometer is an electromechanical device that measures translation or linear acceleration. These measurements may be static, such as sensing tilt angles or orientation with respect to gravity. While FSRs are also capable of detecting static forces such as weight, modern accelerometer sensors have a larger measuring range, have higher resolution, and can measure dynamic forces caused by movement or vibration. Accelerometers are widely used in many applications, such as consumer electronics and navigation. Due to the low power
requirement of these sensors, they are frequently used in medical and wearable devices as well, since they pose no health or safety risk to the users. Compared to the FSR sensors, accelerometer sensors are also cheaper, smaller, and represent a simple-to-use, networkable technology in the domain of electronics. Figure 3.3 illustrates the system level diagram and physical hardware of SleepSmart 2.0.

![Figure 3.3: SleepSmart 2.0 system level diagram (left) and the physical components (right)](image)

The accelerometers used for the SleepSmart 2.0 system are LIS3DH sensors, manufactured by ST Microelectronics. It is a low-power, high performance, tri-axial linear accelerometer, capable of Serial Peripheral Interface (SPI) communications. The sensor was factory pre-calibrated for sensitivity, with user selectable scales ranging from ±2g to ±16g. The power rating of each sensor is rated at 3.3V and 20µA when using the sensor at a high-resolution data output operating mode (12-bit), with a sampling rate of 100 Hz.

The PC communicates with the SleepSmart sensors via Teensy 3.6, a USB-based microcontroller development system which features a 32-bit, 180 MHz ARM Cortex-M4 processor. This microcontroller is powered by a USB 2.0 port from the PC (5V, max 500mA), and it subsequently
supplies power to the SleepSmart mattress components via the microchip’s on-board regulators (3.3V, max 250mA).

In the main bus, the microcontroller is instrumented with 8 decoder chips (manufactured by Texas Instruments). As depicted in Figure 3.4, each decoder receives 4 inputs from the microcontroller (3 binary inputs and 1 “enable” input), and outputs “low” or “0” to one of the 6 output wires corresponding to the binary number represented in the inputs, following the 3 to 8-line decoder truth table. The output wires connect to the accelerometer sensors via the chip enable/chip select (CS) lines, where one of the 6 sensor chips that receives the corresponding active-low signal operates and transmits the acceleration data back to the microcontroller via the SPI interface. The Teensy microcontroller then sends the data to the PC using USB serial communication.

![Figure 3-4: Pin diagram for the decoder component.](image)

The SPI interface allows for writing to and reading from the registers of the sensor devices. The serial interface interacts using 4 wires: Chip Select (CS), Serial Port Clock (SPC), Serial Data
Input (SDI/MISO), and Serial Data Output (SDO/MOSI). A wiring diagram of SleepSmart 2.0 conveying SPI and wiring connections is illustrated in Figure 3.5.

Figure 3-5: Wiring diagram in one sensor row (top) and all 8 sensor rows (bottom).
At the accelerometer sensors, a set of power supply decoupling capacitors is placed near the $V_{dd}$ line (see Figure 3.6) to stabilize the voltage supplied to the sensor chip at the regular level (3.3V). As the sensors and decoders are active low, proper activation of the chips (via EN1-8 and CS1-6) was done using zero-ohm resistor links to prevent cross-talk of the signals. Figure 3.7 shows a schematic diagram for the zero-ohm links at Decoder #1 and Sensor #1.

![Figure 3-6: LIS3DH accelerometer sensor schematic.](image)
The connections between the decoders and the microcontroller were done via ribbon cables, while flat flexible cables (FFC) were used to connect the sensors. The ribbon cables were terminated at rectangular header connectors, while the FFCs were adjoined in between sensors via slide-lock connectors. The Teensy microcontroller was soldered to an external breadboard, while the other electronic components were assembled and surface-mounted on printed circuit boards (PCBs). A total of 56 PCBs (8 decoder-mounted PCBs and 48 sensor-mounted PCBs) were fabricated in this build. When assembled with the LIS3DH sensors and all other electronic components, each sensor-mounted PCB has a footprint of approximately 18 mm x 20 mm in length and width, and a height of 2.5 mm. Figure 3.8 below shows a close-up view of the physical sensors.
The longitudinal spacing, or the distance between rows of decoder chips, is 15 cm, while the lateral distance between sensors is 13 cm. The spacings were maintained via a mattress-sensor topper sheet, sewn with row pockets dedicated for the placement of the sensor array. On top of the mattress topper, a waterproof, fitted-sheet was used to wedge and cover the sensor array on a twin-sized mattress (188cm x 97cm in length and width). Velcro strips and 3D printed casings were also used for the decoder-mounted PCBs so that the boards can be housed and adhered to the side of the mattress.

Figure 3-8: Close-up view of a sensor-mounted PCB (left) and a decoder-mounted PCB, attached with a row of sensors (right).
3.2.2 Software and User Interfaces

Figure 3.10 shows the graphical user interface (GUI) developed in MATLAB that functions as the audio player interface for the study protocol. The interface displays the progress of the movement protocol through table annotations, while also providing button access to data logging and preliminary processing. While the audio player interface operates, the corresponding video data are streamed and recorded at the local PC machine at the same time.

Figure 3-9: Mattress topper for the SleepSmart sensor array (left), together with a waterproof-fitted layer (right).
Figure 3-10: Graphical user interface (GUI) of the audio player and the video stream previewing the study setup.
To interface with the accelerometers, the sensors were programmed on an Arduino integrated development environment (IDE), leveraging open-source sensor API libraries from Adafruit. Custom scripts were also written to perform other functions such as the construction of the serial communication data pipeline, initialization and error detection for the LIS3DH sensors. The serial output of the data (ASCII format) is logged into a text file with the corresponding timestamps via the PuTTY terminal, an open-source SSH and telnet client. The data are then parsed into MATLAB and Python for further analysis.

3.2.3 Video Device

Video recording is used as the ground truth in this experiment. A webcam (Logitech HD Webcam C270) was mounted approximately 2 meters above the mattress to record video of the whole bed. RGB images were recorded at a resolution of 640 x 480, with a rate of 30 frames per second. To allow for a quantitative measure of body movements using video, the subjects wore 8 bands of different colors on the head, arms, legs, feet, and torso. The actual movement intervals were estimated by tracking the trajectory of the color bands and further validated with manual, visual confirmation (see Section 3.4.1). The figure below shows an example of the camera’s view, and the worn cloth bands in the experiment.
3.3 Movement Classes Definition

Different movement descriptions have been previously adopted to analyze the distribution and type of movements during sleep. In this work, the movements will be classified into 3 domains: Domain A, Domain B, and Domain C.

3.3.1 Domain A

Adapting from movement descriptions in other work [12][103], movements were divided into 3 separate classes in this domain. Class 1 refers to large postural movements. A movement example of this class is shown in the figure below.

Figure 3-12: An example of a Class 1 movement (moving from right to left).
Class 2 refers to isolated movements of the head or upper limbs. A movement example for this class is shown in Figure 3.13.

![Figure 3-13: An example of a Class 2 movement (straightening left arm).](image)

Class 3 refers to leg movements. These leg movements can be related with periodic limb movements in sleep (PLMS) [12]. Two examples are illustrated in Figure 3.14.

![Figure 3-14: Two examples of Class 3 movements: flexing both ankles (top) and bending right leg (bottom).](image)
A summary of the movements and their corresponding classes is shown in Table 3.3.

Table 3-3: Summary of the movement descriptions and their corresponding classes in Domain A

<table>
<thead>
<tr>
<th>Class</th>
<th>Movement Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Large postural movements</td>
<td>Move from back to right  &lt;br&gt; Move from back to left  &lt;br&gt; Move from left to right &lt;br&gt; Move from right to left &lt;br&gt; Move from right to back &lt;br&gt; Move from left to back</td>
</tr>
<tr>
<td>2: Movements from the head and upper limbs</td>
<td>Straighten left arm &lt;br&gt; Bend left arm &lt;br&gt; Straighten right arm &lt;br&gt; Bend right arm &lt;br&gt; Turn head to the right &lt;br&gt; Turn head to the left</td>
</tr>
<tr>
<td>3: Leg movements</td>
<td>Bend left leg &lt;br&gt; Straighten left leg &lt;br&gt; Flex both ankles (while facing right) &lt;br&gt; Straighten both legs (while facing right) &lt;br&gt; Bend both legs (while facing right) &lt;br&gt; Bend right leg &lt;br&gt; Straighten right leg &lt;br&gt; Flex both ankles (while facing left) &lt;br&gt; Straighten both legs (while facing left) &lt;br&gt; Bend both legs (while facing left)</td>
</tr>
</tbody>
</table>

3.3.2 Domain B

The classification in Domain B labels movement events to specific body locations: head, torso and limbs, left arm, right arm, left leg, right leg, both legs, and both feet. Similar classification frameworks were utilized in other studies [72][103]. The classes and their corresponding movements are defined in Table 3.4.
### Table 3-4: Summary of the movement descriptions and their corresponding classes in Domain B

<table>
<thead>
<tr>
<th>Class Description</th>
<th>Movement Description</th>
</tr>
</thead>
</table>
| Class 1: Head     | Turn head to the right  
|                   | Turn head to the left |
| Class 2: Torso and limbs | Move from back to right  
|                   | Move from back to left  
|                   | Move from left to right  
|                   | Move from right to left  
|                   | Move from right to back  
|                   | Move from left to back  |
| Class 3: Left Arm | Straighten left arm  
|                   | Bend left arm |
| Class 4: Right arm | Straighten right arm  
|                   | Bend right arm |
| Class 5: Left leg | Bend left leg  
|                   | Straighten left leg |
| Class 6: Right leg | Bend right leg  
|                   | Straighten right leg |
| Class 7: Both legs | Straighten both legs (while facing right)  
|                   | Bend both legs (while facing right)  
|                   | Straighten both legs (while facing left)  
|                   | Bend both legs (while facing left) |
| Class 8: Both Feet | Flex both ankles (while facing left)  
|                   | Flex both ankles (while facing right) |

