Detecting passive fatigue with brain-computer interface in drowsy driving and stroke rehabilitation

Foong, Ruyi

2019

Foong, R. (2019). Detecting passive fatigue with brain-computer interface in drowsy driving and stroke rehabilitation. Doctoral thesis, Nanyang Technological University, Singapore

<https://hdl.handle.net/10356/107596>

<https://doi.org/10.32657/10356/107596>

Detecting Passive Fatigue with Brain-Computer Interface in Drowsy Driving and Stroke Rehabilitation

FOONG RUYI

School of Computer Science and Engineering

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

20/09/19

.....
Date

.....
Foong Ruyi

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

20/9/19

.....

Date

Chai

.....

Quek Hiok Chai

Authorship Attribution Statement

This thesis contains material from 2 papers published in the following peer-reviewed journals where I was the first author.

Chapters 3 and 4 are published as Foong R., Ang K.K., Zhang Z., & Quek C. (2019). An iterative cross-subject negative-unlabeled learning algorithm for quantifying passive fatigue. *Journal of Neural Engineering*. doi: 10.1088/1741-2552/ab255d

The contributions of the co-authors are as follows:

- Dr Ang provided the direction for this study.
- Dr Zhang provided guidance in the algorithm development and implementation.
- I co-designed the experimental design, performed the data analysis, algorithm development and implementation, and prepared the manuscript drafts.
- All co-authors revised the manuscript.

Chapter 5 is published as Foong R., Ang K.K., Quek C., Guan C., Phua K.S., Kuah C.W.K., Deshmukh V.A., Yam L.H.L., Rajeswaran D.K., Tang N., Chew E., & Chua K.S.G. (2019). Assessment of the Efficacy of EEG-based MI-BCI with Visual Feedback and EEG Correlates of Mental Fatigue for Upper-Limb Stroke Rehabilitation, *IEEE Transactions of Biomedical Engineering*. doi: 10.1109/TBME.2019.2921198

The contributions of the co-authors are as follows:

- The principle and co-investigators are Dr Chua, Dr Ang, Prof Guan, Dr Chua and Dr Chew who provided direction for the study.
- Mr Phua, Mr Kuah, Mr Deshmukh, Mr Lam, Dr Rajeswaran and Dr Tang carried out the data collection and stroke rehabilitation.
- I analyzed the clinical outcomes and EEG data, and wrote the manuscript drafts.
- All co-authors revised the manuscript.

20/09/2019

.....
Date

.....
Foong Ruyi

Acknowledgements

This Ph.D. journey was a deeply humbling one, in more ways than one, and would not have been possible without the support of many.

The work done in this thesis was only possible due to constant and patient guidance from my supervisors, Dr Ang Kai Keng and Associate Professor Quek Hiok Chai. Thank you for always keeping your door open to me whether the problems were big or small, and looking out for me as a whole person and not just another research student. Your time and efforts are very much appreciated and valued by this undeserving student.

This thesis' projects have been successful only because of the support from many staff and students in the Neural and Biomedical Technology Department (NBT) in the Institute for Infocomm Research (I2R). Thank you for being subjects, collecting data, entertaining my questions, giving advice and solving problems I faced. I especially thank Aung Aung Phyto Wai for help with the experiments, and Dr Zhang Zhuo for help with the algorithms. Thank you for graciously rendering help for little in return from this lowly student.

For the opportunity to take up this Ph.D. program, I thank the School of Computer Science and Engineering (SCSE) in Nanyang Technological University (NTU) and the Agency for Science, Technology and Research (A*STAR). The A*STAR Graduate Scholarship and sureties Ms Lee Wai Lan and Mr Robert Lam have made this financially possible as well. Thank you for believing in me.

For making the administrative aspects of this Ph.D. journey smooth, I thank the scholarship officers at A*STAR Graduate Academy, Ms Agnes Loi at NBT, Ms Janice Lee at A*STAR, Mr Kesavan Asaithambi at Computational Intelligence Lab (CIL), and the staff at SCSE Graduate Office. Thank you for patiently helping this frazzled student.

Having fellow travelers on this Ph.D. road, who share similar technical and personal struggles, definitely made this journey more bearable. For that I thank the other students in the Brain-Computer Interface (BCI) lab, as well as other of Prof Quek's students in the CIL lab, and other Ph.D. students who are friends from outside these labs. The ones who have graduated are inspirations, and I hope that my graduation brings hope to those who have yet to finish as well. Thank you for listening and relating to my research woes, providing technical insight, and encouraging me.

To the many friends and brothers and sisters in Christ, thank you for giving a listening ear even though you probably could not relate to many of my struggles. Thank you for not judging my flaws but encouraging me on this long journey, and even keeping me accountable to be productive. The meet-ups we had were probably instrumental in keeping me sane as well. For those who belong to Christ, I treasure especially your prayers and spiritual support, reminding me to trust in God and rely on Him above all.

To my dearest family – my father, mother and sister, thank you for unconditionally loving and accepting me, for celebrating my highs and enduring my lows. Thank you for accommodating my flaws and way of doing things. Thank you too for your prayers, hugs, encouragement and support to this underserving one.

Finally, I give praise and thanks to the only living God. All of the above people are in my life only because of Him, and all of the work below is accomplished only with His strength and wisdom. Thank you for Your faithful love even when I was faithless, and for humbling me and drawing me to Yourself and others through this Ph.D. journey. To You the only wise God be glory forevermore through Jesus Christ! Amen (Romans 16:27).

Table of Contents

Abstract.....	xii
List of Figures	xiv
List of Tables.....	xvi
List of Abbreviations	xvii
Chapter 1 A Brief Introduction to Mental State Detection with Brain-Computer Interface	1
1.1 Definition of Brain-Computer Interface (BCI)	1
1.2 BCI Applications.....	1
1.3 Motivation of Thesis	3
1.4 Organization of Thesis	4
Chapter 2 BCI Mental State Detection in Drowsy Driving and Stroke Rehabilitation.....	5
2.1 Definition of Fatigue	5
2.2 Types of Fatigue.....	6
2.3 Overview of Passive Fatigue Detection with EEG-based BCI in Drowsy Driving..	7
2.3.1 Detection of Fatigue.....	7
2.3.2 EEG Correlates of Fatigue	7
2.3.3 Challenges of Fatigue Detection in Drowsy Driving	13
2.4 Overview of Fatigue Detection in EEG-based MI-BCI Stroke Rehabilitation	14
2.4.1 EEG-based MI-BCI with Only Visual Feedback	14
2.4.2 Currently Unexplored: Fatigue in EEG-based MI-BCI Stroke Rehabilitation	18
2.5 Open Problems, Challenges and Research Objectives.....	18

Chapter 3 A Two-Step Iterative Cross-Subject Negative-Unlabeled Learning Algorithm for Detecting and Quantifying Fatigue	21
3.1 Objective	21
3.2 Introduction to Negative-Unlabeled Learning	21
3.2.1 Machine Learning Approaches	22
3.2.2 Negative-Unlabeled Learning in the Drowsy Driving Context	23
3.2.3 Positive-Unlabeled Learning	24
3.3 Proposed Two-Step Iterative Negative-Unlabeled Learning Algorithm	25
3.3.1 Step One: Identifying the Most Fatigued Block	Error! Bookmark not defined.
3.3.2 Step Two: Computing the Fatigue Score.....	28
3.3.3 Comparison with Current PU Learning Algorithms	29
3.4 Summary	29
Chapter 4 Detecting and Quantifying Passive Fatigue in a Drowsy Driving Study with Dry EEG Electrodes	31
4.1 Objective	31
4.2 Materials and Methods.....	31
4.2.1 Experimental Set-up and Protocol	31
4.2.2 Preprocessing and Segmentation of EEG	33
4.2.3 EEG Feature Extraction.....	34
4.2.4 Negative-Unlabeled Learning Algorithm Parameters.....	36
4.2.5 Comparison of Proposed Algorithm with Baselines	37

4.2.5.1 Assumptions of Driving and Alert Data	37
4.2.5.2 Identification of the Most Fatigued Block	39
4.2.6 Fatigue Score and its Correlation with Beta Band Power	40
4.3 Results	40
4.3.1 Comparison of Proposed Algorithm with Baselines	41
4.3.2 Convergence of the Proposed Algorithm	43
4.3.3 Fatigue Score and its Correlation with Beta Band Power	44
4.4 Discussion	44
4.4.1 Comparison of Proposed Algorithm with Baselines	46
4.4.1.1 Assumptions of Driving and Alert Data	44
4.4.1.2 Identification of the Most Fatigued Block	47
4.4.1.3 Other Machine Learning Approaches	48
4.4.2 Convergence of the Proposed Algorithm	49
4.4.3 Fatigue Score and its Correlation with Beta Band Power	49
4.5 Summary	50
Chapter 5 Assessment of the Efficacy of EEG-based MI-BCI with Visual Feedback and the Presence of Mental Fatigue in Upper-Limb Stroke Rehabilitation	53
5.1 Objective	53
5.2 Materials and Methods	53
5.2.1 The nBETTER system	54
5.2.2 Clinical Protocol	55
5.2.2.1 Flow of Clinical Trial	55

5.2.2.2 Subject Inclusion and Exclusion Criteria	56
5.2.2.3 MI-BCI Sessions.....	57
5.2.2.3.1 MI-BCI screening and calibration sessions	58
5.2.2.3.2 nBETTER supervised session.....	59
5.2.2.3.3 nBETTER therapy session.....	59
5.2.2.3.4 Standard arm therapy session.....	59
5.2.3 Retrospective Comparison Groups	59
5.2.4 Patient Demographics.....	60
5.2.5 Clinical Efficacy of the nBETTER Intervention.....	61
5.2.6 Correlation of Beta Band Power and BCI Performance.....	62
5.2.7 Application of Proposed Negative-Unlabeled Learning Algorithm	63
5.3 Results.....	64
5.3.1 Patient Demographics.....	64
5.3.2 Clinical Efficacy of the nBETTER Intervention.....	64
5.3.3 Correlation of Beta Band Power and BCI Performance.....	66
5.3.4 Application of Proposed Negative-Unlabeled Learning Algorithm	67
5.4 Discussion	68
5.4.1 Patient Demographics.....	69
5.4.2 Clinical Efficacy of the nBETTER Intervention.....	69
5.4.3 Correlation of Beta Band Power and BCI Performance.....	70
5.4.4 Application of Proposed Negative-Unlabeled Learning Algorithm	71
5.5 Summary	73

Chapter 6 Contributions, Limitations and Future Work.....	75
6.1 Contributions	75
6.1.1 Proposed NU Learning Algorithm.....	75
6.1.2 Drowsy Driving Study	76
6.1.3 Efficacy and Fatigue Study with nBETTER Intervention	77
6.2 Limitations	78
6.2.1 Proposed NU Learning Algorithm.....	78
6.2.2 Drowsy Driving Study	79
6.2.3 Efficacy and Fatigue Study with nBETTER Intervention	80
6.3 Future Work.....	81
6.3.1 Proposed NU Learning Algorithm.....	81
6.3.2 Drowsy Driving Study	82
6.3.3 Efficacy and Fatigue Study with nBETTER Intervention	83
6.3.4 Towards Passive BCI (pBCI)	84
Bibliography.....	87
Author's Publications	109

Abstract

A recent topic of interest for electroencephalography(EEG)-based brain-computer interface (BCI) is mental state monitoring. Mental states such as fatigue, frustration and attention have been shown to affect healthy users' BCI performance. However, ground truth of mental states is often difficult to ascertain. Instead, surrogate measures such as ratings or reaction times have been used. Additionally, mental state monitoring has yet to be applied in the medical domain. Currently, a handful of studies have reported the mental state of fatigue in motor imagery (MI)-BCI stroke rehabilitation only via subject ratings, but not EEG data. Hence, there is a need to develop algorithms that can detect mental states despite a lack of ground truth. In particular, this thesis focuses on detecting the mental state of passive fatigue induced by monotony in healthy subjects in drowsy driving and in stroke survivors using MI-BCI stroke rehabilitation.

A two-step iterative cross-subject negative-unlabeled (NU) learning algorithm is proposed to address the lack of ground truth of passive fatigue in drowsy driving studies. In the first step, subjects' alert and unlabeled driving EEG data are iteratively used to label each subject's most fatigued driving block. The second step then used the alert and newly-labeled fatigued blocks to compute subjects' fatigue score.

A drowsy driving experiment is conducted to verify effectiveness of the NU learning algorithm. Dry EEG MuseTM headband data is collected from 29 healthy subjects. The results showed that the algorithm converged in 7 iterations and yielded a high mean accuracy of $93.77\% \pm 8.15\%$ in detecting fatigue in a cross-subject manner. The results also showed that the fatigue score is significantly negatively correlated with relative EEG beta band power, an indicator of passive fatigue. This suggested that high

fatigue score is associated with low beta band power and therefore low brain arousal. Hence the proposed algorithm was able to quantify passive fatigue well.

In the medical domain, a clinical trial is conducted with Neurostyle Brain Exercise Therapy Towards Enhanced Recovery (nBETTER), an EEG-based MI-BCI with only visual feedback. Subjects' Fugl-Meyer Motor Assessment (FMA) score is measured to assess clinical efficacy. Also, BCI performance is correlated with relative EEG beta band power and the proposed NU learning algorithm is applied as well to investigate the presence of fatigue. The results showed significant FMA score gains, but are comparable to a retrospective control group. The results also showed a significant positive correlation of BCI performance with beta band power. Additionally, the proposed NU learning algorithm yielded fatigue scores that are non-significantly negatively correlation with BCI performance. Together, these suggested that fatigue might be present, resulting in the poorer BCI performance and comparable efficacy to the control group. This postulated a disadvantage of MI-BCI systems in inducing fatigue due to the monotony of performing MI repetitively.

These studies have presented the potential of BCI to play a fatigue monitoring role, by presenting a cross-subject algorithm to detect and quantify fatigue from alert data in a drowsy driving setting, and finding the possibility of fatigue during a MI-BCI intervention.

List of Figures

1.1 BCI is a closed loop system and has a wide range of applications (modified from Wolpaw and Wolpaw (2012)).	2
2.1 Different forms of exertion leads to different types of fatigue.	6
3.1 Machine learning approaches.	22
4.1 The experimental set-up with the hitting game and driving simulation interfaces.	32
4.2 Comparison of across-subject accuracies of the proposed algorithm (PA) against that of baselines alert block (AB), only alert (OA), fatigue driving (FD), reaction time (RT), no iteration with block 9 (NI9), no iteration with block 18 (NI18), same block with block 9 (SB9) and same block with block 18 (SB18). Accuracy is the percentage of correctly classified samples. Baselines associated with NU and one-class learning algorithms are in magenta while baselines associated with supervised learning are in blue.	42
4.3 Convergence of the mean error across subjects over 7 iterations in the first step of the PA. The error is the average of the probability estimates of samples belonging to the other class. The stopping criterion is a change in the mean error of less than 0.01%, which took the PA 7 iterations to achieve. A final mean error of $8.46 \pm 7.56\%$ is achieved.	43
4.4 Fatigue scores as calculated in the second step of the PA are smoothed. (a) This displays the average across all subjects, where a general increase over driving blocks is seen. (b) S4 displays a slight dip in their score in the later blocks. (b) S5 shows a more obvious trend of increasing fatigue score.	45
4.5 Repeated measures correlation of mean score against mean beta band power for each of the 29 subjects' 18 driving blocks. A significant negative correlation is found.	46
5.1 Subject using the nBETTER system.	54
5.2 CONSORT diagram showing the flow of the study from enrolment to analysis.	56
5.3 Trials in the MI-BCI screening and calibration sessions and in the nBETTER supervised and therapy sessions.	58
5.4 Illustration of the subject-specific model update process, where the model is updated after every supervised session and used for the same day's therapy session and the subsequent supervised session.	60
5.5 Boxplot showing the improvement in FMA scores at post-intervention (Week 6) and follow-ups (Week 12 and 24) for the nBETTER, SAT and BCI-HK groups.	65
5.6 Plot of repeated measures correlations of relative EEG beta band power against BCI performance in terms of classification accuracy. Significant positive correlations are found in frontal and central, but not parietal-occipital, brain regions.	67

5.7 Fatigue score over 18 nBETTER supervised sessions: (a) S6's score, and (b) S10's score.....	68
5.8 Pearson's correlation of subjects' mean fatigue score against BCI accuracy of nBETTER supervised sessions. A negative correlation is found.....	69

List of Tables

2.1 Summary of Passive Fatigue Driving Simulation Studies	9
2.2 Summary of Recent EEG-Based MI-BCI RCT	16
4.1 Comparison of PA with Baselines	38
4.2 Comparison of PA with Baselines in Terms of Accuracy.....	41
5.1 Similarities and Differences Between the nBETTER, SAT and BCI-HK Groups' Protocol	61
5.2 Demographics and Baseline Characteristics of the nBETTER, SAT and BCI-HK Groups.....	62
5.3 Improvement in FMA Score.....	65
6.1 Summary of Thesis' Contributions	75

List of Abbreviations

AAA	American Automobile Association
AB	Alert Block
ANOVA	Analysis Of Variance
BCI	Brain-Computer Interface
BCI-HK	Brain-Computer Interface-Haptic Knob
CNS	Central Nervous System
CONSORT	Consolidated Standards of Reporting Trials
CSP	Common Spatial Pattern
CT	Computed Tomography
EEG	Electroencephalography
EOG	Electrooculography
FBCSP	Filter Bank Common Spatial Pattern
FD	Fatigue Driving
FFT	Fast Fourier Transform
FMA	Fugl-Meyer Motor Assessment
HR	Heart Rate
MAS	Modified Ashworth Scale
MI	Motor Imagery
MRI	Magnetic Resonance Imaging
nBETTER	Neurostyle Brain Exercise Therapy Towards Enhanced Recovery
NI	No Iteration
NU	Negative-Unlabeled
PA	Proposed Algorithm
pBCI	Passive Brain-Computer Interface
PU	Positive-Unlabeled
RBF	Radial Basis Function

RCT	Randomized Controlled Trials
SAT	Standard Arm Therapy
SB	Same Block
SVM	Support Vector Machine
TMS	Transcranial Magnetic Stimulation
VAS	Visual Analog Scale

Chapter 1

A Brief Introduction to Mental State Detection with Brain-Computer Interface

1.1 Definition of Brain-Computer Interface (BCI)

In the first international meeting on brain-computer interface (BCI) technology in 2000, BCI was defined as “a communication system that does not depend on the brain’s normal output pathways of peripheral nerves and muscles” Wolpaw *et al.* (2000). A more recent definition in 2012 states that “A BCI is a system that measures CNS activity and converts it into artificial output that replaces, restores, enhances, supplements, or improves natural CNS output and thereby changes the ongoing interactions between the CNS and its external or internal environment” (Wolpaw & Wolpaw, 2012). It is therefore a closed-loop system starting from the acquisition of brain signals from the user, followed by the extraction of features from the signals and the translation of those features into commands which then result in feedback to the user (see Figure 1.1). There are various brain signals that can be obtained via different neuroimaging methods, as detailed in recent reviews (Nicolas-Alonso & Gomez-Gil, 2012; Ramadan & Vasilakos, 2017; Tiwari *et al.*, 2018). A widely used brain signal is electroencephalography (EEG), due to its non-invasiveness, inexpensiveness and portability. It also provides high temporal resolution for the detection of transient brain activity. Its low risk of use allows for the ease of recruiting subjects for research as well. Thus, this thesis focuses on EEG-based BCI.

1.2 BCI Applications

BCI can be used for a range of applications (refer to Figure 1.1, modified from

Wolpaw and Wolpaw (2012)). It may replace a lost function such as movement for a disabled person, by commanding a wheelchair. It can also restore lost functions by controlling machines that can help stimulate nerves of the lost function. It may enhance the CNS output, for example, by detecting drowsiness and then alerting the user to wake up. It may also supplement the CNS output by controlling a third arm of sorts. It can improve CNS output as well by encouraging neuroplasticity of lost neuropathways, such as in stroke patients.

Figure 1.1 BCI is a closed loop system and has a wide range of applications (modified from Wolpaw and Wolpaw (2012)).

1.3 Motivation of Thesis

Early research in BCI focused more on replacing and restoring lost CNS functions (Wolpaw & Wolpaw, 2012). It then naturally moved on to neurorehabilitation to improve the CNS of patients with brain injury, as evidenced by several reviews (Dobkin, 2007; Mak & Wolpaw, 2009; McFarland & Wolpaw, 2017; Shih *et al.*, 2012; Silvoni *et al.*, 2011; Soekadar *et al.*, 2015; van Dokkum *et al.*, 2015). However, the most recent upcoming area of BCI technology is in the area of enhancing CNS output, or generally using BCI in non-medical applications. Of particular interest is the detection of user mental state (Abdulkader *et al.*, 2015; Blankertz *et al.*, 2016; Blankertz *et al.*, 2010; Millán *et al.*, 2010; Moghimi *et al.*, 2012; Ozkan & Kahya, 2018; Ramadan & Vasilakos, 2017; Roy *et al.*, 2013; Roy & Frey, 2016; Tiwari *et al.*, 2018; van Erp *et al.*, 2012; Voznenko *et al.*, 2016). This has mainly been researched in non-medical areas, such as driving (Chuang *et al.*, 2010; Hashemi *et al.*, 2014), or BCI tasks (Käthner *et al.*, 2014; Myrden & Chau, 2015, 2016, 2017) such as motor imagery (Rozado *et al.*, 2015) or cursor control (Curran & Stokes, 2003).

One main challenge in the detection of mental states is the determination of the ground truth. For example, in the drowsy driving context, the labeling of the fatigued mental state is done manually via surrogate measures (Lin *et al.*, 2012). As such, these have a certain degree of subjectivity or intrusiveness into subjects' mental state (Arico *et al.*, 2018; Dasari *et al.*, 2017).

In addition to this, such investigations of mental states have not yet taken place in the early applications of BCI, i.e. in medical applications. This is a gap that requires bridging as medical-related BCI research have often pointed out the need to monitor user mental states such as their attention (van Dokkum *et al.*, 2015), fatigue (Mak & Wolpaw, 2009), motivation (Ahn & Jun, 2015) and emotion (Pillette *et al.*, 2018). For instance, in

the BCI stroke rehabilitation setting, a few studies have pointed out that fatigue could affect patients' BCI performance (Frolov *et al.*, 2017b; Leamy *et al.*, 2014a; Prasad *et al.*, 2010), but did not use EEG to investigate further.

These are thus the impetus for this thesis. In particular, the focus is on detecting the mental state of passive fatigue brought on by monotony in driving situations, as well as in the motor imagery (MI) BCI stroke rehabilitation setting.

1.4 Organization of Thesis

Chapter 2 provides an overview of passive fatigue and its presence in drowsy driving and stroke rehabilitation.

Chapter 3 presents a novel two-step iterative cross-subject negative-unlabeled learning algorithm for quantifying passive fatigue in the drowsy driving context.

Chapter 4 reports the experimental set-up that the proposed algorithm is applied in and the results obtained for detecting and quantifying passive fatigue with dry EEG electrodes.

Chapter 5 presents the clinical trial of an EEG-based MI-BCI stroke intervention with only visual feedback and the results of its clinical efficacy and presence of passive fatigue.

Chapter 6 summarizes the contributions and limitations of this thesis. It also discusses future work of using BCI in mental state monitoring in driving and stroke rehabilitation.

Chapter 2

BCI Mental State Detection in Drowsy Driving and Stroke Rehabilitation

2.1 Definition of Fatigue

A definition of fatigue is proposed by Phillips (2015):

“Fatigue is a suboptimal psychophysiological condition caused by exertion. The degree and dimensional character of the condition depends on the form, dynamics and context of exertion. The context of exertion is described by the value and meaning of performance to the individual; rest and sleep history; circadian effects; psychosocial factors spanning work and home life; individual traits; diet; health, fitness and other individual states; and environmental conditions. The fatigue condition results in changes in strategies or resource use such that original levels of mental processing or physical activity are maintained or reduced.”

This definition takes into account the multidimensionality of fatigue, namely: the psychological, physiological, and performance output aspects. The psychological aspect of fatigue refers to subjective experience, which were the negative effects subjects felt when fatigued. Fatigue is also manifested physiologically. For example, subjects' biosignals such as their brain signals and heart rate are different when they are fatigued. The performance output aspect refers to the performance of subjects when coping with fatigue. For example, there is performance deterioration when a fatigued person can not react to emergency situations. By understanding these various effects of fatigue, effective countermeasures can then be formulated to counter them.

