

UNIVERSITY OF SHEFFIELD

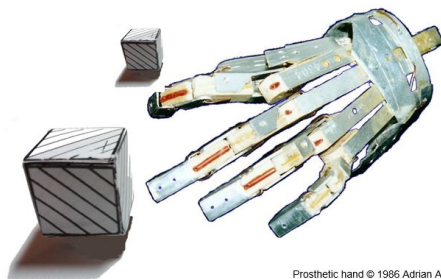
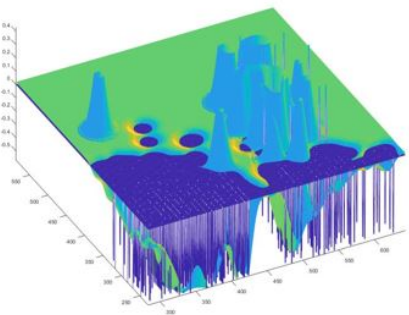
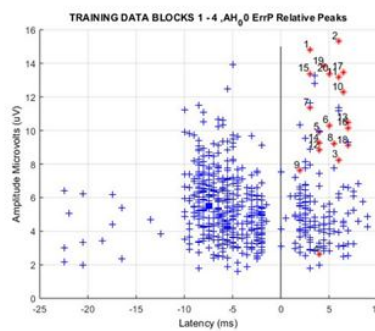
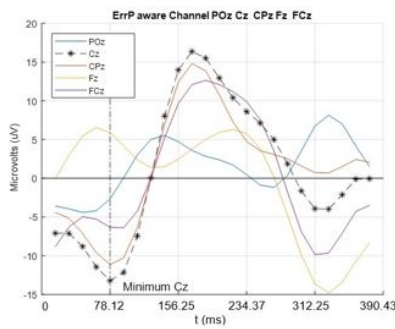
AUTOMATIC CONTROL AND SYSTEMS ENGINEERING

MASTER OF PHILOSOPHY THESIS

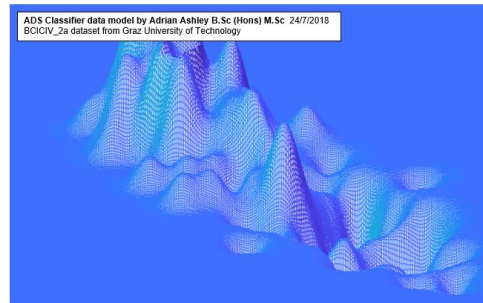
Improving classification of error related potentials using novel feature extraction and classification algorithms for an assistive robotic device



January 26, 2021



Prosthetic hand © 1986 Adrian Ashley



Author
Mr. ADRIAN ASHLEY

Supervisor
Dr. Mahnaz ARVANEH

Second supervisor
Professor Lyudmila S
MIHAYLOVA

Acknowledgements

I'm a mature self-funded student returning back to ACSE University of Sheffield from 1994. I would like to thank my parents Cynthia and Everal Ashly for the opportunity to study for more future qualifications such as a self-funded PhD after this present self-funded MPhil. I would also like to thank the administrative staff Matthew Ham, Renata Ashton and technical support in ACSE by Darren Fox. I would like to thank Dr. Eric Lacey for providing the experimental data and the participants therein. The University of Sheffield Maths and Statistics support (MASH) staff provided me with excellent support Dan Wilson, Mustabsar Awais. I would like to acknowledge the support and guidance using the SPSS software, in particular, ANOVA the paired T-test analysis [1] advice provided by Dr. Basile Marquier and Ellen Marshall, Mathematics and statistic tutor at Maths and Stats Help (MASH) which enabled me to prove the statistical advantages of the novel feature selection method called relative peak feature (RPF). I attended training and would like to thank Norbert Gyenge, instructor for the Python programming. Ken Deeley, Ph.D. Senior Training Engineer from The MathWorks Ltd who delivered the parallel computing with MATLAB @training.

Abstract

We evaluated the proposed feature extraction algorithm and the classifier, and we showed that the performance surpassed the state of the art algorithms in error detection. Advances in technology are required to improve the quality of life of a person with a severe disability who has lost their independence of movement in their daily life. Brain-computer interface (BCI) is a possible technology to re-establish a sense of independence for the person with a severe disability through direct communication between the brain and an electronic device. To enhance the symbiotic interface between the person and BCI its accuracy and robustness should be improved across all age groups. This thesis aims to address the above-mentioned issue by developing a novel feature extraction algorithm and a novel classification algorithm for the detection of erroneous actions made by either human or BCI. The research approach evaluated the state of the art error detection classifier using data from two different age groups, young and elderly. The performance showed a statistical difference between the aforementioned age groups; therefore, there needs to be an improvement in error detection and classification. The results showed that my proposed relative peak feature (RPF) and adaptive decision surface (ADS) classifier outperformed the state of the art algorithms in detecting errors using EEG for both elderly and young groups. In addition, the novel classification algorithm has been applied to motor imagery to improve the detection of when a person imagines moving a limb. Finally, this thesis takes a brief look at object recognition for a shared control task of identifying utensils in cooperation with a prosthetic robotic hand.

Contents

1	Introduction and Motivation	8
1.1	Background	8
1.2	Motivation	9
1.3	Aims and objectives	9
1.3.1	EEG-based error detection in adults and elderly	10
1.3.2	Error Driven Learning	10
1.3.3	A novel classifier for improving motor imagery	10
1.3.4	Object recognition for Kitchen utensil identification	11
1.4	Impact to the scientific community and wider society	13
1.5	Publications from this work	13
2	Literature Review	14
2.1	Introduction	14
2.2	Brain-computer interfaces (BCI)	14
2.2.1	Recording brain activity using EEG	15
2.2.2	EEG electrode positions	15
2.2.3	Different types of BCI	16
2.2.3.1	P300	16
2.2.3.2	Steady State Evoked Potential (SSEP)	16
2.2.3.3	Motor Imagery-BCI (MI-BCI)	17
2.2.4	BCI Components	17
2.2.4.1	BCI pre-processing	17
2.2.4.2	BCI feature extraction	17
2.2.4.3	Feature characteristics and properties	19
2.2.5	BCI Classifiers	19
2.2.5.1	K-means clustering algorithm (K-MC) and Multi-layered perceptrons	19
2.2.5.2	K-Nearest Neighbour algorithm (k-NN)	20
2.2.5.3	Bayesian Network (BN)	20
2.2.5.4	Artificial Neural Network (ANN)	21
2.2.5.5	Multi-layered perceptron (MLP)	21
2.2.5.6	Back Propagation Neural Network (BPNN)	21
2.2.5.7	Convolution Neural Network(CNN) Deep Learning (DL)	21
2.2.5.8	Support Vector Machine (SVM)	21
2.2.5.9	Linear Discriminant Analysis (LDA)	22
2.2.5.10	Interval Type-2 Fuzzy Logic System (IT2FLS)	22
2.2.5.11	Conclusion classifiers	22
2.2.5.12	Challenges of BCI classifiers	22
2.2.6	Non-stationary EEG signals	22
2.2.7	Other problems using EEG, BCI and ErrP	22
2.3	Error detection using brain signals	23
2.3.1	Error-related brain signals	23

2.3.2	Error-related brain signals elicited by observation	24
2.3.3	Error detection process	24
2.3.4	Error detection using time and frequency domain features	24
2.3.5	Error detection for P300 BCI using real-time Bayesian dynamic stopping framework	25
2.3.6	Error detection with a double error-related potential (ErrP) detection . . .	25
2.3.7	Anticipating an error	25
2.4	Error Driven Learning	25
2.4.1	Introduction	25
2.4.2	Recognising an 'Interaction ErrP' to improve a Brian computer interface .	26
2.4.3	Methods of corrective actions	26
2.4.4	Methods of learning form mistakes	26
2.4.5	Reinforcement learning	26
2.4.6	Learning algorithms	27
2.5	Shared control strategy	28
2.5.1	Introduction	28
2.5.2	Interactive control strategies BCI	28
2.5.3	Prosthetic hand restoring the sense of touch	29
2.5.4	Prosthetic hand deep learning image recognition	29
2.5.5	Adaptive hierarchical brain-computer interface	29
2.6	Summary	29
3	Deficiencies of state of the art error detection algorithms between young and elderly	30
3.1	Introduction	30
3.2	Methodology	31
3.2.1	Sampling	31
3.2.2	Pre-processing	31
3.2.3	Feature extraction and classifications	31
3.3	Experiment	32
3.3.1	Participants	32
3.3.2	Experiment Protocol	32
3.3.2.1	Methodology	32
3.4	Results	37
3.4.1	Results of the statistical analysis	37
3.4.2	Young / Old participants, Experiment 1	37
3.4.3	Young / Elderly participants, Experiment 2	38
3.5	Summary	39
3.6	Discussion	39
4	Improving EEG-based error detection using proposed novel relative peak features (RPF)	40
4.1	Introduction	40
4.2	Methodology	41
4.2.1	Participants and the task	41
4.2.2	Preprocessing	41
4.2.3	Feature extraction and classifications	42
4.3	Relative peak features (RPF)	42
4.4	Results	44
4.4.1	Classifier comparison results using temporal features	44
4.4.2	Classifier comparison results using relative peak features	44
4.5	Summary	44
4.6	Results of classifier comparisons using a novel feature extraction method	46
4.7	Introduction	46