### 3.3.3 Domain C

The classification in Domain C equates classes to the movements’ descriptions (e.g., class 15 = flex both ankles). Some classes in this domain overlap with labels defined in other works [102][104]. In the movement set protocol, there are a total of 22 unique movements, which correspond to 22 labels as summarized in Table 3.5. While certain movement groups occasionally share a common movement action (such as bend both legs or flex both ankles), the subject’s postural state while performing the action is still being considered, therefore deeming that instance of movement event as a unique label in this work (bend both legs while facing right).
<table>
<thead>
<tr>
<th>Class</th>
<th>Movement Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Move from back to right</td>
</tr>
<tr>
<td>2</td>
<td>Move from back to left</td>
</tr>
<tr>
<td>3</td>
<td>Move from left to right</td>
</tr>
<tr>
<td>4</td>
<td>Move from right to left</td>
</tr>
<tr>
<td>5</td>
<td>Move from right to back</td>
</tr>
<tr>
<td>6</td>
<td>Move from left to back</td>
</tr>
<tr>
<td>7</td>
<td>Straighten left arm</td>
</tr>
<tr>
<td>8</td>
<td>Bend left arm</td>
</tr>
<tr>
<td>9</td>
<td>Straighten right arm</td>
</tr>
<tr>
<td>10</td>
<td>Bend right arm</td>
</tr>
<tr>
<td>11</td>
<td>Turn head to the right</td>
</tr>
<tr>
<td>12</td>
<td>Turn head to the left</td>
</tr>
<tr>
<td>13</td>
<td>Bend left leg</td>
</tr>
<tr>
<td>14</td>
<td>Straighten left leg</td>
</tr>
<tr>
<td>15</td>
<td>Flex both ankles (while facing right)</td>
</tr>
<tr>
<td>16</td>
<td>Straighten both legs (while facing right)</td>
</tr>
<tr>
<td>17</td>
<td>Bend both legs (while facing right)</td>
</tr>
<tr>
<td>18</td>
<td>Bend right leg</td>
</tr>
<tr>
<td>19</td>
<td>Straighten right leg</td>
</tr>
<tr>
<td>20</td>
<td>Flex both ankles (while facing left)</td>
</tr>
<tr>
<td>21</td>
<td>Straighten both legs (while facing left)</td>
</tr>
<tr>
<td>22</td>
<td>Bend both legs (while facing left)</td>
</tr>
</tbody>
</table>
3.4 Data Processing and Analysis

3.4.1 Data Segmentation

Each movement event performed in the protocol varies in terms of the time when the movement is initiated \( (t_{\text{begin}}) \), and the time when the movement has ended \( (t_{\text{end}}) \). The intervals for each movement need to be determined to acquire analysis windows for feature extraction.

As indicated in the study protocol, audio cues were provided to guide subjects when they perform the movements. These audio cues provide a coarse segmentation for each movement during a trial. The subjects were also asked to wear color bands on specific body locations. The positioning of the colored bands corresponds to respective parts of the body: the forehead, torso, wrist, lower legs, and feet. Using these markers and referring to techniques used in other studies [12][37], a procedure for segmentation using video is developed to quantitatively identify the performed movements. The intervals were estimated using color thresholding, which segments image RGB using the Image Processing Toolbox from MATLAB [106].

Prior to segmenting, the thresholder is constructed by pre-selecting the colors corresponding to each color bands as the thresholding factor. A coarse segmentation is provided by the timestamps generated by the audio cues, and the video frame data is then fed through the thresholder using the color space and range for each RGB channel corresponding to the specific band’s color. Each frame is divided into regions of interest prior to color thresholding (e.g. cropping the lower half of the frame to identify head movement events). The function then identifies location candidates in the video frame with similar color ranges and returns the segmented mask of these locations. The Euclidean distances were calculated between the location candidates over time, and the corresponding movement interval was estimated when the distance reaches a given threshold.
The result of the generated boundary is then visually reviewed and corrected. In many cases, tracking of the color bands was not feasible due to poor fitting of the band and discomfort. In these events, the intervals were acquired manually by analyzing the video recording frames for the movement event. Manual segmentation using video has been conducted in other works [12][72][107]. The corresponding frame intervals are then synchronized to the raw data from the mattress. Events from outside the audio movement intervals are excluded from this analysis.
Figure 3-16: Tri-axial acceleration data from the mattress sensors displayed with Domain A labels.

Figure 3-17: Tri-axial acceleration data from the mattress sensors displayed with Domain B labels.
3.4.2 Pre-Processing

The system recorded mattress sensor data and header data (such as sensor number, trial number, time stamps, and frame number) at an average sampling rate of 75 Hz. To obtain error-free data for the classification, all the data were exposed to a selection of pre-processing operations before classification. Custom MATLAB scripts were developed to read in all the data log files from a user-specified folder and then process the files one by one.

A script was developed to identify segments of the trial data and removed major outliers present in the acceleration data. Outliers are defined as data that are above or below ± 2 g thresholds. To address variation in the system’s sampling rate, a resampling operation was also carried out to balance the uniformity of the data. Using built-in MATLAB functions, the data were
resampled at 50 Hz to reduce memory and computational workload. The data were also passed through a 4th order, two-way, low-pass Butterworth filter, which is a commonly used filtering technique in studies related to human movements [50]. A cutoff frequency of 11 Hz was selected since the frequency content of the signal for the movement set performed is concentrated below 10 Hz as indicated in Figure 3.19. Similar studies assessing body movements have also reported cutoff values between 5-11 Hz [50][108].

![Figure 3-19: Frequency spectrum of the z-axis acceleration signal.](image)

### 3.4.3 Feature Extraction

Features were computed and extracted from the accelerometer data in consideration of the following parameters: movement duration, movement intensity, and movement spatial
displacements. The choice of features was motivated by evaluating the acceleration signal changes corresponding to the movement events. Similar feature definitions were also employed in other studies involving human motion [12][50][103].

Table 3-6: Features computed from the time-series acceleration data. The features are defined as $x^{(i)}$, ..., $x^{(n)}$, where $n$ represents the total number of features.

<table>
<thead>
<tr>
<th>Number of Feature</th>
<th>Feature Description</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-192</td>
<td>Variance of linear acceleration (x, y, z, vector magnitude) [109]</td>
<td>$f_1 = \sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$, where $\mu = \frac{\sum_{i=1}^{N}x_i}{N}$</td>
</tr>
<tr>
<td>193-384</td>
<td>Peak linear acceleration (x, y, z, vector magnitude) [110]</td>
<td>$f_2 = (\max(x_1, ..., x_i))^2$</td>
</tr>
<tr>
<td>385</td>
<td>Movement Duration [103]</td>
<td>$f_3 = \arg(t_{end}) - \arg(t_{initial})$</td>
</tr>
<tr>
<td>386 – 577</td>
<td>Instance of peak acceleration (x, y, z, vector magnitude)</td>
<td>$f_4 = \arg \max(f_2)$</td>
</tr>
</tbody>
</table>

3.4.4 Classifier Algorithms

The datasets collected in the study were limited in size (up to 330 movements per subject) and labelled with different classes similar to the type and number of movements that are typically observed during the night [54][111]. Given the higher dimensionality of the dataset, some considerations were made with regards to the choice of learning algorithm used. The questions to be answered are: 1) (O2) out of 4 of the selected machine learning algorithms, which is the most effective to map the relationship of the test observations to the target classes? 2) (O3) How much training data is needed to achieve good prediction accuracies in consideration with different classification domains?

To answer these questions, this study employed four learning algorithms: k-Nearest Neighbor (k-NN), Random Forests (RF), a Naive Bayes (NB) classifier, and Support Vector Machines (SVM) for this classification problem. These algorithms were used for data mining in medicine
and sleep studies [101][112][110][103][88][113], and have also excelled in other work that have utilized higher dimensional datasets [114][115].

Random forests, also known as random decision forests, is an ensemble method that can be used to perform classification and regression tasks [116]. Ensemble methods aggregates predictions by using multiple learning models to achieve better predictive results. Through the structure of uncorrelated decision trees, each tree considers a random selection of input features (called feature bagging) and finally reaches to the best possible decision via majority voting. For classification, the prediction representation for random forest can be defined as the following [116]:

\[
c_{rf}^N(X) = \text{majority vote} \{c_N(X)\}_1^N
\]  

In Equation 3.1, \(X\) is a new feature, \(N\) is the total number of trees in the forest, and \(c_{rf}^N(X)\) is the result of the class prediction based on the majority of votes from the trees.

![Figure 3-20: A simplified representation of the random forest method using majority voting to predict classes.](image-url)
k-NN is an instance-based learning algorithm that uses the principle of close proximities or similarities between data observations in the feature space as the basis for classification [117]. The proximity or the similarity between points is determined using a metric. A few examples of the many different metrics are the Euclidean distance or the Cosine similarity, which is a measure that computes the cosine of the angle between two vectors, as represented in Figure 3.21 and Equation 3.2.

\[
Simularity = \cos(\theta) = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| \cdot |\vec{b}|}
\]  

(3.2)

In Equation 3.3 [118], the k-NN classifier first identifies instances of data points (up to a positive number \(K\)) that are closest or more similar to the test instance \(x_0\) using the selected metric. Subsequently, the class, \(j\) corresponding to each labelled point were summed and represented as a fraction of the total number of points. The final prediction for the class of the test instance were made by considering the class with the largest probability.
\[ c(x_0) = \arg\max_c P(c \mid x_0) \]
\[ = \arg\max_c \left( \frac{1}{K} \sum_{i \in n_0} I(j_i = c) \right) \] (3.3)

Support Vector Machines (SVM), is a classifier defined by a separating hyperplane [117][118]. For example, the hyperplane is a line for a linearly-separable, two-dimensional dataset. The algorithm determines an optimal hyperplane using the labelled training data, and subsequently classifies new test data by considering which side of the hyperplane a test instance lies in. The classifier requires solving for the following optimization problem [116]:

\[
\max_{\alpha} \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} c^{(i)} c^{(j)} \alpha_i \alpha_j K(x_i, x_j) \\
\text{s.t. } \alpha_i \geq 0, \ i = 1, \ldots, m \\
\sum_{i=1}^{m} \alpha_i c^{(i)} = 0 
\] (3.4) (3.5) (3.6)

Once the parameters for \( \alpha \) and \( \beta \) are estimated using the training data and the class labels, the decision function represented in Equation 3.7 [118] can be employed in order to classify the test observation. Equation 3.7 calculates a quantity that depends on the inner product between \( x \) and the points \( x_i \) in the training set:

\[ g(x) = \beta_0 + \sum_{i \in S} \alpha_i K(x, x_i) \] (3.7)
For non-linear class boundaries, the data can be mapped into a higher-dimension feature space using a kernel approach. Subsequently, the optimal hyperplane can be determined and constructed in this space [118]. Examples of common kernel functions are represented in the equations below [118]. Instead of the standard linear kernel (Equation 3.8), using the polynomial kernel function (Equation 3.9) of degree $d$, where $d$ is a positive integer, would yield a much more flexible decision boundary for the classifier.

\[
K(x_i, x_i') = \langle x^{(i)}, x^{(j)} \rangle
\]  

(3.8)

\[
K(x_i, x_i') = (1 + \langle x^{(i)}, x^{(j)} \rangle)^d
\]  