2.2 Types of Fatigue

As per Phillips' definition, the manifestations of fatigue depends on the “form, dynamics and context of exertion”. Such exertions can be task-related or sleep-related. Matthews *et al.* (2012) provides a taxonomy of fatigue which differed from Phillips in how the constructs are grouped, but has similar terminology. The task may be physical or mental (Hockey, 2012; Matthews *et al.*, 2012; Phillips, 2015). It has also been postulated that physical and mental fatigue may share similar cognitive mechanisms (Hockey, 2012). Different tasks result in different kinds of fatigue; active fatigue is caused by overload situations whereas passive fatigue is caused by performing repetitive simple tasks (Desmond & Hancock, 2001). A brief review of studies in this area is found in Neubauer *et al.* (2012). It is important to identify the type of fatigue caused in order to develop and deploy effective countermeasures (May & Baldwin, 2009).

Figure 2.1 illustrates the forms of exertion and the types of resulting fatigue. The focus of this thesis is on task-related passive fatigue as experienced by subjects driving under monotonous scenarios, and performing MI during BCI stroke rehabilitation. In both contexts, subjects perform repetitive mental tasks that could lead to passive fatigue.

Figure 2.1 Different forms of exertion leads to different types of fatigue.

2.3 Overview of Passive Fatigue Detection with EEG-based BCI in Drowsy Driving

Drowsy driving is dangerous and may be fatal for road users. It has been estimated to be accountable for 20% of fatal crashes (ETSC, 2001; Noy *et al.*, 2011). In the United States, a recent study by the American Automobile Association (AAA) Foundation for Traffic Safety estimates that drivers were drowsy in 9.5% of crash cases (Owens *et al.*, 2018). Passive fatigue can lead to a state of drowsiness when driving in monotonous scenarios (Gimeno *et al.*, 2006). This is thus a pressing safety issue for drivers (Saxby *et al.*, 2013) especially during automated driving (Vogelpohl *et al.*, 2018).

2.3.1 Detection of Fatigue

There are various methods to detect driver fatigue or drowsiness. Khushaba *et al.* (2013) has reported three main approaches: 1) video-based methods that used image processing to detect features such as head nods; 2) vehicle behavior features such as lane variability; and 3) neurophysiological signals such as EEG. A more comprehensive review of these approaches is available in Balandong *et al.* (2018). A major disadvantage of the video-based and vehicle behavior approaches is that fatigue is only detected when drowsy behavior has already been manifested. In contrast, the approach using neurophysiological signals has the potential to detect fatigue before the onset of drowsy behavior. As drowsy behavior is unsafe for the driver, it is more advantageous to detect fatigue before it is manifested. Thus in many studies and also in this thesis, BCI with the neurophysiological signal EEG is used to detect fatigue.

2.3.2 EEG Correlates of Fatigue

There are many studies on the EEG correlates of passive driver fatigue. Lal and Craig (2002) and Craig *et al.* (2012) have comprehensively reviewed and studied the changes in the power of various frequency bands; delta (0 to 4 Hz), theta (4 to 8 Hz),

alpha (8 to 12 Hz) and beta (12 to 30 Hz). Borghini *et al.* (2014) have also reviewed the EEG frequency band power changes that occurred with fatigue, along with electrooculography (EOG) and heart rate (HR) features. The general finding is that beta band power decreased (Borghini *et al.*, 2012; Jap *et al.*, 2009; Zhao *et al.*, 2012) and theta band power increased (Huang *et al.*, 2016) with passive fatigue. The changes in theta power may however be due to compensatory mechanisms to the fatiguing task (Wascher *et al.*, 2014). Alpha band power also reportedly increases with fatigue (Zhao *et al.*, 2012) but may decrease as well (Foong *et al.*, 2015; Jap *et al.*, 2009).

Other EEG correlates studies include frequency band power ratios (Jap *et al.*, 2009), alpha spindles (Lawhern *et al.*, 2013; Simon *et al.*, 2011; Sonleitner *et al.*, 2012), autoregressive model coefficients (Chai *et al.*, 2017; Zhao *et al.*, 2011b), wavelet packet energy (Zhao *et al.*, 2011a), functional connectivity (Dimitrakopoulos *et al.*, 2018), inter-hemispheric coherence (Jap *et al.*, 2010) and entropies (Hu, 2017) with complexity measures (Chi *et al.*, 2014). Various machine learning methods have been applied for classification and regression of passive fatigue in drowsy driving studies. Classification studies have focused on the prediction of alert and one or more fatigue states (Khushaba *et al.*, 2013), while regression studies have estimated driving performance (Lin *et al.*, 2005a; Lin *et al.*, 2005b) or fatigue level (Liu *et al.*, 2016; Wu *et al.*, 2016). Some of these machine learning methods include support vector machines (SVM) (Yeo *et al.*, 2009; Zhao *et al.*, 2011b), various neural networks (Hajinoroozi *et al.*, 2016; Lin *et al.*, 2012; Liu *et al.*, 2016), and decision trees (Hu & Min, 2018), amongst others. The ensemble approach has been applied (Hu, 2017; Hu & Min, 2018) and transfer learning frameworks proposed as well (Wei *et al.*, 2018; Wu *et al.*, 2016). Most proposed methods are offline, except for a few (Huang *et al.*, 2016; Wu *et al.*, 2016). Table 2.1 provides a summary of these passive fatigue driving studies.

Table 2.1
Summary of Passive Fatigue Driving Simulation Studies

Study	Driving simulation paradigm	No of subj.	EEG features	Contribution	Determination of “ground truth”	Cross-subject model
Lin <i>et al.</i> (2005a)	45min of monotonous driving with lane departures	5	EEG power	<i>Regression</i> PCA with correlation analysis and linear regression model	Deviation from lane	x
Lin <i>et al.</i> (2005b)	2 sessions of: 45min of monotonous driving with lane departures	5	EEG power	<i>Regression</i> ICA with correlation analysis and linear regression model	Deviation from lane	x
Pal <i>et al.</i> (2008)	60min of monotonous continuous driving	13	Theta and alpha power	<i>Regression</i> Mahalanobis distance from multivariate distribution model	Deviation from lane Time-on-task (start of driving is considered alert)	✓
Jap <i>et al.</i> (2009)	10-15min of alert driving 60min of monotonous continuous driving	52	Frequency band power ratios	<i>Statistical analysis</i> ANOVA analysis of 10 time-points of monotonous session compared to alert	-	x
Yeo <i>et al.</i> (2009)	90min of monotonous continuous driving	20	Characteristics of peaks in alpha and beta bands	<i>Classification</i> SVM	Inspection of EEG	✓
Jap <i>et al.</i> (2010)	10-15min of alert driving 60min of monotonous continuous driving	24	Inter-hemispheric coherence	<i>Statistical analysis</i> ANOVA analysis of 5 time-points of monotonous session compared to alert session	-	x

Table 2.1 (cont.)

Study	Driving simulation paradigm	No of subj.	EEG features	Contribution	Determination of “ground truth”	Cross-subject model
Zhao <i>et al.</i> (2011b)	150min of monotonous continuous driving	13	Autoregressive model coefficients	<i>Classification</i> KPCA for feature reduction and 3-class SVM	Time-on-task (start, middle and end of driving)	✓
Zhao <i>et al.</i> (2011a)	90min of monotonous continuous driving	15	Wavelet packet energy	<i>Classification</i> KPCA for feature reduction and 2-class SVM	Time-on-task (start and end of driving)	✓
Lin <i>et al.</i> (2012)	60min of monotonous driving with lane departures at 5-10s intervals	6	Alpha and theta powers from occipital IC	<i>Regression</i> Self-organizing neural fuzzy inference network	Reaction time to lane departure events	✓
Sonnleitner <i>et al.</i> (2012)	60min of driving with emergency braking, and secondary auditory and visuomotor tasks	28	Alpha spindles Alpha and gamma band powers	<i>Statistical analysis</i> ANOVA analysis across 6 blocks of time	-	x
Zhao <i>et al.</i> (2012)	90 minutes of monotonous continuous driving	13	Relative frequency band powers	<i>Statistical analysis</i> Paired t-test	Time-on-task (start and end of driving)	x
Khushaba <i>et al.</i> (2013)	25min of alert driving 60min of monotonous continuous driving	31	Time domain and autoregressive features	<i>Classification</i> Uncorrelated fuzzy neighborhood preserving analysis with any classifier	Observation of subject’s fatigue	✓

Table 2.1 (cont.)

Study	Driving simulation paradigm	No of subj.	EEG features	Contribution	Determination of “ground truth”	Cross-subject model
Lawhern <i>et al.</i> (2013)	70min of monotonous driving with perturbations	2	Alpha spindles	<i>Classification</i> Detection of alpha spindles with autoregressive modelling	Expert rating of spindles	x
Chi <i>et al.</i> (2014)	120min of continuous driving	20	Entropies and complexity measures	<i>Classification</i> 4-class multilayer perceptron neural network	Observation of subject’s fatigue and patterns of entropies	✓
Wu <i>et al.</i> (2016)	60-90min of monotonous driving with lane departures at 5-10s intervals	16	Theta band power	<i>Regression</i> Online weighted adaptation regularization for regression with source domain selection (transfer learning framework)	Drowsiness index based on reaction time	✓
Liu <i>et al.</i> (2016)	90min of monotonous driving with lane departures at 5-10s intervals	20	Frequency band powers	<i>Regression</i> Recurrent self-evolving fuzzy neural network	Reaction time to lane departure events	✓
Huang <i>et al.</i> (2016)	2 sessions of: 4min of alert driving 40-60min of monotonous driving with lane departures at 16-20s intervals	12	Alpha power	<i>Regression</i> Threshold of alpha power based on reaction time to lane departure events	Reaction time to lane departure events	x
Hajinoroozi <i>et al.</i> (2016)	90min of monotonous driving with lane departures at 8-12s intervals	47	Frequency bands and their characteristics	<i>Classification</i> Channel-wise convolutional neural network with restricted Boltzmann machine	Reaction time to lane departure events	✓

Table 2.1 (cont.)

Study	Driving simulation paradigm	No of subj.	EEG features	Contribution	Determination of “ground truth”	Cross-subject model
Chai <i>et al.</i> (2017)	Lane-keeping driving with reaction time task till subject is observed to be fatigued	43	Autoregressive model coefficients	<i>Classification</i> Sparse-deep belief network	Observation of subject’s fatigue Time-on-task (start and end of driving)	✓
Hu (2017)	10min of alert driving Monotonous continuous driving until subject self-reports fatigue	22	Entropies	<i>Classification</i> Gradient boosting decision tree (ensemble)	Self-reported fatigue Time-on-task (last 5min of alert and monotonous driving)	✓
Dimitrakopoulou <i>et al.</i> (2018)	60min of monotonous continuous driving	20	Functional connectivity	<i>Classification</i> SVM	Time-on-task (start and end of driving)	✓
Hu and Min (2018)	20min of alert driving Monotonous continuous driving till subject is observed to be fatigued	28	Entropies	<i>Classification</i> AdaBoost classifier (ensemble)	Observation of subject’s fatigue Time-on-task (start and end of driving)	✓
Wei <i>et al.</i> (2018)	90min of monotonous driving with lane departures at 6-10s intervals	37	Frequency band powers	<i>Regression</i> Weighted fusion of SVR models of other subjects (transfer learning framework)	Drowsiness index based on reaction time	✓

PCA – principal component analysis, ICA – independent component analysis, ANOVA – Analysis of variance, KPCA – kernel principal component analysis, IC – independent component, SVR – support vector regression

2.3.3 Challenges of Fatigue Detection in Drowsy Driving

One challenge is that while it is relatively easy to collect alert data (Wei *et al.*, 2018), it is difficult to label the fatigue state or quantify the fatigue level precisely in drowsy driving experiments. Hence the ground truth of passive fatigue is not easily obtained. As seen in Table 2.1, labeling of the fatigue class was mostly done manually. It could be either labeled by raters via observations (Chai *et al.*, 2017) or labeled by scales (Li & Chung, 2015), such as the Wierwille scale (Wierwille & Ellsworth, 1994). It could also be labeled based on driving performance such as reaction time to driving simulation events (Lin *et al.*, 2012; Tan *et al.*, 2015) or lane variability in the driving simulation (Lin *et al.*, 2005a). A third way of labeling is to use the time-on-task factor, whereby the middle or end of the driving simulation is considered fatigued (Zhao *et al.*, 2011b). Such manual labeling undermines the reliability of using supervised learning methods.

Using unsupervised learning is a way to resolve this problem with the assumption that an alert model can be built from the early minutes of the experiment (Pal *et al.*, 2008). Another way is to use semi-supervised learning whereby the model is able to learn from both labeled and unlabeled data. This is thus a more promising research direction because there is more flexibility in detecting and monitoring fatigue when little labeled data is available.

The second challenge is to develop subject-independent models. As seen in Table 2.1, studies often utilized cross-subject models, as such models were robust against inter-subject variability (Wei *et al.*, 2018). For instance, Wu *et al.* (2016) utilized the transfer learning framework and Wei *et al.* (2018) showed that it was feasible to quantify the fatigue of a new subject using a database of other subjects' alert and fatigue data. By doing so, new subjects' calibration time could be reduced as well (Wei *et al.*, 2018). Thus it is promising to investigate cross-subject fatigue detection models.

A third challenge is to develop more practical real-life implementations using dry EEG sensors (Foong *et al.*, 2017). The fatigue driving studies referred to so far are conducted with multi-channel gel-based EEG electrodes. However, for real-life implementation, it is more convenient to use only a few dry EEG sensors that can be worn easily by drivers. Motivated by this, a few groups had developed wireless wearable EEG fatigue detection systems. Lin *et al.* (2010) positioned the electrode at the Oz position and used an unsupervised method proposed by Pal *et al.* (2008) to detect fatigue. Lin *et al.* (2014) subsequently went on to develop another such system with novel EEG electrodes at the occipital region and used support vector regression (SVR) to predict subjects' response to lane-departure events in their driving simulation. Li and Chung (2015) put electrodes at the O1 and O2 positions and also used a gyroscope. They extracted features from EEG and gyroscope signals for a SVM to detect fatigue. It is thus feasible to use dry EEG electrodes to investigate fatigue detection in drowsy driving.

2.4 Overview of Fatigue Detection in EEG-based MI-BCI Stroke Rehabilitation

Thus far, mental state detection and monitoring has been investigated in the non-medical domain. However, it can also be applied in the medical domain (Mak & Wolpaw, 2009). One such instance is when MI-BCI is employed for stroke rehabilitation.

2.4.1 EEG-based MI-BCI with Only Visual Feedback

Stroke is a major cause of mortality and neuro-disability (Mukherjee & Patil, 2011). It causes much economic and social burden, and is exacerbated by aging population, globalization and urbanization (Mendis, 2013). In Singapore, stroke is the 4th leading cause of death and among the top causes of hospitalization (Venketasubramanian & Chen, 2008). Nearly half of the stroke survivors remain permanently disabled and have a poorer quality of life. Current interventions for stroke survivors ranged from traditional

physical therapy, to the use of technology such as robotic aided therapies, virtual reality enabled therapies, and BCI (Asín Prieto *et al.*, 2014). Preliminary studies have shown that stroke patients are able to operate non-invasive EEG-based MI-BCI as effectively as healthy subjects (Ang *et al.*, 2008a; Ang *et al.*, 2008b, 2011). MI, the mental practice of movements, has been shown to activate similar cortical patterns as actual movements (Claflin *et al.*, 2015; Zich *et al.*, 2015) and is a potential facilitator of neuroplasticity (Monge-Pereira *et al.*, 2017; Remsik *et al.*, 2016; Várkuti *et al.*, 2013). MI-BCI may also be used for patients with severely compromised motor control or residual motor abilities (Guger *et al.*, 2017; Molinari *et al.*, 2016). MI-BCI is not complete without a form of feedback for the user to regulate their brain activity; feedback can be visual or kinesthetic (Vukelić & Gharabaghi, 2015).

A number of randomized controlled trials (RCT) involving various EEG-based MI-BCI systems have been conducted in recent years (Ang *et al.*, 2015a; Ang *et al.*, 2015b; Ang *et al.*, 2014; Frolov *et al.*, 2017b; Kim *et al.*, 2016; Li *et al.*, 2014; Ono *et al.*, 2014; Pichiorri *et al.*, 2015; Ramos-Murguialday *et al.*, 2013). Table 2.2 tabulates their details. Most of these BCI systems are coupled with other interventions such as orthosis and stimulations. Few investigate the effect from the BCI component with just visual feedback (Ang & Guan, 2013). Prasad *et al.* (2010) first showed that EEG-based BCI with visual neurofeedback and physical practice was able to help 5 chronic stroke patients improve functional outcome measures. Subsequently, Pichiorri *et al.* (2015) showed that for 1 chronic and 13 subacute stroke patients, BCI-supported MI training resulted in significantly better outcomes than performing MI without BCI support. Both studies lacked follow-up assessment though it would have been interesting to see if there were any differences in motor scores as time passed. Hence it is worthwhile to investigate the effects of MI-BCI with engaging visual feedback post-intervention and at

Table 2.2
Summary of Recent EEG-Based MI-BCI RCT

RCT	Intervention and Control Groups	Group Size and Type (S / C)	Intervention Sessions	Main Results Post-Intervention (excluding follow-ups)	Significant intergroup differences?
Ramos-Murguialday et al. (2013) <i>Kinesthetic feedback</i>	<i>Intervention:</i> Orthosis was moved upon successful MI detection <i>Control:</i> Orthosis was moved at random	16 C 16 C	~18 sessions over 4 weeks	cFMA +3.4* cFMA +0.35	Not reported
Ang et al. (2015a) <i>Visual and kinesthetic feedback</i>	<i>Intervention:</i> MIT-Manus robot was used with the BCI system and moved upon successful MI detection <i>Control:</i> MIT-Manus robot was used without the BCI system	11 C 15 C	12 sessions x 1.5h over 4 weeks	UE-FMA +6.3* rBSI negatively correlated with FMA UE-FMA +4.5*	No
Ang et al. (2014) <i>Visual and kinesthetic feedback</i>	<i>Intervention:</i> HK was used with the BCI system and moved upon successful MI detection <i>Control 1:</i> HK was used without the BCI system <i>Control 2:</i> No HK or BCI was used, only SAT was administered	6 C 8 C 7 C	18 sessions x 1.5h over 6 weeks	UE-FMA +7.2* UE-FMA +7.3* UE-FMA +4.9*	Yes, between Intervention and Control 2 at mid-intervention and follow-ups
Ono et al. (2014) <i>Visual or kinesthetic feedback</i>	<i>Intervention:</i> Orthosis was moved upon successful MI detection <i>Control:</i> No orthosis was used, only visual feedback was provided	2 S + 4 C 6 C	12-20 sessions x 1h	Some in Intervention group improved SIAS finger function whereas none in Control group did	Not reported
Li et al. (2014) <i>Visual, auditory and FES feedback</i>	<i>Intervention:</i> FES feedback was delivered upon successful MI detection <i>Control:</i> No BCI system was used, FES treatment was given	7 S 7 S	24 sessions x 1-1.5h over 8 weeks	FMA +12.7* ARAT +18* FMA +6.7* ARAT +7.6*	Yes, for ARAT at weeks 6-8 during intervention

Table 2.2 (cont.)

RCT	Intervention and Control Groups	Group Size and Type (S / C)	Intervention Sessions	Main Results Post-Intervention (excluding follow-ups)	Significant intergroup differences?
Pichiorri <i>et al.</i> (2015) <i>Visual feedback</i>	<i>Intervention:</i> BCI-supported MI training was conducted <i>Control:</i> MI training with no BCI assistance was conducted	13 S + 1 C 14 S	12 sessions x 0.5h over 4 weeks	FMA +13.6* FMA +6.5*	Yes
Ang <i>et al.</i> (2015b) <i>Visual and kinesthetic feedback</i>	<i>Intervention:</i> tDCS was given and MIT-Manus robot was used with the BCI system and moved upon successful MI detection <i>Control:</i> Sham tDCS was given and MIT-Manus robot was used with the BCI system and moved upon successful MI detection	10 C 9 C	10 sessions x 1h over 2 weeks	UE-FMA +0.9 UE-FMA+2.8	Yes, EEG laterality coefficients of Intervention group were significantly higher than those of Control group
Kim <i>et al.</i> (2016) <i>FES feedback</i>	<i>Intervention:</i> AOT training with FES feedback triggered by successful MI and minimum attention level <i>Control:</i> No BCI system was used, only conventional therapy was administered	15 C 15 C	20 sessions x 0.5h over 4 weeks	UE-FMA +7.8* UE-FMA +2.9*	Yes
Frolov <i>et al.</i> (2017b) <i>Visual and kinesthetic feedback</i>	<i>Intervention:</i> Exoskeleton was moved upon successful MI detection <i>Control:</i> Exoskeleton was moved according to a repetitive scheme	25 S 30 C 9 S 10 C	10 sessions x 0.5h	UE-FMA +3.0* UE-FMA +7.5* UE-FMA +5.0 UE-FMA +3.0	Not reported

Group type – S: Subacute (1-6 months from onset of stroke), C: Chronic (>6 months from onset of stroke)

Interventions – HK: Haptic Knob, SAT: Standard Arm Therapy, FES: Functional Electrical Stimulation, tDCS: Transcranial Direct Current Stimulation, AOT: Action Observational Training

Main Results – cFMA: Combined Fugl-Meyer Assessment, UE-FMA / FMA: Upper Extremity Fugl-Meyer Motor Assessment, rBSI: Revised Brain Symmetry Index, SIAS: Stroke Impairment Assessment Set, ARAT: Action Research Arm Test

Symbols – *significant results ($p < 0.05$)

follow-ups.

2.4.2 Currently Unexplored: Fatigue in EEG-based MI-BCI Stroke Rehabilitation

A caveat of using MI-BCI that has largely been unexplored in the stroke rehabilitation setting, is the mental fatigue that is induced due to performing mental tasks repetitively (Myrden & Chau, 2015). Mental fatigue can be detected from EEG (Barwick *et al.*, 2012) and is reported to be closely related to BCI performance (Myrden & Chau, 2015). However, no study thus far has investigated EEG for mental fatigue experienced by stroke survivors using EEG-based MI-BCI therapy.

Instead, studies have only reported subjective reports of measures of fatigue. One study used the visual analog scale (VAS) to assess fatigue in the patients subjectively and reported the trend that greater VAS fatigue scores seemed to correspond with larger BCI performance variability (Prasad *et al.*, 2010). Another study made passing remarks that there might be a link between fatigue and a session's cross-validation results based on the subjects' subjective qualitative report of being fatigued (Leamy *et al.*, 2014a). An anecdotal example of self-reported fatigue is in the study by Frolov *et al.* (2017b) where inter-trial or inter-session intervals had to be increased due to the patient's fatigue. Hence, if an EEG-based MI-BCI system is able to detect or predict mental fatigue objectively from EEG, it could provide a better rehabilitation experience by customizing break-times accordingly for example (Frolov *et al.*, 2017b).

2.5 Open Problems, Challenges and Research Objectives

In summary, the literature review has revealed some open problems and challenges in detecting passive fatigue in the areas of drowsy driving and stroke rehabilitation. In drowsy driving, there is a lack of passive fatigue ground truth. There is also a need to move towards dry EEG that can be used in future real-life implementations. In stroke rehabilitation, the efficacy of using just the EEG-based MI-BCI with only visual

feedback, and without any haptic feedback is not yet fully explored. In addition, the presence of fatigue during the MI-BCI rehabilitation sessions has yet to be investigated via EEG.

To address these challenges, this thesis thus has the following research objectives:

1) To develop a semi-supervised learning algorithm for fatigue detection, 2) to detect fatigue in a dry EEG drowsy driving experiment, and 3) to study the efficacy of an EEG-based MI-BCI rehabilitation system with only visual feedback as well as investigate the presence of fatigue. This adds to the passive fatigue literature in both drowsy driving and stroke rehabilitation areas and is a start to bridging the gap between mental state monitoring and medical BCI interventions.

Chapter 3

A Two-Step Iterative Cross-Subject Negative-Unlabeled Learning Algorithm for Detecting and Quantifying Fatigue¹

3.1 Objective

The objective of this chapter is to present a two-step iterative cross-subject negative-unlabeled (NU) learning algorithm for detecting and quantifying fatigue, and compare it to other machine learning approaches. The proposed algorithm addresses the need for cross-subject models and the lack of fatigue ground truth in the drowsy driving context. Part of this chapter has been published by Foong *et al.* (2019b).