4.7.1	Statistical Analysis Comparisons	46
4.7.2	Elderly participants classifier comparison results using temporal features	47
4.7.3	Elderly participants classifier comparison results using relative peak features	48
4.7.4	Young participants classifier comparison results using temporal features .	49
4.7.5	Young participants classifier comparison results using relative peak features	50
4.8	Results of classifier comparisons using a novel feature extraction method	51
4.9	Conclusion	51
4.10	Results of the performance gains summary	52
4.11	Summary	52
5	Proposed novel adaptive decision surface (ADS) classifier	54
5.1	Introduction	54
5.2	Methodology	56
5.2.1	Electroencephalograph EEG feature classification process	57
5.2.2	Adaptive decision surface (ADS) mathematical definition	57
5.2.2.1	Preprocessing	61
5.2.2.2	Feature extraction and classifications	62
5.2.3	Adaptive Decision Surface(ADS) Classifier	64
5.2.3.1	Building the adaptive decision surface classifier	65
5.2.3.2	Building 'Likelihood' bias into the adaptive decision surface ADS	65
5.2.3.3	Building the adaptive decision surface ADS	66
5.2.3.4	Optimising and calibrating the ADS decision level	66
5.2.3.5	Adapting the surface of the ADS	67
5.3	Experiment data	67
5.3.1	Participants	67
5.4	Results	68
5.4.1	Results from the adaptive decision surface ADS	68
5.5	Summary	69
5.6	Discussion future investigations	69
6	Motor imagery based BCI and performance comparison of 23 classifiers	71
6.1	Introduction	71
6.1.1	Re-establish a sense of empowerment	71
6.1.1.1	Functional regions of the human brain	71
6.1.1.2	Locations for motor imagery in the human brain	72
6.1.2	Motor Imagery Brain Computer Interface (MI-BCI) Components	74
6.1.2.1	MI-BCI Electrode channel location	75
6.1.2.2	Feature extraction using band power	75
6.1.2.3	MI-BCI left and right hand detection performance	76
6.1.2.4	MI-BCI Improving performance with Spatial filters	77
6.1.2.5	MI-BCI basic spatial filters	77
6.1.2.6	MI-BCI supervised spatial filters	78
6.1.2.7	MI-BCI When classification goes wrong	79
6.2	Methodology	80
6.3	Classifier performance experiments using 23 motor imagery classifiers	81
6.4	Results from the 23 motor imagery classifiers	82
6.5	Summary of the 23 motor imagery classifier comparisons	82
6.6	A novel three dimensional probability-based classifier for improving motor imagery-based BCI	82
6.7	Introduction	82
6.8	Methodology	82

6.8.1	MI-BCI Improving performance with Spatial filters	82
6.8.2	Motor Imagery Adaptive decision surface(MI-ADS) mathematical defini- tion	82
6.8.3	Building 'Likelihood' bias into the ADS model for motor imagery	84
6.8.4	A constructed motor imagery adaptive decision surface model	85
6.8.5	Motor imagery test data	85
6.8.6	Adapting the surface of the ADS	86
6.9	Summary	87
6.9.1	Results from the motor imagery adaptive decision surface MI ADS	87
6.10	Discussion	88
6.11	MI ADS Classifier Results	90
7	Shared control image decoding and tactile LiDAR of kitchen utensils for an assistive prosthetic hand	91
7.1	Introduction	91
7.2	Methodology	93
7.2.1	Initial image processing stage	93
7.2.2	Tactile LiDAR signals	103
7.3	Summary	103
7.3.1	Kitchen utensil tactile LiDAR scanning	104
8	Conclusions and Future Work	106
8.1	Conclusions	106
8.2	Future work	111
8.2.1	Future development the absorption adaptability function of the (ADS)	111
8.2.2	Future development reverse adaptation function of the (ADS)	111
8.2.3	Future development a computational inference reasoning CIR	111
8.2.4	Future development of a method for prior computational inference	111
8.2.5	Future work self-organising surface management	112
8.2.6	Future development an adaptive decision surface data threshold	112
8.2.7	Future work on error driven learning algorithm	113
8.2.8	Future work on a novel shared control strategy algorithm	113
A	Participants Dataset	116
A.1	Participants	116
A.1.1	Young participants data table	116
A.1.2	Elderly participants data table	117
A.1.3	Participants age bar chart	117
B	Participants relative peak feature(RPF) datasets	118
B.1	Dataset elderly participants relative peak features(RPF)	118
B.1.1	Elderly participants 1 and 2, showing their relative peak features(RPF)	119
B.1.2	Elderly participants 3 and 4, showing their relative peak features(RPF)	120
B.1.3	Elderly participants 5 and 6, showing their relative peak features(RPF)	121
B.1.4	Elderly participants 7 and 8, showing their relative peak features(RPF)	122
B.1.5	Elderly participants 9 and 10, showing their relative peak features(RPF)	123
B.1.6	Elderly participants 11 and 12, showing their relative peak features(RPF)	124
B.1.7	Elderly participants 13 and 14, showing their relative peak features(RPF)	125
B.1.8	Elderly participants 15 and 16, showing their relative peak features(RPF)	126
B.1.9	Elderly participants 9 and 10, showing their relative peak features(RPF)	127
B.1.10	Elderly participants 11 and 12, showing their relative peak features(RPF)	128
B.1.11	Elderly participants 13 and 14, showing their relative peak features(RPF)	129
B.1.12	Elderly participants 15 and 16, showing their relative peak features(RPF)	130

B.2	Dataset young participants relative peak features(RPF)	131
B.2.1	Young participants 1 and 2, showing their relative peak features(RPF)	132
B.2.2	Young participants 3 and 4, showing their relative peak features(RPF)	133
B.2.3	Young participants 5 and 6, showing their relative peak features(RPF)	134
B.2.4	Young participants 7 and 8, showing their relative peak features(RPF)	135
B.2.5	Young participants 9 and 10, showing their relative peak features(RPF)	136
B.2.6	Young participants 11 and 12, showing their relative peak features(RPF)	137
B.2.7	Young participants 13 and 14, showing their relative peak features(RPF)	138
B.2.8	Young participants 15 and 16, showing their relative peak features(RPF)	139
B.2.9	Young participants 17 and 18, showing their relative peak features(RPF)	140
B.2.10	Young participants 19 and 20, showing their relative peak features(RPF)	141
B.2.11	Young participants 21 and 22, showing their relative peak features(RPF)	142
B.2.12	Young participants 23 and 24, showing their relative peak features(RPF)	143
B.2.13	Young participants 25 and 26, showing their relative peak features(RPF)	144
B.2.14	Young participant 27, showing their relative peak features(RPF)	145
C	Standard deviation and variance	146
C.1	Standard deviation and variances	147
C.1.1	Elderly participants Standard deviation and variance	147
C.1.2	Young participants Standard deviation and variance	148
D	Results Tables	149
D.1	Participants	150
D.1.1	Elderly participants classifier using temporal features results tables	150
D.1.2	Elderly participants classifier using relative features results tables	151
D.1.3	Young participants classifier using temporal features results tables P2-1	152
D.1.4	Young participants classifier using temporal features results tables P2-2	153
D.1.5	Young participants classifier using relative peak features results tables P2-1	154
D.1.6	Young participants classifier using relative peak features results tables P2-2	155
E	Results of 23 motor imagery classifier	156
F	Results of classifier comparisons using a novel feature extraction method	167
F.0.1	Results of the statistical analysis using a subspace discriminant classifier using young participants	168
F.0.2	Tests of Within-Subjects Effects	168
F.0.3	Results of the statistical analysis using a subspace discriminant classifier using elderly participants	170
F.0.4	Tests of Within-Subjects Effects	170
F.0.5	Results of the statistical analysis using LDA classifier using elderly participants	172
F.0.6	Tests of Within-Subjects Effects and tests Within-Subjects Contrasts	172
F.0.7	Results of the statistical analysis using LDA classifier using young participants	174
F.0.8	Tests of Within-Subjects Effects and tests Within-Subjects Contrasts	174
F.0.9	Results of the statistical analysis using SVM Linear classifier using young participants	176
F.0.10	Tests of Within-Subjects Effects and tests Within-Subjects Contrasts	176
F.0.11	Results of the statistical analysis using SVM Linear classifier using elderly participants	178
F.0.12	Tests of Within-Subjects Effects and tests Within-Subjects Contrasts	178