(3.9)

Naïve Bayes (NB) is a classification technique based on the Bayes’ Theorem that assumes the predictors are independent between each other. In other words, the Naïve Bayes classifier assumes that given a class $C_i$, the features $X_i$ are independent [117]. The Bayes theorem can be represented as:

Figure 3-22: Simplified representation of a linear (left) and a non-linear decision boundary (right) using different kernel functions.
\[ P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (3.10) \]

In Equation 3.10, \( P(C) \) is the prior probability of the class, \( P(X) \) is the prior probability of the feature, \( X \). \( P(X|C) \) is the likelihood of the feature, \( X \) occurring given class, \( C \) and \( P(C|X) \) is the posterior probability of class, \( C \) occurring. The decision function for a feature vector \( X \) can be represented in Equation 3.11 [119].

\[ c(X) = \arg\max_c P(c) \prod_{i=1}^{n} P(x_i|c) \quad (3.11) \]

When continuous features are used, the conditional probabilities \( P(x_i|c) \) can be computed by using probability distribution methods, such as the Gaussian kernel density estimation (Equation 3.12 and 3.13), which has a similar representation to the Gaussian distribution, except that the estimated density is averaged over a number of kernels [120]. In the equations below, \( n \) is the sample size, \( K \) is the kernel function, and \( h \) is the bandwidth, also called the smoothing parameter.

\[ P_{kde}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{r}{x-x_i}\right) \quad (3.12) \]

\[ K(r) = \frac{1}{\sqrt{2\pi}} e^{-\frac{r^2}{2h^2}} \quad (3.13) \]
3.5 Summary

In this chapter, the research objectives and hypotheses explored in this work were detailed. The study procedures and movement protocols were described in Section 3.1.3. In Section 3.2, the materials and sensors used in this study were described. The mattress-based device was introduced in Section 3.2.1 (O1). The classification domains (Domain A, Domain B, Domain C) were defined in Section 3.3. Section 3.4.1 outlined the techniques used for data segmentation using video, while Section 3.4.2 and 3.4.3 describes the post-processing steps applied to the mattress sensor data. Section 3.4.4 provides an overview of the algorithms used in this study (RF, k-NN, SVM, NB).
Chapter 4: Results

This chapter starts by providing a description of the number of samples for each class in the 3 classification domains. Additionally, two variants of the cross-validation method were described when dealing with subject-dependent and subject-independent model assessments. The hyperparameters associated with each learning algorithm and the model performance measure were outlined in this section as well. The classification accuracies for each domain are detailed in Section 4.1, Section 4.2, and Section 4.3.

The main goal of this work is to classify movements into three different domains. The main features of the domains (earlier described in Chapter 3.3) are summarized below:

- Domain A: major postural movements, isolated movements, and leg movements, corresponding to a total of 3 classes.
- Domain B: movements of specific body areas: the head, torso and limbs, left arm, right arm, left leg, right leg, both legs, and both feet. This corresponds to a total of 8 classes in this domain.
- Domain C: unique movement descriptions, which equates to a total of 22 classes.

The total number of movements in each domain is provided in Table 4.1. As erroneous movements could occur during the audio cue intervals, there are minor differences in the number of movements per class between subjects due to the exclusion of such movement events. In this study, each participant performed the movement protocol described in Section 3.3 for 15 iterations. The data corresponding to each trial were segmented and arranged according to the labels defined in the respective domains.
Table 4-1: Number of movement samples in each class for the domains.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Class</th>
<th>Total movements per subject for each class</th>
<th>Total movements per trial for each class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain A</td>
<td>Class 1, 2, 3</td>
<td>90</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Class 4, 5, 6, 7</td>
<td>150</td>
<td>10</td>
</tr>
<tr>
<td>Domain B</td>
<td>Class 1, 2, 3, 4, 5, 6, 7</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Class 8</td>
<td>90</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Class 7</td>
<td>60</td>
<td>4</td>
</tr>
<tr>
<td>Domain C</td>
<td>Class 1-22</td>
<td>15</td>
<td>1</td>
</tr>
</tbody>
</table>

This study investigates the prediction accuracies of the learning algorithms using two cross-validation approaches: a) cross-trial validation for subject-specific model assessments, and b) cross-subject validation (leave-one-out) for subject-independent model assessments.

![Figure 4-1: Illustration of the cross-trial validation approach. Data from 1 trial were held out as a testing set while using the remaining data from 14 trials as the training set.](image)

In subject-dependent models, cross-trial validation was employed in all classification domains to evaluate the performance of the learning algorithms and to investigate the data consistency between the 15 trials. The procedure of this approach was illustrated in Figure 4.1. For each subject, the data was segmented into 15 trials. Data from one trial were used as the testing data,
while the remaining 14 trial datasets were then used for training. This continues for a total of 15 iterations, until every other trial was held out and used as a test set.

![Figure 4-2: Illustration of the cross-subject validation approach. Data from 1 subject was held out as a testing set while data from the remaining 9 subjects were used as the training set.](image)

To estimate the generalizability of the learning algorithms in a subject-independent approach, cross-subject validation [121] was implemented for all domains using the subject-wise segmented data. As illustrated in Figure 4.2, in one iteration, data from one subject were held out as a test set, while the remaining data from nine other subjects were used for training. Similarly, this continues for a total of 10 iterations until every other subject dataset was used as a test set.

The classifiers used in this study were the RF, k-NN, SVM, and NB algorithms. For the Random Forest classifier, 200 trees were used in the forest. To train the k-Nearest Neighbor learning model, k = 10 neighbors with the cosine similarity metric were used. Support Vector Machines with quadratic kernels were used using a one-vs-one multi-class configuration. For the Naïve Bayes classifier, a Gaussian kernel density estimation was utilized. The hyperparameters selected for the algorithms were referenced from other studies [101][122][120][110], and the values were experimented using a pilot data set (different from the data utilized in this study).
To compare the overall effectiveness of the learning models, the classification accuracies were computed using the proportion of the test samples, $s_c$, that are correctly classified over all samples, $S$ [12][123]. The model accuracies reported in the subsequent sections were the average of the accuracies reported from all iterations in both of the cross-validation approaches (Equation 3.14) [118].

\[ A = \frac{s_c}{S} \]

\[ A_{ave} = CV(n) = \frac{1}{n} \sum_{i=1}^{n} A_i \]  

(3.14)

### 4.1 Domain A Classification

The classes in Domain A consisted of a tri-class system: class 1, which accounts for major postural movements, class 2 for isolated movements, and class 3 for leg movements. The prediction accuracies for Domain A for both subject-dependent and subject-independent models are presented in Figure 4.3.
4.2 Domain B Classification

There were 8 classes in Domain B, corresponding to movements performed at a specific body location: head, torso and limbs, left arm, right arm, left leg, right leg, both legs, and both feet. The
classification accuracies for Domain B for both subject-dependent and subject-independent models are presented in Figure 4.4.

![Bar chart showing classification accuracies for Domain B.]

**Figure 4-4:** Average classification accuracies and minimum-maximum accuracy ranges for subject-dependent and subject-independent models for Domain B.

For subject-dependent models, the highest classification accuracy was acquired using Random Forest (98.00%), followed by SVM (95.58%), NB (92.07%), and the k-NN algorithm (85.77%). For subject-independent models, similarly, the highest classification accuracy was obtained using RF (92.02%), followed by SVM (90.66%), NB (77.51%), and k-NN (72.18%). Like Domain A, the best prediction accuracies were attained when using the RF algorithm, while the k-NN algorithm showed the lowest classification accuracies in both subject-dependent and subject-independent models.
4.3 Domain C Classification

The classes in Domain C consist of 22 labels that correspond to the description of the movements. For instance, *bend left leg*, and *straighten left leg* are considered as two distinct classes. The list and details of the movement set are described in Section 3.2.

Each trial contains only one sample from each of the classes (from Class 1 to Class 22). This amounts to a total of 15 samples from each class from each subject. The prediction accuracies for Domain C for both subject-dependent and subject-independent models are presented in Figure 4.5.

![Figure 4-5: Average classification accuracies and minimum-maximum ranges for subject-dependent and subject-independent models for Domain A.](image)

For subject-dependent models, the highest classification accuracy was attained using Random Forest (94.14%), followed by SVM (88.53%), NB (80.48%), and k-NN (77.44%). For subject-
independent models, the highest classification accuracies are as follows: SVM (81.63%), Random Forest (80.7%), NB (55.49%), and k-NN (54.89%).

In this domain, SVM outperforms the RF algorithm in the subject-independent models by a slight margin. For subject-dependent models, the best prediction accuracies were achieved when using the RF algorithm. The k-NN algorithm showed the lowest classification accuracies in both subject-dependent and subject-independent models.

4.4 Summary

In this chapter, we summarized the number of movements performed in each class for the domains. Using the movement observation samples, we presented two variants of the cross-validation method to assess the learning algorithms with subject-dependent and subject-independent considerations. We also outlined the hyperparameters used for each classifier algorithm for the dataset and presented the classification accuracies for the learning models in all domains. Table 4.2 summarizes the classification accuracies of the learning models in all domains.

| Classifier | Domain A | | Domain B | | Domain C |
|------------|----------|----------|----------|----------|
|            | Subject Dependent | Subject Independent | Subject Dependent | Subject Independent | Subject Dependent | Subject Independent |
| RF         | 98.59%    | 96.89%    | 98.00%    | 92.02%    | 94.14%    | 80.70%    |
| SVM        | 98.19%    | 96.05%    | 95.58%    | 90.66%    | 88.53%    | 81.63%    |
| NB         | 94.19%    | 88.85%    | 92.07%    | 77.51%    | 80.48%    | 55.49%    |
| k-NN       | 88.42%    | 80.56%    | 85.77%    | 72.18%    | 77.44%    | 54.89%    |

From the table above, the classification accuracies for all algorithms exhibit a decreasing trend when going from Domain A to Domain C for subject-dependent models. This trend can also be seen when comparing subject-dependent to subject-independent models. In all domains and model variants, RF demonstrated the highest prediction accuracies (except for Domain C where SVM
performed slightly better in the subject-independent model), and k-NN exhibits lower accuracies in all models.
Chapter 5: Discussion

This chapter aims to provide additional observations on the work reported in previous chapters. Section 5.1 summarizes the performance accuracies in Chapter 4 and provides complementary insights to the reported metrics using other performance measures: precision, recall, and the confusion matrices. Confusion matrix is a classification table consisting of counts from the true positives, true negatives, false positives, and the false negatives. Using these observations, precision, recall, and specificity can be defined by the following equations [123]. These measures were commonly used to evaluate performance of the classifier models in other studies [103][110].

\[
\text{Precision} = \frac{TP}{TP + FP} \quad (5.1)
\]

\[
\text{Recall (Sensitivity)} = \frac{TP}{TP + FN} \quad (5.2)
\]

\[
\text{Specificity} = \frac{TN}{FP + TN} \quad (5.3)
\]

In Section 5.2, the research objective O3 was explored. The study investigated the impact of different training set sizes on model performance in subject-dependent and -independent approaches. Section 5.3 identifies the limitations associated with the work described related to this thesis.