3.2 Introduction to Negative-Unlabeled Learning

Collection of data on mental states is not easy as it is difficult to induce the mental state or determine the ground truth on when the mental state has been achieved. It is often easier to collect a baseline data instead. For example, in the case of drowsy driving, it is easier to collect and label the alert state (Wei et al., 2018), rather than the fatigue mental state, as fatigue may not be induced in all subjects or it is difficult to determine when fatigue has occurred. Thus, there is a need for algorithms that learn from the class that is not of interest, in order to find the class that is of interest. This is precisely what NU learning algorithms do. The following sub-sections hence present a brief overview of the types of machine learning algorithms, before explaining the application of NU learning in

¹ **Foong, R.**, Ang, K. K., Zhang, Z., & Quek, C. (2019b). An iterative cross-subject negative-unlabeled learning algorithm for quantifying passive fatigue. *Journal of Neural Engineering*. doi:10.1088/1741-2552/ab255d

the drowsy driving context. The final sub-section then describes positive-unlabeled (PU) learning, the more well-known counterpart to NU learning.

3.2.1 Machine Learning Approaches

The goal of machine learning is to extract patterns from data (Bagherzadeh & Asil, 2018). To achieve this, there are three main machine learning approaches: supervised, unsupervised, and semi-supervised. These are summarized in Figure 3.1.

Supervised learning utilizes labeled data from both positive and negative classes to build a model that discriminates between the classes (Kotsiantis, 2007). This approach thus does well when the labels are accurate. However, if labels are not certain or provided, unsupervised learning is utilized instead to build a model that learns the underlying distribution of the data (Hofmann, 2001). In cases where there are some labeled data as well as unlabeled data, semi-supervised learning is then employed (Zhu, 2008) to enhance the model that would otherwise be just built with supervised or

Figure 3.1 Machine learning approaches.

unsupervised machine learning algorithms.

In the above scenarios, it is assumed that data from all classes are available. However, there are cases where the labeled data comes from only one class. These are thus the two special subsets of semi-supervised learning. One subset is the one-class learning (Khan & Madden, 2014) where there is some labeled data from one class and a large amount of unlabeled data of unknown classes. In this approach, only the labeled data is utilized to learn a model that subsequently identifies which of the unlabeled data do not belong to the class.

The other special subset of semi-supervised learning is PU and NU learning. They are similar to one-class learning in that there is only labeled data from one class and a large amount of unlabeled data with unknown labels. If the labeled data is from the positive class, it is termed PU learning. Conversely if the labeled data is from the negative class, it is termed NU learning (Niu *et al.*, 2016). PU and NU learning differ from one-class learning in that information in the unlabeled data is utilized in learning a model to discriminate between the classes.

As such, it can be inferred that the machine learning approach to be used in a study depends on the data available. For example, if there is sufficient and reliable labeled data, supervised learning is suitable. If instead there is labeled data from only one class as well as unlabeled data, one-class or PU and NU learning is more suitable.

3.2.2 Negative-Unlabeled Learning in the Drowsy Driving Context

In the drowsy driving context, the fatigued data is the class of interest and is therefore denoted as the positive class, while alert data is then the negative class. To label alert (negative class) and fatigued (positive class) data from the driving data, many studies have used ratings (Chai *et al.*, 2017; Li & Chung, 2015), driving performance (Lin

et al., 2005a; Lin *et al.*, 2012; Tan *et al.*, 2015) or the time-on-task factor (Zhao *et al.*, 2011b). The supervised learning approach is then utilized. However, these labels are based on surrogate measures and thus may not reflect the ground truth of when passive fatigue is present. Thus, the supervised learning approach may not be suitable.

To address this problem, unsupervised learning algorithms that treat the driving data as unlabeled had been developed. For example, the fatigue (positive class) state was identified based on the distances of the unlabeled driving data's features from that of the baseline alert (negative class) model (Pal *et al.*, 2008). This is also somewhat similar to the one-class learning approach as well. However, unsupervised learning and one-class learning neglect to utilize the useful information in the unlabeled data (Bekker & Davis, 2018). Thus instead, the NU learning approach is more suitable, as it is able to learn from both baseline alert (negative class) data as well as unlabeled driving data.

3.2.3 Positive-Unlabeled Learning

As PU learning is more prevalent in literature than NU learning, it is introduced here before presenting the proposed NU learning algorithm. Its definition by Li and Liu (2005) is as follows:

“Given a set of examples of an particular class P (called the positive class) and a set of unlabeled examples U , which contains both class P and non-class P (called the negative class) instances, the goal is to build a binary classifier to classify the test set T into two classes, positive and negative, where T can be U .”

PU learning and its applications have been growing over the years. It is often used in document or text classification (Lee & Liu, 2003; Li & Liu, 2003; Liu *et al.*, 2003). It has also found applications in time series classification (Nguyen *et al.*, 2011; Nguyen *et*

al., 2012), derivation of protein interaction networks (Kılıç & Tan, 2012), inference of drug interaction (Hameed *et al.*, 2017), identification of disease gene (Yang *et al.*, 2012), and remote sensing in terms of anomaly detection (Li *et al.*, 2011). Refer to Bekker and Davis (2018) and Kılıç and Tan (2012) for comprehensive surveys of various PU learning algorithms, as well as Niu *et al.* (2016) for vigorous theoretical treatments and comparisons of PU, NU and supervised learning. In this thesis, a new application of NU learning in drowsy driving is proposed.

There are two main methods of PU learning (Kılıç & Tan, 2012). The first method is a two-step framework whereby firstly, features from the labeled positive set is used to iteratively extract reliable negative examples from the unlabeled set via a machine learning method tailored to the application. Subsequently, the machine learning method is applied again to identify the classes of the remaining unlabeled examples. The second method is the construction of a one-step estimator that makes use of the knowledge obtained from the positive and unlabeled sets.

3.3 Proposed Two-Step Iterative Negative-Unlabeled Learning Algorithm

A two-step iterative NU learning algorithm similar to the two-step PU learning algorithm (Kılıç & Tan, 2012) is proposed. The first step uses the available class and unlabeled data to iteratively extract the other class. Following which, the second step then applies the trained model on the rest of the unlabeled data. As a classifier that gives probabilistic estimates is desired, a SVM with cost parameter C and a radial basis function (RBF) kernel with width parameter γ is used. The definition of the kernel acting on the i^{th} and j^{th} feature vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ is shown by equation (3.1),

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2\right) \quad (3.1)$$

where $\|\mathbf{x}_i - \mathbf{x}_j\|$ is the Euclidean distance.

The SVM is implemented with the LIBSVM software package available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm> (Chang & Lin, 2011). The optimized parameters C^* and γ^* are found via a grid search method and 10-fold cross-validation with all subjects' data as a whole dataset. During the grid search, weighted SVM is used as the unlabeled data is treated as the positive class and has its cost parameter value compensated according to its class size relative to the negative class. The grid search and NU learning algorithm are presented as a pseudo-code in Algorithm 3.1, and the

Algorithm 3.1 Determination of positive class blocks and computation of positive class score

Input: Number of blocks in unlabeled data, R
Number of samples in each subject's negative class, L
Number of samples in each subject's unlabeled data, G , where $G \leftarrow L \times R$
Number of features, F
Number of subjects, S
Set of blocks, $\mathcal{R} \leftarrow \{1, \dots, R\}$
Set of negative samples, $\mathcal{L} \leftarrow \{1, \dots, L\}$
Set of unlabeled samples, $\mathcal{G} \leftarrow \{1, \dots, G\}$
Set of features, $\mathcal{F} \leftarrow \{1, \dots, F\}$
Set of subjects, $\mathcal{S} \leftarrow \{1, \dots, S\}$
Negative class feature set, $\mathcal{N} \leftarrow \{n_{l,f,s} \mid l \in \mathcal{L}, f \in \mathcal{F}, s \in \mathcal{S}\}$
Unlabeled data feature set, $\mathcal{U} \leftarrow \{u_{g,f,s} \mid g \in \mathcal{G}, f \in \mathcal{F}, s \in \mathcal{S}\}$

Output: Positive class blocks, $\mathcal{B}$
Mean accuracy across subjects, A
Positive class score, P

Parameters: Initial block labeled as positive, λ
Stopping criterion, η

```
// Initialization
Outer loop counter,  $i \leftarrow 0$ 
Inner loop counter,  $j \leftarrow 0$ 
// Grid search for optimal cost and width parameters (respectively  $C$  and  $\gamma$ ) with SVM  $\Pi$  that outputs accuracy
for  $v \in \{0, 0.1, \dots, 2\}$  do
   $i \leftarrow i + 1$ 
  for  $\omega \in \{0, 0.1, \dots, 2\}$  do
     $j \leftarrow j + 1$ 
    Cost parameter for negative class,  $C \leftarrow 10^v$ 
    Cost parameter for unlabeled data,  $D \leftarrow C / R$ 
    Width parameter of RBF kernel,  $\gamma \leftarrow 10^\omega$ 
    Cross-validation data,  $\mathcal{V} \leftarrow \mathcal{N} \cup \mathcal{U}$ 
    Number of cross-validation folds,  $k$ 
    Accuracy,  $\{a_{i,j}\} \leftarrow \Pi(k, \mathcal{V}, C, D, \gamma)$ 
  end
end
 $\mathcal{A} \leftarrow \cup \{a_{i,j}\}$ 
 $(i^*, j^*) \leftarrow \operatorname{argmax}_{i,j} \mathcal{A}$ 
 $v^* \leftarrow (i^* - 1) / 10$ 
 $\omega^* \leftarrow (j^* - 1) / 10$ 
Optimal cost parameter,  $C^* \leftarrow 10^{v^*}$ 
Optimal width parameter,  $\gamma^* \leftarrow 10^{\omega^*}$ 
```

algorithm's application in drowsy driving is described in the following sub-sections.

3.3.1 Step One: Identifying the Most Fatigued Block

As it is possible to use other subjects' alert and driving data to find a subject's fatigue data (Wei *et al.*, 2018), a cross-subject SVM model is trained whereby all but one subject are training subjects. The remaining subject is then the test subject. For training

Algorithm 3.1 (cont.)

// Step one: Determination of $\mathcal{B}$ with SVM Ψ that outputs probabilistic estimates and accuracy

// Initialization

Outer loop counter, $i \leftarrow 1$

Positive class blocks, $\mathcal{W} \leftarrow \{w_{s,j} = \lambda \mid s \in \mathcal{S}, j = 1\}$

Accuracy of positive class blocks, $\Gamma \leftarrow \{\gamma_{s,j} = 0 \mid s \in \mathcal{S}, j = 1\}$

Average probabilistic estimates of positive class blocks, $\mathcal{P} \leftarrow \{p_{s,j} = 0 \mid s \in \mathcal{S}, j = 1\}$

Mean of accuracy, positive class blocks across subjects, $\mathbf{T} \leftarrow \{\tau_j = 0 \mid j = 1\}$

Mean of average probabilistic estimates, $\mathcal{M} \leftarrow \{m_j = 0 \mid j = 1\}$

Difference in mean from previous iteration, $\Delta \leftarrow \{\delta_j = 1 \mid j = 1\}$

// Iteration to determine $\mathcal{B}$

while $\{\delta_j \mid j = i\} > \eta$ **do**

for $z \in \mathcal{S}$ **do**

 Accuracy, $\mathbf{K} \leftarrow \emptyset$

 Average probabilistic estimates, $\mathbf{M} \leftarrow \emptyset$

for $r \in \mathcal{R}$ **do**

 Negative class test data, $\mathcal{H} \leftarrow \{n_{l,f,s} \mid (n_{l,f,s} \in \mathcal{N}) \wedge (s = z)\}$

 Unlabeled test data, $\mathcal{J} \leftarrow \{u_{g,f,s} \mid (u_{g,f,s} \in \mathcal{U}) \wedge ((r-1)*L+1 \leq g \leq r*L) \wedge (s = z)\}$

 Test data, $\mathcal{T} \leftarrow \mathcal{H} \cup \mathcal{J}$

 Test labels, $\mathcal{L} \leftarrow \{0, \dots, 0\}^L \cup \{1, \dots, 1\}^L$

 Negative class train data, $\mathcal{I} \leftarrow \{n_{l,f,s} \mid (n_{l,f,s} \in \mathcal{N}) \wedge (s \neq z)\}$

 Unlabeled train data, $\mathcal{K} \leftarrow \emptyset$

for $k \in (\mathcal{S} \setminus z)$

 Unlabeled data block, $\mathcal{V} \leftarrow \{w_{s,j} \mid (w_{s,j} \in \mathcal{W}) \wedge (s = k) \wedge (j = i)\}$

 Unlabeled train data for subject k , $\mathcal{O} \leftarrow \{u_{g,f,s} \mid (u_{g,f,s} \in \mathcal{U}) \wedge ((r-1)*L+1 \leq g \leq r*L) \wedge (s = k)\}$

$\mathcal{K} \leftarrow \mathcal{K} \cup \mathcal{O}$

end

 Train data, $\mathcal{X} \leftarrow \mathcal{I} \cup \mathcal{K}$

 Train labels, $\mathcal{Z} \leftarrow \{0, \dots, 0\}^{(S-1)*L} \cup \{1, \dots, 1\}^{(S-1)*L}$

$\mathbf{K} \leftarrow \mathbf{K} \cup \{\kappa_q = \Psi(\mathcal{X}, \mathcal{Z}, \mathcal{T}, \mathcal{L}, C^*, \gamma^*) \mid q = r\}$

 Probabilistic estimates, $\Phi \leftarrow \{\phi_{q,l} = \Psi(\mathcal{X}, \mathcal{Z}, \mathcal{T}, \mathcal{L}, C^*, \gamma^*) \mid q = r, l \in \mathcal{L}\}$

$\mathbf{M} \leftarrow \mathbf{M} \cup \{\mu_q = \frac{\sum_{l=1}^L \phi_{q,l}}{L} \mid (\phi_{q,l} \in \Phi) \wedge (q = r) \wedge (l \in \mathcal{L})\}$

end

$\mathcal{W} \leftarrow \mathcal{W} \cup \{w_{s,j} = \arg\max_q \mathbf{M} \mid (s = z) \wedge (j = i+1)\}$

$\Gamma \leftarrow \Gamma \cup \{\gamma_{s,j} = \kappa_q \mid (\kappa_q \in \mathbf{K}) \wedge (q = w_{s,j}) \wedge (w_{s,j} \in \mathcal{W}) \wedge (s = z) \wedge (j = i+1)\}$

$\mathcal{P} \leftarrow \mathcal{P} \cup \{p_{s,j} = \mu_q \mid (\mu_q \in \mathbf{M}) \wedge (q = w_{s,j}) \wedge (w_{s,j} \in \mathcal{W}) \wedge (s = z) \wedge (j = i+1)\}$

end

$i \leftarrow i + 1$

$\mathbf{T} \leftarrow \mathbf{T} \cup \{\tau_j = \frac{\sum_{z=1}^S \gamma_{z,j}}{S} \mid (\gamma_{z,j} \in \Gamma) \wedge (j = i) \wedge (z \in \mathcal{S})\}$

$\mathcal{M} \leftarrow \mathcal{M} \cup \{m_j = \frac{\sum_{z=1}^S p_{z,j}}{S} \mid (p_{z,j} \in \mathcal{P}) \wedge (j = i) \wedge (z \in \mathcal{S})\}$

$\Delta \leftarrow \Delta \cup \{\delta_j = m_j - m_{j-1} \mid (m_j \in \mathcal{M}) \wedge (j = i)\}$

end

$\mathcal{B} \leftarrow \{b_s = w_{s,j} \mid (w_{s,j} \in \mathcal{W}) \wedge (s \in \mathcal{S}) \wedge (j = i)\}$

$\mathbf{A} \leftarrow \{\alpha = \tau_j \mid (\tau_j \in \mathbf{T}) \wedge (j = i)\}$

subjects, their alert data is designated as the negative class while for their driving blocks, the λ block is designated as the positive class in the first iteration. Subsequently, the trained model is evaluated on the test subject's alert data and each driving block data in turn. The evaluation produces probabilistic estimates of samples in the test data belonging in the respective negative (alert) or positive (fatigue) class, as well as the accuracy of the model on the test data. These estimates are computed by the LIBSVM software, using a modified Platt's scaling (Lin *et al.*, 2007). The average score across samples in each driving block is then computed. The block that results in the highest average score for the test subject is then selected for use as the positive block in the next iteration. In each iteration, this is done for all subjects being the test subject, i.e. there is 1 test subject and 28 train subjects, and this is repeated for all 29 subjects as the test subject in turn.

The mean of subjects' selected block's score is then computed. Following which, the algorithm repeats the process with the driving blocks that are newly-labeled as positive. These iterations stop when the change in the mean across subjects' selected block's scores is less than the stopping criterion η . The first output from this step is $\mathcal{B}$, a set containing the subject-specific blocks, one per subject, that most likely belong in the positive class and are interpreted as the most fatigued blocks. The second output from this step is the mean accuracy $\mathcal{A}$ across subjects of their selected blocks.

3.3.2 Step Two: Computing the Fatigue Score

Using the final selected blocks $\mathcal{B}$ from step one as the positive (fatigue) class data, and the alert data as the negative class data, a new supervised SVM model is computed for each subject in a leave-one-subject-out fashion. This means that each subject is the test subject in turn, and all other subjects' alert and subject-specific fatigued blocks are used as training data. The test data is the test subject's driving data, and thus

Algorithm 3.1 (cont.)

// Step two: Use of $\mathcal{B}$ to obtain positive class score $\mathbf{P}$ for unlabeled data $\mathcal{U}$

```
for  $z \in \mathcal{S}$  do
  Test data,  $\mathcal{Q} \leftarrow \{u_{g,f,s} \mid (u_{g,f,s} \in \mathcal{U}) \wedge (s = z)\}$ 
  Test labels,  $\mathcal{G} \leftarrow \{1, \dots, 1\}^G$ 
  Negative class train data,  $\mathcal{J} \leftarrow \{n_{l,f,s} \mid (n_{l,f,s} \in \mathcal{N}) \wedge (s \neq z)\}$ 
  Unlabeled train data,  $\mathcal{K} \leftarrow \emptyset$ 
  for  $k \in (\mathcal{S} \setminus z)$  do
    Unlabeled data block,  $\mathcal{V} \leftarrow \{b_s \mid (b_s \in \mathcal{B}) \wedge (s = k)\}$ 
    Unlabeled train data for subject  $k$ ,  $\mathcal{O} \leftarrow \{u_{g,f,s} \mid (u_{g,f,s} \in \mathcal{U}) \wedge ((\mathcal{V} - 1) * L + 1 \leq g \leq \mathcal{V} * L) \wedge (s = k)\}$ 
     $\mathcal{K} \leftarrow \mathcal{K} \cup \mathcal{O}$ 
  end
  Train data,  $\mathcal{Y} \leftarrow \mathcal{J} \cup \mathcal{K}$ 
  Train labels,  $\mathcal{Z} \leftarrow \{0, \dots, 0\}^{(S-1)*L} \cup \{1, \dots, 1\}^{(S-1)*L}$ 
   $\mathbf{P} \leftarrow \{\rho_{s,l} = \Psi(\mathcal{Y}, \mathcal{Z}, \mathcal{Q}, \mathcal{G}, C^*, \gamma^*) \mid s = z, l \in \mathcal{L}\}$ 
end
```

the model outputs the fatigue score for the test subject's driving data, where the fatigue score is the probabilistic estimate of each sample belonging to the positive (fatigue) class.

3.3.3 Comparison with Current PU Learning Algorithms

Other than the class of interest, there are differences between the proposed NU learning algorithm and the current PU learning algorithms. They are that: 1) the unlabeled data used in the proposed algorithm comprised of multiple blocks, 2) the proposed algorithm also took into consideration which data belonged to which subject, and at the end of the first step, only one block per subject is labeled as the positive class, and 3) the iterative process of identifying the subject-specific positive block required a block to be initialized as the positive class. These differences are due to the different contexts; the PU learning algorithms were mainly developed in the text classification context (Lee & Liu, 2003; Li & Liu, 2003; Liu *et al.*, 2003) where the dataset was treated as a whole with no subject or time elements, whereas in the drowsy driving EEG context, it is taken into consideration that there are different subjects with EEG that change over time.

3.4 Summary

The NU learning approach has been found suitable for the drowsy driving context. A novel two-step iterative NU learning algorithm that used a alert (negative class) data

and driving (unlabeled) data to find fatigue (positive class) data in a cross-subject fashion has thus been proposed.

Chapter 4

Detecting and Quantifying Passive Fatigue in a Drowsy Driving Study with Dry EEG Electrodes²

4.1 Objective

The objective of this chapter is to apply the NU learning algorithm proposed in Chapter 3 to detect and quantify passive fatigue in a drowsy driving experiment. A small amount of alert data and large amount of unlabeled driving data is collected from subjects using commercially available dry EEG Muse™ headband. Part of this chapter has been published by Foong *et al.* (2019b).

4.2 Materials and Methods

Ethics approval from the Institution's Domain Specific Review Board, National Healthcare Group, Singapore, has been obtained for this study. Subjects are duly informed and have given their consent before participating in the experiment.

The following sub-sections describe the experimental set-up and protocol, the preprocessing and segmentation of the EEG data, and the EEG features extracted. They also present the parameters of the NU learning algorithm, baseline comparisons for the algorithm, and correlation of the fatigue score obtained by the algorithm with relative EEG beta band power.

4.2.1 Experimental Set-up and Protocol

This experimental set-up is first reported by Foong *et al.* (2017). A driving

² **Foong, R.**, Ang, K. K., Zhang, Z., & Quek, C. (2019b). An iterative cross-subject negative-unlabeled learning algorithm for quantifying passive fatigue. *Journal of Neural Engineering*. doi:10.1088/1741-2552/ab255d

simulation environment with a desktop computer, a Logitech gaming seat (Playseat® Evolution TopGear) and Thrustmaster steering wheel (T500 RS) is set up. A Tobii eye-tracker and Microsoft Kinect are utilized to respectively capture eye gaze and video data. EEG is concurrently collected with 10-bit resolution at 256 Hz with the commercially available Muse™ headband, which has dry EEG electrodes placed at the Fp1, Fp2, TP9 and TP10 positions. This is similar to previous studies which have used commercially available headsets with pre-frontal channels to successfully detect fatigue (Morales *et al.*, 2017; Ogino & Mitsukura, 2018; Rohit *et al.*, 2017). Figure 4.1a shows the whole set-up as well as a subject performing the experiment.

Alert (negative) EEG data is collected via a target hitting game which took approximately 7 minutes. In this game, subjects had to look out for and click on red dots using a mouse as they randomly appeared (Figure 4.1b shows the game's graphical user interface). The game consisted of 10 quick rounds and is designed to be short and engaging to ensure that subjects are alert.

Unlabeled driving EEG data is collected via an hour-long driving simulation (see Figure 4.1a for the simulation interface). Subjects faced a monotonous straight highway driving scene and performed the repetitive task of monitoring a virtual car from a first-

Figure 4.1 The experimental set-up with the hitting game and driving simulation interfaces.

person view. They were to turn the wheel to steer the virtual car when it deviated from the straight path. The deviations were random in direction (left or right) and occurrence (random intervals of 5 to 10s). Subjects had a maximum of 5s to respond. This car deviation protocol was similar to the lane-departure events conducted by Lin *et al.* (2005a). The objective of the monotony in scene and task was to induce passive fatigue. Subjects' sleepiness rating are collected every 10 minutes via the 9-point Karolinska Sleepiness Scale (KSS) (Akerstedt & Gillberg, 1990).

A total of 31 healthy subjects participated in this study. After scanning through the recorded videos for any obvious artifacts, the data from 29 subjects are analyzed while the data from the remaining 2 subjects are excluded due to recording errors. The 29 subjects are aged 38.9 ± 9.3 years old and consisted of 9 females and 20 males.

4.2.2 Preprocessing and Segmentation of EEG

EEG data from the MuseTM headset (Fp1, Fp2, TP9 and TP10 channels) are first preprocessed with a notch filter to remove the 50 Hz line noise. It is then filtered with a fourth-order Butterworth bandpass filter of 0.3 to 64 Hz. Subsequently, a fast ensemble empirical mode decomposition method (Wang *et al.*, 2014) is applied, followed by a canonical correlation analysis publicly available at https://github.com/germangh/eeglab_plugin_aar (De Clercq *et al.*, 2006), in order to reduce muscle noise contamination (Whitham *et al.*, 2007). Thereafter, an algorithm by Chang *et al.* (2016) which is publicly available at <http://cone.hanyang.ac.kr/BioEST/Eng/EyeblinkMaster.zip>, and that utilized multi-window summation of first derivatives within a window, is used to detect eye blink regions. The contaminated signals within those regions are then replaced with cleaned signals, computed via a multi-window-size moving average.