F.0.13	Paired T-test Results of the statistical analysis using SVM Linear classifier with young and elderly participants	180
F.0.13.1	Independent samples test summary	180
F.0.13.2	T-Test Box Plot	181
F.0.13.3	T-Test Profile Plot	182
F.0.14	LDA Classifiers Independent samples test summary	183
F.0.14.1	LDA T-Test results	183
F.0.15	Logistic regression classifiers Independent samples test summary	184
F.0.15.1	Logistic regression classifiers T-Test significance results	184
F.0.16	KNN Coarse classifiers Independent samples test summary	185
F.0.16.1	KNN Coarse classifiers T-Test significance results	185
F.0.17	Subspace discriminant classifier Independent samples test summary	186
F.0.17.1	Subspace discriminant classifiers T-Test significance results	186
G	Results of the performance gains summary	187
G.0.1	Results of the performance gains summary	188
G.0.2	Comparing the standard deviation and variance	193
H	Mobility and empowerment	196
H.1	Re-establish a sense of empowerment	196
H.1.1	Disabled persons daily life	196
H.1.2	Understanding the problem and the difficulties	196
I	Statistical Analysis of Variance (ANOVA)	198
I.0.1	Statistical Analysis ANOVA Test 1, 13/01/2017	198
I.0.2	SPSS Software Settings for statistical analysis of feature extraction methods	200
J	LDA Methodology	202
J.1	LDA Methodology	202
J.1.0.1	Verification of the eigenvector-eigenvalues calculations	203
K	Error Awareness Sustained Attention Task	204
K.1	Error Awareness Sustained Attention Task	204
K.1.1	Experiment Objectives	204
K.1.1.1	Experiment Configuration	204
K.1.1.2	Methodology	204
L	Error Selection Menu	207
	References	210

List of Tables

3.1	Experiment preprocessing configuration	31
3.2	Tests of Within-Subjects Contrasts Experiment 1	37
3.3	Tests of Within-Subjects Contrasts Test Experiment 2	38
4.1	Experiment preprocessing configuration	42
5.1	Experiment Configuration	61
5.2	Young and elderly classifier results	68
5.3	ADS Classifier with relative peak features performance	69
5.4	Young Participants	69
5.5	Elderly Participants	69
8.1	Young and elderly classifier results	108
F.1	Tests of Within-Subjects Effects	168
F.2	Mauchly's Test of Sphericity	168
F.3	Tests of Within-Subjects Contrasts	168
F.4	Tests of Within-Subjects Effects	170
F.5	Mauchly's Test of Sphericity	170
F.6	Tests of Within-Subjects Contrasts	170
F.7	Tests of Within-Subjects Effects and tests Within-Subjects Contrasts	172
F.8	Tests of Within-Subjects Effects and tests Within-Subjects Contrasts	174
F.9	Tests of Within-Subjects Effects and tests Within-Subjects Contrasts	176
F.10	Tests of Within-Subjects Effects and tests Within-Subjects Contrasts	178
H.1	Sample of daily life functions	196
K.1	Experiment Configuration	205

Chapter 1

Introduction and Motivation

1.1 Background

There is a need to improve the quality of life for a person with a disability who has lost their independence of movement in their daily life. Their debilitating condition maybe a severe motor disability or limb amputation. The loss of independence has consequential effects on the person's health and social well-being in their community. Imagine for one moment the transformation from being able-bodied fit and healthy to a person with disability due to a serious illness such as stroke or paralysis after a serious accident. For example, you once were able to pick up an object in front of you and place it back on the table a simple basic task for the able-bodied. A person with paralysis or loss of a limb can only look at the object on the table with no ability to move the object. The brain-computer interface (BCI) is a well-known method to re-established a sense of independence for a person with disability.

This research will have an incremental impact on the scientific community by developing a novel feature extraction algorithm, and developing a novel classification algorithm for error detection and classification using EEG signals. The research approach evaluated the state of the art error detection classifier using data from two different age groups young and elderly. The performance showed a statistical difference between the aforementioned age groups therefore there needs to be an improvement in error detection and classification. A novel relative peak feature selection method was developed which removed the statistical difference between the two groups, subsequently improve the performance of the error detection classifier.

A disability can occur to any person at any stage in life or even born with a disability or following a serious illness. Debilitating conditions such as amyotrophic lateral sclerosis, spinal cord injury, brain-stem strokes will inhibit a person from interacting with their environment around them [2],[3], Fig 1.1. Brain-computer interface (BCI) enables a direct interaction of the human brain with the external environment [4]. To enable this interaction the BCI will record brain signals from inside the persons' brain.

BCI will analyse the brainwave signals and interpret them to determine the desired actions of an external device such as a prosthetic hand [5], or moving a cursor on a computer screen [6]. Most of the BCI systems use electroencephalograph (EEG) to acquire brain signals [7]. During the analysis of EEG signals there may be an occurrence of an error-related potential (ErrP) [8] [9],[6] which is a reaction response to an erroneous event. Moreover, there is also a performance monitoring and adaptive behaviour exhibited with the person brain [10] in addition to the ErrP. Usually, these ErrP events, are evoked when the desired actions of the person have not been fulfilled [9, 11]. It would provide a significant advantage to develop BCI that detects evoked ErrP accurately, and corrects the error and update the system accordingly. In addition, a BCI could be further improved by adapting the parameters of the system such that the likelihood of the same error in the future is reduced. A person with a disability will have a more empowering

experience with a BCI that responds and adapts to the person such that the interactions are more intuitive [12].

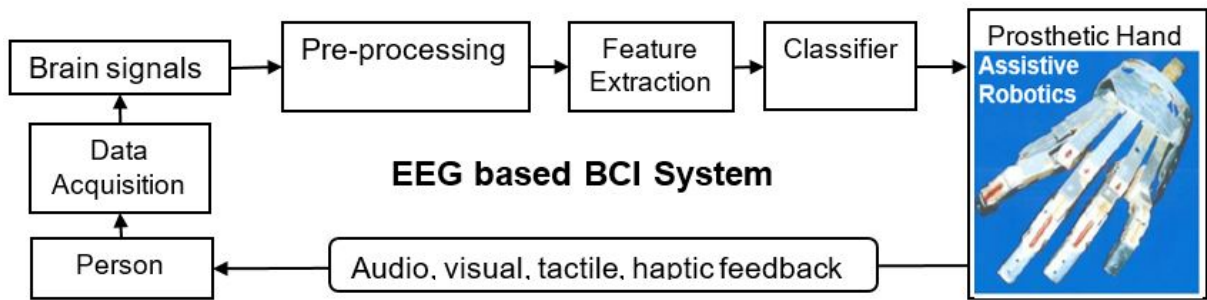


Figure 1.1: A schematic diagram of an EEG based BCI System

1.2 Motivation

The motivation behind this thesis is to improve the quality of life for a person who has lost the ability to interact with their environment [13] due to severe motor disability. Imagine for one moment to be a person with no upper arm movement. What kind of assistance could make their life more enabling and restore a sense of hope and independence[14]? The proposed BCI is non-invasive and previous research [15, 9, 16, 3] has show the possibility of developing methodologies for an assistive robotic device by this modality[17, 18]. In addition, BCI using EEG is affordable and portable to envisage practical applications.

Firstly how do we know their choice and adapt BCI to the person to improve accuracy? Secondly, who is going to be there to understand their choice? Moreover, how can an assistive robotic device interact with the person in a cooperative shared task[19]?

The anticipated research will hopefully have an impact on the scientific community by developing novel algorithms and methods for error detection and classification. Such that new methods of error-driven learning will establish new strategies for adapting the feature selection and classification. The research investigated the shared control strategy such as, at what point should a semi-autonomous robotic device switch to a fully autonomous and back? What would be the strategy, and the implications?

1.3 Aims and objectives

This research aimed to improve the classification accuracy of BCI in the elderly and young participants. The aims and objectives of this research work presented in this MPhil thesis were to improve the accuracy in terms of the following objectives:

- 1) Improving the accuracy of EEG-based error detection Please see 1.3.1
- 2) Improving the accuracy of EEG-based motor imagery BCI. Please see 1.3.3
- 3) Shared control strategy such that the prosthetic robotic hand may recognise kitchen utensils? In order to support user interaction with less workload. Please see 1.3.4

To address error detection, we propose improving EEG-based error detection using a novel feature extraction method. Thus in chapter 4, we propose a novel feature extraction method called the relative peak features (RPF). Further, we combine the aforementioned proposed feature extraction method with a proposed novel classification algorithm in chapter 5 called the adaptive decision surface ADS Classifier.

To address the accuracy in motor imagery, we propose to use common spatial filters combined with the adaptive decision surface (ADS) Classifier. Spatial filters are used to reduce

unnecessary spatial EEG electrical activity and highlight a particular location of interest. In addition, the spatial filter will maximize the signal to noise ratio such that accuracy of EEG-based communication will be improved shown in [20]. The classification process will benefit from the improved EEG signal with a more accurate classification. Please see chapter 6.6 a novel three dimensional probability-based classifier for improving motor imagery-based BCI

To address the objective of recognising kitchen utensils; I developed a method discovered by Hubel Wiesel as shown in [21]. Please see chapter 7 Shared control image decoding and tactile LiDAR of kitchen utensils for an assistive prosthetic hand.