5.1 Comparison of Techniques Used

To evaluate the trained models, variants of the cross-validation method were used and described in Chapter 4. Cross-validation is an efficient sample re-use method to estimate prediction error and
accuracies [116][121]. This technique is also widely used in other studies in the sleep literature [12][101][124].

As seen in Figure 5.1, by using conventional “hold-out” subsets for training and testing, the model accuracies can be highly variable, depending on precisely which observations are included in the training set, and which observations are included in the testing set [118][125]. To study the generalizability of the system to new input data, this study has employed cross-trial validation for subject-dependent models and cross-subject validation for subject-independent models. The procedures of these approaches were described in Chapter 4.

In this study, the dataset used exhibited a low ratio of number of observations to the number of features. For the classification problem, four classifiers were chosen: RF, SVM, NB and k-NN. A ranking of the learning models in this application could be inferred from the classification accuracies reported in Chapter 4. RF has the best classification accuracies with an accuracy value of 93.39% averaged across all domains and approaches. Its performance is followed by SVM, with
an averaged accuracy value of 91.77%, while NB and k-NN have averaged values of 81.43% and 76.54%, respectively.

The RF algorithm, an extension of the decision tree ensembles (via bagging), is suitable for use in high-dimensional datasets [122][126], and it is relatively insensitive to outliers and noise [127]. Random Forests also uses a technique called feature bagging to reduce model variance and decorrelate the decision trees [127], ranks the importance of predictors [115], and eliminates unnecessary features during training [101]. Similarly, SVM proved to be effective in classifying high-dimensional data and continuous features, although a larger sample set might be needed to reach the maximum prediction accuracy [117][114], which can be inferred in Figure 5.2. The figure shows the accuracies of the RF and SVM model under varying training set sizes. The procedure for the evaluation of the model performances using different training sizes was detailed in Section 5.2. Overall, since RF and SVM are the better performing algorithms, we speculate that higher variance algorithms such as decision tree methods and SVMs, which could generate more complex models to better approximate data variations [117] would be best suited for our classification function.

Figure 5-2: The differences in performance between RF and SVM models under different training set sizes.
NB algorithms operate under a naïve assumption that the feature sets are independent and can be represented by a single probability distribution [117]. As a high-bias classifier, it is prone to under-fitting [117], and this characteristic can be observed through a relatively bigger decline in accuracies in the NB models in domains with higher number of classes when compared to the RF and SVM models. Under-fitting occurs when a model is not complex enough to capture the patterns of the data [116]. Having features that are correlated could also affect performance as this violates the assumption of Naïve Bayes [117]. Moreover, decision trees and the NB classifier generally work in opposite profiles. When one algorithm is more accurate, the other is not [117]. This is consistent with the differences in performance measures with the RF algorithm. The performance of the k-NN algorithm can be affected by irrelevant features [117]. In higher-dimensional feature spaces, a decline in the performance of local methods (such as the k-NN algorithm) can be expected since they perform predictions using the concept of close proximities. In richer feature spaces, the labelled points are spread out further, and this results in a phenomenon that the given test instance has no nearby neighbor. This is also known as the curse of dimensionality [118].

Based on the model performances reported in Chapter 4 and the rationales discussed here, subsequent evaluations and analyses in this chapter will be performed with RF models.
Figure 5-3: Confusion matrices for Domain A (top) and Domain B (bottom) in cross-trial approaches.
Figure 5.3 illustrates the confusion matrices for the aggregated predicted samples in Domain A (left) and Domain B (right) from the cross-trial approach described in Section 4. For Domain A, we see that the most frequent prediction error for the RF classifier occurs in Class 2. In other words, the classifier mistakenly identified isolated movements (movements occurring at the head or the limbs) as leg movements. One possible reason for this confusion was that certain limb movements (such as *turn head to the right*) and certain leg movements (such as *flex both ankles*) have similar intensity levels. Nevertheless, the recall for Class 2 remained high (95%), while for Class 1 and Class 3, at least 99% recall was observed for both classes. The overall recall and precision for Domain A were 98.16% and 98.66%, respectively.

For Domain B, the lowest recalls were observed for both Class 5 and Class 6 (93.6%). The classifier mistakenly identified individual movements occurring at the left leg or right leg with movements occurring with both legs. Again, the characteristics of the features could be difficult to distinguish as both movements occur at similar spaces on the mattress. The overall recall and precision for Domain B were 96.94% and 97.77%, respectively.
Figure 5-4: Confusion matrix for Domain C in the cross-trial approach.
The confusion matrix for Domain C is shown in Figure 5.4. Class 7 and Class 9 showed relatively low recall rates (87.7% and 83.8%, respectively). In both classes, the confusions were similar, in that the *straighten arm* movements were confused with their opposing counterparts (*bending arms*). While these movements were visually noticeable (larger movements), we theorized that the corresponding signal profile might not be fully picked up by the sensor as the body segments were not directly in contact with the sensors. For this domain, the overall recall and precision of the classification were 93.76% and 93.79%, respectively.
Figure 5-5: Confusion matrices for Domain A (top) and Domain B (bottom) in cross-subject approaches.
Figure 5.5 illustrates the confusion matrices for the predicted samples in Domain A (left) and Domain B (right) for the cross-subject approach in Section 4. For both domains, we see similar confusion trends (Class 2 in Domain A, and Class 5 and 6 in Domain B) with the cross-trial approach. The overall recall and precision for Domain A were 96.83% and 97.31%, respectively, while for Domain B, the values were 90.18% for recall and 91.56% for precision.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1-movement</td>
<td>86.0%</td>
<td>13.1%</td>
<td>8.7%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>2-movement</td>
<td>12.6%</td>
<td>7.3%</td>
<td>9.1%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>3-movement</td>
<td>2.2%</td>
<td>7.3%</td>
<td>9.1%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>4-movement</td>
<td>1.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>5-movement</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

**Figure 5-6:** Confusion matrices for Domain C for the cross-subject approach.
For Domain C, the overall recall and precision were 80.69% and 80.85%. In Figure 5.6, recall values lower than 80% were: Class 5, Class 6, Class 7, Class 9, Class 12, Class 14, Class 15, and Class 19. Similar to the cross-trial results for Domain C, the movements were mostly confused with the opposing actions (bend right arm and straighten right arm). Binning the movements into a common class (such as bending or straightening right arm) would yield better accuracy measures. However, further experimentation with new features that primarily address the sequence of the movement events would be needed to perform finer-grained differentiation of the movements. Table 5.1 summarizes the precision and recall values for the domains in both model approaches. Due to the random selection implementation in the RF algorithm, minor variations of the accuracies of the RF were observed. As differences were insignificant, accuracy values reported below were taken from Table 4.2 in Section 4.4.

| Table 5-1: Performance measures for the RF model in all domains. |
|---------------------------------|-----------------|-----------------|-----------------|
|                                 | Domain A        | Domain B        | Domain C        |
|                                 | Subject Dependent | Subject Dependent | Subject Dependent | Subject Independent | Subject Independent | Subject Independent |
| Accuracy                        | 98.59%          | 96.90%          | 98.00%          | 92.02%          | 94.14%          | 80.70%          |
| Recall                          | 98.16%          | 96.83%          | 96.94%          | 90.18%          | 93.76%          | 80.69%          |
| Precision                       | 98.66%          | 97.31%          | 97.77%          | 91.56%          | 93.79%          | 80.85%          |

According to the confusion matrices and the additional performance measures reported, we deduced that the type of data, whether collected from the same subject or from a different subject, has little effect on the classification accuracies in Domain A (98.38% vs. 97.04%) and Domain B (97.62% vs. 91.96%). However, for Domain C, a noticeable drop in model accuracy (93.77% vs. 80.72%) was exhibited when data from other subjects were used for training. The decline in performance suggests that using subject-dependent data, or a mixture of data from both sources,
would be ideal to attain comparable accuracies in Domain C. As the study has only acquired data from 10 subjects, more training data could be favorable for the subject-independent approach. The effect of varying training sizes will be investigated in the next section. Class-wide and class-specific recall, precision, and specificity for both approaches were listed in Appendix B.

### 5.2 Effect of Training Set Sizes

In the previous chapter, we assessed the model performances when considering subject-dependent and subject-independent approaches. We have consistently observed better prediction accuracies in subject-dependent models when compared to subject-independent ones. In real world training sets, the availability of datasets could be scarce and must be built progressively [114]. It would be worthwhile to gauge the flexibility of the learning models when using training data from the same subject, or from other subjects to build up the classifier. With this information, we seek to investigate the effect of training set size on the classification accuracy for the RF model using: a) data from the same subject, and b) data from other subjects.

![Figure 5-7: Iterative removal of trial data from the available training set in subject-dependent datasets.](image-url)
To evaluate the performance of the algorithms in subject-dependent models, we first split the data using 3/5 as a training set, and 2/5 as a testing set in the first iteration [12]. Then, training data were removed one trial at a time until data amounting to one trial remained in the training set. The classifier was retrained in every iteration. This procedure (illustrated in Figure 5.7) was repeated until the testing was performed for all subjects. For each iteration, the average prediction accuracy was averaged among all subjects and illustrated in the figure below.

Figure 5-8: The corresponding model accuracies when the training set sizes were varied in each domain for subject-dependent datasets.
A baseline accuracy of 84% was established for evaluation. This value is based on metrics reported in other studies involving classification work on voluntary or abnormal nocturnal body movements [109][110]. Figure 5.8 exhibits the corresponding classification accuracies for the domains when the sizes of the training samples were varied. Domain A consists of the tri-class system (described in Section 3.3). In Domain A, we see an accuracy difference of 3% when we increased the training samples used to 44 for the domains. This is not considered to be a major accuracy jump. Even when using 22 samples (which is approximately equivalent to 8 training samples per class), the model was able to acquire an accuracy of 94%. From the results shown, we can deduce that by using 22 training samples (~8 training samples per class) from the same subject, high accuracies can be attained for Domain A.