To further remove remaining artifacts, Krigolson *et al.* (2017) MuseTM study's criteria of rejecting data points with values greater than 100 μ V and gradients more than 40 μ V per sample is used. The 2 bipolar signals are then obtained by subtracting TP9 from Fp1 and Fp2.

After these preprocessing, the EEG signal from each of the 6 10min runs in the driving simulation was segmented into 3 equal and non-overlapping blocks. Thus, every subject had 1 block of negative class data, comprised of the alert data from the hitting game, as well as 18 blocks of unlabeled data from the driving simulation.

4.2.3 EEG Feature Extraction

On each continuous EEG epoch, 2s sliding windows with 1s overlap are used to extract relative frequency band powers. 8 frequency band powers with bounds 0.3 to 4 Hz, 4 to 6 Hz, 6 to 8 Hz, 8 to 12 Hz, 12 to 18 Hz, 18 to 30 Hz, 30 to 50 Hz and 50 to 64 Hz are chosen for their importance in drowsy driving, and are mostly matched to various common brainwaves such as delta, theta, alpha and beta (Zhang & Guan, 2017).

Fast Fourier transform (FFT) is performed on the EEG signal to obtain the power spectrum P . The total power P_{total} is then calculated by summing the power spectrum. This is shown in equation (4.1),

$$P_{total} = \sum_{f=F_{min}}^{F_{max}} P(f) \quad (4.1)$$

where $P(f)$ is the power of frequency f , F_{min} is the lowest cut-off frequency of 0.3 Hz, and F_{max} is the highest cut-off frequency of 64 Hz. Subsequently, the k^{th} relative frequency band power P_k is computed by equation (4.2),

$$P_k = \frac{\sum_{f=f_k}^{f_{k+1}} P(f)}{P_{total}} \quad (4.2)$$

where f_k and f_{k+1} are respectively the lower and upper bound of the k^{th} frequency band.

After calculating the relative frequency band powers, their upper and lower envelopes are computed. This is first done by finding the peaks and troughs. The set p of the windows containing peaks are found via equation (4.3),

$$p = \arg \min_{2 \leq i \leq W} (\text{sgn}(x_{i+1} - x_i) - \text{sgn}(x_i - x_{i-1})) \quad (4.3)$$

where x_i is the value of a feature in the i^{th} window and W is the total number of windows. The set t of the windows containing troughs is found in a similar manner as shown in equation (4.4) below.

$$t = \arg \max_{2 \leq i \leq W} (\text{sgn}(x_{i+1} - x_i) - \text{sgn}(x_i - x_{i-1})) \quad (4.4)$$

The peaks and troughs are then the values of the band power at the windows in p and t respectively. This is followed by a cubic spline with not-a-knot end condition between the peaks and troughs to obtain the respective upper and lower envelopes. The cubic spline is implemented with MATLAB® function *interp1*.

The Gaussian parameters, specifically the mean and standard deviation, of the 8 relative frequency band powers and their envelopes over 5 windows are then calculated via equations (4.5) and (4.6) respectively,

$$M = \frac{1}{E} \sum_{i=1}^E y_i \quad (4.6)$$

$$S = \sqrt{\frac{1}{E-1} \sum_{i=1}^E |y_i - M|^2} \quad (4.5)$$

where M and S are respectively the mean and standard deviation across E windows, y_i is the value of a feature in the i^{th} window, and $E = 5$. As there are 2 Gaussian parameters

each for the 8 relative frequency band powers and their 2 envelopes, there are thus a total of $2 \times 3 \times 8 = 48$ features per bipolar signal. As there are two bipolar signals, there are therefore a total of 96 features.

These features are used as they have been proven to be useful in a sleep detection study by Zhang and Guan (2017) for utilizing the periodic nature of EEG as well as the Gaussian distribution of the spectral features for a better feature representation. As passive fatigue is related to sleepiness and drowsiness, these features are suitable for use in this study.

After extracting the features, to ensure a balanced classification, the minimum sample size across all subjects' alert and driving block data is found. Only that minimum number of samples is kept. Also, to ensure a decent sample size for classification, subjects are to be rejected if they had alert or driving block data with fewer than 30 samples. No subjects ended up being rejected as all fulfilled this criterion, and each block contained 36 samples. This results in a sufficient sample size ($36 \times 28 = 1008$ samples) per class for training each subject's SVM model.

4.2.4 Negative-Unlabeled Learning Algorithm Parameters

As per Algorithm 3.1 described in Section 3.3, a grid search is first conducted to find the optimized parameters C^* and γ^* . During the grid search, weighted SVM is used, where the unlabeled data is treated as the positive class and had a cost parameter value that was $1/18^{\text{th}}$ that of the negative class (alert) data to reflect the differences in class sizes. C^* and γ^* are found to be 10^1 and $10^{0.4}$ respectively.

In addition, the block pseudo-labeled as the positive (fatigue) class in the first iteration is set to $\lambda = 9$ since subjects provided post-experiment feedback that they started to be fatigued in the middle of the driving simulation.

Also, an appropriate value for the stopping criterion to terminate the iterative first step of the NU learning algorithm is selected as $\eta = 0.0001$. This small value is selected to ensure that the convergence of the algorithm was stable. As per Section 3.3.1, the output from this step of the algorithm is the cross-subject accuracy results, as well as the subject-specific most fatigued driving block $\mathcal{B}$.

4.2.5 Comparison of Proposed Algorithm with Baselines

To compare the accuracies of the Proposed Algorithm (PA) versus the baselines, a MATLAB® code from Öner and Deveci Kocakoç (2017) is first used to perform the Shapiro-Wilk test (Ghasemi & Zahediasl, 2012) to test the normality of the accuracies. The distributions of the accuracies are found to be not normal. As such, paired two-sided Wilcoxon signed-rank tests are then conducted to test the hypothesis that the difference in accuracies came from a distribution with zero median. The multiple comparisons are corrected with the Holm-Bonferroni method (Holm, 1979) in R. Table 4.1 shows the comparison of the PA with baselines as detailed in the following sub-sections.

4.2.5.1 Assumptions of Driving and Alert Data

In this study, the PA used the hitting game data as alert data. In other studies, the alert model had often been built on data from the early minutes of the driving simulation (Chai *et al.*, 2017; Hu & Min, 2018; Huang *et al.*, 2016). To compare with such current methods, *what if a similar assumption of alertness in the first driving block is made?* To address this question, an Alert Block (AB) baseline is computed, whereby the first driving block is used as the alert data. It then underwent the same process as in the PA, where it iteratively labeled the most fatigued block per subject. Based on these labeled blocks $\mathcal{B}_{AB}$, the accuracy is calculated accordingly. It is also noted that the AB baseline uses the NU learning approach.

Table 4.1
Comparison of PA with Baselines

	Alert Data	Most Fatigued Driving Block		Machine Learning Approach	Labeling Methods from Drowsy Driving Literature
		Training	Testing		
PA	Hitting game	$\mathcal{B}$		NU	-
AB	1 st driving block	$\mathcal{B}_{AB}^{\#}$		NU	Alert driving (Chai <i>et al.</i> , 2017; Hu & Min, 2018; Huang <i>et al.</i> , 2016)
OA	Hitting game	$\mathcal{B}_{OA}^{\wedge}$		One-class	-
FD	Hitting game	All blocks	$\mathcal{B}$	Supervised	-
RT	Hitting game	Driving block with slowest reaction time	$\mathcal{B}$	Supervised	Reaction time (Lin <i>et al.</i> , 2012; Liu <i>et al.</i> , 2016; Wei <i>et al.</i> , 2018; Wu <i>et al.</i> , 2016)
NI9	Hitting game	9 th driving block	$\mathcal{B}$	Supervised	Time-on-task (Dimitrakopoulos <i>et al.</i> , 2018; Zhao <i>et al.</i> , 2011a; Zhao <i>et al.</i> , 2011b)
NI18	Hitting game	18 th driving block	$\mathcal{B}$	Supervised	
SB9	Hitting game	9 th driving block		Supervised	
SB18	Hitting game	18 th driving block		Supervised	

$\mathcal{B}_{AB}^{\#}$ is the most fatigued driving blocks, one per subject, identified by the first step of the PA in an iterative manner in the AB condition, i.e. when the 1st driving block is used as alert data

$\mathcal{B}_{OA}^{\wedge}$ is the most fatigued driving blocks, one per subject, identified by the one-class SVM in the OA condition, i.e. when only the alert hitting game data is used to find the most fatigued driving blocks

Another feature of the PA is that it used both negative class (i.e. the hitting game alert data) and unlabeled data to find the most fatigued driving block. *What if only the alert data is used?* This then is the Only Alert (OA) baseline, where a one-class SVM (Schölkopf *et al.*, 2001) with default parameters is implemented in LIBSVM (Lin *et al.*, 2007). The one-class SVM learns from only the alert data to find the most fatigued subject-specific driving blocks $\mathcal{B}_{OA}$ and the final accuracy across subjects is calculated accordingly. It is noted that the OA baseline uses the one-class learning approach, and is similarly used by Mordelet and Vert (2014) to benchmark their PU learning algorithm.

Additionally, the driving data are assumed to contain both alert and fatigue data. As such, the driving data are considered unlabeled data, and subsequently the PA found $\mathcal{B}$, the subject-specific blocks labeled as fatigue data, one block per subject. *What if all driving blocks are assumed to be fatigued instead?* To address this question, a Fatigue Driving (FD) baseline is computed, whereby all 18 blocks of driving data is considered

fatigue data in the training models. The cost parameter C is compensated accordingly for the different class sizes. For a fair comparison, $\mathcal{B}$ is used to test the accuracy of the trained models. It is also noted that the FD baseline used the supervised learning approach for unbalanced classes, and is similarly used by Kılıç and Tan (2012) as a benchmark for their PU algorithms.

4.2.5.2 Identification of the Most Fatigued Block

In this study, PA has identified most fatigued blocks without ground truth. In other studies, surrogate measures for the ground truth of fatigue have been used, such as reaction time (Lin *et al.*, 2012; Liu *et al.*, 2016; Wei *et al.*, 2018; Wu *et al.*, 2016). *What if reaction time is used to identify most fatigued blocks instead?* This is then the Reaction Time (RT) baseline, where the most fatigued block for each subject is the driving block with the slowest mean reaction time to the car deviation events. This is used in training, and for a fair comparison, testing is done on the $\mathcal{B}$ blocks found by the PA. As this baseline differs from the FD baseline only in the training data, they can be compared as well. It is also noted that the RT baseline used the supervised learning approach.

To identify $\mathcal{B}$, the subject-specific most fatigued block, the PA used an iterative process. As other studies have used a fixed time block for all subjects as fatigued, such as in the middle or end of the simulation, based on the time-on-task assumption (Dimitrakopoulos *et al.*, 2018; Zhao *et al.*, 2011a; Zhao *et al.*, 2011b), *was this iterative process necessary?* To address this question, No Iteration (NI) baselines are computed whereby the same block per subject is used for training instead of $\mathcal{B}$. As subjects often feedback that they felt fatigued in the middle as well as at the end of the driving simulation, NI baselines with block 9 and block 18 are computed, respectively termed NI9 and NI18. For a fair comparison with PA, the cross-subject models for the NI baselines are tested on $\mathcal{B}$. A comparison between NI9 and NI18 will be made as well. In

addition, the NI baselines will be compared against the FD and RT baselines since they differed only in the training data. It is noted that the NI baselines used the supervised learning approach.

Additionally, $\mathcal{B}$ is subject-specific. *Was it necessary for the fatigued blocks to be subject-specific instead of using the same block for each subject?* To address this question, Same Block (SB) baselines are computed, whereby the same block for all subjects is used for training and testing. Similar to the NI baselines, the SB baselines are computed with blocks 9 and 18, respectively termed SB9 and SB18. Other than comparing against the PA, SB9 and SB18 will be compared with each other as well. Additionally, the SB baselines will be compared to their counterpart NI baselines as they differed only in the testing data. It is noted that the SB baselines used the supervised learning approach.

4.2.6 Fatigue Score and its Correlation with Beta Band Power

The fatigue score for each subject is computed by the second step of the PA. To analyze the fatigue scores, they are correlated to the EEG beta (12 to 30 Hz) band power as a reduction in beta band power is associated with decreased alertness and brain arousal (Borghini *et al.*, 2012; Jap *et al.*, 2009; Zhao *et al.*, 2012). The means of the relative EEG beta band power and fatigue score in each driving block are obtained and a repeated measures correlation (Bakdash & Marusich, 2017) is performed in R.

4.3 Results

The following sub-sections present the accuracies of the PA and baselines, as well as the fatigue scores of two subjects and the correlation between fatigue scores and beta band power.

4.3.1 Comparison of Proposed Algorithm with Baselines

The PA is compared with various baselines: AB, OA, FD, RT, NI9, NI18, SB9, and SB18 in terms of accuracy, which is the percentage of correctly classified samples. Table 4.2 presents the mean and standard deviation of the across-subject accuracies of PA and the baselines while Figure 4.2 presents the comparison in boxplots. Table 4.2 also shows the significance of the difference in their medians, as tested by a paired two-sided Wilcoxon signed-rank test.

The results in Table 4.2 showed that the averaged accuracy of PA ($93.77\% \pm 8.15\%$) significantly outperformed AB ($76.96\% \pm 11.19\%$) with $p = 1.01 \times 10^{-4}$, OA

Table 4.2
Comparison of PA with Baselines in Terms of Accuracy

	PA	AB	OA	FD	RT	NI9	NI18	SB9	SB18
PA	$93.77\% \pm 8.15\%$	***	***	***	*	*	N.S.	***	**
AB		$76.96\% \pm 11.19\%$	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
OA			$74.52\% \pm 12.78\%$	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
FD				$68.92\% \pm 9.84\%$	***	***	***	N.A.	N.A.
RT					$89.51 \pm 8.78\%$	N.S.	*	N.A.	N.A.
NI9						$90.80\% \pm 7.42\%$	*	***	N.A.
NI18							$92.86\% \pm 7.43\%$	N.A.	*
SB9								$81.56\% \pm 12.21\%$	*
SB18									$88.12\% \pm 8.06\%$

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ where p is the Holm-Bonferroni-adjusted p -value, N.S. = Non-significant, N.A. = Non-applicable

($74.52 \pm 12.78\%$) with $p = 6.59 \times 10^{-5}$, FD ($68.92\% \pm 9.84\%$) with $p = 4.28 \times 10^{-5}$, , RT ($89.51 \pm 8.78\%$) with $p = 0.016$, NI9 ($90.80\% \pm 7.42\%$) with $p = 0.016$, SB9 ($81.56\% \pm 12.21\%$) with $p = 6.06 \times 10^{-5}$ and SB18 ($88.12\% \pm 8.06\%$) with $p = 0.0013$. PA was also better than NI18 ($92.86\% \pm 7.43\%$), but not significant with $p = 0.47$.

The results showed that the FD baseline is found to perform significantly worse than RT, NI9 and NI18, all with $p = 4.28 \times 10^{-5}$, while the RT baseline was significantly worse than NI18 but not NI9 with $p = 0.012$ and $p = 0.47$ respectively. In addition, NI9 was significantly worse than NI18 with $p = 0.026$, but was significantly better than SB9

Figure 4.2 Comparison of across-subject accuracies of the proposed algorithm (PA) against that of baselines alert block (AB), only alert (OA), fatigue driving (FD), reaction time (RT), no iteration with block 9 (NI9), no iteration with block 18 (NI18), same block with block 9 (SB9) and same block with block 18 (SB18). Accuracy is the percentage of correctly classified samples. Baselines associated with NU and one-class learning algorithms are in magenta while baselines associated with supervised learning are in blue.

with $p = 9.61 \times 10^{-4}$. Also, SB9 was significantly worse than SB18 with $p = 0.026$, while SB18 was significantly worse than NI18 with $p = 0.017$. All above reported p-values are Holm-Bonferroni-adjusted.

4.3.2 Convergence of the Proposed Algorithm

Figure 4.3 shows the convergence of the PA in 7 iterations, with a final error across subjects of $8.46\% \pm 7.56\%$. The stopping criterion used is a change of less than 0.01% in the mean error across subjects, where the error is the average of the probabilistic estimates of samples belonging to the other class.

Figure 4.3 Convergence of the mean error across subjects over 7 iterations in the first step of the PA.

The error is the average of the probability estimates of samples belonging to the other class. The stopping criterion is a change in the mean error of less than 0.01%, which took the PA 7 iterations to achieve. A final mean error of $8.46 \pm 7.56\%$ is achieved.

4.3.3 Fatigue Score and its Correlation with Beta Band Power

Figure 4.4 shows the smoothed fatigue scores averaged across all subjects as well as two example subjects, S4 and S5. The results showed that though there was much fluctuation, there was a general increase in the fatigue score as time passed and subjects completed more blocks. The averaged fatigue score does not increase greatly due to variability in the fatigue score across subjects. This is illustrated by S4 and S5. S5's fatigue score generally increases over time whereas S4's fatigue score slightly dipped near the end.

Figure 4.5 shows the repeated measures correlation of the fatigue score to the relative EEG beta band power. The results showed that there was a significant negative correlation ($R = -0.112$, $p < 0.05$).

4.4 Discussion

The following sub-sections discuss the results of the comparison of the PA with baselines, convergence of the PA, and the fatigue score and its correlation with relative EEG beta band power.

4.4.1 Comparison of Proposed Algorithm with Baselines

The results in Table 4.2 showed that there were significant differences between the PA and the various baselines. A discussion on these differences is detailed in the following sub-sections.

4.4.1.1 Assumptions of Driving and Alert Data

The PA significantly outperformed the AB baseline. The mean accuracy across subjects for the AB baseline was 76.96%, which indicated that the most fatigued blocks chosen by AB were somewhat similar to the first driving block such that the classifier was not able to differentiate the classes with high accuracy. This might be due to some

(a)

(b)

(c)

Figure 4.4 Fatigue scores as calculated in the second step of the PA are smoothed. (a) This displays the average across all subjects, where a general increase over driving blocks is seen. (b) S4 displays a slight dip in their score in the later blocks. (b) S5 shows a more obvious trend of increasing fatigue score.

Figure 4.5 Repeated measures correlation of mean score against mean beta band power for each of the 29 subjects' 18 driving blocks. A significant negative correlation is found.

subjects being fatigued near the beginning of the driving simulation. In contrast, the PA which utilized a short engaging game to ensure subjects' alertness achieved a mean accuracy of 93.77%. Thus, in contrast to studies which assume alertness in the early part of the driving simulation (Chai *et al.*, 2017; Hu & Min, 2018; Huang *et al.*, 2016), this study's results showed the merit of conducting a short engaging simulation or game before collecting monotonous driving simulation data, in order to better obtain EEG features of the alert state and detect the fatigue state.

Additionally, the PA significantly outperformed the OA baseline. This shows that OA's model which is built on only one class' data is not as effective as PA's that utilizes information in the unlabeled data as well. Another possible reason is that the OA model is sensitive to outliers and free parameters' tuning (Li *et al.*, 2011). However, it is also noted

that while this has been observed in this study, it may not be true in other applications, as discussed by Mordelet and Vert (2014) that PU learning is not always superior to one-class SVM in all contexts.

The PA also significantly outperformed the FD baseline. This suggested that the driving data did not contain only fatigued data, but also alert data. Hence this validated the motivation for modelling the driving data as unlabeled data and using NU learning to find and label the blocks most likely to belong to the positive (fatigue) class.

4.4.1.2 Identification of the Most Fatigued Block

In addition, the PA significantly outperformed the RT baseline (Lin *et al.*, 2012; Liu *et al.*, 2016; Wei *et al.*, 2018; Wu *et al.*, 2016). While the RT baseline achieved a decent mean accuracy of 89.51%, the PA outperforming it suggested that the surrogate measure of reaction time might not be the best way to gauge when subjects are fatigued. It also suggested that the PA is able to detect fatigue well despite the lack of ground truth or surrogate measures of fatigue. Thus, future studies might not need to include secondary tasks to collect surrogate measures as it might be intrusive to the subjects' mental state (Arico *et al.*, 2018). Instead, fatigued states may be detected by the PA via EEG.

The results in Table 4.2 showed that PA's accuracy was significantly better than NI9's, though both PA and NI9 are not significantly better than NI18. In addition, the NI baselines also achieved significantly higher accuracies than their SB counterparts. Altogether, these indicated that the performance dropped when the same block for subjects is assumed as the most fatigued block. This suggested that the subject-specific fatigue blocks $\mathcal{B}$ are well-identified, also justifying the need for an iterative process of identifying $\mathcal{B}$.

In addition, NI18 and SB18 had significantly higher accuracy than NI9 and SB9 respectively. This suggested that subjects were generally more fatigued at the end than in the middle. Despite this, the PA which initiated the iterative process with the 9th driving block was able to identify the most fatigued block at the end of the iterations, since the PA significantly outperformed the NI9 and SB baselines.

Table 4.2 also shows that the RT and NI baselines performed significantly better than the FD baseline. This was probably due to the FD baseline using all the driving data as fatigue data when it contained both alert and fatigue data. Thus the RT and NI baselines which only used one driving block per subject performed better. It is further observed that the RT baseline was not significantly different from the NI9 baseline but is significantly worse than NI18. This suggests that reaction time may not be a good surrogate measure; instead the time-on-task factor may be stronger.

Overall, all these validated the ability of the iterative method in the PA to identify the subject-specific fatigued blocks rather than assuming that all blocks or the same block for all subjects as fatigued as other studies have done (Dimitrakopoulos *et al.*, 2018; Zhao *et al.*, 2011a; Zhao *et al.*, 2011b), or even when the initial block was not the most ideal.

4.4.1.3 Other Machine Learning Approaches

It is noted that the PA and AB baseline utilized the NU learning approach, while the OA baseline utilized the one-class learning approach, and the other baselines utilized the supervised learning approach.

The PA significantly outperforming the OA baseline shows that there is useful information in the unlabeled driving data that the one-class learning approach neglects. The PA significantly outperforming all baselines that use the supervised learning approach, except the NI18 baseline, also shows that the labeling of fatigue by rating,

driving performance, or time-or-task are not ideal as they may not reflect the ground truth of fatigue. This is thus an advantage of NU learning over supervised learning in the drowsy driving context where it is difficult to accurately label fatigue data.

It is also observed that the supervised learning approach is comparable to the PA for the NI18 baseline, when the last block of driving data was labeled as fatigued for train subjects. This suggests a disadvantage of the PA compared to supervised learning, in that the iterative process may not be necessary when the fatigued state is relatively well-labeled. This disadvantage may be overcome by one-step NU learning algorithms in future investigations. However, the comparable performance of the PA to the supervised learning NI18 baseline shows that it can work as well as supervised learning though the fatigue labels are not present.

4.4.2 Convergence of the Proposed Algorithm

Figure 4.3 presents the convergence of the PA. The results showed that it terminated in 8 iterations when the change in the mean score across subjects was less than 0.01%. It is observed that the final error was not the minimum across the iterations. This was due to the stopping criterion that did not take into account the direction of the change in mean score so as to avoid local extremes and ensure a final stable result. As the final mean error was low at 8.86% across subjects and achieved at the 8th iteration, the PA is demonstrated to be able to converge and work well enough.

4.4.3 Fatigue Score and its Correlation with Beta Band Power

The results in Figure 4.4 showed that there was a general increase in the fatigue score as time passed and subjects completed more driving blocks. Subjects also showed fluctuations within the general trend, suggesting that passive fatigue is not a mental state experienced for long periods of time, but is experienced at different levels over time. This might be reflective of the microsleeps (Lin *et al.*, 2005a) or lapses (Peiris *et al.*, 2011) that

other similar drowsy driving experiments had observed. However, there are also subject-specific differences. It is observed that the increasing trend is stronger and more evident for S5 than S4. Furthermore, in contrast to S5, S4 seemed to have started out the experiment with some fatigue and experienced more fluctuations. These observations showed that subjects might be similar in that they experienced increasing fatigue, but might also have subject-specific differences.

To investigate the validity of the PA, the repeated measures correlation of fatigue score with relative EEG beta band power, is computed. As shown in Figure 4.5, there was a significant negative correlation. This meant that a higher fatigue score is associated with a lower beta band power, which according to literature (Borghini *et al.*, 2012; Jap *et al.*, 2009; Zhao *et al.*, 2012) was indicative of lower alertness. This thus validated the fatigue score produced by the PA.