1.3.1 EEG-based error detection in adults and elderly

As we get older brain signals attenuate; therefore, the BCI needs to accommodate for these brain signal changes. It has been proposed to develop a novel algorithms to improve the accuracy and speed of detecting an error signal on-line using EEG. An evaluation of the error detection accuracy in the elderly and adult participants was carried out using the state of the art error detection algorithm. The results of the algorithm were evaluated using statistical methods such as the analysis of variants (ANOVA). The results showed a need for error detection accuracy improvements particularly in the elderly since they have an attenuated ErrP as widely shown in neuroscience papers [22, 23].

1.3.2 Error Driven Learning

When a BCI can learn from its misclassification's, it will inevitably be more accurate in the future classifying events; therefore, It has been proposed to use the detected errors, as an adaptive signal to improve the performance of the error-driven learning system. Thus it will require developing a novel methodology for enhancing a future BCI system using the detected error signals. It is anticipated that there will be a requirement to accomplish this adaptation in real time with significant computational agility. Error-driven learning maybe comprising of the following steps;

1) Error correction

A method of BCI error correction such that an error can be reversed and corrected therefore resolving a bad adaptation.

2) Adaptation

Adaptive decision surface (ADS) classifier is proposed in Chapter 5. For the implementation of the ADS, it has been hypothesized that the BCI system has the ability to decide independently the evolution strategy to correct the previous error, such that after adaptation the BCI system is less likely to repeat the same mistake given the same scenario of inputs [18, 24, 24, 16, 18].

3) Shared control strategy

A shared control strategy, such that reinforcement-learning methods should be incorporated into the BCI system and the robotic assistive device. Furthermore, combining the developed algorithms from the above objectives into a single BCI system [24, 16] will enable a more accurate system. Error detection could be integrated into the shared control strategy as part of the reinforcement learning such that when an error is detected, the learning algorithm is updated. Please see the proposed novel adaptive decision surface (ADS) classifier in chapter 5

1.3.3 A novel classifier for improving motor imagery

Motor imagery is a type of EEG-based BCI that detects a person imagining moving their limbs. A motor imagery BCI based assistive robotics solution has the potential to empower the upper mobility independence of a person with a disability. The objective will be to compare the classification performance of well-established classifiers with a novel prototype classifier. I develop an adaptive decision surface ADS classifier with the future objective to augment an

assistive robotic prosthetic hand to open and close to grasp an object in cooperation with LiDAR sensors. Please see chapter 6.6 A novel three dimensional probability-based classifier for improving motor imagery-based BCI

1.3.4 Object recognition for Kitchen utensil identification

In order to reduce the workload on the person with a disability, it would be advantageous for an assistive robotic device such as a prosthetic hand to recognise the kitchen utensil. Furthermore, it would be an even greater advantage when the prosthetic hand could recognise the utensil and have knowledge of how to use it in terms of the functionality of the objects.

A disability can occur to any person at any stage in life. The debilitating conditions such as spinal cord injury, brain-stem strokes, accidents or illness resulting in upper limb amputation will inhibit a person from interacting with their environment.

Brain-computer interface (BCI) enables a direct interaction of the human brain with the external environment. Moreover, an assistive robotic device with the capability to interact with the person's environment by recognising kitchen utensils using image processing and Light Detection and Ranging (LiDAR) has been implemented. Moreover, this additional shared control capability with further reduce the workload on the user. (For more information please see chapter 7)

The assistive robotic device has the ability to retrieve an object for the person with a disability. Moreover, the prosthetic hand will be aware of how to handle the objects. Fig 1.2 shows a selection of kitchen utensils.

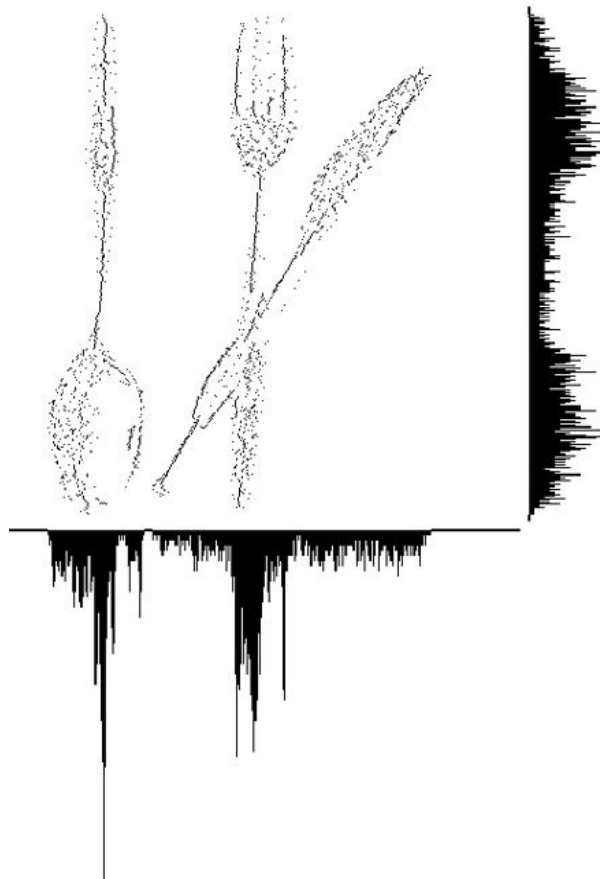


Figure 1.2: Cutlery

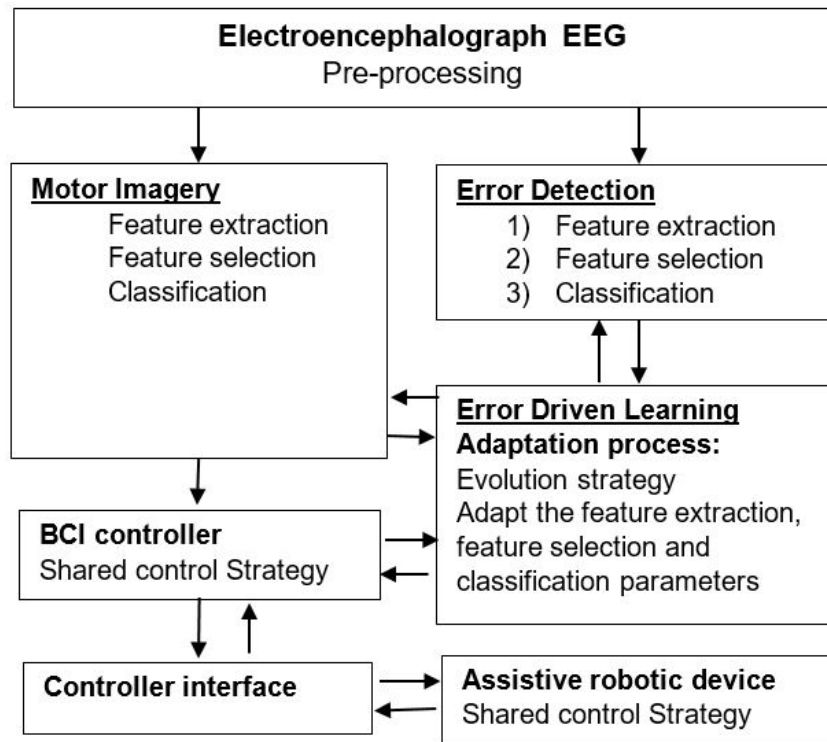


Figure 1.3: A schematic diagram of the envisaged adaptive brain-computer interface system

An overview of the proposed adaptive computer interface process flow is presented in Fig 1.3 schematic diagram. The process flow stages are not limited to the following stages:

- **Electroencephalograph EEG:** Electrodes connect to specific channels acquire specific brain-wave activity
- **Motor imagery algorithm:** Acquires the persons imagined movements
- **Error detection algorithm:** Captures errors when the person is aware of the error
- **Error driven learning:** Evolve a strategy to adapt the feature selection and classification parameter
- **BCI Controller:** Coordinate the shared control strategy between the BCI and the assistive robot
- **Controller interface:** Provide the managed connection between the BCI and the assistive device
- **Shared control strategy:** Manage the autonomous and semi-autonomous modes of operation
- **Assistive robotic device:** Cooperative robotic prosthetic hand, Fig 1.4



Figure 1.4: Robotic prosthetic hand (RPH) ©1986 Mr A L Ashley

1.4 Impact to the scientific community and wider society

This research will impact on the scientific community by providing a novel method of feature extraction in chapter 4 and error detection classification in chapter 5. I envisaged that the wider impact on society should enhance independence for a better quality of life for people with disabilities?

1.5 Publications from this work

The results of this thesis has been published in two IEEE conference papers:

The first IEEE accepted conference paper publication authors: Adrian Ashley, Mahnaz Arvaneh, Lyudmila Mihaylova, with the title of **“A Novel Three Dimensional Probability-based Classifier for Improving Motor Imagery-based BCI”** published in: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) date of Conference: 12-17 May 2019 Date Added to IEEE Xplore: 17 April 2019 INSPEC accession Number: 18813183 DOI: 10.1109/ICASSP.2019.8683136

The second IEEE accepted conference paper publication authors: Adrian Ashley, Mahnaz Arvaneh, with the title of **“Improving EEG-Based Error Detection Using Relative Peak Features”** 2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) in conjunction with the 43rd Annual Conference of the Canadian Medical and Biological Engineering Society Montréal, Canada, July 20-24, 2020

Chapter 2

Literature Review

2.1 Introduction

The literature review will establish the state of the art in the area of brain-computer interface in terms of an “*Adaptive learning in brain-robot interaction*”. There is a substantial literature in the subject area of BCI; therefore the literature review will primarily address relevant components such as error detection algorithm and error-driven learning. Firstly the review begins with investigating the different processes used in the error detection algorithm. Secondly, the review will proceed to identify some of the current solutions in the literature. Finally, the review will further investigate error-driven learning and the shared control strategy used in the current literature.