Domain B consists of an 8-class system defining movements at specific body locations. In this domain (center figure), we see a big jump in accuracy (from 84% to 95%) when we increase the training samples from 22 to 66 (roughly equivalent to an increase from 3 to 9 training samples per class). While the starting accuracy (84%) is lower compared to Domain A, the number of classes in Domain B is larger. With only 22 training samples (approximately 3 training samples per class), the model accuracy is comparable with the baseline performance (84%) previously defined.

Domain C consists of a 22-class system defining the movement descriptions. In this domain (right figure), the base accuracy starts at a low measure (57%), but this is expected due to the relatively larger number of classes present in this domain. An increase of 30% in accuracy (from 57% to 87%) can be observed when the training samples used for this domain is increased from 22 to 110, which is equivalent to the increase from 1 to 5 training samples per class. We conclude
that with at least 88 training samples (4 samples per class) for this domain, we expect to achieve classification accuracies higher (87%) than the baseline defined.

To evaluate the effect of training set size in subject-independent models, we first split the dataset using 9/10 as a training set and 1/10 as a testing set in the first iteration. Training data were removed one subject at a time, until data amounting to one subject remained in the training set. The classifier was retrained at every iteration, and the procedure (illustrated in Figure 5.9) was repeated for all 10 subjects. For each iteration of training data removal, the average prediction accuracy was averaged for all subjects and illustrated in the figure below. Compared to the number of training samples in the subject-dependent approach, the training datasets were significantly larger since all trial data from one subject (15 trials per subject) were being used to train the classifier. This corresponds to a total of ~330 observations per subject.
Figure 5.10 demonstrates the corresponding classification accuracies for the domains when the sizes of the training samples were varied. For Domain A (left figure), the learning model attained good accuracy (93%) when using data acquired from one subject. The subsequent increments in accuracy were not large when more data from additional subjects were used for training (from 93% to 97%).

For Domain B (center figure), the classifier was able to achieve accuracy comparable to the baseline performance (84%) when data from 2 subjects were used to train the classifier. We can
also observe that the accuracy steadily increased when data from more subjects were used to train the classifier (from 84% to 92%). The increments were not significant when training data from more than 5 subjects were used.

For Domain C (right figure), low base accuracy was observed (54%) when data from one subject were used to train the classifier. The classifier achieved an accuracy of 81% when data from all 9 subjects were used to train the classifier (2970 total training samples and 135 training samples per class). This is lower than the baseline accuracy. More subject data would be needed to evaluate if the RF model would perform better given the extra training sets. In Figure 5.11, we further partitioned the training set (subject-independent) into 15 trials, using data from the remaining subject.
**Figure 5-11**: Remaining subject data further partitioned into 15 trials to train the classifier.

In the left figure, we see that even when an average of 8 sample per class from different subject were used for training, baseline accuracy was able to be achieved for Domain A. For Domain B and Domain C, the classification accuracy decreases as a smaller portion of the data was allocated for training. Table 5.2 summarizes the minimum number of training samples per class required to achieve values higher or equal to the baseline accuracy defined (84%).
Table 5-2: Minimum training samples per class to attain higher than baseline accuracy for all domains.

<table>
<thead>
<tr>
<th></th>
<th>Domain A</th>
<th></th>
<th>Domain B</th>
<th></th>
<th>Domain C</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subject</td>
<td>Subject</td>
<td>Subject</td>
<td>Subject</td>
<td>Subject</td>
<td>Subject</td>
</tr>
<tr>
<td>Training Data</td>
<td>1 trial</td>
<td>1 subject</td>
<td>1 trial</td>
<td>2 subjects</td>
<td>4 trials</td>
<td>9 subjects</td>
</tr>
<tr>
<td>Accuracy</td>
<td>94%</td>
<td>88%</td>
<td>84%</td>
<td>84%</td>
<td>87%</td>
<td>81%</td>
</tr>
</tbody>
</table>

5.3 Comparison to Existing Studies

This study has employed machine learning techniques and classified movement events into 3 domains with varying label granularities using a mattress-based approach. While optimal performances (at least 91% classification accuracy) have been demonstrated when using non-contact devices, a majority have only categorized sleep positions [74][73][99]. The positions were categorized into classes mainly consisting of static configurations of poses (facing left, right, or supine with legs bent). One study has demonstrated good performances when classifying movements (up to 98% precision and recall) [72], but the work has only categorized movements into very specific classes (such as leg movements). In our work, this resembled one of the classes in Domain A (Class 3 – leg movements). In this class, our device shown an average precision value of 99.3%. This is comparable to the performance of one other study utilizing non-contact devices (98%). In addition, we demonstrated capabilities of further segmenting these movements into Domain B (left leg, right leg, both legs), and Domain C (bend left leg, straighten left leg, etc.). Suboptimal performances when line of sight is obstructed were also reported [72][74][99], and in some cases, these devices required elaborated or sophisticated setups that would be inappropriate to manage in a home-based setting [73][72]. Nevertheless, non-contact devices have several advantages over mattress-based methods. They are not in contact with the sleeper, easier to interpret (video footage), and in some cases might be cheaper in terms of costs to implement.
A few mattress-based studies have demonstrated success in classification with granular labels. In this work, Domain A was adapted from a classification framework in one other study [12]. They have attained classification accuracies of 80.3% for subject-independent models and 85.8% for subject-dependent models. In our work, we have attained 96.90% and 98.59% accuracies for subject independent and subject dependent models, using the similar classification framework (Domain A). In the context of subject-dependent models, they indicated a minimum training size of 10 samples per class to attain good accuracies (81.5%), while we have shown that an average of 8 training samples per class were needed to attain a 94% classification accuracy in a similar classification domain. Furthermore, we have shown comparable performance (88%) when using the subject-independent approach (using data from other subjects to train the test subject’s data). While we were able to achieve better performances in our work, there were differences in terms of the study protocol, algorithm, data preparation and device design. For instance, they have utilized a clustering algorithm (Gaussian Mixture Models) for classification, and deployed load cells under bed posts for measurements [12]. Our study, on the other hand, has included additional movements (such as flexion of ankles), considered only predefined movements, collected less data, and employed variants of cross-validations when building the classifier.

Other studies that have utilized load cells have classified movements into 6 [102] and 9 classes [103]. One study has classified postural movements (such as move from right to supine) with an accuracy of 79.7%. They have also classified stationary positions and incorporated sitting states (move from supine to sit) [102]. In our work, we have demonstrated an accuracy of 99.7% using similar movement sets and class definitions (Domain A: Class 1 – major postural movements). In addition to major postural movements, we have also incorporated various type of movements (up to 22) in our protocol and extended the classification work for these motor events. Another study
has classified 35 different types of movements into nine classes and have achieved an average accuracy of 90% [103]. Their movements consisted of more specific movements (*putting hands on chest*), and they have employed a multi-level binary decision tree that implements SVM model in each node for classification [103]. The classes defined in their study have some overlap with the Domain B definitions in our study, where an average accuracy of 98% was achieved in our work in comparison. We have also managed to accomplish higher recall and precision for class-specific measures, and also introduced a new domain for classification in terms of unique descriptors. One other study has used a 6-axis inertial measurement unit [104] to classify patient movement events. Most of their movement sets consisted of awake activities (such as *enter bed, lie down, sit up*). In their study, they have demonstrated a 90% accuracy using a 12-class system, but demonstrated suboptimal performances when distinguishing more granular movements, such as *roll left, roll right*, with precision and recall ranges from 40% to 58% [104]. In their study, they have also investigated the effect of training set size in the context of 50 patients and utilized mostly a subject-independent approach when building the classifier. They have reported a limitation with previous studies (with the exception of [12]), in regards to using certain data resampling methods for training (such as using a subset of data from all subject for training and testing), questioning the degree of effectiveness in their approaches.

### 5.4 Summary

In our study, we have compared and underlined the performances between cross-trial and cross-subject approaches. While the procedure for cross-trial in our work has resemblance to the limitation described in the study above (e.g. splitting data from the same subject to be used for training and testing), we demonstrated that both approaches (cross-trial and cross-subject
 validations) have comparable accuracies. We also showed that different approaches have different minimum training size requirements to reach comparable accuracies (Section 5.2). For cross-subject analysis, none of the data from the test subjects were used to train the classifier. The purpose for this is that in the event of data scarcity [12], we wanted to highlight the flexibility of potentially using data from the same subject (cross-trial), and/or data from other subject (cross-subject) to attain classification accuracies that are better, or still comparable to one another.

Overall, we have demonstrated better and more granular classification capabilities (H1, H2, H3) without the need for extensive data processing, when using the SleepSmart system (O1). We have also investigated the effect of training set sizes in subject-independent and subject-dependent approaches (O3), and demonstrated smaller sample requirements for training in both approaches when compared to another study [12]. The main reason for this improvement could be due to the greater spatial sensing resolution from the SleepSmart sensor array. More sensors were used in our work when compared to other mattress-based studies (4 load cells). Based on these results, we reason that using a mattress-based array approach would yield better effectiveness in movement-based assessments and classifications, when compared to other device types (non-contact devices, load cells). Nevertheless, there are some drawbacks with mattress array implementations, the most prominent one being the relatively higher costs of such devices when compared to other research systems. In summary, we would like to point out that all methods have their advantages and disadvantages [128] and there is not one encompassing solution for all problems in the context of sleep. The specific application should be carefully considered, and device synergies [128] should be explored to enable for versatile and more robust solutions towards sleep assessments.
5.5 Limitations

As indicated in Section 4, there were some imbalances in the classes for Domain A and Domain B. For instance, in Domain A, Class 1 has a total sample size of 90 and Class 3 has a total sample size of 150. This study has used accuracy as the main metric for comparison between learning algorithms. Although the degree of imbalance is smaller compared to another study [109], the skewed class proportion might lead to misleading comparisons of overall effectiveness for the algorithms, especially when an overall metric such as accuracy is used [123].

Additionally, no feature selection methods were performed for the features computed in Section 3. While the dimension of the feature space is much lower compared to other studies [114], RF might still have an unfair advantage as the algorithm inherently performs feature selection [101]. Preliminary evaluations of the RF models have estimated that variance, $f_1$ and peak linear acceleration, $f_2$ were more important predictors to the models. A good extension of this work would be to re-evaluate the performance of NB, k-NN, SVM, and RF when feature selection or dimensionality reduction techniques such as Principal Component Analysis (PCA) or Forward Feature Selection are utilized.

This study extracted features while leveraging true movement interval information that was determined during data segmentation (see Section 3.4.1). In our work, an algorithm for the estimation of movement intervals was not proposed, and this resulted in analysis windows that were minimally affected by errors in our studies. Other studies have employed the use of estimated [102][103] and a combination of true and estimated interval windows when developing their classifier [12]. While one of the study have reported statistically insignificant classifier performance differences when using true intervals for the analyses windows [12], it would be
important to evaluate potential effects on performance when using true and estimated movement intervals for our system.