4.5 Summary

In this study, a quick and engaging target hitting game is designed to collect alert labeled EEG data from 29 healthy subjects. Unlabeled driving EEG data is then collected via a monotonous hour-long driving simulation and divided into 18 equal non-overlapping blocks. The mean and standard deviation of 8 band powers and their upper and lower envelopes from two bi-polar channels are then extracted as features. Following which, the proposed two-step iterative NU learning algorithm yielded a high accuracy of $93.77\% \pm 8.15\%$ across subjects, which was significantly better than various baselines. This demonstrated the PA's capability in identifying the subject-specific most fatigued driving blocks as well as the value of doing so. The results also suggested that there was merit in obtaining alert data from a quick engaging game prior to monotonous driving simulations. Furthermore, the results showed that the fatigue score computed by the PA was significantly negatively correlated with relative EEG beta band power that is known

to decrease in the presence of passive fatigue. This demonstrated the PA's effectiveness in scoring subjects' fatigue level. These results thus show promise in the PA's potential to detect fatigue from EEG data. Moving forward, an online detection and warning system could be developed to alert drivers experiencing passive fatigue from monotonous driving.

Chapter 5

Assessment of the Efficacy of EEG-based MI-BCI with Visual Feedback and the Presence of Mental Fatigue in Upper-Limb Stroke Rehabilitation³

5.1 Objective

The objective of this study is to investigate the efficacy of using the Neurostyle Brain Exercise Therapy Towards Enhanced Recovery (nBETTER) system together with standard arm therapy for upper-limb rehabilitation of subacute and chronic hemiparetic stroke survivors, and explore the aspect of mental fatigue in MI-BCI stroke rehabilitation. Though fatigue may negatively affect stroke patients' BCI performance due to monotonously performing MI, it has yet to be investigated. Part of this chapter has been published by Foong *et al.* (2019a).

5.2 Materials and Methods

This study has obtained ethics approval from the Institution's Domain Specific Review Board, National Healthcare Group, Singapore. The trial is registered under NCT02765334 in ClinicalTrials.gov. Participants have given their informed consent before enrolling in the study.

The following sub-sections describe the nBETTER system, the clinical protocol, retrospective comparison groups and patient demographics. They also describe the measurement of clinical efficacy of the nBETTER system, computation of the correlation

³ **Foong, R.**, Ang, K. K., Quek, C., Guan, C., Phua, K. S., Kuah, C. W. K., . . . Chua, K. S. G. (2019a). Assessment of the Efficacy of EEG-based MI-BCI with Visual Feedback and EEG Correlates of Mental Fatigue for Upper-Limb Stroke Rehabilitation. *IEEE Transactions on Biomedical Engineering*, 1-1. doi:10.1109/TBME.2019.2921198

of the relative EEG beta band power to BCI performance and application of the proposed NU learning algorithm on this dataset.

5.2.1 The nBETTER system

The nBETTER system is a portable internet-connected device that detected the imagination of movement of the stroke-affected limb. It is an EEG-based MI-BCI that provided visual feedback for exercising the brain towards better recovery after stroke. The nBETTER system was conceptualized in continuation of the study by Ang *et al.* (2014). This study improved Ang *et al.*'s BCI system's gameplay in an effort to increase engagement with the subject, but did not include the haptic knob in order to study only the effect of BCI without haptic feedback (Ang & Guan, 2013). Figure 5.1 shows a subject using the nBETTER system.

The nBETTER system collected EEG data using the Neurostyle acquisition hardware (see <http://neuro-style.com/> for more details) with 24 unipolar Ag/AgCl channels placed in the international 10-20 system positioning: F3, F4, FC3, FC4, C3, C4, CP3, CP4, P3, P4, FT7, FT8, T3, T4, TP7, TP8, Fz, Oz, FCz, Cz, CPz, Pz, A1 and A2. These electrode positions are used as they provide full-head overage, including the motor

Figure 5.1 Subject using the nBETTER system.

cortex, which is relevant for MI detection (Pfurtscheller *et al.*, 2006). The EEG is digitally sampled at 256 Hz with a resolution of 24 bits for voltage ranges of $\pm 300\text{mV}$. The impedance level is ensured to be below $5\text{ k}\Omega$ before proceeding with the experiment.

5.2.2 Clinical Protocol

The following sub-sections describe the flow of the clinical trial, subject inclusion and exclusion criteria and MI-BCI sessions conducted during the clinical trial.

5.2.2.1 Flow of Clinical Trial

This multisite clinical trial is conducted over ~ 1.5 year period from February 2016 to September 2017, involving stroke survivors recruited from Tan Tock Seng Hospital and National University of Singapore. Figure 5.2 shows the consolidated standards of reporting trials (CONSORT) flow chart of the clinical trial.

Subjects who fulfilled the inclusion criteria underwent a 40-minute MI-BCI screening session, and those who achieved a BCI accuracy above 57.5% are recruited into the study. In total, 13 stroke patients are recruited, 7 from Tan Tock Seng Hospital, and 6 from National University of Singapore. They then completed a 40-minute MI-BCI calibration session that was identical to the screening session. This is followed by 6 weeks of thrice-weekly intervention sessions supervised by an occupational therapist and bioengineer. Each of the 18 intervention sessions consisted of a 60-minute nBETTER supervised and therapy session, and a 30-minute standard arm therapy session with appropriate rest breaks.

Out of the 13 recruited subjects, 2 dropped out and 11 completed the study. One subject dropped out due to social reasons prior to starting the study intervention, and the other dropped out due to a second stroke occurring 2 days after his 5th nBETTER training

Figure 5.2 CONSORT diagram showing the flow of the study from enrolment to analysis.

session (Figure 5.2). This is deemed not to be directly related to the nBETTER therapy. Thus, a total of 11 subjects underwent the calibration session and the nBETTER sessions.

5.2.2.2 Subject Inclusion and Exclusion Criteria

This study included subjects aged 21-80 years old with first-ever clinical stroke, either ischaemic or haemorrhagic, diagnosed through computed tomography (CT) or magnetic resonance imaging (MRI). Subjects were to have a stroke duration of 3 to 24 months, and a Fugl-Meyer Motor Assessment (FMA) score (Fugl-Meyer *et al.*, 1975) of 10–50 for their affected upper limb. Subjects also had to have the ability to pay attention, maintain supported sitting for 1.5 hours continuously, give their own consent and understand simple instructions. Finally, subjects were to achieve a BCI accuracy above

the chance level of 57.5% in the screening MI-BCI session (Ang *et al.*, 2014).

Subjects are excluded if they had recurrent clinical stroke, or other conditions that were likely to affect their participation in the experiment, such as severe aphasia or inattention; unresolved sepsis; postural hypotension; end stage renal failure; a life expectancy of less than a year due to malignancy or neurodegenerative disorder; a history of epilepsy, severe depression or active psychiatric disorder; and hemispatial neglect or visual impairment. Other exclusion criteria included contraindications to transcranial magnetic stimulation (TMS) such as pregnancy, orthodontics, cardiac pacemakers, metal implants, or cranial surgery. Subjects were also not to have contraindications to repetitive arm training, such as severe spasticity with modified Ashworth scale (MAS) (Bohannon & Smith, 1987) > 2 in any region, pain with visual analog scale (VAS) (Price *et al.*, 1999) $> 4/10$, or fixed joint contractures. Subjects are also excluded if they had poor skin conditions, or skull defects affecting the fit of the EEG cap.

5.2.2.3 MI-BCI Sessions

For all MI-BCI sessions, subjects are instructed to perform upper-arm kinesthetic MI of their affected side for the MI trials; they are specifically instructed to repeatedly “imagine to move” the stroke-affected arm and hand forward to reach for an imaginary target in front of them. They are also advised to keep still and minimize any body movements. For the Idle trials, subjects are instructed to relax, keep still and stare straight at the computer screen.

MI is detected using the Filter Bank Common Spatial Pattern (FBCSP) algorithm, described in detail by Ang *et al.* (2012). Specifically, the EEG signal was band-passed into several frequency bands within 4 to 40 Hz. This removed eye artifacts which were maximal at frequencies below 4 Hz (Fatourechi *et al.*, 2007). The common spatial pattern (CSP) algorithm is then used for each frequency band to detect the event-related

desynchronization/synchronization that occurs when MI is performed. The best CSP features are then selected by mutual information methods to build a subject-specific model. Finally, a fisher linear discriminant classifier is used to detect MI. The BCI accuracy was thus the classification accuracy output of the FBCSP in detecting MI.

The following sub-sections describe the various MI-BCI sessions and standard arm therapy sessions in the clinical protocol. Figure 5.3 shows the trials in the various MI-BCI sessions.

5.2.2.3.1 MI-BCI screening and calibration sessions

The screening and calibration sessions were identical. Subjects performed MI and Idle tasks as instructed. In each session, subjects completed a total of 4 runs, each consisting of 20 MI trials and 20 Idle trials. Each trial lasted about 12s (Figure 5.3a). Accuracies of these sessions are obtained via a 10 by 10-fold cross validation. The purpose of the screening session was to ensure that the subject was able to perform

(a) A trial in the MI-BCI screening and calibration sessions. In each of these sessions, 4 runs, each consisting of 20 MI and 20 Idle trials, are performed.

(b) A trial in the nBETTER supervised and therapy sessions. In each supervised session, 1 run of 20 MI and 20 Idle trials is performed. In each therapy session, 4 runs of 40 MI trials each are performed.

Figure 5.3 Trials in the MI-BCI screening and calibration sessions and in the nBETTER supervised and therapy sessions.

kinesthetic MI which the system was able to detect. Thus, as per the study by Ang *et al.* (2014), only those whose accuracies were above the chance level of 57.5% are recruited. The calibration session is then used by the system to train a subject-specific model to detect MI performed by the subject in subsequent sessions.

5.2.2.3.2 nBETTER supervised session

In the nBETTER supervised session, subjects performed both the MI and Idle tasks but with the nBETTER interface which provided visual feedback; as subjects successfully performed MI, a picture is slowly revealed. In this session, subjects completed 1 run consisting of 20 MI and 20 Idle trials. Each trial lasted about 13s, including 1s of game play providing visual feedback (Figure 5.3b). The purpose of the supervised session was to tune the subject-specific model built during the calibration session. Figure 5.4 illustrates how the model is updated for each subject. Further details and analysis of this adapted model can be found in a paper by Zhang *et al.* (2018).

5.2.2.3.3 nBETTER therapy session

The nBETTER therapy session was similar to the nBETTER supervised session, except that subjects performed only the MI trials for a total of 4 runs, each consisting of 40 MI trials. The purpose of the nBETTER therapy session was to promote neuroplasticity in the brain towards recovery of movement on the affected side.

5.2.2.3.4 Standard arm therapy session

The standard arm therapy session was modelled along neurodevelopmental techniques and included passive and active-assisted mobilization, abnormal tone management, functional reach and grasp-release training, and arm ergometer exercise.

5.2.3 Retrospective Comparison Groups

This study's clinical protocol is designed to be similar to the clinical study by Ang *et al.* (2014). In Ang *et al.*'s study, the BCI-Haptic knob (HK) group had similarly undergone 60 minutes of BCI intervention and 30 minutes of standard arm therapy. In

Figure 5.4 Illustration of the subject-specific model update process, where the model is updated after every supervised session and used for the same day’s therapy session and the subsequent supervised session.

addition, Ang *et al.*’s standard arm therapy (SAT) control group which underwent 90 minutes of standard arm therapy per session is used as a retrospective control group for this study. Table 5.1 shows the similarities and differences between the groups’ protocol.

5.2.4 Patient Demographics

Table 5.2 shows the demographics and baseline characteristics of the nBETTER group along with the BCI-HK and SAT groups in the study by Ang *et al.* (2014) used for retrospective comparison. A one-way analysis of variance (ANOVA) is conducted to test

Table 5.1
Similarities and Differences Between the nBETTER, SAT and BCI-HK Groups' Protocol

	nBETTER Group	SAT Group (Ang <i>et al.</i> , 2014)	BCI-HK Group (Ang <i>et al.</i> , 2014)
Pre-Intervention			
MI-BCI Screening	√	√	√
MI-BCI Calibration	√	-	√
Intervention Session			
Session frequency	3 sessions per 6 consecutive weeks		
Classifier adaptation	√	-	-
60 minutes of intervention	nBETTER: 160 MI trials with visual feedback	SAT: repetitive task training	BCI-HK: 120 MI trials with haptic feedback
30 minutes of physical therapy	Therapist-assisted arm mobilization		

for intergroup differences in the demographics and baseline characteristics.

5.2.5 Clinical Efficacy of the nBETTER Intervention

Clinical efficacy is measured by the FMA score as it is a reliable assessment of stroke patients' upper extremity strength (Alt Murphy *et al.*, 2015). The FMA score is measured at 4 time points: baseline (Week 0), post-intervention (Week 6), and follow-up weeks (Week 12 and Week 24).

To test the normality of each group's FMA scores, the Shapiro-Wilk test (Ghasemi & Zahediasl, 2012) is used. The MATLAB® implementation of this test by Öner and Deveci Kocakoç (2017) is utilized.

To analyze the effects due to intervention and over time, a two-way mixed-design ANOVA is conducted. Intervention was the between-subject factor with 3 levels (nBETTER, SAT, BCI-HK), and time was the within-subject repeated measure factor with 4 levels (Weeks 0, 6, 12 and 24). Subsequently, post-hoc analysis for the time effect is performed on each group's FMA score at time-points of Weeks 6, 12 and 24 against that at baseline (Week 0). This is done via one-sample paired t-tests with Bonferroni correction for multiple comparisons. These statistical tests are carried out in MATLAB®.

Table 5.2
Demographics and Baseline Characteristics of the nBETTER, SAT and BCI-HK Groups

	nBETTER Group	SAT Group (Ang <i>et al.</i> , 2014)	BCI-HK Group (Ang <i>et al.</i> , 2014)
N	11	7	6
Age (years)	55.2 $\pm$ 11.0	58.0 $\pm$ 19.3	54.0 $\pm$ 8.9
Gender, N (%)			
Male	6 (54.5%)	4 (57.1%)	4 (66.7%)
Female	5 (45.5%)	3 (42.9%)	2 (33.3%)
Stroke Type, N (%)			
Infarction	8 (72.7%)	5 (71.4%)	2 (33.3%)
Hemorrhagic	3 (27.3%)	2 (28.6%)	4 (66.7%)
Paresis Side, N (%)			
Right	8 (72.7%)	3 (42.9%)	4 (66.7%)
Left	3 (27.3%)	4 (57.1%)	2 (33.3%)
Stroke Nature, N (%)			
Cortical	3 (27.3%)	3 (42.9%)	1 (16.7%)
Subcortical	8 (72.7%)	4 (57.1%)	5 (83.3%)
Stroke Period, N (%)			
Subacute (1-6 months post-stroke)	2 (18.2%)	0 (0.0%)	0 (0.0%)
Chronic (>6 months post-stroke)	9 (81.8%)	7 (100.0%)	6 (100.0%)
Post-stroke duration (days)	333.7 $\pm$ 179.6	455.4 $\pm$ 109.6	285.7 $\pm$ 64.0
Baseline Characteristics (Week 0)			
FMA	35.5 $\pm$ 12.6	23.4 $\pm$ 14.0	30.7 $\pm$ 17.2

5.2.6 Correlation of Beta Band Power and BCI Performance

As mentioned in Section 2.3.2, it is known in the literature that beta activity decreases in the presence of fatigue (Ahn *et al.*, 2016; Borghini *et al.*, 2012; Jap *et al.*, 2009; Tanaka *et al.*, 2012; Zhao *et al.*, 2012). To understand how fatigue affects the BCI performance, 3-second pre-cue EEG relative beta band power is investigated for correlation with the nBETTER supervised sessions' accuracies.

Before computing the EEG beta activity, the raw EEG is preprocessed channel-wise. Firstly, Wang *et al.*'s fast ensemble empirical mode decomposition method (Wang *et al.*, 2014) with noise level 0.4, 10 ensembles and 8 intrinsic mode functions is performed. Subsequently, canonical correlation analysis (De Clercq *et al.*, 2006) implemented with publicly available code is used with default parameters (https://github.com/germangh/eeglab_plugin_aar). This is followed by a next-nearest

neighbor Laplacian with the available channels. These steps are performed altogether to reduce muscle noise contamination in the EEG (Fitzgibbon *et al.*, 2015; McFarland, 2015; Whitham *et al.*, 2007). The beta activity is then computed as follows: a 3-second pre-cue EEG segment (see Figure 5.3b) is first extracted from each trial; total power in 4 to 50 Hz is then computed as it removes artifacts in the low and high frequency range; subsequently beta frequency (12 to 30 Hz) band power is calculated for each trial and normalized by the total power; finally relative EEG beta band power is expressed in decibels.

Relative EEG beta band power is then averaged among the channels in each brain region [frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal-occipital (P3, Pz, P4, Oz)]. To analyze the correlation between the relative EEG beta band power and BCI performance with consideration to each subject having 18 supervised sessions, a repeated measures correlation (Bakdash & Marusich, 2017) is computed in R for each brain region.

5.2.7 Application of Proposed Negative-Unlabeled Learning Algorithm

To further investigate the passive fatigue experienced by subjects during the nBETTER supervised sessions, the proposed NU algorithm in Section 3.3 is applied on this study's data. In the absence of alert data from the stroke dataset, the driving dataset will be used as train data instead. Thus, all driving subjects' alert and subject-specific fatigue driving data obtained from the driving study are used to train a SVM model, with the same parameters as in Section 4.2.4.

EEG data from the nBETTER supervised sessions are then pre-processed in the same manner as the driving EEG, as described in Section 4.2.2. However, due to the channels available in this stroke study, instead of the bipolar channels Fp1-TP9 and Fp2-TP9 used in the driving study, bipolar channels F3-TP7 and F4-TP7 are used. Subsequently, the same features are extracted as detailed in Section 4.2.3. After pre-

processing and feature extraction, 6 subjects had data from all 18 sessions, 4 subjects 17 sessions, and 1 subject 16 sessions. These then form the test data.

Thereafter, the SVM model outputs the fatigue score for the 11 stroke patients' nBETTER supervised sessions. Following which, the Pearson's correlation of the mean fatigue score against the BCI accuracy of the supervised sessions is computed.

5.3 Results

The following sub-sections report the results of the patient demographics, the clinical efficacy of the nBETTER intervention, and correlation of relative EEG beta band power to BCI performance, subjects' fatigue scores and their correlation with BCI performance.

5.3.1 Patient Demographics

Using a one-way ANOVA, no significant intergroup differences in the demographics and baseline characteristics are found: age ($F_{(2,21)} = 0.16$, $p = 0.86$), gender ($F_{(2,21)} = 0.11$, $p = 0.90$), stroke type ($F_{(2,21)} = 1.45$, $p = 0.26$), paresis side ($F_{(2,21)} = 0.79$, $p = 0.47$), stroke nature ($F_{(2,21)} = 0.51$, $p = 0.61$), stroke period ($F_{(2,21)} = 1.26$, $p = 0.30$), post-stroke duration ($F_{(2,21)} = 2.64$, $p = 0.10$), and baseline FMA score ($F_{(2,21)} = 1.61$, $p = 0.22$).

5.3.2 Clinical Efficacy of the nBETTER Intervention

A retrospective comparison is performed with the SAT and BCI-HK groups in the study by Ang *et al.* (2014). All groups' FMA scores at all time points are tested to be normal. Table 5.3 shows the mean and standard deviation of each group's FMA score gains relative to Week 0. Additionally, Figure 5.5 shows a boxplot for each group's FMA score gains. The results in Figure 5.5 showed that the median of the nBETTER and BCI-

HK groups' FMA improvement increased over time while the SAT group's decreased. In addition, the nBETTER and BCI-HK groups also had smaller variance in FMA score improvement compared to the SAT group.

Table 5.3
Improvement in FMA Score

Group	Baseline FMA (Week 0)	Improvement in FMA score relative to Week 0		
		Week 6	Week 12	Week 24
nBETTER	35.5 ± 12.6	4.4 ± 3.6*	3.5 ± 6.7	5.8 ± 5.6*
SAT	23.4 ± 14.0	4.9 ± 4.1	3.6 ± 5.5	3.6 ± 5.9
BCI-HK	33.0 ± 16.2	7.2 ± 2.3*	8.2 ± 2.9*	9.7 ± 2.9*

*significant improvement, $p < 0.05$

FMA Score Improvement Relative to Week 0

Figure 5.5 Boxplot showing the improvement in FMA scores at post-intervention (Week 6) and follow-ups (Week 12 and 24) for the nBETTER, SAT and BCI-HK groups.

Using a two-way mixed-design ANOVA, no significant interaction effect ($F_{(6,63)} = 1.59, p = 0.17$) is found. There was also no significant intervention effect between groups ($F_{(2,21)} = 1.89, p = 0.18$). However, there was a significant main effect within groups ($F_{(3,63)} = 20.25, p = 2.64 \times 10^{-9}$) over time. Thus, separate post-hoc analysis within each group with one-sample paired t-tests is performed for time points Weeks 6, 12 and 24 compared to Week 0. Bonferroni correction is used for multiple comparisons.

The results in Table 5.3 showed that relative to Week 0, the nBETTER group demonstrated significant FMA score gains at Week 6 [(M = 4.4, SD = 3.6), $t_{(10)} = -4.07, p = 0.0068$] and Week 24 [(M = 5.8, SD = 5.6), $t_{(10)} = -3.46, p = 0.019$] but not Week 12 [(M = 3.5, SD = 6.7), $t_{(10)} = -1.72, p = 0.35$]. Relative to Week 0, the SAT group had no significant FMA score gains at Week 6 [(M = 4.9, SD = 4.1), $t_{(6)} = -3.10, p = 0.063$], Week 12 [(M = 3.6, SD = 5.5), $t_{(6)} = -1.71, p = 0.42$] and Week 24 [(M = 3.6, SD = 5.9), $t_{(6)} = -1.60, p = 0.48$]. The BCI-HK group had significant FMA score gains at Week 6 [(M = 7.2, SD = 2.3), $t_{(5)} = -7.58, p = 0.0019$], Week 12 [(M = 8.2, SD = 2.9), $t_{(5)} = -6.83, p = 0.0031$] and Week 24 [(M = 9.7, SD = 2.9), $t_{(5)} = -6.83, p = 0.0014$]. All t-test p-values reported above are Bonferroni-corrected.

5.3.3 Correlation of Beta Band Power and BCI Performance

Figure 5.6 shows the repeated measures correlations between relative EEG beta band power and BCI performance, in terms of classification accuracy, of 18 supervised sessions per subject. The results showed that there were significant positive correlations in the frontal and central, but not in the parietal-occipital, brain regions [frontal (F3, Fz, F4), $R = 0.251, p < 0.01$; central (C3, Cz, C4), $R = 0.181, p < 0.05$; parietal-occipital (P3, Pz, P4, Oz), $R = 0.033, p > 0.05$].

Figure 5.6 Plot of repeated measures correlations of relative EEG beta band power against BCI performance in terms of classification accuracy. Significant positive correlations are found in frontal and central, but not parietal-occipital, brain regions.

5.3.4 Application of Proposed Negative-Unlabeled Learning Algorithm

Figure 5.7 shows the fatigue scores for S6's and S10's nBETTER supervised sessions. S10 seemed to experience similar fatigue levels in each session while S6 has varied fatigue levels in different sessions.

To further analyze the fatigue score, the Pearson's correlation of the mean fatigue score against the BCI accuracy of the supervised sessions is computed. Figure 5.8 shows the negative correlation, though it is not significant ($R = -0.468, p = 0.146 > 0.05$).

(a)

(b)

Figure 5.7 Fatigue score over 18 nBETTER supervised sessions: (a) S6's score, and (b) S10's score.

5.4 Discussion

The following sub-sections discuss the results of the patient demographics, the clinical efficacy of the nBETTER intervention, the correlation of relative EEG beta band power to BCI performance, subjects' fatigue scores and their correlation with BCI performance, and relate them to current literature.

Figure 5.8 Pearson's correlation of subjects' mean fatigue score against BCI accuracy of nBETTER supervised sessions. A negative correlation is found.

5.4.1 Patient Demographics

The one-way ANOVA did not reveal any significant intergroup differences in the demographics or baseline (Week 0) FMA score. However, the small and unequal sizes of the groups warrant that the results be interpreted with caution. For instance, the difference in the groups' stroke duration might have affected the FMA analysis as it showed a trend ($p = 0.1$), though it was not significant possibly due to the small and unequal group sizes.

5.4.2 Clinical Efficacy of the nBETTER Intervention

This single-arm multisite clinical study has investigated the clinical efficacy, in terms of FMA score improvement, of using nBETTER, an EEG-based MI-BCI system, for stroke rehabilitation. Retrospectively comparing with the SAT and BCI-HK groups of a previous study by Ang *et al.* (2014), no significant interaction between intervention and time is found. No significant intervention effects are found either, meaning that all groups

performed comparably, though the two BCI intervention groups had only about 33% of the conventional therapy duration of the SAT group.