Classifiers, in general, require a significant amount of training data in order for the classifier to differentiate between the different classes. As a result of detecting the aforementioned error-related potential provided the brain-computer interface, with an opportunity to learn and adapt. Therefore following the adaptation of the brain-computer interface the error-related potential is less likely to occur in the future. The consequences of using insufficient training data will result in classification errors. [15].

There are practical considerations to consider during the acquisition of the EEG signals. The EEG signals are in the magnitude of microvolts and subsequently are operating and subjected to muscle artefacts, irrelevant but concurrent neural activity and non-stationarity of EEG signals [25]. In addition, the electrodes are subjected to a physically changing environment. Firstly the physical movement of the head may result in electrodes moving by sliding on the scalp, hence affecting the signal properties. Secondly, artefacts are generated from muscles may overpower or produce interference with the EEG signals. Moreover, the person’s physiological signals may change.

The amplitude of an EEG signal is specifically reduced in the elderly [23], [22] The strength of certain EEG signals is different when compared between age groups. For example, the amplitude of an ErrP is reduced as the person ages [22]. As a consequence of this reduced amplitude detecting the occurrence of certain EEG signals such as an ErrP is more challenging, firstly in terms of training a classifier, secondly maintaining the accuracy of the brain-computer interface.

2.2 Brain-computer interfaces (BCI)

A person is given the scenario of losing all forms of conventional human interaction such as movement and speech. What possible system can extend the limitations of their boundaries to

which they find their existence confined? The person's brain activity can be detected and interpreted to establish some basic form of limited interaction. The detectable brainwave electrical patterns can be associated with specific actions that the person used to do in the past. Opening and closing their right or left hand will be associated with a particular electrical activity pattern detectable with electrodes over the top surface of the scalp. These discernible patterns can be associated with a cognitive process in terms of thought or physical movement, such as thinking about left or right arm movements. In real-time the BCI must capture the brain's electrical activity and accurately classify the discernible patterns and associate that pattern with a cognitive process. BCI has the potential for a person to interact with their environment and feel less isolated. BCI will enable a new method for that person to communicate by classifying these discernible patterns which may be associated with a cognitive process [26, 27, 28, 29]. Other applications of BCI, such as game interface, BCI speller and mobility controlling a wheelchair.

2.2.1 Recording brain activity using EEG

Electrodes are placed on the scalp at a specified position to record brain signals from the person. The signals from the electrodes on the scalp are connected to the EEG recording machines. A trained operator may determine the brain activity from the EEG recordings.

2.2.2 EEG electrode positions

EEG uses electrodes positioned on the top scalp such that electrical activity can be detected at certain locations on the person's scalp. Fig 2.1 shows the different areas of the brain, such as the 'Primary motor cortex' associated with physical movement. The same area shows changes in electrical activity when a person imagines a movement.

Moreover brain oscillations in specific frequency bands are categorised as delta: δ (0.5 - 4 Hz), theta: θ (4 - 8 Hz), mu: μ (7.5 - 12.5) voluntary movement such as purposefully picking up a cup of tea and all the motor actions instigated by the person, alpha: α (8 - 12Hz), beta: β (12 - 30 Hz), gamma: γ (> 30 Hz) [30], [31]

There is an international standard for positioning the electrodes called the 10-20 system. For example, in order to detect an error-related potential (ErrP), the following electrode locations are the most informative Fz, FCz, Cz, CPz, Pz, POz.

Motor and Sensory Regions of the Cerebral Cortex

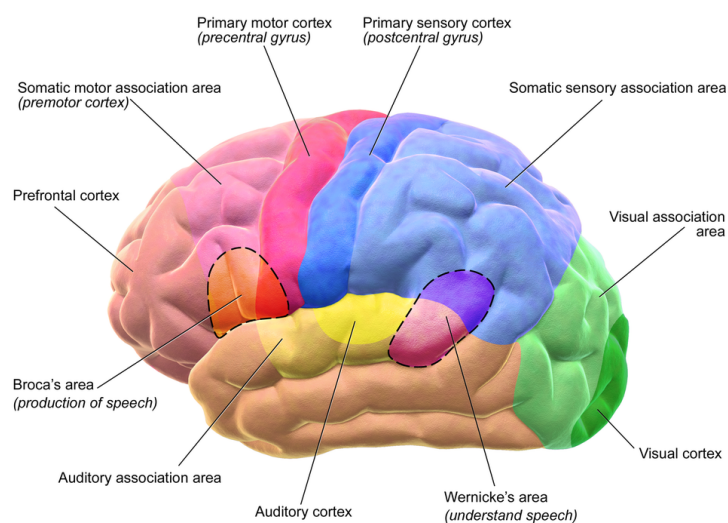


Figure 2.1: Brain Motor Sensory [32]

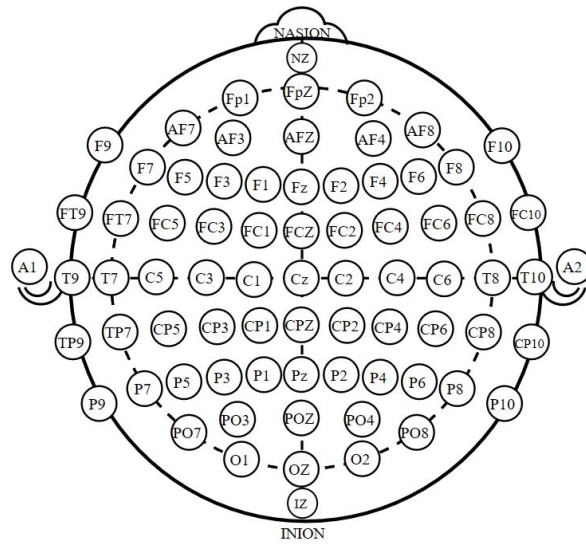


Figure 2.2: International 10-20 system for EEG [33]

2.2.3 Different types of BCI

BCI has three main configurations, such as Invasive brain-computer interface which involves implanting a device into the brain. The signal strength will be very high; however, there is a potential build-up of scar tissue [26]. Moreover, there is a high risk of infection and the bodies immune system reacting to the inserted device. In a partially invasive brain-computer interface, the device is implanted under the skull but on the surface of the brain. The signals have a reduced amplitude and a risk of infection. Non-invasive brain-computer interface does not involve surgery, so this method has the least risk of infection [26]. This project will focus on the non-invasive brain-computer interface using an EEG to acquire the brain signals because there is less risk of infection and the equipment is mobile.

2.2.3.1 P300

Communication is possible by using a device called the P300 speller by generating text that a computer may use the command instructions. Thus establishing a communications path between a person with a disability and the wider environment and other people. There is an event-related potential (ERP) known as the P300 (P3) wave [34] occurring in parieto-central of a person's skull. The P300 event occurs 300 ms after the presentation of an infrequent but relevant stimuli. There are two components of P3 which are P3a and P3b. The P3a associated with a persons attention generating a positive potential located in the frontal-central area of the scalp occurring 250-280 ms. The P3b is also a positive potential occurs about 300 ms. The Oddball paradigms have been used to elicit the P3a and P3b of the P300 whereby a sequence of visual stimuli for example letters are presented to the person. The person was required to target letter 'T' the oddball stimuli. The P3b is detectable 300 ms after the anticipated target 'T' oddball stimuli; however, the P3b may occur in the range between 250ms to 500ms.

Li et al [35] presented a paper for an EEG based BCI system cursor control which combined a Mu/Beta rhythm and P300 potentials. Dinteren et al [36] reviewed the development of the P300 over a lifespan. The paper hypothesized that the latency and amplitude such that the P300 possible indicated brain efficiency. Dal Seno et al [37] proposed and online detection of P300 and error potentials in a BCI speller.

2.2.3.2 Steady State Evoked Potential (SSEP)

Mischner et al [38] proposed using multi-modal frequency tagging for a BCI. The person is presented with stimuli which have a predefined frequency which evokes a response with the

brain that matches the same frequency as the stimuli. The stimuli could be either of the following. Steady-state somatosensory evoked potential (SSSEP) evoked by vibration on the skin of the person the frequency of the vibration can be detected in the EEG signals [39]. Auditory steady-state response (ASSR) detectable in the auditory cortex when presented with an audio signal [40]. Steady-state visual evoked potential (SSVEP) detectable in the persons' visual cortex when stimulated with a flashing display a predefined frequency from different targets. Yin et al [41] proposed a combined SSVEP and P300 BCI system for a more accurate BCI speller. Please see section 2.2.3.1 for more information about P300.