The study did not investigate whether external factors such as bed size and type would affect the performance of the system. This study also did not account for individualistic factors such as weight and height. This generalized assumption of subject-wide differences can affect the evaluation for the effect of training set size in the subject-independent model approach. While other studies [109][90] have demonstrated that such factors do not have an effect on their model performances, it would be important to evaluate this effect on a mattress-based device. Our findings in this study are based on experimental data collected from healthy subjects under lab-controlled conditions, where subjects performed movements that were designed to simulate common movements during sleep. As a result, the dataset collected were relatively “noise-free” since the simulated motions were discrete (movements performed in fixed intervals). To investigate if these results would hold in realistic settings, conducting studies in sleep environments and conditions would be required.
Chapter 6: Conclusion

In this thesis, we presented a mattress-based sensor device and set of algorithms for the classification of body movements in bed using an accelerometer array. In this work, we have demonstrated the performance of SleepSmart in characterizing movements into three classification domains: (H1) Domain A, the 3-class system defining major postural movements, isolated movements, and leg movements, (H2) Domain B, an 8-class system that corresponds to movements at the head, torso and limbs, left arm, right arm, left leg, right leg, both legs, and both feet, (H3) Domain C, a 22-class system that describes individual movement events. The modelling performances were compared (O2) and the corresponding effect of different training set sizes (O3) was investigated in the best performing algorithms.

In Chapter 3, we detailed the protocol for data collection. In the protocol, subjects were asked to perform a predefined movement set that are common during sleep. We also introduced the study setup which includes the SleepSmart mattress sensors (O1), video devices, and recording applications. We established definitions for the domains and classes. Techniques used for video data segmentation and data preprocessing were outlined. We also defined the features extracted from the processed data while also providing some intuitions of the learning algorithms used in this study.

In Chapter 4, variants of the cross-validation technique for subject-dependent and independent models were described, and the hyperparameters for the learning algorithm were defined. The prediction accuracies for all domains were detailed in this chapter. We have determined that: (O2) the best performing models in our study were acquired using the RF algorithm. With the RF models, we have demonstrated optimal classification accuracies (summarized in Section 4.4) in all classification domains (H1, H2, H3) for both subject-dependent and subject-independent
approaches, with one exception. The RF model had lower than baseline accuracy (80.70% compared to 84%) in Domain C when using subject-independent data.

In Chapter 5, a comparison of the techniques such as the cross-trial, cross-subject validation approaches, and the learning algorithms used for this study were described. We also investigated the performance of the RF models under varying amounts of training data (O3) for each classification domain and inferred optimal values for training samples per class for both subject-dependent and subject-independent approaches in Table 5.2.

6.1 Recommendations and Future Work

This thesis provided several directions in which the SleepSmart mattress-based device and its associated work could be pursued further. The study explored the classification of voluntary movements that were predefined to simulate common nocturnal motor patterns. As the majority of this work was focused on classification, a valuable addition would be to incorporate techniques (such as rule-based or learning methods) for the detection of movements. The time intervals for the movement events can be estimated and the corresponding model performances can be evaluated using these parameters. Subsequently, the performance of the detector could be evaluated with the true boundaries acquired from video labelling methods as detailed in this study, or with other works [37][109]. The estimation of physiological signals such as heart rate and breathing rate would be a valuable addition, and further inferences of common sleep parameters (such as total sleep time, apneic events or periodic limb movement events) could be made using information acquired from body movements and biosignals.

As this study was conducted in a lab setting, it would be meaningful to conduct sleep studies in other settings, whether in clinical or home-based environments. Testing the SleepSmart mattress
sensors concurrently with clinical or commonly used devices (such as the PSG or actigraph) would inform the viability of the mattress system to perform the desired assessments. As indicated in Section 5.3, it would be worthwhile to investigate the classification capabilities of movements for different types of devices (such as the previous studies that utilized video methods [72][74]), using the similar domain convention defined in Section 3.3.

The participants recruited from this study ranges from 17 to 21 years old. It would be important to compare the performance of the system when testing with younger, older, or even clinical populations such as children with sleep-related movement disorders. We could also evaluate the effects of individualistic factors (such as height and weight) on model performance when using a mattress-based device.

The current build for the SleepSmart mattress used small discrete PCBs placed on the mattress for the sensor measurements. A smaller-scale prototype for SleepSmart was tested using flexible PCBs. While the form factor was improved, the increased manufacturing costs and the suboptimal robustness of the hardware connections due to moving contacts made the implementation difficult. As comfortable and non-intrusive monitoring is key to population with tactile sensitivity, it would be worthwhile to revisit design, material, and component choices to better improve the form of the SleepSmart mattress-based sensor.

In this work, we have demonstrated comparable model performances in all classification domains by extracting kinematic and standard statistical features from the sensor data. As indicated by the lower model performances in Domain C, one of the next step would be to introduce new feature sets (such as time-frequency features [12]), apply feature selection techniques, and evaluate whether such implementations would have a positive effect on the classifier.
In this thesis, we have demonstrated good classification accuracies when using the SleepSmart system to characterize body movements into three domains consisting of coarse-grained (less descriptive) and fine-grained (more descriptive) classes in subject-dependent and subject-independent models. In both models, we have exhibited accuracies comparable or higher than the baseline, even when using small training sizes (as low as 3 training instances per class). Besides needing only a small portion of data for training, one other approach to build the subject-dependent model could be to pre-allocate a short session dedicated for model training, where the subject or patient performs simulated movements depending on the desired application to build a classifier model unique to the specific individual. We could also explore various types of model such as height, weight, age, or even cohort-dependent models, tuned for a specific application. Compared to clinical PSG which can be tedious to score [101][33] and has a tendency to underestimate movement indices [37], the mattress-based sensor array is capable of providing comprehensive, descriptive and automated classifications of body movements while addressing the prohibitive factors on the effectiveness of the PSG such as the intrusive monitoring [36], inaccessibility [31], and high resource requirements of PSG studies [101]. While PSG remains the most reliable and widely-used clinical tool in sleep studies, there are some inherent problems associated with the gold-standard tool as described in this thesis and in other work. With more testing and validations, we believe that the SleepSmart system could be a viable alternative for movement assessments in home-based or even clinical settings during sleep, and the ultimate goal would be to alleviate the burden associated with the limitations of PSG and to complement the work of medical professionals in sleep assessments.
Bibliography


[18] D. L. Picchietti and A. S. Walters, “Moderate to severe periodic limb movement disorder


[34] E. J. Pino, A. A. Moran, A. Dorner De La Paz, and P. Aqueveque, “Validation of non-


2194, no. c, pp. 1–1, 2016.


101

Electrical and Computer Engineering (CCECE), 2015, pp. 784–791.


S. Saeb, L. Lonini, A. Jayaraman, D. C. Mohr, and K. P. Kording, “The need to approximate the use-case in clinical machine learning,” *Gigascience*, vol. 6, no. 5, May
2017.


Appendices

Appendix A - Methods, Questionnaires, Consents, Recruitment Form.

Appendix A included the study advertisements used to recruit participants, consent forms, and the questionnaires in the studies in this thesis.

A.1 Study Recruitment Flyer

Tired? Restless?
Relax for science!

About: The CARIS & Sleep/Wake-Behavior Research labs are evaluating non-invasive technologies (video and mattress sensors) in detecting sleep/wake-behaviors and movement patterns.

Who: Healthy adolescents and young adults (14yrs +).

What: Complete sets of resting movement patterns on the SleepSmart mattress & videography system in a 60-90 minute session.

Participants will be compensated $10 for their time.

Sign up/more information:
A.2 Study Demographics Questionnaire

Date: ___________________________ Subject: ___________________________

1. Sex (circle one): Male Female
2. Age: ___________________________

3. Height: ___________________________
4. Weight: ___________________________

SleepSmart-NDN Questionnaire v4   December 5, 2018
A.3 Consent Form

THE UNIVERSITY OF BRITISH COLUMBIA

SleepSmart-NDN

Informed Consent Form

Principal Investigator
Name: Hendrik F. Machiel (Mike) Van der Loos
Position title: Associate Professor
Organization: UBC Department of Mechanical Engineering
Mailing address: 6250 Applied Science Lane
Phone: <omit> Email: <omit>

Co-Applicants
Name: Osman Ipsiroglu
Position title: Clinical Associate Professor
Organization: UBC Department of Paediatrics, Faculty of Medicine
Mailing Address: <omit>
Phone: <omit>; Email: <omit>

Name: Yi Jui Lee
Position title: MASc. Candidate, Research Assistant
Organization: UBC Department of Mechanical Engineering
Mailing Address: 6250 Applied Science Lane
Phone: <omit> Email: <omit>

Contact Person:
Please contact Osman Ipsiroglu (phone: <omit> Email: <omit>) or Hendrik F. Machiel (Mike) Van der Loos (office phone: <omit>, cell phone: <omit>, email: <omit>), in the event of any unusual occurrences or difficulties related to this research.

Funding Agency
This study is being funded by NeuroDevNet, a Canadian Network of Centers of Excellence (NCE), as part of the TotTech initiative (Tangible, Organizing, and Therapeutic Technologies to engage Children). More information on NeuroDevNet can be found at http://www.neurodevnet.ca/.
Introduction

We invite you to take part in a research study being conducted by Mike Van der Loos, who is a professor at the University of British Columbia, and his colleagues. Your participation in this study is voluntary and you may withdraw from the study at any time. The study is described below. This description tells you about the risks, inconvenience, or discomfort which you might experience. Participating in the study will likely not benefit you directly, but we might learn things that will benefit others. You should discuss any questions you have about this study with Dr. Van der Loos or the other investigators present.

Your Participation is Voluntary

Your participation is voluntary. You have the right to refuse to participate in this study. If you decide to participate, you may still choose to withdraw from the study at any time without any negative consequences to the medical care, education, or other services to which you are entitled or are presently receiving.

If you wish to participate in this study, you will be asked to sign this form.

Please take time to read the following information carefully and to discuss it with your family, friends, and doctor before you decide.

Purpose of the Study

The purpose of this study is to collect preliminary data from the SleepSmart device, Kinect 2.0 video system & by electromyography (EMG). SleepSmart is a sheet fitted over a mattress pad that collects information about body temperature, movement, and position. The Kinect 2.0 video system collects visual movement pattern data and the EMG device collects muscle movement through two electrode pads placed on the lower legs. This preliminary phase will be used to further the development of new hardware and software for a future version of SleepSmart, Kinect and EMG devices, which will ultimately be used as a diagnostic tool for paediatric sleep disorders in conjunction with video-recording of movement patterns.