The results in Table 5.3 and Figure 5.5 also showed that the nBETTER group had lesser FMA gains and consistency than the BCI-HK group, though it had the same amount of BCI intervention. This is likely due to the lack of kinesthetic feedback in the nBETTER system compared to the BCI-HK system (Ono *et al.*, 2014). However, the nBETTER system might be able to treat stroke patients who are excluded from robotic arm training protocols due to not fulfilling the criteria required for using robotic systems. Other advantages are that the nBETTER system is highly portable and accommodates a variety of supported postures.

5.4.3 Correlation of Beta Band Power and BCI Performance

A major challenge in the use of EEG-based MI-BCI in stroke patients is the cognitive pre-requisite of sustained attention and potential fatigue induction during training, which might limit its suitability to the majority of stroke patients and reduce its efficacy. Thus in this study, the presence of fatigue via an EEG correlate of fatigue, the relative beta band power, is investigated.

The repeated measures correlation between relative EEG beta band power and the BCI performance, in terms of classification accuracy, is computed across the 18 supervised sessions for the frontal, central and parietal-occipital brain regions. Significant positive correlations are found at the frontal and central areas. The lack of significance of the negative correlation in the parietal-occipital brain region might be due to the occipital region that reacted to visual stimuli (Kropotov, 2016). However, the significant correlations from the other regions showed the possibility that beta power is low when low BCI accuracy is achieved.

From the drowsy driving literature, a lower beta activity might suggest lower brain arousal and the presence of fatigue (Ahn *et al.*, 2016; Borghini *et al.*, 2012; Jap *et al.*, 2009; Tanaka *et al.*, 2012; Zhao *et al.*, 2012). Thus, this study's result indicated that fatigue might be present and a cause of poorer BCI performance. This might have affected the efficacy of the nBETTER group and resulted in it having comparable FMA gains to the SAT group. This suggested a disadvantage of EEG-based MI-BCI systems in inducing fatigue due to the monotony of performing MI.

5.4.4 Application of Proposed Negative-Unlabeled Learning Algorithm

To further investigate the presence of passive fatigue in stroke patients, the NU algorithm proposed in Section 3.3 is used to obtain fatigue scores for patients' nBETTER supervised sessions. Driving subjects' alert and most fatigued driving data as found in Section 4.3.1 are used to train an SVM model, which is then used on the stroke patients' data to obtain their fatigue scores. The results showed that there are subject-specific differences, as well as a trend of the mean fatigue score against BCI accuracy, though insignificant.

Having subject-specific differences in the patients' fatigue scores is similar to the driving case (see Section 4.3.3) and is unsurprising, as different patients will respond differently to the challenge of using EEG-based MI-BCI. This is also seen in the anecdotal example given by Frolov *et al.* (2017b) where one patient had to be given more break time than others due to the fatigue experienced.

To investigate if there is a correlation between fatigue score and BCI accuracy, a Pearson's correlation is computed. The result showed a negative correlation of the mean fatigue score against BCI accuracy for the 11 patients. This suggested a general trend that subjects with low BCI accuracy had high fatigue scores and vice versa. This potentially showed that fatigue is negatively affecting patients' BCI performance. This is in line with

postulations from Prasad *et al.* (2010) and Leamy *et al.* (2014a) who found that BCI performance decreased as subjective fatigue ratings increased.

However, the correlation value found in this study is not significant. This might be due to the differences in the driving and stroke datasets. Firstly, the stroke dataset not only contained subject-specific differences, but also additional non-stationarity due to patients performing the nBETTER supervised sessions on different days (Krauledat, 2008). Thus while the proposed negative-unlabeled algorithm was able to overcome subject-specific differences in the driving dataset, it might not be robust enough to overcome this additional non-stationarity in the stroke dataset. Secondly, pre-frontal channels were used in the driving dataset while frontal channels were used in the stroke dataset. Moreover, dry electrodes were used in the driving study, as opposed to the wet electrodes in the stroke study. These electrode differences may have affected the results. Thirdly, the SVM model was trained and tested on different datasets which involved different tasks. Additionally, the driving dataset consisted of healthy subjects while the stroke dataset consisted of stroke patients whose EEG may be different from healthy subjects (Leamy *et al.*, 2014b; Wolf *et al.*, 2017).

Furthermore, there were possibly other mental states involved, such as frustration (Myrden & Chau, 2017). Myrden and Chau (2015) also found that some frustration might motivate healthy subjects to perform better, while a certain fatigue level might be indicative of focus on the task at hand. This was also subjectively reported in another study, where stroke patients experienced fatigue due to concentration of attention (Frolov *et al.*, 2017a). This meant that subjects could be frustrated or fatigued to a certain extent without negatively affecting their BCI performance. Thus, the BCI accuracy might not strongly reflect fatigue levels, resulting in an insignificant correlation with the fatigue score.

Despite all these, the correlation found in this study is a preliminary result that indicated that fatigue might possibly be affecting patients' BCI performance negatively during MI-BCI rehabilitation. Thus, fatigue and other mental states should be further investigated in this context, and future BCIs can be adapted to accommodate such mental states as well, in order to enhance patients' BCI rehabilitation experience.

5.5 Summary

A single-arm multisite clinical trial is conducted with the nBETTER system, which was a non-invasive EEG-based MI-BCI system that provided visual feedback for stroke rehabilitation. This study aimed to investigate the clinical efficacy of using nBETTER via FMA score as well as the presence of fatigue during the intervention via correlation of relative EEG beta band power and BCI performance.

Clinical efficacy is measured using FMA score improvement. The nBETTER group demonstrated significant improvement at Weeks 6 and 24, better than the SAT control group, though not better than the BCI-HK intervention group from the previous similar study by Ang *et al.* (2014). However, all groups have comparable FMA gains.

In addition, the presence of fatigue is investigated via the relative EEG beta band power as well as the proposed NU learning algorithm. A significant positive repeated measures correlation with the nBETTER supervised sessions' accuracies is found in the frontal and central areas, while the mean fatigue score is negatively correlated with BCI performance, though not significant. These suggest that fatigue may be causing reduced BCI performance. This poorer BCI performance might also be a reason for the comparable efficacies of the nBETTER and SAT groups.

Overall, this study demonstrated that nBETTER as an EEG-based MI-BCI, together with conventional therapy, was effective and safe for use within a rehabilitation

setting. In addition, the EEG fatigue analysis motivate further research to investigate and mitigate the effects of fatigue in stroke rehabilitation; there is potential for BCI to play an additional fatigue or mental state monitoring role to enhance the clinical efficacy and personalization of stroke rehabilitation for stroke survivors according to their detected or predicted fatigue or mental state.

Chapter 6

Contributions, Limitations and Future Work

6.1 Contributions

The contributions in this thesis are summarized in Table 6.1. The following subsections elaborate on each contribution in more detail.

Table 6.1
Summary of Thesis' Contributions

Part	Contribution	Described in	Problems addressed	Results in
(a)	Proposed NU Learning Algorithm	Section 3.3	Lack of fatigue ground truth and cross-subject models	Section 4.3
(b)	Drowsy Driving Study	Section 4.2	Practical implementation using dry EEG	Section 4.3
(c)	Efficacy and Fatigue Study with nBETTER Intervention	Section 5.2	Clinical efficacy and fatigue in MI-BCI with only visual feedback for stroke rehabilitation	Section 5.3

6.1.1 Proposed NU Learning Algorithm

In the area of EEG-based BCI applications in mental state monitoring, the ground truth has thus far been difficult to ascertain. Specifically in the investigation of passive fatigue in drowsy driving, labeling the fatigue state or quantifying the fatigue level had thus often been achieved via surrogate measures such as subjective ratings or reaction time instead. These manual labeling of the driving EEG data undermined the reliability of using supervised learning methods. Thus, a two-step iterative cross-subject NU learning algorithm is proposed in this thesis to handle the labeling and quantification of the fatigue state automatically from a large amount of unlabeled driving EEG data, as well as a small amount of labeled alert data.

The PA used a RBF-kernel SVM classifier that gave probabilistic estimates. A grid search is utilized to determine the optimal parameters of the SVM cost parameter C and the RBF width parameter γ . Following which, the first step of the PA utilized labeled alert data and one unlabeled driving block λ per subject initially designated as fatigue data. Using the probabilistic estimates, it then iteratively searched for each subject's most fatigued driving block $\mathcal{B}$ in a cross-subject manner. In this step, a stopping criterion η is used to ensure convergence of the iterations. Subsequently, the second step of the PA used $\mathcal{B}$ as the fatigue data as well as alert data to quantify subjects' fatigue score in the driving data in a cross-subject manner.

The PA adds to the drowsy driving literature by employing NU learning, a special subset of semi-supervised learning. This does not require the manual labeling and supervised learning mostly used in current literature. By doing so, the PA can be generally applied to other contexts where the ground truth of mental states is difficult to obtain as well.

6.1.2 Drowsy Driving Study

The PA is then applied in a drowsy driving study to label and quantify subjects' passive fatigue. In this study, a small amount of labeled alert data and large amount of unlabeled driving data is collected from 29 subjects using commercially available MuseTM headband with dry pre-frontal EEG electrodes. The alert data is collected via a short and engaging target hitting game. Subsequently, unlabeled driving data is collected via an hour-long driving simulation in a monotonous highway environment with car deviation events every 5 to 10s.

Before using the PA, the alert and unlabeled driving EEG data are preprocessed to remove muscle and eye artifacts. After preprocessing, features are extracted from the EEG, namely the mean and standard deviation of 8 band powers and their upper and

lower envelopes over 5 windows. A grid search is then conducted and the optimized SVM parameters C^* and γ^* are found to be 10^1 and $10^{0.4}$ respectively. Other parameters are set to $\lambda = 9$ and $\eta = 0.0001$. Using these parameters, the first step of the PA then iteratively found $\mathcal{B}$, the most fatigued block of each subject, converging within 8 iterations with a final high accuracy of $93.77\% \pm 8.15\%$ across subjects. This performed significantly better than various baselines that assumed a different alert or fatigued data, showing the validity and effectiveness of the PA. The second step of the PA then quantified subjects' fatigue score in the driving EEG. Using a repeated measures correlation, the fatigue score is then shown to be significantly negatively correlated with relative EEG beta band power. This suggested that the fatigue score is in line with current literature where lower beta band power is associated with passive fatigue (Jap *et al.*, 2009; Zhao *et al.*, 2012).

This drowsy driving study has thus provided empirical evidence for the potential of the PA to detect and quantify fatigue from dry EEG electrodes in a cross-subject manner without manual labeling of the driving EEG data. This study thus contributes to dry EEG drowsy driving literature towards future real-time implementation.

6.1.3 Efficacy and Fatigue Study with nBETTER Intervention

In this thesis, the clinical efficacy and presence of fatigue are investigated in nBETTER, an EEG-based MI-BCI system with visual feedback for stroke rehabilitation. The data is collected from a multisite clinical trial with 11 stroke survivors, and is retrospectively compared to a SAT control group and BCI-HK intervention group from a previous similar study (Ang *et al.*, 2014). A SVM model trained on the driving dataset using the proposed NU learning algorithm is also applied on the stroke dataset to compute their fatigue score.

The clinical efficacy of the nBETTER group is shown by its significant FMA score gains at Weeks 6 and 24. This is better than the SAT control group, but not better

than the BCI-HK group. Nevertheless, the ANOVA analysis also revealed that all three groups performed comparably.

In addition, the results showed a significant positive correlation between relative EEG beta band power and the nBETTER group's BCI performance, as well as a non-significant negative correlation between the mean fatigue score and BCI accuracy, suggesting that fatigue might have been present due to monotony of performing MI, and caused poorer BCI performance. This poorer BCI performance might have been a contributing cause to the nBETTER group having comparable FMA performance to the SAT group. This showed the need and potential of BCI to play an additional role of fatigue or mental state monitoring in stroke rehabilitation.

This study contributes to the EEG-based MI-BCI stroke rehabilitation literature, as it is preceded only by Prasad *et al.* (2010) and Pichiorri *et al.* (2015) in using the MI-BCI with only visual feedback. This study also had a larger number of chronic patients than the previous studies, as well as follow-up assessments. Furthermore, this study is the first to use EEG data and a proposed NU learning algorithm to analyze the fatigue experienced by patients during the MI-BCI stroke rehabilitation. Previous studies had only took note of fatigue via ratings or anecdotes (Frolov *et al.*, 2017b; Leamy *et al.*, 2014a; Prasad *et al.*, 2010). Thus this study is a start to bridging the gap between mental state monitoring and medical BCI interventions.

6.2 Limitations

The following sub-sections describe the limitations of the various studies conducted in this thesis.

6.2.1 Proposed NU Learning Algorithm

A limitation of the PA was that it computed subjects' fatigue score offline and not

in real-time as the algorithm required all alert and driving data from subjects in order to find the subject-specific most fatigued driving block. However, this offline implementation is necessary to first verify the usefulness of the algorithm at this preliminary stage. Furthermore, after this initial offline training, the final model obtained can be easily used for obtaining fatigue scores in real-time for new test subjects.

6.2.2 Drowsy Driving Study

In this study, 29 subjects' data are analyzed. While this was not a large enough number to study any effects due to demographics or other characteristics, it was sufficiently large for a proof-of-concept of the PA.

This study also collected EEG data with only the MuseTM headset, and thus only had EEG features from pre-frontal electrodes. Furthermore, other modalities such as video and eye gaze information are collected, but are yet to be studied due to time constraint. However, the algorithm was still able to detect and quantify fatigue well. This showed the feasibility of using dry pre-frontal EEG electrodes that can be easily used in real-life settings for detecting and quantifying fatigue.

Additionally, in this study only correlation with beta band power is performed. This is due to drowsy driving literature showing consistent results in the beta frequency range, of decreased beta band power in the presence of passive fatigue (Borghini *et al.*, 2012; Jap *et al.*, 2009; Zhao *et al.*, 2012). Alpha band power might increase or decrease, and changes in theta band power might be confounded by coping mechanisms of the subject in reaction to fatigue (Wascher *et al.*, 2014). Furthermore, as pre-frontal electrodes are used, and beta waves are generally more evident in the frontal brain region (Satapathy *et al.*, 2019), it is thus used as the only EEG correlate of fatigue in this study.

6.2.3 Efficacy and Fatigue Study with nBETTER Intervention

A limitation in this clinical study is that only 11 stroke patients completed the study, due to the difficulties in recruiting subjects. This number of subjects is not large enough to be representative of the stroke population. However, based off the FMA gains of the retrospective control group in the study by Ang *et al.* (2014), for a statistical power of 80%, a sample size of 6 is required for the nBETTER group. Thus, having 11 subjects is sufficient for this preliminary clinical study with the nBETTER system.

Another limitation with this clinical study is that it is not a true RCT as it did not have a control group due to time and budget constraints. Instead, the SAT control group and BCI-HK intervention group from a previous study (Ang *et al.*, 2014) are used for retrospective comparison.. However, since the study by Ang *et al.* (2014) has similar subject inclusion and exclusion criteria as well as similar protocol, it is a fair enough comparison for a preliminary clinical study of the nBETTER system.

A limitation with this study's EEG analysis was that only the relative beta band power is used as an EEG correlate of fatigue. However, the beta power might be indicative of other mental states such as frustration as well (Myrden & Chau, 2017). Furthermore, no explicit data for fatigue detection are collected. For instance, no scales are used to obtain subjective fatigue ratings, or any engaging games designed to collect alert data. In addition, the fatigue score obtained by the proposed NU learning algorithm is not significantly correlated with the BCI accuracy. This may be due to the differences between the driving and stroke datasets, as well as the presence of other mental states, as discussed in Section 5.4.4. However, the preliminary findings generally suggest that it is possible that fatigue was present and might be causing poorer BCI performance and a lack of significant difference in efficacy compared to the SAT control group.

6.3 Future Work

The following sub-sections discuss the future work of the studies conducted in this thesis, some of which address the limitations presented in Section 6.2.

6.3.1 Proposed NU Learning Algorithm

Idea 1: Improving robustness and usage in real-time

To detect and quantify passive fatigue, other methods can be explored with the NU learning framework. For example, ensembles (Hu, 2017; Nguyen *et al.*, 2012) with the same base classifier (Claesen *et al.*, 2015) or different base classifiers (Yang *et al.*, 2014) can be utilized. Co-training, a technique in semi-supervised learning which makes use of the different information provided by different features (Blum & Mitchell, 1998), can also be explored. These techniques may make the NU learning algorithm more robust (Hu, 2017).

In addition, online unsupervised or semi-supervised machine learning techniques can be investigated in order to modify the NU learning algorithm to monitor fatigue levels continuously on a single-trial basis in real-time (Myrden & Chau, 2017). This will make the NU learning algorithm more applicable in real-life scenarios.

Idea 2: Applying to other mental states

In general, this algorithm can be applied to detect other mental states as well, where a small amount of negative class data can be collected and there exists unlabeled data from which the desired mental state is to be found. One example is the mental state of stress (Xia *et al.*, 2018) where a short relaxation task can be used to collect negative data and a longer experiment inducing stress can be used to collect unlabeled data from which the stress level is to be quantified.

6.3.2 Drowsy Driving Study

Idea 3: Reducing dependence on surrogate measures of fatigue

In researching passive fatigue in driving studies, the ground truth has often been determined through surrogate measures such as reaction time to secondary tasks (Lin *et al.*, 2012). However, they may be intrusive to the driver's mental state (Arico *et al.*, 2018; Dasari *et al.*, 2017). In addition, this thesis' proposed NU learning algorithm has shown that it is feasible to detect and quantify passive fatigue without surrogate measures from secondary tasks. As such, moving forward, drowsy driving simulations may not need to be designed with secondary tasks for gauging fatigue level. Other algorithms that can quantify fatigue without dependence on surrogate measures can be further proposed and explored as well.

Idea 4: Investigating other headsets and brain areas

In this study, only pre-frontal electrodes are utilized. However, headsets with electrodes in other brain regions can be investigated as well. For instance, passive fatigue had been known to affect other brain areas, such as the occipital area (Pal *et al.*, 2008). In addition, there may be interactions between the different cortical areas (Myrden & Chau, 2017; Qi *et al.*, 2019). Thus the PA and drowsy driving experiment can be conducted with other headsets and brain areas for further validation.

Idea 5: Exploring other EEG features

In terms of features extracted from the EEG data for fatigue detection, other features such as entropy (Mu *et al.*, 2017) and brain connectivity (Qi *et al.*, 2019) can be explored. In addition, features from other modalities such as the video and eye gaze information collected in this study can be extracted and compared. Such additional features can be used to develop a multi-modal fatigue detection system as well.

Idea 6: Using public datasets and expanding own dataset

In this study, the proposed NU learning algorithm was benchmarked against several baselines. For further comparison of this proposed algorithm with other known methods, drowsy driving public datasets with EEG data such as the one recently released by Cao *et al.* (2019) can be utilized.

The driving data collected in this thesis can also act as a database; new subject data from future similar studies can adapt this dataset's trained model to detect fatigue in their unlabeled data (Wei *et al.*, 2018). More subject data can also be collected for further exploration of whether more data will make the NU learning algorithm more robust. It can also be investigated whether the NU learning algorithm is able to handle session-to-session non-stationarity for the same subject.

Idea 7: Conducting field trials

To move towards real-life implementations to benefit drivers, the quality of dry EEG hardware can be further improved, especially under real-world settings where vibration and movement is common and will cause larger noise in the recordings. Thereafter, effective countermeasures such as warning systems can be developed accordingly (May & Baldwin, 2009) to ensure that drivers have an appropriate alertness level. Ultimately, field trials will be conducted to validate the effectiveness of such fatigue detection and warning systems.

6.3.3 Efficacy and Fatigue Study with nBETTER Intervention

Idea 8: Conducting future clinical trials

Future clinical studies with larger and balanced intervention and control groups can be conducted to further investigate the effectiveness of using nBETTER, an EEG-based MI-BCI with only visual feedback, in stroke rehabilitation. Such studies can also explore how nBETTER both with and without robotic haptic feedback can aid in

neuroplasticity and stroke recovery better than conventional methods (McFarland & Wolpaw, 2017).

Idea 9: Enhancing fatigue analysis in stroke datasets

In this study, the collected driving dataset is used to train a model for testing on the stroke dataset. In future, public drowsy driving datasets such as that by Cao *et al.* (2019) can be used instead as well. Additionally, to further investigate fatigue experienced by the stroke survivors in using MI-BCI stroke rehabilitation, data such as rating scales (Prasad *et al.*, 2010) can be used to subjectively assess their fatigue. A short engaging game that they can easily participate in can be designed to collect alert data as well. The NU learning algorithm can then be applied directly on the stroke dataset to detect any fatigued states that the patients experience during MI-BCI rehabilitation.

Idea 10: Furthering mental state monitoring in BCI medical applications

In future, mental state monitoring in BCI medical applications can be investigated in general, as it is a strategic approach to overcoming performance variation (Ahn & Jun, 2015). Other mental states that require attention in the medical domain include attention (van Dokkum *et al.*, 2015), motivation (Ahn & Jun, 2015) and emotion (Pillette *et al.*, 2018). Further on, learning companions that monitor mental states and adapt accordingly (Pillette *et al.*, 2018) can be developed to enhance the stroke rehabilitation experience even more.

6.3.4 Towards Passive BCI (pBCI)

Idea 11: Adding feedback of fatigue level

The focus of this thesis has been on mental state monitoring in both non-medical and medical domains. Such mental state research is the first basic and essential component of passive BCI (pBCI) (Krol *et al.*, 2018), which detects user mental states in order to aid the user (George & Lécuyer, 2010; Zander *et al.*, 2014; Zander & Kothe,

2011; Zander *et al.*, 2010). As such, the works in this thesis contribute to the literature on pBCI, particularly in the mental state of passive fatigue.

To further these contributions, feedback to the BCI user such as the drowsy driver or stroke patient on their fatigue level is to be provided in order to enrich the human-machine interaction (Arico *et al.*, 2018; Krol *et al.*, 2018; Zander *et al.*, 2014; Zander & Kothe, 2011). By providing feedback such as audio warnings (Huang *et al.*, 2016), haptic vibrations (Arimitsu *et al.*, 2007) or visual indicators (Galarza *et al.*, 2018) of the user's fatigue, it can influence the user's mental state and close the feedback loop to the BCI system. Currently there lacks literature and focus on the feedback aspect of pBCI (Krol *et al.*, 2018).

Idea 12: Investigating other mental states

In addition, passive fatigue is but one mental state amongst many. Other mental states investigated in literature include active fatigue (Desmond & Hancock, 2001), frustration (Myrden & Chau, 2015), attention (Myrden & Chau, 2015; Roy & Frey, 2016) and stress (Xia *et al.*, 2018). These may be experienced by drowsy drivers and stroke patients as well. Thus, there is a need for future experiments to investigate the presence and effects of these other mental states in driving and stroke rehabilitation. The proposed NU learning algorithm may also be applied to explore how it can be utilized for detecting mental states in general and not just passive fatigue.

Idea 13: Integrating with active BCI

Further on, these pBCI can be integrated with active BCI. This is especially meaningful in the medical domain where many current BCI systems are active BCI that require users to perform certain mental activities to achieve desired outcomes. By monitoring and adapting to users' mental states, these BCI systems can then improve the detection of their intention and produce the desired outcomes despite changes in their

mental states. An example in the stroke rehabilitation context would be to develop an active MI-BCI rehabilitation system that can both detect MI ('active' signal) and fatigue ('passive' signal), and use the latter to better inform the former. Investigation along this line will also shed light on whether the 'passive' signal interferes with the 'active' signals instead of aiding it (George & Lécuyer, 2010), and also has the potential to validate the measured user state (Krol *et al.*, 2018) without intruding into the user's main task (Arico *et al.*, 2018).

Idea 14: Applying to emotional states

Additionally, all these future work in pBCI can be applied to emotional states (Arico *et al.*, 2018; Roy & Frey, 2016) as well. Thus, pBCI can potentially cater to a wide range of mental and emotional states to better aid the driver and stroke patient.