2.2.3.3 Motor Imagery-BCI (MI-BCI)

When a person imagines moving their left or right hand or opening and closing their hand without actually physically moving their hand, the brain's electrical activity can be detected by EEG signals and thus use motor imagery BCI to capture the person's desire to open or close their hands. The possibilities therein provide the practical solution for a prosthetics hand to fulfil the persons desire to physically open and close a prosthetic hand using motor imagery BCI. In this thesis we will use a novel algorithm to enhance a motor imagery brain-computer interface MI-BCI [35], [42]

2.2.4 BCI Components

Please see figure 1.1 for an overview of the BCI components comprising of; data acquisition, brain signals, pre-processing, feature extraction, classifier.

2.2.4.1 BCI pre-processing

We need to do pre-processing because EEG is very noisy, and many artefacts need to be removed to improve the signal to noise ratio; this will improve the classification process. Sampling the EEG brain waves from the set of electrodes configured using the 10-20 system 2.2.2. Selecting EEG channels such as 'Cz' depend on the information required. Re-sampling adjusts the level of detail from the original signal. Bandpass filters are used to select the frequency band of interest, for example, band-pass zero phase filter 1 to 10Hz. EEG segment extraction selects the capture window or frame such as the start and stop time in (ms). Artefact rejection removes unwanted signals such as muscle movement or eye blinking, for example, reject any value greater than 30 microvolts [43]. Baseline correction captures the EEG data 200 (ms) before the capture window to adjust the amplitude to a new baseline for the EEG channel of interest. Motor imagery preprocessing, for example, may have the following steps. The EEG is band-passed filtered from 8Hz to 13Hz to extract the μ band which is detected over the motor cortex [44]

2.2.4.2 BCI feature extraction

The ability to identify patterns within the EEG signals will require a decision about which features will represent an error. The specific pattern can be associated with certain events such as motor imagery or detecting errors. It is beneficial to improve the speed and accuracy of motor imagery or detecting errors for a BCI system. Feature extraction is a substantial part of the motor imagery and error detection algorithms [45].

There are different types of features that can be used in BCI systems such as using:

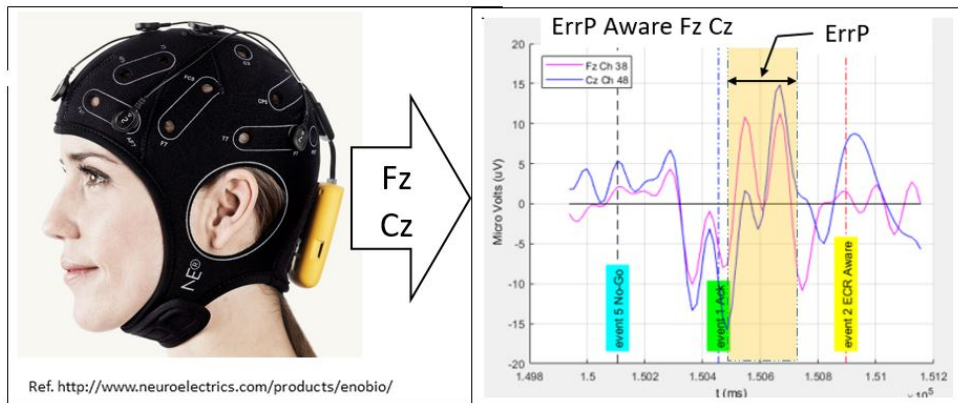


Figure 2.3: ErrP Fz Cz

Amplitude: The amplitude of the EEG signals Fig 2.3, has been used as a feature in BCI [46]. However, in [22, 23] showed the age of a participant has an effect on the amplitude of the EEG signals during an error. Subsequently detecting this feature during an error will be more challenging. In [22] suggest the error negativity (Ne) and error positivity (Pe) is consistently smaller in terms of the amplitude in the elderly participants.

Wavelets: Wavelet transforms used in signal processing are in the frequency-time domain was first proposed by Haar in 1909 [47]. Haar Wavelet was applied to Iris recognition of image features [48]. Wavelets are related to harmonic analysis and have been applied in BCI, for example, Morlet wavelet was applied to the data set [23], please see Fig 6.6 in the frequency-time domain.

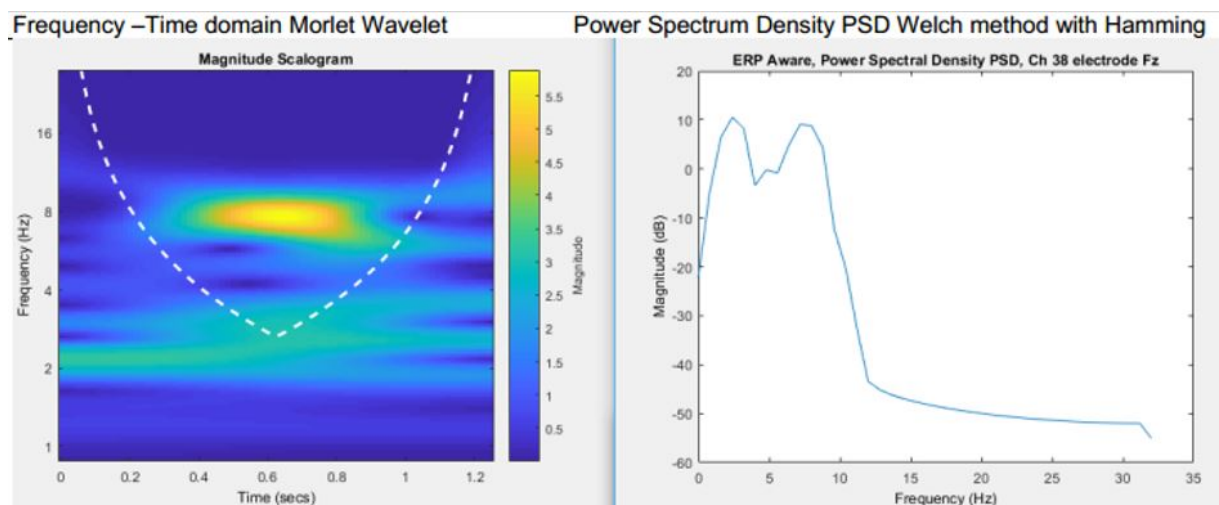


Figure 2.4: Morlet wavelet ErrP Fz Cz by A Ashley

In addition, the following has also been used as a feature, power bands (PB) **Autoregressive (AR) and adaptive autoregressive (AAR)** features used in statistics and signal processing it is a representation of a random process such that the output or the AR model depends on previous values generated. AR model is a random probability distribution or stochastic difference equation that can be analysed statistically. In [28] used AR features from the EEG data and then used a Bayesian logistic regression model to classify them in a motor imagery BCI.

Common Spatial Patterns (CSP):

Common spatial pattern (CSP) used in signal processing to transform multivariate signals to subcomponents for adding, such that a maximum variance between two windows is achieved. In the application of BCI, the CSP linear transform is a projected multi-channel EEG data into a low-dimensional spatial-subspace with a projection matrix. The prime objective is to increase the variance in one class and minimising the variance in the other class achieved by spatial filters. Ghaheri et al [49] showed a method using a temporal windowing CSP for motor imagery classification such as LDA to discriminated between classes. [42]

2.2.4.3 Feature characteristics and properties

The following should be considered when selecting a feature for a classification algorithm:

EEG Channel selection: Deciding which channels to use for feature selection?

Feature integrity: How to avoid information loss from the features

Time-locked windows: The specifications of the time locked window

Time-stamp : The time event associated with the feature to match the occurrence of the brain pattern

Multiple features: Can the feature be combined with other features from other EEG Channels?

Feature reliability: Is the feature reliable between different subjects?

Algorithm selection: Is the feature compatible with classification algorithm

Quantity of features: How many features are needed for accurate classification?

Signals to noise ratio: EEG signals are noisy and it will affect the captured features

Outliers: During feature acquisition outliers maybe captured

Training data : The training data may be small which will affect the quantity and quality of the features, Subsequently effecting the classifier

Increased dimensionality: Usually, more that EEG channels are used with multiple time segments are concatenated [45]

non-stationarity: The feature is non-stationary and vary within the session and between sessions data [45]

2.2.5 BCI Classifiers

One of the main components for BCI is the ability to recognise a particular cognitive process associated with an electrical patterns within the EEG waveforms. There are various machine learning classifiers widely used in BCI. A brief description of some of the main classification methods will follow [19, 50].

Lotte et al [45] reviewed classification algorithms for EEG BCI and suggest two different types of classifiers such as regression [51] a statistical process to analyse the relationships of variables or classification [28] used a neural network for categorising new vectors and making a decision about which class this new vector belongs during a motor imagery task.

Lotte et al [45] suggest that classification performance depends upon the feature extraction 2.2.4.2 and the classification algorithm.