Study Design

The research will take place in the CARIS lab, room X015 of the ICICS building, 2366 Main Mall at the UBC Vancouver campus. The experiment could run for up to 2 hours for one session. The research will collect quantitative data associated with healthy volunteers who might have Restless Legs Syndrome (i.e., volunteers with or without quick jerky movements in wakefulness and/or sleep). Each session will begin with the collection of basic demographic data. You will be asked to (1) sit face-to-face with a professional for a clinical interview (10-15 minutes), which will end with the Suggested Clinical Immobilization Test (5 minutes), (2) sit on a bed fitted with the SleepSmart system at a 45 degree angle (30-45
minutes), (3) wear colored cloth bands and lie on the bed fitted with the SleepSmart system and perform sets of pre-defined movements (60-90 minutes).

Heart and respiration rates may be recorded using standard consumer-grade monitoring devices. Video recordings will be taken using a Kinect 2.0 video system, a video camera that captures video images at 30 frames per second and performs vision analysis on the images in real-time. Data regarding body temperature, position, and muscle movement will also be collected. At the end of the data collection period, you may be asked some questions about your experience, a process that will take no longer than five minutes.

**Who can participate in this study?**

We are seeking adolescents and young adult volunteers who are willing to test the SleepSmart system. They should be:

i. At least 14 years old
ii. Generally healthy (no heart or breathing problems)
iii. Ambulatory
iv. Able to communicate in spoken and written English

**Who should not participate in this study?**

Individuals who have:

i. Pre-existing heart disorders or diseases
ii. Pre-existing breathing disorders or diseases
iii. A diagnosed sleep disorder, except for mild forms of Restless Legs Syndrome and Periodic Limb Movement Disorder which don’t require therapeutic interventions

**How many participants will take part in this study?**

We are aiming to recruit up to 20 healthy volunteers for this stage of the study.

**Who is conducting the research?**

The study is being conducted by Dr. Van der Loos and colleagues listed on the title page of this form.

**Potential Conflict of Interest**

This patent was assigned to Stanford University, but the university has decided not to make periodic maintenance payments to the USPTO since 2010 due to the lack of any licensing activity related to the patent. Hence the patent is not enforceable at this time.

**Potential Risks**

The physical, mental, and psychological risks associated with this study are minimal. You will be asked to sit and lie on a bed for two 30-45 minute periods. If at any point you become uncomfortable or are unwilling to continue, you are entitled to stop data collection and exit the study at any time.

**Potential Benefits**

It is possible that resting for 30-45 minutes may refresh you. You may also benefit from the knowledge that you are helping contribute to the development of a device that will assist in paediatric sleep diagnostics.

**After the Study is Finished**

With your permission, you may be contacted in the future regarding your participation in other studies or phases of this project. At that time you can refuse to participate and your name will be removed from future correspondence.

If you would be interested in receiving more information about these future studies, please check the appropriate box at the end of this form.

If you are interested in the results of the study, you can contact the researchers and they will provide you with information about the results from this study.

**Confidentiality and Anonymity**

All collected data will be kept in a secure, locked room at UBC for at least five years after the end of this study. We will blur identifying features in videos and frames obtained from videos that will be presented in publications. Access to the original videos and data will be restricted to the investigators, our research partners at the Austrian Institute of Technology (Vienna, Austria) who will analyze the data, and our working group of researchers and clinicians who will convene at the German Society for Sleep Medicine & Research in Regensburg, Germany (March 2017), to review the analyzed data. We will use the collected data only in relation to this particular study. Information from these sessions will be used to design the software and hardware for the next phase of the project. Also, manuscripts based on the findings will be submitted to scientific journals for publication. In the event that quotes from a discussion are used, there will be no information included that could identify the speaker or the client, and you will not be identifiable in any report.
Your confidentiality will be respected. However, research records and health or other source records identifying you may be inspected in the presence of the Investigator or a designate by representatives of the UBC Clinical Research Ethics Board for the purpose of monitoring the research. No information or records that disclose your identity will be published without your consent, nor will any information or records that disclose your identity be removed or released without your consent unless required by law.

You will be assigned a unique study number as a participant in this study. Only this number will be used on any research-related information collected about you during the course of this study, so that your identity [i.e. your name or any other information that could identify you] as a participant in this study will be kept confidential. Information that contains your identity will remain only with the Principal Investigator and/or designate. The list that matches your name to the unique study number that is used on your research-related information will not be removed or released without your consent unless required by law.

The raw video data (captured using a Kinect 2.0 video system), questionnaire data, notes, and physiological signal data will be retained by the study PI for 5 years after publication. The videos will be stored on a local password protected drive and backed up on encrypted, password-protected DVDs that will be stored in a locked filing cabinet in the CARIS lab, a secure, limited-access area which is locked at all times. The video data will be retained by the study PI for 5 years after publication. The local password-protected drives on which the data are stored (including the video data) will then be professionally erased or destroyed.

The Kinect video stream and collected data will be analyzed by our research partners at the Austrian Institute of Technology. The data will be transferred using a secure server that is located at the lab and encrypts and password protects the data. From the transfer server, the data is moved to a secure storage system, which is located in a protected server room at the lab. The storage system is encrypted, and not accessible from outside of the lab. For the purposes of the working group session, the video data will be encrypted and then transported by the PI/Co-Pi on a password protected USB stick to the German Society for Sleep Medicine & Research in Regensburg, Germany (March 2017); only the research team will have access to it. The data will be kept by the research partners 5 years after publication in case review of the data will be necessary after publication and will then be professionally erased or destroyed.

Please note that any study related data sent outside of Canadian borders may increase the risk of disclosure of information because the laws in those countries dealing with protection of information may not be as strict as in Canada. However, all study related data that might be transferred outside of Canada will be coded (this means it will not contain your name or personal identifying information) before leaving the study site. By signing this consent form, you are consenting to the transfer of your information to organizations located outside of Canada, including the Austrian Institute of Technology.
Your rights to privacy are legally protected by federal and provincial laws that require safeguards to insure that your privacy is respected and also give you the right of access to the information about you that has been provided to the sponsor and, if need be, an opportunity to correct any errors in this information. Further details about these laws are available on request to the researchers.

**What if you still have some questions?**

If you have any questions or desire further information about this study before or during participation, you can contact Osman Ipsioglu at <omit> or <omit>, or the study Principal Investigator, Machiel (Mike) Van der Loos at <omit> or <omit>.

**What happens if something goes wrong?**

Signing this consent form in no way limits your legal rights against the sponsor, investigators, or anyone else, and you do not release the researchers or participating institutions from their legal and professional responsibilities.

**Do you have to participate?**

Your participation is completely on a volunteer basis. There are no penalties if you do not wish to participate. If you do volunteer, you have the right to withdraw at any time, for any reason, without penalty. Similarly, the researchers have the right to terminate this research project at any time.

If you do not wish to participate, you do not have to provide any reason for the decision.

You will receive $10 cash as remuneration for your time. Participants will be asked to sign and receive their remuneration at the conclusion of the experiment.

**Problems or Concerns**

If you have any concerns or complaints about your rights as a research participant and/or your experiences while participating in this study, contact the Research Participant Complaint Line in the University of British Columbia Office of Research Ethics by e-mail at RSIL@ors.ubc.ca or by phone at <omit> (Toll Free: <omit>). Please reference the study number (H15-01090) when calling so the Complaint Line staff can better assist you.
SleepSmart-NDN

Informed Consent Form: Adults
My signature on this consent form means:

- I have read and understood the information in this consent form.
- I have had enough time to think about the information provided.
- I have been able to ask for advice if needed.
- I have been able to ask questions and have had satisfactory responses to my questions.
- I understand that all of the information collected will be kept confidential and that the results will only be used for scientific purposes.
- I understand that my participation in this study is voluntary.
- I understand that I am completely free at any time to refuse to participate or to withdraw from this study at any time, and that this will not change the quality of care that I receive.
- I understand that I am not waiving any of my legal rights as a result of signing this consent form.
- I understand that there is no guarantee that this study will provide any benefits to me.
- I will receive a signed copy of this consent form for my own records.
- I consent to participate in this study.

__________________________________        _________________________
Printed Name,                          Signature,                   Date
Participant                             Participant

In addition and separately, I agree to allow my comments to be quoted in reports or publications. If a quote is used, there would be nothing in the quote that could identify me.

__________________________________        _________________________
Printed Name,                          Signature,                   Date
Participant                             Participant

Person Obtaining Consent:

__________________________________        Signature      Date
Printed Name                          Signature      Date
Study Role: __________________________

OPTIONAL INFORMATION:

Yes, please contact me: [ ] with information about participating in future studies
Phone/email:

__________________________________________________________________________________

_______________________  _____________________  ________________
Printed Name,  Signature,  Date
Principal Investigator  Principal Investigator
Appendix B - Summary of Performance Measures

B.1 Overall Performance Measures

Due to feature bagging implementations in the RF models, the average values exhibited in the tables listed below might have minor differences to the accuracy values reported in Chapter 4.

Table 6-1: Overall performance measures in all classification domains and models.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Recall (Sensitivity)</th>
<th>Precision</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subject Dependent</td>
<td>Subject Independent</td>
<td>Subject Dependent</td>
<td>Subject Independent</td>
</tr>
<tr>
<td>Domain A</td>
<td>0.9816</td>
<td>0.9683</td>
<td>0.9866</td>
<td>0.9731</td>
</tr>
<tr>
<td>Domain B</td>
<td>0.9694</td>
<td>0.9018</td>
<td>0.9777</td>
<td>0.9156</td>
</tr>
<tr>
<td>Domain C</td>
<td>0.9376</td>
<td>0.8069</td>
<td>0.9379</td>
<td>0.8085</td>
</tr>
<tr>
<td>Average</td>
<td>0.9629</td>
<td>0.8923</td>
<td>0.9674</td>
<td>0.8991</td>
</tr>
</tbody>
</table>

B.2 Subject Dependent Performance Measures

Table 6-2: Domain A performance measures for subject-dependent models.