Bibliography

- Abdulkader, S. N., Atia, A., & Mostafa, M.-S. M. (2015). Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal*, 16(2), 213-230. doi:10.1016/j.eij.2015.06.002
- Ahn, M., & Jun, S. C. (2015). Performance variation in motor imagery brain-computer interface: A brief review. *Journal of Neuroscience Methods*, 243, 103-110. doi:10.1016/j.jneumeth.2015.01.033
- Ahn, S., Nguyen, T., Jang, H., Kim, J. G., & Jun, S. C. (2016). Exploring neurophysiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data. *Frontiers in Human Neuroscience*, 10, 219. doi:10.3389/fnhum.2016.00219
- Akerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, 52(1-2), 29-37. Retrieved from <https://www.tandfonline.com/loi/ines20>
- Alt Murphy, M., Resteghini, C., Feys, P., & Lamers, I. (2015). An overview of systematic reviews on upper extremity outcome measures after stroke. *BMC Neurology*, 15, 29. doi:10.1186/s12883-015-0292-6
- Ang, K. K., Chin, Z. Y., Wang, C., Guan, C., & Zhang, H. (2012). Filter bank common spatial pattern algorithm on BCI Competition IV datasets 2a and 2b. *Frontiers in Neuroscience*, 6, 39. doi:10.3389/fnins.2012.00039
- Ang, K. K., Chua, K. S. G., Phua, K. S., Wang, C., Chin, Z. Y., Kuah, C. W. K., . . . Guan, C. (2015a). A randomized controlled trial of EEG-based motor imagery brain-computer interface robotic rehabilitation for stroke. *Clinical EEG and Neuroscience*, 46(4), 310-320. doi:10.1177/1550059414522229
- Ang, K. K., & Guan, C. (2013). Brain-computer interface in stroke rehabilitation. *Journal of Computing Science and Engineering*, 7(2), 139-146. doi:10.5626/JCSE.2013.7.2.139
- Ang, K. K., Guan, C., Chua, K. S., Ang, B. T., Kuah, C., Wang, C., . . . Zhang, H. (2008a). *A clinical evaluation on the spatial patterns of non-invasive motor imagery-*

based brain-computer interface in stroke. Paper presented at the Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, Vancouver, BC, Canada. doi:10.1109/iembs.2008.4650129

Ang, K. K., Guan, C., Chua, K. S., Ang, B. T., Kuah, C. W., Wang, C., . . . Zhang, H. (2008b). *A clinical evaluation of non-invasive motor imagery-based brain-computer interface in stroke*. Paper presented at the Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, Vancouver, BC, Canada. doi:10.1109/iembs.2008.4650130

Ang, K. K., Guan, C., Chua, K. S., Ang, B. T., Kuah, C. W., Wang, C., . . . Zhang, H. (2011). A large clinical study on the ability of stroke patients to use an EEG-based motor imagery brain-computer interface. *Clinical EEG and Neuroscience*, 42(4), 253-258. doi:10.1177/155005941104200411

Ang, K. K., Guan, C., Phua, K. S., Wang, C., Zhao, L., Teo, W. P., . . . Chew, E. (2015b). Facilitating effects of transcranial direct current stimulation on motor imagery brain-computer interface with robotic feedback for stroke rehabilitation. *Archives of Physical Medicine and Rehabilitation*, 96(3 Suppl), S79-87. doi:10.1016/j.apmr.2014.08.008

Ang, K. K., Guan, C., Phua, K. S., Wang, C., Zhou, L., Tang, K. Y., . . . Chua, K. S. G. (2014). Brain-computer interface-based robotic end effector system for wrist and hand rehabilitation: Results of a three-armed randomized controlled trial for chronic stroke. *Frontiers in Neuroengineering*, 7, 30. doi:10.3389/fneng.2014.00030

Arico, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., & Babiloni, F. (2018). Passive BCI beyond the lab: Current trends and future directions. *Physiological Measurement*, 39(8), 08tr02. doi:10.1088/1361-6579/aad57e

Arimitsu, S., Sasaki, K., Hosaka, H., Itoh, M., Ishida, K., & Ito, A. (2007). Seat belt vibration as a stimulating device for awakening drivers. *IEEE/ASME Transactions on Mechatronics*, 12(5), 511-518. doi:10.1109/TMECH.2007.905704

Asín Prieto, G., Cano-de-la-Cuerda, R., López-Larraz, E., Metrot, J., Molinari, M., & van Dokkum, L. E. H. (2014). Emerging Perspectives in Stroke Rehabilitation. In J. L. Pons

- & D. Torricelli (Eds.), *Emerging Therapies in Neurorehabilitation* (pp. 3-21). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Bagherzadeh, J., & Asil, H. (2018). A review of various semi-supervised learning models with a deep learning and memory approach. *Iran Journal of Computer Science*. doi:10.1007/s42044-018-00027-6
- Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, 8, 456. doi:10.3389/fpsyg.2017.00456
- Balandong, R. P., Ahmad, R. F., Saad, M. N. M., & Malik, A. S. (2018). A review on EEG-based automatic sleepiness detection systems for driver. *IEEE Access*, 6, 22908-22919. doi:10.1109/ACCESS.2018.2811723
- Barwick, F., Arnett, P., & Slobounov, S. (2012). EEG correlates of fatigue during administration of a neuropsychological test battery. *Clinical Neurophysiology*, 123(2), 278-284. doi:10.1016/j.clinph.2011.06.027
- Bekker, J., & Davis, J. (2018). Learning from positive and unlabeled data: A survey. *arXiv preprint arXiv:1811.04820*.
- Blankertz, B., Acqualagna, L., Dähne, S., Haufe, S., Schultze-Kraft, M., Sturm, I., . . . Müller, K.-R. (2016). The Berlin brain-computer interface: Progress beyond communication and control. *Frontiers in Neuroscience*, 10, 530-530. doi:10.3389/fnins.2016.00530
- Blankertz, B., Tangermann, M., Vidaurre, C., Fazli, S., Sannelli, C., Haufe, S., . . . Müller, K.-R. (2010). The Berlin brain-computer interface: Non-medical uses of BCI technology. *Frontiers in Neuroscience*, 4, 198-198. doi:10.3389/fnins.2010.00198
- Blum, A., & Mitchell, T. (1998). *Combining labeled and unlabeled data with co-training*. Paper presented at the Proceedings of the Eleventh Annual Conference on Computational Learning Theory, Madison, Wisconsin, USA. doi:10.1145/279943.279962
- Bohannon, R. W., & Smith, M. B. (1987). Interrater reliability of a modified Ashworth scale of muscle spasticity. *Physical Therapy*, 67(2), 206-207. Retrieved from <https://academic.oup.com/ptj>

- Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2012). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*(0). doi:10.1016/j.neubiorev.2012.10.003
- Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience and biobehavioral reviews*, 44, 58-75. doi:10.1016/j.neubiorev.2012.10.003
- Cao, Z., Chuang, C.-H., King, J.-K., & Lin, C.-T. (2019). Multi-channel EEG recordings during a sustained-attention driving task. *Scientific Data*, 6(1), 19. doi:10.1038/s41597-019-0027-4
- Chai, R., Ling, S. H., San, P. P., Naik, G. R., Nguyen, T. N., Tran, Y., . . . Nguyen, H. T. (2017). Improving EEG-based driver fatigue classification using sparse-deep belief networks. *Frontiers in Neuroscience*, 11(103), 103. doi:10.3389/fnins.2017.00103
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1-27. doi:10.1145/1961189.1961199
- Chang, W.-D., Cha, H.-S., Kim, K., & Im, C.-H. (2016). Detection of eye blink artifacts from single prefrontal channel electroencephalogram. *Computer Methods and Programs in Biomedicine*, 124, 19-30. doi:10.1016/j.cmpb.2015.10.011
- Chi, Z., Hong, W., & Rongrong, F. (2014). Automated detection of driver fatigue based on entropy and complexity measures. *IEEE Transactions on Intelligent Transportation Systems*, 15(1), 168-177. doi:10.1109/TITS.2013.2275192
- Chuang, C., Lai, P., Ko, L., Kuo, B., & Lin, C. (2010). *Driver's cognitive state classification toward brain computer interface via using a generalized and supervised technology*. Paper presented at the International Joint Conference on Neural Networks (IJCNN). doi:10.1109/IJCNN.2010.5596835
- Claesen, M., De Smet, F., Suykens, J. A., & De Moor, B. (2015). A robust ensemble approach to learn from positive and unlabeled data using SVM base models. *Neurocomputing*, 160, 73-84. doi:10.1016/j.neucom.2014.10.081

- Clafflin, E. S., Krishnan, C., & Khot, S. P. (2015). Emerging treatments for motor rehabilitation after stroke. *The Neurohospitalist*, 5(2), 77-88. doi:10.1177/1941874414561023
- Craig, A., Tran, Y., Wijesuriya, N., & Nguyen, H. (2012). Regional brain wave activity changes associated with fatigue. *Psychophysiology*, 49(4), 574-582. doi:10.1111/j.1469-8986.2011.01329.x
- Curran, E. A., & Stokes, M. J. (2003). Learning to control brain activity: A review of the production and control of EEG components for driving brain-computer interface (BCI) systems. *Brain and Cognition*, 51(3), 326-336. doi:10.1016/S0278-2626(03)00036-8
- Dasari, D., Shou, G., & Ding, L. (2017). ICA-Derived EEG Correlates to Mental Fatigue, Effort, and Workload in a Realistically Simulated Air Traffic Control Task. *Frontiers in Neuroscience*, 11(297). doi:10.3389/fnins.2017.00297
- De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W., & Van Huffel, S. (2006). Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. *IEEE Transactions on Biomedical Engineering*, 53(12 Pt 1), 2583-2587. doi:10.1109/TBME.2006.879459
- Desmond, P. A., & Hancock, P. A. (2001). Active and passive fatigue states. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload and fatigue* (pp. 455-465). Mahwah, NJ: Erlbaum.
- Dimitrakopoulos, G. N., Kakkos, I., Dai, Z., Wang, H., Sgarbas, K., Thakor, N., . . . Sun, Y. (2018). Functional connectivity analysis of mental fatigue reveals different network topological alterations between driving and vigilance tasks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(4), 740-749. doi:10.1109/TNSRE.2018.2791936
- Dobkin, B. H. (2007). Brain-computer interface technology as a tool to augment plasticity and outcomes for neurological rehabilitation. *The Journal of Physiology*, 579(Pt 3), 637-642. doi:10.1113/jphysiol.2006.123067
- ETSC. (2001). *The role of driver fatigue in commercial road transport crashes*. Brussels.

- Fatourech, M., Bashashati, A., Ward, R. K., & Birch, G. E. (2007). EMG and EOG artifacts in brain computer interface systems: A survey. *Clinical Neurophysiology*, 118(3), 480-494. doi:10.1016/j.clinph.2006.10.019
- Fitzgibbon, S. P., DeLosAngeles, D., Lewis, T. W., Powers, D. M. W., Whitham, E. M., Willoughby, J. O., & Pope, K. J. (2015). Surface Laplacian of scalp electrical signals and independent component analysis resolve EMG contamination of electroencephalogram. *International Journal of Psychophysiology*, 97(3), 277-284. doi:10.1016/j.ijpsycho.2014.10.006
- Foong, R., Ang, K. K., & Quek, C. (2017). *Correlation of reaction time and EEG log bandpower from dry frontal electrodes in a passive fatigue driving simulation experiment*. Paper presented at the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). doi:10.1109/EMBC.2017.8037360
- Foong, R., Ang, K. K., Quek, C., Guan, C., Phua, K. S., Kuah, C. W. K., . . . Chua, K. S. G. (2019a). Assessment of the Efficacy of EEG-based MI-BCI with Visual Feedback and EEG Correlates of Mental Fatigue for Upper-Limb Stroke Rehabilitation. *IEEE Transactions on Biomedical Engineering*, 1-1. doi:10.1109/TBME.2019.2921198
- Foong, R., Ang, K. K., Quek, C., Guan, C., & Phyo Wai, A. A. (2015). *An analysis on driver drowsiness based on reaction time and EEG band power*. Paper presented at the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). doi:10.1109/EMBC.2015.7320244
- Foong, R., Ang, K. K., Zhang, Z., & Quek, C. (2019b). An iterative cross-subject negative-unlabeled learning algorithm for quantifying passive fatigue. *Journal of Neural Engineering*. doi:10.1088/1741-2552/ab255d
- Frolov, A. A., Húsek, D., Biryukova, E. V., Bobrov, P. D., Mokienko, O. A., & Alexandrov, A. (2017a). Principles of motor recovery in post-stroke patients using hand exoskeleton controlled by the brain-computer interface based on motor imagery. *Neural Network World*, 27(1), 107. doi:10.14311/nnw.2017.27.006
- Frolov, A. A., Mokienko, O., Lyukmanov, R., Biryukova, E., Kotov, S., Turbina, L., . . . Bushkova, Y. (2017b). Post-stroke rehabilitation training with a motor-imagery-based

- brain-computer interface (BCI)-controlled hand exoskeleton: A randomized controlled multicenter trial. *Frontiers in Neuroscience*, 11, 400. doi:10.3389/fnins.2017.00400
- Fugl-Meyer, A. R., Jaasko, L., Leyman, I., Olsson, S., & Steglind, S. (1975). The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. *Scandinavian Journal of Rehabilitation Medicine*, 7(1), 13-31. Retrieved from <https://europepmc.org/>
- Galarza, E. E., Egas, F. D., Silva, F. M., Velasco, P. M., & Galarza, E. D. (2018). *Real time driver drowsiness detection based on driver's face image behavior using a system of human computer interaction implemented in a smartphone*. Paper presented at the Proceedings of the International Conference on Information Technology & Systems (ICITS 2018), Cham. doi:10.1007/978-3-319-73450-7_53
- George, L., & Lécuyer, A. (2010). *An overview of research on "passive" brain-computer interfaces for implicit human-computer interaction*. Paper presented at the International Conference on Applied Bionics and Biomechanics (ICABB), Venice, Italy. Retrieved from <https://hal.inria.fr/inria-00537211>
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486-489. doi:10.5812/ijem.3505
- Gimeno, P. T., Cerezuela, G. P., & Montanes, M. C. (2006). On the concept and measurement of driver drowsiness, fatigue and inattention: Implications for countermeasures. *International Journal of Vehicle Design*, 42(1-2), 67-86. doi:10.1504/ijvd.2006.010178
- Guger, C., Coyle, D., Mattia, D., De Lucia, M., Hochberg, L., Edlow, B. L., . . . Annen, J. (2017). Trends in BCI Research I: Brain-computer interfaces for assessment of patients with locked-in syndrome or disorders of consciousness. In C. Guger, B. Allison, & M. Lebedev (Eds.), *Brain-Computer Interface Research: A State-of-the-Art Summary 6* (pp. 105-125). Cham: Springer International Publishing.
- Hajinoroozi, M., Mao, Z., Jung, T.-P., Lin, C.-T., & Huang, Y. (2016). EEG-based prediction of driver's cognitive performance by deep convolutional neural network.

Signal Processing: Image Communication, 47, 549-555.
doi:10.1016/j.image.2016.05.018

Hameed, P. N., Verspoor, K., Kusljic, S., & Halgamuge, S. (2017). Positive-unlabeled learning for inferring drug interactions based on heterogeneous attributes. *BMC Bioinformatics*, 18, 140. doi:10.1186/s12859-017-1546-7

Hashemi, A., Saba, V., & Resalat, S. N. (2014). Real time driver's drowsiness detection by processing the EEG signals stimulated with external flickering light. *Basic and Clinical Neuroscience*, 5(1), 22-27. Retrieved from <http://bcn.iums.ac.ir/>

Hockey, G. R. J. (2012). Challenges in fatigue and performance research. In G. Matthews, P. A. Desmond, C. Neubauer, & P. A. Hancock (Eds.), *The Handbook of Operator Fatigue* (pp. 45-62). Farham, Surrey, England: Ashgate Publishing Limited.

Hofmann, T. (2001). Unsupervised Learning by Probabilistic Latent Semantic Analysis. *Machine Learning*, 42(1), 177-196. doi:10.1023/A:1007617005950

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6(2), 65-70. Retrieved from <http://www.jstor.org.ezlibproxy1.ntu.edu.sg/stable/4615733>

Hu, J. (2017). Automated detection of driver fatigue based on adaboost classifier with EEG signals. *Frontiers in Computational Neuroscience*, 11, 72. doi:10.3389/fncom.2017.00072

Hu, J., & Min, J. (2018). Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model. *Cognitive Neurodynamics*, 1-10. doi:10.1007/s11571-018-9485-1

Huang, K.-C., Huang, T.-Y., Chuang, C.-H., King, J.-T., Wang, Y.-K., Lin, C.-T., & Jung, T.-P. (2016). An EEG-based fatigue detection and mitigation system. *International Journal of Neural Systems*, 26(04), 1650018. doi:10.1142/S0129065716500180

Jap, B. T., Lal, S., & Fischer, P. (2010). Inter-hemispheric electroencephalography coherence analysis: Assessing brain activity during monotonous driving. *International Journal of Psychophysiology*, 76(3), 169-173. doi:10.1016/j.ijpsycho.2010.03.007

- Jap, B. T., Lal, S., Fischer, P., & Bekiaris, E. (2009). Using EEG spectral components to assess algorithms for detecting fatigue. *Expert Systems with Applications*, 36(2), 2352-2359. doi:10.1016/j.eswa.2007.12.043
- Käthner, I., Wriessnegger, S. C., Müller-Putz, G. R., Kübler, A., & Halder, S. (2014). Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain-computer interface. *Biological Psychology*, 102(Supplement C), 118-129. doi:10.1016/j.biopsycho.2014.07.014
- Khan, S., & Madden, M. (2014). One-Class Classification: Taxonomy of Study and Review of Techniques. *The Knowledge Engineering Review*, 1-30. doi:10.1017/S026988891300043X
- Khushaba, R. N., Kodagoda, S., Lal, S., & Dissanayake, G. (2013). Uncorrelated fuzzy neighborhood preserving analysis based feature projection for driver drowsiness recognition. *Fuzzy Sets and Systems*, 221(0), 90-111. doi:10.1016/j.fss.2012.12.003
- Kılıç, C., & Tan, M. (2012). Positive unlabeled learning for deriving protein interaction networks. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 1(3), 87-102. doi:10.1007/s13721-012-0012-8
- Kim, T., Kim, S., & Lee, B. (2016). Effects of action observational training plus brain-computer interface-based functional electrical stimulation on paretic arm motor recovery in patient with stroke: A randomized controlled trial. *Occupational Therapy International*, 23(1), 39-47. doi:10.1002/oti.1403
- Kotsiantis, S. (2007). Supervised Machine Learning: A Review of Classification Techniques. *Informatica*, 31(3), 249-268. Retrieved from <http://www.informatica.si/index.php/informatica>
- Krauledat, M. (2008). *Analysis of Nonstationarities in EEG Signals for Improving Brain-Computer Interface Performance*. (Doctoral dissertation), Technische Universität Berlin. Retrieved from <https://depositonce.tu-berlin.de/> (10.14279/depositonce-1817)
- Krigolson, O. E., Williams, C. C., Norton, A., Hassall, C. D., & Colino, F. L. (2017). Choosing Muse: Validation of a low-cost, portable EEG system for ERP research. *Frontiers in Neuroscience*, 11, 109. doi:10.3389/fnins.2017.00109

- Krol, L. R., Andreessen, L. M., & Zander, T. O. (2018). Passive brain-computer interfaces: A perspective on increased interactivity. In C. S. Nam, A. Nijholt, & F. Lotte (Eds.), *Brain-Computer Interfaces Handbook: Technological and Theoretical Advances* (pp. 69-86). Boca Raton, FL, USA: CRC Press.
- Kropotov, J. D. (2016). Chapter 2.3 - Beta and Gamma Rhythms. In J. D. Kropotov (Ed.), *Functional Neuromarkers for Psychiatry* (pp. 107-119). San Diego: Academic Press.
- Lal, S. K. L., & Craig, A. (2002). Driver fatigue: Electroencephalography and psychological assessment. *Psychophysiology*, 39(3), 313-321. doi:10.1017/S0048577201393095
- Lawhern, V., Kerick, S., & Robbins, K. A. (2013). Detecting alpha spindle events in EEG time series using adaptive autoregressive models. *BMC Neuroscience*, 14(1), 101. doi:10.1186/1471-2202-14-101
- Leamy, D. J., Kocijan, J., Domijan, K., Duffin, J., Roche, R. A. P., Commins, S., . . . Ward, T. E. (2014a). An exploration of EEG features during recovery following stroke – implications for BCI-mediated neurorehabilitation therapy. *Journal of NeuroEngineering and Rehabilitation*, 11, 9. doi:10.1186/1743-0003-11-9
- Leamy, D. J., Kocijan, J., Domijan, K., Duffin, J., Roche, R. A. P., Commins, S., . . . Ward, T. E. (2014b). An exploration of EEG features during recovery following stroke – implications for BCI-mediated neurorehabilitation therapy. *Journal of NeuroEngineering and Rehabilitation*, 11(1), 9. doi:10.1186/1743-0003-11-9
- Lee, W. S., & Liu, B. (2003). *Learning with positive and unlabeled examples using weighted logistic regression*. Paper presented at the Proceedings of the Twentieth International Conference on International Conference on Machine Learning, Washington, DC, USA. Retrieved from <https://dblp.org>
- Li, G., & Chung, W.-Y. (2015). A context-aware EEG headset system for early detection of driver drowsiness. *Sensors (Basel, Switzerland)*, 15(8), 20873-20893. doi:10.3390/s150820873
- Li, M., Liu, Y., Wu, Y., Liu, S., Jia, J., & Zhang, L. (2014). Neurophysiological substrates of stroke patients with motor imagery-based brain-computer interface

- training. *International Journal of Neuroscience*, 124(6), 403-415. doi:10.3109/00207454.2013.850082
- Li, W., Guo, Q., & Elkan, C. (2011). A positive and unlabeled learning algorithm for one-class classification of remote-sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 49(2), 717-725. doi:10.1109/TGRS.2010.2058578
- Li, X.-L., & Liu, B. (2005). *Learning from positive and unlabeled examples with different data distributions*. Paper presented at the Machine Learning Knowledge Discovery Databases, Berlin, Heidelberg. doi:10.1007/11564096_24
- Li, X., & Liu, B. (2003). *Learning to classify texts using positive and unlabeled data*. Paper presented at the Proceedings of the 18th International Joint Conference on Artificial Intelligence, Acapulco, Mexico. Retrieved from <https://www.ijcai.org/>
- Lin, C.-T., Chang, C.-J., Lin, B.-S., Hung, S.-H., Chao, C.-F., & Wang, I.-J. (2010). A real-time wireless brain-computer interface system for drowsiness detection. *IEEE Transactions on Biomedical Circuits and Systems*, 4(4), 214-222. doi:10.1109/TBCAS.2010.2046415
- Lin, C.-T., Chuang, C.-H., Huang, C.-S., Tsai, S.-F., Lu, S.-W., Chen, Y.-H., & Ko, L.-W. (2014). Wireless and wearable EEG system for evaluating driver vigilance. *IEEE Transactions on Biomedical Circuits and Systems*, 8(2), 165-176. doi:10.1109/TBCAS.2014.2316224
- Lin, C.-T., Wu, R.-C., Jung, T.-P., Liang, S.-F., & Huang, T.-Y. (2005a). Estimating driving performance based on EEG spectrum analysis. *EURASIP Journal on Advances in Signal Processing*, 2005(19), 521368. doi:10.1155/ASP.2005.3165
- Lin, C.-T., Wu, R.-C., Liang, S.-F., Chao, W.-H., Chen, Y.-J., & Jung, T.-P. (2005b). EEG-based drowsiness estimation for safety driving using independent component analysis. *Circuits and Systems I: Regular Papers, IEEE Transactions on*, 52(12), 2726-2738. doi:10.1109/TCSI.2005.857555
- Lin, F.-C., Ko, L.-W., Chuang, C.-H., Su, T.-P., & Lin, C.-T. (2012). Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 59(9), 2044-2055. doi:10.1109/TCSI.2012.2185290