2.2.5.1 K-means clustering algorithm (K-MC) and Multi-layered perceptrons

Orhan et al [52] presents a method for the classification of EEG signals by using K-means clustering together with a multi-layer perceptron(MLP). The authors investigated the impact of the k-means clustering algorithm on the multilayer perceptron neural network accuracy for detecting epileptic events. The MLP will be further described below in 2.2.5.5. The process of k-means clustering is given a set of input vectors. Place random centroids around the input data. With each input data find the nearest euclidian distance to the centroid. Associate the data point to the nearest centroid. Determine the new positioning of the centroid based on the means of the cluster of data point associated to the centroid. This process is repeated until the

movement of the centroid is constant. The k-means clustering algorithm is fast but sensitive to noise and outliers in the data and uses numeric values only [53]

2.2.5.2 K-Nearest Neighbour algorithm (k-NN)

The k-NN is used in pattern recognition it is non-parametric with applications for classification and regression i.e estimating the variables in terms of dependent and independent variable relationships. If the k-NN is used for classification the output will be the class membership. Generally the processes is as follows. Given an unknown object its classification will be determined by the vote generated by the neighbours please see Fig 2.5

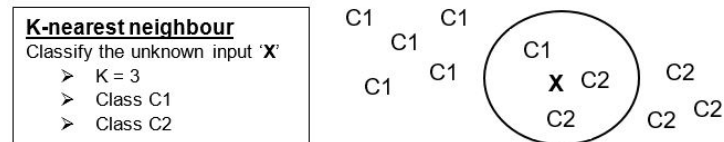


Figure 2.5: k-NN Example

With $k=3$ the algorithm will look at the 3 nearest neighbours to 'x' to determine which class to assign 'c' Fig 2.5. The value of k must be odd and not a multiple of class numbers such that the resulting vote is decisive.[54, 55, 19] In the case of k-NN where $k=1$ a 'Voronoi' partition is used to divided regions. The size of the regions is a function of the distance between the vector features. [56] Suggests methods of fast coverage for a Centroidal Voronoi tessellation (CVT).

In [57] a motor imagery classifier performed a maximum classification accuracy of 72% using k-NN.

2.2.5.3 Bayesian Network (BN)

A probabilistic graphical model can be used to represent a set conditional dependencies which can be represented by a directed acyclic graph (DAG) see Fig 2.6 and their random variables are the nodes of the DAG. Such that a probabilistic relationship in a given event can be established in a probabilistic graphical model. This probabilistic model is also known as the joint distribution. The application of the Bayesian network was used in a study for machine learning [54]. For example, a Bayesian network can be used to determine the relationships between cognitive processes or EEG features. Such an event or stimuli the Bayesian network will compute the probability of a specific EEG feature. The probabilistic model is also known as Bayes network, belief network, Bayesian model, and probabilistic directed acyclic graphical modal. [54] suggests that the Bayesian network is useful for combining prior knowledge. Moreover, [58] presented a paper about precise probabilistic revisions of beliefs used in an expert system. Further [59] showed a novel multi-variate discretization approach for a Bayesian network structured learning in robot grasping. Bayesian networks have been used to recognise emotions using EEG signals proposed in [60], by comparing the calculations from standard values in the Bayesian network. [61] Presented a paper about the concept of affordance which is used in robotics for making decisions about objects. Moreover, the paper used recent advances in probabilistic programming using an iCub robot.

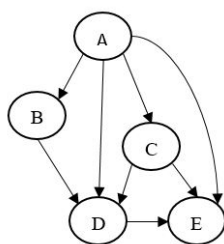


Figure 2.6: Directed acyclic graph DAG

2.2.5.4 Artificial Neural Network (ANN)

Classification of EEG signals has been presented in [62] using wavelet transforms (WT), a classification method using ANNs and logistic regression(LR). ANNs require training data to set the weights of the individual artificial neurons [63]. ANNs are able to classify non-linear transformations in order to classify the data [19]. The multi-layered perceptron is an ANN comprising of many layers of ANNs. [64]

2.2.5.5 Multi-layered perceptron (MLP)

The configuration of the MPL is as follows: an input layer then one or more hidden layers then an output layer. The output of each layer forms the input to the next layer. The input feature vector is applied to the input layer and the classification of the input feature vector is determined by the output layer. [64]. MLP can classify any number of classes and have been applied to BCI classification such as binary [65] ,multi-class [66] MLP, synchronous [67] and asynchronous BCI [68],[45].

2.2.5.6 Back Propagation Neural Network (BPNN)

BPNN Propagation forward to acquire the output values, compute the difference error. Propagate the activation back using the training pattern target to establish the difference between the output and the hidden neurons. The adjustment of the weight is updated in the following back pass process. Compute the weights for the hidden layer to the output layer. Compute the weights for the input layer to the hidden layer. Motor imagery classifier performance [57] maximum classification accuracy obtained using BPNN is 80%.

2.2.5.7 Convolution Neural Network(CNN) Deep Learning (DL)

The CNN is comprising of a number of layers such as Convolutions layers, Pooling layers and fully connected layers on the multilayer neural network. The input data to the CNN is usually 2D for example speech patterns or images. In [69] applied CSP using CNN, another example of CNN application is modelling grasp motor imagery. In [70] investigated the application of deep learning to transform motor imagery to robotic grasping. In addition, the paper also looked at a probabilist framework.

2.2.5.8 Support Vector Machine (SVM)

Vapnik et al [71] was the first to propose the Support Vector Machines (SVM) which finds the maximum separation between classes. A radial basis function can be used to generate a non-linear class separation. In [57] a motor imagery classifier performance a maximum classification accuracy of 78.57% using SVM.

2.2.5.9 Linear Discriminant Analysis (LDA)

Ronald A. Fisher formulated the Linear Discriminant in 1936. LDA is similar to PCA, in addition, the data is assumed to be normally 'Gaussian' distributed [72] explained five steps of processing with LDA. ERP Classification Blankertz, et al [73] presented clear tutorial steps for ERP classification. The main requirement for using LDA is that the data is divided into two groups or classes subset of each class is labelled training data and the remainder will be used for testing the trained LDA

2.2.5.10 Interval Type-2 Fuzzy Logic System (IT2FLS)

IT2FLS for Handling Uncertainty Effects in Brain-Computer Interface Classification of Motor Imagery Induced EEG Patterns [74] showed an increase in performance by comparing with other types of classifiers. Such that IT2FLS had an improved performance compared to T1FFLS, LDA, KFD with Gaussian Kernel, linear SVM and non-linear SVM with a homoscedastic Gaussian Kernel.

2.2.5.11 Conclusion classifiers

There is a multitude of classifiers with each having their own advantages and disadvantages. Moreover, all the classifiers may perform differently when data is presented from different age groups. For example, detecting error-related potentials in the young and the elderly. Lotte et al [45] showed that linear classifiers can work as good as a non-linear classifier in BCI.

2.2.5.12 Challenges of BCI classifiers

In the process of selecting features 2.2.4.2 and training classifiers there are a number of problems to consider which effect pattern recognition [45].

Outliers: This can occur during acquisitions in terms of electrical noise EEG has a low signal to noise ratio. Consequently, spurious data could be captured as a feature.

Over-training: The classifier will become less generalised when presented with a narrow diversity of data. Moreover using a small training data set can also limit the classifiers ability to generalise.

curse-of-dimensionality: The quantity of data used to describe the different classes will increase exponentially with dimensionality of feature vectors [45, 75, 76]

bias-variance trade-off: Lotte et al [45] suggests that the lowest classification error can be attained when the variance and bias is low.

2.2.6 Non-stationary EEG signals

The definition of 'non-stationarity' is that the characteristics of the time series such as mean, variance and spectral properties change with time [25]. The electroencephalogram as a time series [25] has a number of characteristics that must be considered. More over using non-linear dynamic systems has made it possible to analyse brain activity [25]. EEG signals are non-stationary and present a high dimensionality concluded by Layne et al [77]. Consequently, any classifiers must be able to accommodate the non-stationarity EEG signals such that the classification accuracy is maintained.

2.2.7 Other problems using EEG, BCI and ErrP

- **Connection leads breaking**

From the user's point of view different people may find it difficult to wear the EEG equipment from a cosmetic point of view. Can it be designed to blend into a hat for example?

- **Classification accuracy**

Accurately classifying an error-related potential ErrP is not robust, and misclassification can occur. Generating a cascade of erroneous commands.

- **Indeterminate state of decision deadlock**

The scenario when the control system is unable to decide which path to take for a mobile robot and the users, input from the BMI is not available. Will a hypothetical mobile robot wait indefinitely in a system decision deadlock?

- **System induced false error-related potential ErrP**

The brain-computer interface may induce a false ErrP in the user. The scenario is best described in a navigation grid. The computer system may navigate a mobile robot on a path that the users are observing and perceiving the moves of the mobile robot as an error and hence evoke an ErrP. The movements of the mobile robot may be misinterpreted as an error in movement and the user may not be aware of a different paths that the computer has chosen based on a different object such as using a more efficient route.

2.3 Error detection using brain signals

The design of a more accurate error detection algorithm must take into account the details of the characteristics of the error. A brief explanation of the error-related signals will follow to define these characteristics.