<table>
<thead>
<tr>
<th>Class</th>
<th>Recall (Sensitivity)</th>
<th>Precision</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Major Postural Movements</td>
<td>0.9966</td>
<td>0.9921</td>
<td>0.9970</td>
</tr>
<tr>
<td>2: Isolated Movements</td>
<td>0.9549</td>
<td>0.9953</td>
<td>0.9983</td>
</tr>
<tr>
<td>3: Leg Movements</td>
<td>0.9933</td>
<td>0.9724</td>
<td>0.9763</td>
</tr>
<tr>
<td>Average</td>
<td>0.9816</td>
<td>0.9866</td>
<td>0.9905</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.9838</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 6-3: Domain B performance measures for subject-dependent approach

<table>
<thead>
<tr>
<th>Class</th>
<th>Recall (Sensitivity)</th>
<th>Precision</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Head</td>
<td>0.9966</td>
<td>0.9933</td>
<td>0.9993</td>
</tr>
<tr>
<td>2: Torso and Limbs</td>
<td>0.9977</td>
<td>0.9910</td>
<td>0.9966</td>
</tr>
<tr>
<td>3: Left Arm</td>
<td>0.9522</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>4: Right Arm</td>
<td>0.9594</td>
<td>0.9930</td>
<td>0.9993</td>
</tr>
<tr>
<td>5: Left Leg</td>
<td>0.9362</td>
<td>0.9489</td>
<td>0.9949</td>
</tr>
<tr>
<td>6: Right Leg</td>
<td>0.9362</td>
<td>0.9587</td>
<td>0.9959</td>
</tr>
<tr>
<td>7: Both Leg</td>
<td>0.9866</td>
<td>0.9395</td>
<td>0.9857</td>
</tr>
<tr>
<td>8: Both Feet</td>
<td>0.9899</td>
<td>0.9966</td>
<td>0.9996</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.9693</strong></td>
<td><strong>0.9776</strong></td>
<td><strong>0.9964</strong></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
<td><strong>0.9762</strong></td>
</tr>
<tr>
<td>Class</td>
<td>Recall (Sensitivity)</td>
<td>Precision</td>
<td>Specificity</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>----------------------</td>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>1-Move from back to right</td>
<td>0.9400</td>
<td>0.9463</td>
<td>0.9974</td>
</tr>
<tr>
<td>2-Move from back to left</td>
<td>0.9466</td>
<td>0.9726</td>
<td>0.9987</td>
</tr>
<tr>
<td>3-Move from left to right</td>
<td>0.9530</td>
<td>0.8987</td>
<td>0.9948</td>
</tr>
<tr>
<td>4-Move from right to left</td>
<td>0.9657</td>
<td>0.9591</td>
<td>0.9980</td>
</tr>
<tr>
<td>5-Move from right to back</td>
<td>0.9591</td>
<td>0.9724</td>
<td>0.9987</td>
</tr>
<tr>
<td>6-Move from left to back</td>
<td>0.9396</td>
<td>0.9523</td>
<td>0.9977</td>
</tr>
<tr>
<td>7-Straighten left arm</td>
<td>0.8767</td>
<td>0.9343</td>
<td>0.9971</td>
</tr>
<tr>
<td>8-Bend left arm</td>
<td>0.9115</td>
<td>0.9305</td>
<td>0.9968</td>
</tr>
<tr>
<td>9-Straighten right arm</td>
<td>0.8378</td>
<td>0.8857</td>
<td>0.9948</td>
</tr>
<tr>
<td>10-Bend left arm</td>
<td>0.9054</td>
<td>0.8815</td>
<td>0.9942</td>
</tr>
<tr>
<td>11-Turn head to the right</td>
<td>0.9127</td>
<td>0.8831</td>
<td>0.9942</td>
</tr>
<tr>
<td>12-Turn head to the left</td>
<td>0.9000</td>
<td>0.9246</td>
<td>0.9964</td>
</tr>
<tr>
<td>13-Bend left leg</td>
<td>0.9666</td>
<td>0.9354</td>
<td>0.9968</td>
</tr>
<tr>
<td>14-Straighten left leg</td>
<td>0.9396</td>
<td>0.9396</td>
<td>0.9971</td>
</tr>
<tr>
<td>15-Flex both ankles (while facing right)</td>
<td>0.9530</td>
<td>0.9220</td>
<td>0.9961</td>
</tr>
<tr>
<td>16-Straighten both legs (while facing right)</td>
<td>0.9664</td>
<td>0.9795</td>
<td>0.9990</td>
</tr>
<tr>
<td>17-Bend both legs (while facing right)</td>
<td>0.9797</td>
<td>0.9863</td>
<td>0.9993</td>
</tr>
<tr>
<td>18-Bend right leg</td>
<td>0.9533</td>
<td>0.9166</td>
<td>0.9958</td>
</tr>
<tr>
<td>19-Straighten right leg</td>
<td>0.9391</td>
<td>0.9391</td>
<td>0.9971</td>
</tr>
<tr>
<td>20-Flex both ankles (while facing left)</td>
<td>0.9600</td>
<td>0.9350</td>
<td>0.9968</td>
</tr>
<tr>
<td>21-Straighten both legs (while facing left)</td>
<td>0.9333</td>
<td>0.9523</td>
<td>0.9977</td>
</tr>
<tr>
<td>22-Bend both legs (while facing left)</td>
<td>0.9866</td>
<td>0.9866</td>
<td>0.9993</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.9375</strong></td>
<td><strong>0.9379</strong></td>
<td><strong>0.9970</strong></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td><strong>0.9377</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### B.3 Subject-Independent Performance Measures

<table>
<thead>
<tr>
<th>Class</th>
<th>Recall (Sensitivity)</th>
<th>Precision</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Major Postural Movements</td>
<td>0.9877</td>
<td>0.9778</td>
<td>0.9916</td>
</tr>
<tr>
<td>2: Isolated Movements</td>
<td>0.9381</td>
<td>0.9823</td>
<td>0.9937</td>
</tr>
<tr>
<td>3: Leg Movements</td>
<td>0.9792</td>
<td>0.9593</td>
<td>0.9652</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.9683</strong></td>
<td><strong>0.9731</strong></td>
<td><strong>0.9835</strong></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
<td><strong>0.9704</strong></td>
</tr>
</tbody>
</table>

**Table 6-6: Domain B performance measures for subject-independent approach**

<table>
<thead>
<tr>
<th>Class</th>
<th>Recall (Sensitivity)</th>
<th>Precision</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Head</td>
<td>0.9699</td>
<td>0.9797</td>
<td>0.9980</td>
</tr>
<tr>
<td>2: Torso and Limbs</td>
<td>0.9877</td>
<td>0.9735</td>
<td>0.9899</td>
</tr>
<tr>
<td>3: Left Arm</td>
<td>0.9079</td>
<td>0.9500</td>
<td>0.9953</td>
</tr>
<tr>
<td>4: Right Arm</td>
<td>0.9223</td>
<td>0.9286</td>
<td>0.9929</td>
</tr>
<tr>
<td>5: Left Leg</td>
<td>0.7785</td>
<td>0.8689</td>
<td>0.9882</td>
</tr>
<tr>
<td>6: Right Leg</td>
<td>0.7685</td>
<td>0.8208</td>
<td>0.9832</td>
</tr>
<tr>
<td>7: Both Leg</td>
<td>0.9265</td>
<td>0.8473</td>
<td>0.9626</td>
</tr>
<tr>
<td>8: Both Feet</td>
<td>0.9530</td>
<td>0.9562</td>
<td>0.9956</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.9018</strong></td>
<td><strong>0.9156</strong></td>
<td><strong>0.9882</strong></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
<td><strong>0.9196</strong></td>
</tr>
</tbody>
</table>
Table 6-7: Domain C performance measures for subject-independent approach

<table>
<thead>
<tr>
<th>Class</th>
<th>Recall (Sensitivity)</th>
<th>Precision</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Move from back to right</td>
<td>0.8400</td>
<td>0.8571</td>
<td>0.9933</td>
</tr>
<tr>
<td>2-Move from back to left</td>
<td>0.8533</td>
<td>0.9209</td>
<td>0.9965</td>
</tr>
<tr>
<td>3-Move from left to right</td>
<td>0.8389</td>
<td>0.7102</td>
<td>0.9837</td>
</tr>
<tr>
<td>4-Move from right to left</td>
<td>0.8767</td>
<td>0.7901</td>
<td>0.9891</td>
</tr>
<tr>
<td>5-Move from right to back</td>
<td>0.7959</td>
<td>0.8731</td>
<td>0.9946</td>
</tr>
<tr>
<td>6-Move from left to back</td>
<td>0.7785</td>
<td>0.8593</td>
<td>0.9939</td>
</tr>
<tr>
<td>7-Straighten left arm</td>
<td>0.6575</td>
<td>0.7680</td>
<td>0.9907</td>
</tr>
<tr>
<td>8-Bend left arm</td>
<td>0.8435</td>
<td>0.7515</td>
<td>0.9869</td>
</tr>
<tr>
<td>9-Straighten right arm</td>
<td>0.5541</td>
<td>0.6721</td>
<td>0.9872</td>
</tr>
<tr>
<td>10-Bend left arm</td>
<td>0.8378</td>
<td>0.7006</td>
<td>0.9830</td>
</tr>
<tr>
<td>11-Turn head to the right</td>
<td>0.8591</td>
<td>0.7574</td>
<td>0.9869</td>
</tr>
<tr>
<td>12-Turn head to the left</td>
<td>0.7200</td>
<td>0.8571</td>
<td>0.9942</td>
</tr>
<tr>
<td>13-Bend left leg</td>
<td>0.9267</td>
<td>0.8580</td>
<td>0.9926</td>
</tr>
<tr>
<td>14-Straighten left leg</td>
<td>0.6980</td>
<td>0.7172</td>
<td>0.9869</td>
</tr>
<tr>
<td>15-Flex both ankles (while facing right)</td>
<td>0.7718</td>
<td>0.8214</td>
<td>0.9920</td>
</tr>
<tr>
<td>16-Straighten both legs (while facing right)</td>
<td>0.8591</td>
<td>0.8767</td>
<td>0.9942</td>
</tr>
<tr>
<td>17-Bend both legs (while facing right)</td>
<td>0.9324</td>
<td>0.8903</td>
<td>0.9946</td>
</tr>
<tr>
<td>18-Bend right leg</td>
<td>0.8800</td>
<td>0.8516</td>
<td>0.9926</td>
</tr>
<tr>
<td>19-Straighten right leg</td>
<td>0.6757</td>
<td>0.7143</td>
<td>0.9872</td>
</tr>
<tr>
<td>20-Flex both ankles (while facing left)</td>
<td>0.8200</td>
<td>0.8092</td>
<td>0.9907</td>
</tr>
<tr>
<td>21-Straighten both legs (while facing left)</td>
<td>0.7933</td>
<td>0.8207</td>
<td>0.9917</td>
</tr>
<tr>
<td>22-Bend both legs (while facing left)</td>
<td>0.9400</td>
<td>0.9097</td>
<td>0.9955</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.8069</strong></td>
<td><strong>0.8085</strong></td>
<td><strong>0.9908</strong></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td><strong>0.8072</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>