- Lin, H.-T., Lin, C.-J., & Weng, R. C. (2007). A note on Platt's probabilistic outputs for support vector machines. *Machine Learning*, 68(3), 267-276. doi:10.1007/s10994-007-5018-6
- Liu, B., Dai, Y., Li, X., Lee, W. S., & Yu, P. S. (2003). *Building text classifiers using positive and unlabeled examples*. Paper presented at the Third IEEE International Conference on Data Mining, Melbourne, FL, USA. doi:10.1109/ICDM.2003.1250918
- Liu, Y.-T., Lin, Y.-Y., Wu, S.-L., Chuang, C.-H., & Lin, C.-T. (2016). Brain dynamics in predicting driving fatigue using a recurrent self-evolving fuzzy neural network. *IEEE Transactions on Neural Networks and Learning Systems*, 27(2), 347-360. doi:10.1109/TNNLS.2015.2496330
- Mak, J. N., & Wolpaw, J. R. (2009). Clinical applications of brain-computer interfaces: Current state and future prospects. *IEEE Reviews in Biomedical Engineering*, 2, 187-199. doi:10.1109/RBME.2009.2035356
- Matthews, G., Desmond, P. A., & Hitchcock, E. M. (2012). Dimensional models of fatigue. In G. Matthews, P. A. Desmond, C. Neubauer, & P. A. Hancock (Eds.), *The Handbook of Operator Fatigue* (pp. 139-154). Farham, Surrey, England: Ashgate Publishing Limited.
- May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(3), 218-224. doi:10.1016/j.trf.2008.11.005
- McFarland, D. J. (2015). The advantages of the surface Laplacian in brain-computer interface research. *International Journal of Psychophysiology*, 97(3), 271-276. doi:10.1016/j.ijpsycho.2014.07.009
- McFarland, D. J., & Wolpaw, J. R. (2017). EEG-based brain-computer interfaces. *Current Opinion in Biomedical Engineering*, 4, 194-200. doi:10.1016/j.cobme.2017.11.004
- Mendis, S. (2013). Stroke disability and rehabilitation of stroke: World Health Organization perspective. *International Journal of Stroke*, 8(1), 3-4. doi:10.1111/j.1747-4949.2012.00969.x

- Millán, J. d. R., Rupp, R., Müller-Putz, G. R., Murray-Smith, R., Giugliemma, C., Tangermann, M., . . . Mattia, D. (2010). Combining brain–computer interfaces and assistive technologies: State-of-the-art and challenges. *Frontiers in Neuroscience*, 4, 161. doi:10.3389/fnins.2010.00161
- Moghim, S., Kushki, A., Marie Guerguerian, A., & Chau, T. (2012). A review of EEG-based brain-computer interfaces as access pathways for individuals with severe disabilities. *Assistive Technology*, 25(2), 99-110. doi:10.1080/10400435.2012.723298
- Molinari, M., Esquenazi, A., Anastasi, A. A., Nielsen, R. K., Stoller, O., D’Andrea, A., & Calatayud, M. B. (2016). Rehabilitation technologies application in stroke and traumatic brain injury patients. In J. L. Pons, R. Raya, & J. González (Eds.), *Emerging Therapies in Neurorehabilitation II* (pp. 29-64). Cham: Springer International Publishing.
- Monge-Pereira, E., Ibañez-Pereda, J., Alguacil-Diego, I. M., Serrano, J. I., Spottorno-Rubio, M. P., & Molina-Rueda, F. (2017). Use of electroencephalography brain-computer interface systems as a rehabilitative approach for upper limb function after a stroke: A systematic review. *PM&R*, 9(9), 918-932. doi:10.1016/j.pmrj.2017.04.016
- Morales, J. M., Díaz-Piedra, C., Rieiro, H., Roca-González, J., Romero, S., Catena, A., . . . Di Stasi, L. L. (2017). Monitoring driver fatigue using a single-channel electroencephalographic device: A validation study by gaze-based, driving performance, and subjective data. *Accident Analysis & Prevention*, 109, 62-69. doi:10.1016/j.aap.2017.09.025
- Mordelet, F., & Vert, J. P. (2014). A bagging SVM to learn from positive and unlabeled examples. *Pattern Recognition Letters*, 37, 201-209. doi:10.1016/j.patrec.2013.06.010
- Mu, Z., Hu, J., & Min, J. (2017). Driver fatigue detection system using electroencephalography signals based on combined entropy features. *Applied Sciences*, 7(2), 150. doi:10.3390/app7020150
- Mukherjee, D., & Patil, C. G. (2011). Epidemiology and the global burden of stroke. *World Neurosurgery*, 76(6 suppl), S85-S90. doi:10.1016/j.wneu.2011.07.023
- Myrden, A., & Chau, T. (2015). Effects of user mental state on EEG-BCI performance. *Frontiers in Human Neuroscience*, 9, 308. doi:10.3389/fnhum.2015.00308

- Myrden, A., & Chau, T. (2016). Towards psychologically adaptive brain-computer interfaces. *Journal of Neural Engineering*, 13(6), 066022. doi:10.1088/1741-2560/13/6/066022
- Myrden, A., & Chau, T. (2017). A passive EEG-BCI for single-trial detection of changes in mental state. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(4), 345-356. doi:10.1109/TNSRE.2016.2641956
- Neubauer, C., Matthews, G., & Saxby, D. J. (2012). Driver fatigue and safety: A transactional perspective. In G. Matthews, P. A. Desmond, C. Neubauer, & P. A. Hancock (Eds.), *The Handbook of Operator Fatigue* (pp. 365-378). Farham, Surrey, England: Ashgate Publishing Limited.
- Nguyen, M. N., Li, X.-L., & Ng, S.-K. (2011). *Positive unlabeled learning for time series classification*. Paper presented at the Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain. Retrieved from <https://www.ijcai.org/>
- Nguyen, M. N., Li, X.-L., & Ng, S.-K. (2012). *Ensemble based positive unlabeled learning for time series classification*. Paper presented at the International Conference on Database Systems for Advanced Applications, Berlin, Heidelberg. doi:10.1007/978-3-642-29038-1_19
- Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain computer interfaces: A review. *Sensors*, 12(2), 1211-1279. doi:10.3390/s120201211
- Niu, G., Plessis, M. C. d., Sakai, T., Ma, Y., & Sugiyama, M. (2016). *Theoretical comparisons of positive-unlabeled learning against positive-negative learning*. Paper presented at the 30th International Conference on Neural Information Processing Systems, Barcelona, Spain. Retrieved from <https://arxiv.org/list/cs/recent>
- Noy, Y. I., Horrey, W. J., Popkin, S. M., Folkard, S., Howarth, H. D., & Courtney, T. K. (2011). Future directions in fatigue and safety research. *Accident Analysis & Prevention*, 43(2), 495-497. doi:10.1016/j.aap.2009.12.017
- Ogino, M., & Mitsukura, Y. (2018). Portable Drowsiness Detection through Use of a Prefrontal Single-Channel Electroencephalogram. *Sensors (Basel, Switzerland)*, 18(12), 4477. doi:10.3390/s18124477

- Öner, M., & Deveci Kocakoç, İ. (2017). Jmasm 49: A compilation of some popular goodness of fit tests for normal distribution: Their algorithms and matlab codes (matlab). *Journal of Modern Applied Statistical Methods*, 16(2), 547-575. doi:10.22237/jmasm/1509496200
- Ono, T., Shindo, K., Kawashima, K., Ota, N., Ito, M., Ota, T., . . . Ushiba, J. (2014). Brain-computer interface with somatosensory feedback improves functional recovery from severe hemiplegia due to chronic stroke. *Frontiers in Neuroengineering*, 7, 19. doi:10.3389/fneng.2014.00019
- Owens, J., Dingus, T., Guo, F., Fang, Y., Perez, M., McClafferty, J., & Tefft, B. (2018). *Prevalence of drowsy-driving crashes: Estimates from a large-scale naturalistic driving study*. Retrieved from <https://aaafoundation.org>
- Ozkan, N. F., & Kahya, E. (2018). Classification of BCI users based on cognition. *Computational Intelligence and Neuroscience*, 2018, 6315187-6315187. doi:10.1155/2018/6315187
- Pal, N. R., Chuang, C.-Y., Ko, L.-W., Chao, C.-F., Jung, T.-P., Liang, S.-F., & Lin, C.-T. (2008). EEG-based subject- and session-independent drowsiness detection: An unsupervised approach. *EURASIP Journal on Advances in Signal Processing*, 2008(1), 519480. doi:10.1155/2008/519480
- Peiris, M. T., Davidson, P. R., Bones, P. J., & Jones, R. D. (2011). Detection of lapses in responsiveness from the EEG. *Journal of Neural Engineering*, 8(1), 016003. doi:10.1088/1741-2560/8/1/016003
- Pfurtscheller, G., Brunner, C., Schlögl, A., & Lopes da Silva, F. H. (2006). Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage*, 31(1), 153-159. doi:10.1016/j.neuroimage.2005.12.003
- Phillips, R. O. (2015). A review of definitions of fatigue – And a step towards a whole definition. *Transportation Research Part F: Traffic Psychology and Behaviour*, 29(0), 48-56. doi:10.1016/j.trf.2015.01.003
- Pichiorri, F., Morone, G., Petti, M., Toppi, J., Pisotta, I., Molinari, M., . . . Mattia, D. (2015). Brain-computer interface boosts motor imagery practice during stroke recovery. *Annals of Neurology*, 77(5), 851-865. doi:10.1002/ana.24390

- Pillette, L., Jeunet, C., N'kambou, R., N'Kaoua, B., & Lotte, F. (2018). *Towards artificial learning companions for mental imagery-based brain-computer interfaces*. Paper presented at the Workshop sur les “Affects, Compagnons Artificiels et Interactions” (WACAI), Ile de Porquerolles, France. Retrieved from <https://hal.inria.fr/hal-01762612/>
- Prasad, G., Herman, P., Coyle, D., McDonough, S., & Crosbie, J. (2010). Applying a brain-computer interface to support motor imagery practice in people with stroke for upper limb recovery: A feasibility study. *Journal of NeuroEngineering and Rehabilitation*, 7, 60. doi:10.1186/1743-0003-7-60
- Price, C. I. M., Curless, R. H., & Rodgers, H. (1999). Can stroke patients use visual analogue scales? *Stroke*, 30(7), 1357-1361. doi:10.1161/01.STR.30.7.1357
- Qi, P., Ru, H., Gao, L., Zhang, X., Zhou, T., Tian, Y., . . . Sun, Y. (2019). Neural Mechanisms of Mental Fatigue Revisited: New Insights from the Brain Connectome. *Engineering*, 5(2), 276-286. doi:10.1016/j.eng.2018.11.025
- Ramadan, R. A., & Vasilakos, A. V. (2017). Brain computer interface: Control signals review. *Neurocomputing*, 223, 26-44. doi:10.1016/j.neucom.2016.10.024
- Ramos-Murguialday, A., Broetz, D., Rea, M., Läer, L., Yilmaz, Ö., Brasil, F. L., . . . Birbaumer, N. (2013). Brain-machine interface in chronic stroke rehabilitation: A controlled study. *Annals of Neurology*, 74(1), 100-108. doi:10.1002/ana.23879
- Remsik, A., Young, B., Vermilyea, R., Kiekoefer, L., Abrams, J., Elmore, S. E., . . . Prabhakaran, V. (2016). A review of the progression and future implications of brain-computer interface therapies for restoration of distal upper extremity motor function after stroke. *Expert Review of Medical Devices*, 13(5), 445-454. doi:10.1080/17434440.2016.1174572
- Rohit, F., Kulathumani, V., Kavi, R., Elwarfalli, I., Kecojevic, V., & Nimbarte, A. (2017). Real-time drowsiness detection using wearable, lightweight brain sensing headbands. *IET Intelligent Transport Systems*, 11(5), 255-263. doi:10.1049/iet-its.2016.0183
- Roy, R. N., Bonnet, S., Charbonnier, S., & Campagne, A. (2013). *Mental fatigue and working memory load estimation: Interaction and implications for EEG-based passive BCI*. Paper presented at the 35th Annual International Conference of the IEEE

- Engineering in Medicine and Biology Society (EMBC), Osaka, Japan.
doi:10.1109/EMBC.2013.6611070
- Roy, R. N., & Frey, J. (2016). Neurophysiological markers for passive brain–computer interfaces. *Brain–Computer Interfaces 1: Foundations and Methods*, 85-100.
doi:10.1002/9781119144977.ch5
- Rozado, D., Duenser, A., & Howell, B. (2015). Improving the performance of an EEG-based motor imagery brain computer interface using task evoked changes in pupil diameter. *PLoS ONE*, 10(3), e0121262. doi:10.1371/journal.pone.0121262
- Satapathy, S. K., Dehuri, S., Jagadev, A. K., & Mishra, S. (2019). Chapter 1 - Introduction. In S. K. Satapathy, S. Dehuri, A. K. Jagadev, & S. Mishra (Eds.), *EEG Brain Signal Classification for Epileptic Seizure Disorder Detection* (pp. 1-25): Academic Press.
- Saxby, D. J., Matthews, G., Warm, J. S., Hitchcock, E. M., & Neubauer, C. (2013). Active and passive fatigue in simulated driving: Discriminating styles of workload regulation and their safety impacts. *Journal of Experimental Psychology: Applied*, 19(4), 287-300. doi:10.1037/a0034386
- Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the Support of a High-Dimensional Distribution. *Neural Computation*, 13(7), 1443-1471. doi:10.1162/089976601750264965
- Shih, J. J., Krusienski, D. J., & Wolpaw, J. R. (2012). Brain-computer interfaces in medicine. *Mayo Clinic Proceedings*, 87(3), 268-279.
doi:10.1016/j.mayocp.2011.12.008
- Silvoni, S., Ramos-Murguialday, A., Cavinato, M., Volpato, C., Cisotto, G., Turolla, A., . . . Birbaumer, N. (2011). Brain-computer interface in stroke: A review of progress. *Clinical EEG and Neuroscience*, 42(4), 245-252. doi:10.1177/155005941104200410
- Simon, M., Schmidt, E. A., Kincses, W. E., Fritzsche, M., Bruns, A., Aufmuth, C., . . . Schrauf, M. (2011). EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. *Clinical Neurophysiology*, 122(6), 1168-1178.
doi:10.1016/j.clinph.2010.10.044

- Soekadar, S. R., Birbaumer, N., Slutzky, M. W., & Cohen, L. G. (2015). Brain-machine interfaces in neurorehabilitation of stroke. *Neurobiology of Disease*, 83, 172-179. doi:10.1016/j.nbd.2014.11.025
- Sonnleitner, A., Simon, M., Kincses, W. E., Buchner, A., & Schrauf, M. (2012). Alpha spindles as neurophysiological correlates indicating attentional shift in a simulated driving task. *International Journal of Psychophysiology*, 83(1), 110-118. doi:10.1016/j.ijpsycho.2011.10.013
- Tan, Z. X., Foong, R., & Ang, K. K. (2015). *Determining mechanical and electromyographical reaction time in a BCI driving fatigue experiment*. Paper presented at the 10th International Conference on Information, Communications and Signal Processing (ICICS), Singapore. doi:10.1109/ICICS.2015.7459814
- Tanaka, M., Shigihara, Y., Ishii, A., Funakura, M., Kanai, E., & Watanabe, Y. (2012). Effect of mental fatigue on the central nervous system: An electroencephalography study. *Behavioral and Brain Functions*, 8, 48. doi:10.1186/1744-9081-8-48
- Tiwari, N., Edla, D. R., Dodia, S., & Bablani, A. (2018). Brain computer interface: A comprehensive survey. *Biologically Inspired Cognitive Architectures*, 26, 118-129. doi:10.1016/j.bica.2018.10.005
- van Dokkum, L. E. H., Ward, T., & Laffont, I. (2015). Brain computer interfaces for neurorehabilitation – its current status as a rehabilitation strategy post-stroke. *Annals of Physical and Rehabilitation Medicine*, 58(1), 3-8. doi:10.1016/j.rehab.2014.09.016
- van Erp, J., Lotte, F., & Tangermann, M. (2012). Brain-computer interfaces: Beyond medical applications. *Computer*, 45(4), 26-34. doi:10.1109/MC.2012.107
- Várkuti, B., Guan, C., Pan, Y., Phua, K. S., Ang, K. K., Kuah, C. W. K., . . . Sitaram, R. (2013). Resting state changes in functional connectivity correlate with movement recovery for BCI and robot-assisted upper-extremity training after stroke. *Neurorehabilitation and Neural Repair*, 27(1), 53-62. doi:10.1177/1545968312445910
- Venketasubramanian, N., & Chen, C. L. H. (2008). Burden of stroke in Singapore. *International Journal of Stroke*, 3(1), 51-54. doi:10.1111/j.1747-4949.2008.00181.x

- Vogelpohl, T., Kühn, M., Hummel, T., & Vollrath, M. (2018). Asleep at the automated wheel—Sleepiness and fatigue during highly automated driving. *Accident Analysis & Prevention*. doi:10.1016/j.aap.2018.03.013
- Voznenko, T. I., Urvanov, G. A., Dyumin, A. A., Andrianova, S. V., & Chepin, E. V. (2016). The research of emotional state influence on quality of a brain-computer interface usage. *Procedia Computer Science*, 88, 391-396. doi:10.1016/j.procs.2016.07.454
- Vukelić, M., & Gharabaghi, A. (2015). Oscillatory entrainment of the motor cortical network during motor imagery is modulated by the feedback modality. *NeuroImage*, 111, 1-11. doi:10.1016/j.neuroimage.2015.01.058
- Wang, Y.-H., Yeh, C.-H., Young, H.-W. V., Hu, K., & Lo, M.-T. (2014). On the computational complexity of the empirical mode decomposition algorithm. *Physica A: Statistical Mechanics and its Applications*, 400, 159-167. doi:10.1016/j.physa.2014.01.020
- Wascher, E., Rasch, B., Sängler, J., Hoffmann, S., Schneider, D., Rinkenauer, G., . . . Gutberlet, I. (2014). Frontal theta activity reflects distinct aspects of mental fatigue. *Biological Psychology*, 96, 57-65. doi:10.1016/j.biopsycho.2013.11.010
- Wei, C.-S., Lin, Y.-P., Wang, Y.-T., Lin, C.-T., & Jung, T.-P. (2018). A subject-transfer framework for obviating inter- and intra-subject variability in EEG-based drowsiness detection. *NeuroImage*, 174, 407-419. doi:10.1016/j.neuroimage.2018.03.032
- Whitham, E. M., Pope, K. J., Fitzgibbon, S. P., Lewis, T., Clark, C. R., Loveless, S., . . . Willoughby, J. O. (2007). Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clinical Neurophysiology*, 118(8), 1877-1888. doi:10.1016/j.clinph.2007.04.027
- Wierwille, W. W., & Ellsworth, L. A. (1994). Evaluation of driver drowsiness by trained raters. *Accident Analysis & Prevention*, 26(5), 571-581. doi:10.1016/0001-4575(94)90019-1
- Wolf, M. E., Ebert, A. D., & Chatzikonstantinou, A. (2017). The use of routine EEG in acute ischemic stroke patients without seizures: generalized but not focal EEG

- pathology is associated with clinical deterioration. *International Journal of Neuroscience*, 127(5), 421-426. doi:10.1080/00207454.2016.1189913
- Wolpaw, J. R., Birbaumer, N., Heetderks, W. J., McFarland, D. J., Peckham, P. H., Schalk, G., . . . Vaughan, T. M. (2000). Brain-computer interface technology: A review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering*, 8(2), 164-173. doi:10.1109/TRE.2000.847807
- Wolpaw, J. R., & Wolpaw, E. W. (2012). Brain-computer interfaces: Something new under the sun. *Brain-Computer Interfaces: Principles and Practice*, 14. doi:10.1093/acprof:oso/9780195388855.001.0001
- Wu, D., Lawhern, V. J., Gordon, S., Lance, B. J., & Lin, C. T. (2016). Driver drowsiness estimation from EEG signals using online weighted adaptation regularization for regression (OwARR). *IEEE Transactions on Fuzzy Systems*, 25(6), 1522-1535. doi:10.1109/TFUZZ.2016.2633379
- Xia, L., Malik, A. S., & Subhani, A. R. (2018). A physiological signal-based method for early mental-stress detection. *Biomedical Signal Processing and Control*, 46, 18-32. doi:10.1016/j.bspc.2018.06.004
- Yang, P., Li, X.-L., Mei, J.-P., Kwoh, C.-K., & Ng, S.-K. (2012). Positive-unlabeled learning for disease gene identification. *Bioinformatics*, 28(20), 2640-2647. doi:10.1093/bioinformatics/bts504
- Yang, P., Li, X., Chua, H.-N., Kwoh, C.-K., & Ng, S.-K. (2014). Ensemble positive unlabeled learning for disease gene identification. *PLoS ONE*, 9(5), e97079. doi:10.1371/journal.pone.0097079
- Yeo, M. V. M., Li, X., Shen, K., & Wilder-Smith, E. P. V. (2009). Can SVM be used for automatic EEG detection of drowsiness during car driving? *Safety Science*, 47(1), 115-124. doi:10.1016/j.ssci.2008.01.007
- Zander, T. O., Brönstrup, J., Lorenz, R., & Krol, L. R. (2014). Towards BCI-based implicit control in human–computer interaction. In S. H. Fairclough & K. Gilleade (Eds.), *Advances in Physiological Computing* (pp. 67-90). London: Springer London.

- Zander, T. O., & Kothe, C. (2011). Towards passive brain-computer interfaces: Applying brain-computer interface technology to human-machine systems in general. *Journal of Neural Engineering*, 8(2), 025005. doi:10.1088/1741-2560/8/2/025005
- Zander, T. O., Kothe, C., Jatzev, S., & Gaertner, M. (2010). Enhancing human-computer interaction with input from active and passive brain-computer interfaces. In D. S. Tan & A. Nijholt (Eds.), *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction* (pp. 181-199). London: Springer London.
- Zhang, Z., Foong, R., Phua, K. S., Wang, C., & Ang, K. K. (2018). *Modeling EEG-based motor imagery with session to session online adaptation*. Paper presented at the Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, Honolulu, HI, USA. doi:10.1109/EMBC.2018.8512706
- Zhang, Z., & Guan, C. (2017). *An accurate sleep staging system with novel feature generation and auto-mapping*. Paper presented at the International Conference on Orange Technologies (ICOT), Singapore. doi:10.1109/ICOT.2017.8336079
- Zhao, C., Zhao, M., Liu, J., & Zheng, C. (2012). Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator. *Accident Analysis & Prevention*, 45, 83-90. doi:10.1016/j.aap.2011.11.019
- Zhao, C., Zheng, C., Zhao, M., Liu, J., & Tu, Y. (2011a). Automatic classification of driving mental fatigue with EEG by wavelet packet energy and KPCA-SVM. *International Journal of Innovative Computing, Information and Control*, 7, 1157-1168. Retrieved from <http://www.ijicic.org>
- Zhao, C., Zheng, C., Zhao, M., Tu, Y., & Liu, J. (2011b). Multivariate autoregressive models and kernel learning algorithms for classifying driving mental fatigue based on electroencephalographic. *Expert Systems with Applications*, 38(3), 1859-1865. doi:10.1016/j.eswa.2010.07.115
- Zhu, X. (2008). Semi-Supervised Learning Literature Survey. *Comput Sci, University of Wisconsin-Madison*, 2. Retrieved from http://pages.cs.wisc.edu/~jerryzhu/pub/ssl_survey.pdf

Zich, C., Debener, S., Kranczioch, C., Bleichner, M. G., Gutberlet, I., & De Vos, M. (2015). Real-time EEG feedback during simultaneous EEG–fMRI identifies the cortical signature of motor imagery. *NeuroImage*, *114*, 438-447. doi:10.1016/j.neuroimage.2015.04.020

Author's Publications

The following journal papers are published:

1. Assessment of the Efficacy of EEG-based MI-BCI with Visual Feedback and EEG Correlates of Mental Fatigue for Upper-Limb Stroke Rehabilitation, *IEEE Transactions on Biomedical Engineering* (TBME), June 2019.

R. Foong, K. K. Ang, C. Quek, C. Guan, K. S. Phua, C. W. K. Kuah, D. A. Vishwanath, L. H. L. Yam, D. K. Rajeswaran, N. Tang, E. Chew, and K. S. G. Chua

2. An Iterative Cross-Subject Negative-Unlabeled Learning Algorithm for Quantifying Passive Fatigue, *Journal of Neural Engineering* (JNE), June 2019.

R. Foong, K. K. Ang, Z. Zhang, and C. Quek

The following conference papers are published:

1. An Analysis on Driver Drowsiness based on Reaction Time and EEG Band Power, published in the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015.

R. Foong, K. K. Ang, Z. Zhang, C. Quek, C. Guan and A. A. Phyto Wai

2. Determining Mechanical and Electromyographical Reaction Time in a BCI Driving Fatigue Experiment, published in the 10th International Conference on Information, Communications and Signal Processing, 2015.

Z. X. Tan, **R. Foong**, and K. K. Ang

3. Correlation of Reaction Time and EEG Log Bandpower from Dry Frontal Electrodes in a Passive Fatigue Driving Simulation Experiment, published in the 39th Annual International Conference of the IEEE Engineering in Medicine

and Biology Society (EMBC), 2017.

R. Foong, K. K. Ang, Z. Zhang, and C. Quek

4. Modeling EEG-based Motor Imagery with Session to Session Online Adaptation, published in the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018.

Z. Zhang, **R. Foong**, K. S. Phua, C. Wang, and K. K. Ang