2.3.1 Error-related brain signals

To develop an error detection algorithm the literature such as [78, 79] have shown some of the firsts EEG event-related potentials (ERP) during an experiment designed to evoke and (ERP) such as a speed response task.

The error-related potential has been characterised as a negative potential deflection called an error-related negative (ERN) in the frontal-central location electrodes within 100ms [22] of the participant's erroneous response. After the error-related negative (ERN) in the frontal-central location electrode, there may follow an error-related positive event (Pe). The participants' awareness of an error has been associated with this positive deflection (Pe) located in the central-parietal region of the scalp.

The authors highlight a correlation between error-related negative (ERN) events and behavioural adjustments [80, 81, 82]. The authors further suggested above that the error-related negative (ERN) signal is associated with participant monitoring an activity and hence reacting to an erroneous situation, this assertion is supported by [83]. The authors suggest that the context of the task moreover what it means to the participants in terms of the importance has the effect of modulating the amplitude of the ERN ([81, 84]).

A future BCI will have to be adaptive to dynamics in terms of the amplitude of the ErrP signals such that sensitivity of the detection system is independent of the emotional significance to the participant of the task. However, the amplitude could be part of the control policy of the future BCI in-terms of the rapid response time of adjusting the adaptability of the learning algorithm of the future BCI. For example a high amplitude ERN would indicate the error event is of a high emotional significance to the participant and hence a rapid adaptability for the future BCI.

The authors have determined the following according to the literature:

- Event-relational potentials (ErrP) signals are consistent over time [85]
- Event-relational potentials (ErrP) signals are reliable over different tasks [86]

2.3.2 Error-related brain signals elicited by observation

Ferrez [9] investigated the possibility that ErrPs being elicited by a person observing an interface, such that when the interface generates an error in terms of opposing the users' intentions. The authors used a statistical classifier based on the Gaussian classifier which must be trained. The output from the classifier is a statistical estimation based on the probability distribution of the classes. The authors suggest that using information about the cognitive states is beneficial in terms of robot interaction specifically awareness of an erroneous response. Moreover by using the ErrP elicited from the person observing an error in movement by a robot will significantly improve the performance of a brain-computer interface.

2.3.3 Error detection process

The following is a brief overview of the components used during the error detection process. The research activity has sought to improve the accuracy. More details will be provided in proceeding chapters which includes some initial progressive results of this novel concept prototype and feature selection method.

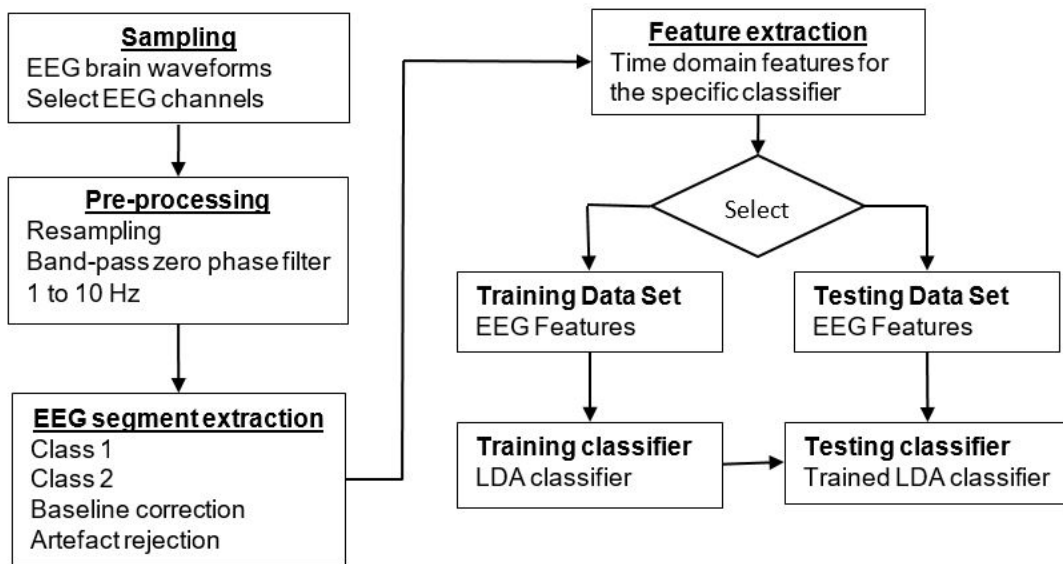


Figure 2.7: A schematic diagram of the state of the art EEG-based error detection algorithm

The generic procedure for extracting features from EEG waveforms is shown in Fig 2.7. Acquire EEG potentials using an EEG Cap with electrodes located by the 10/20 international system. The sampling rate a selected range. Time intervals window starting with a start time after the feedback and ending time after the feedback. A band past zero phase filter of 1 to 10 Hz to isolate the frequencies of interest. The baseline correction interval. The criteria for artefact rejection . Specific features from the EEG waveforms are acquired for the classifier. The features are organised into two sets, such as training and testing data sets for the classifiers. Fig 2.7 A schematic diagram of the state of the art EEG-based error detection algorithm. Fig 5.4 Shows a schematic of prototype adaptive decision surface classifier.

2.3.4 Error detection using time and frequency domain features

Boubchir et al [87] achieved ErrP detection in an SSVEP-based BCI using features from time and and frequency domain and classification using a support vector machine. The objective was to use the ErrP to halt the robot going in the wrong direction and recover a situation. Moreover information in the time and frequency domain of the EEG signals incorporate, Instantaneous Frequency (IF), time and frequency information complexity, Singular Value Decomposition (SVD)

information, energy concentration and sub-bands' energies. The author reported an overall a classification accuracy up to 97% for 50 EEG segments using 2-class support vector machine (SVM) classifier.

2.3.5 Error detection for P300 BCI using real-time Bayesian dynamic stopping framework

Zeyl et al [88] proposed a method of obtaining a P300-confidence with real-time Bayesian dynamic stopping framework. In addition, the authors used stimuli which occurred as a result of epoch and filter delays. The author further used an ensemble of decision trees to combine ErrP and P300-confidence features. The authors explained how the different types of feedback effect the ErrP physiology. The authors reported an average sensitivity of 86.67 % and specificity of 96.59 % and automatic correcting improved the selection accuracy by 13.67 %. The authors suggest the improved performance was as a result of including the P300 confidence in error detection. The P300-based BCI has been successfully used in a BCI Speller [37]. Furthermore the P300-confidence level is used in setting the classification threshold.

2.3.6 Error detection with a double error-related potential (ErrP) detection

Cruz et al [89] proposed a double error-related potential (ErrP) detection to automatically correct erroneous decisions was applied to an ERP-Based BCI Speller. Such that if the wrong automatic correction is applied the ErrP will be subsequently detected. The author report the online average accuracy of the first ErrP was 88.4 % and the second ErrP 84.8 %. Moreover, the automatic correction method improved the system by 5 %.

2.3.7 Anticipating an error

It would be advantageous for a future brain-machine interface to have further insight into a error-related potential. Furthermore, it would reduce the latency by preparing the system for an error-related potential event. The literature suggests a majority brain-machine interfaces are not pre-anticipating such an event.

Steinhauser et al [90] investigated error awareness during an error choice task in terms of the decisions process. Moreover, developing an understanding of how evidence of an error is ascertained. The human event-related potential (ERPs) comprising of a well-established profile specifically known as the error-related negativity (Ne/ERN) and the error positivity (Pe). [90] Hypothesized that part of the decision process involved the Pe. Furthermore, the strength and latency of the suggested accumulated evidence related to the amplitude of Pe. The speed-accuracy trade-off (SAT) was modified during tasks which involved the participant indicating an error. [90] Predicted that a lower amplitude in Pe would occur given a low-speed task compared to a high-speed task. In addition, further analyses in a trial-by-trial approach using a logistic regression approach robustly predicted the occurrence of signalling responses. The authors confirm and support the notion that Pe is involved with the accumulation of evidence for an error prior to error awareness.

Thus, it would be advantageous for the further development of acquiring this build up signal to provide more of an insight into a future BCI system.

2.4 Error Driven Learning

2.4.1 Introduction

An ErrP is evoked when the participant perceives an error made by themselves or by observing an error on a display or by an external device such as a robot.[91] If the ErrP is detected and

decoded then this could potentially allow the BCI to resolve the cause of the ErrP. It would be advantageous to prevent the propagation of an error cascading from the BCI to the assistive device and resolving the cause of the initial error, in addition to learning from the event [92, 91, 37]. Chavarriaga et al [18] investigated the use of error-related potentials in the brain-machine interface by identifying previous research in the past decade. The authors [18] indicated that an ability to identify an error or a mistake is part of adapting a behaviour in nature observed in biological systems such as humans shown in [93]. Decoding an ErrP in a single trial EEG by BCI is a direct opportunity to improve the interface accuracy and performance.

2.4.2 Recognising an 'Interaction ErrP' to improve a Brian computer interface

Ferrez et al [15] investigated the potential of recognising an "Interaction ErrP" such that a Brian computer interface performance can be improved. The authors conducted the experiment to simulate a human-robot interaction with participants who were required to imagine moving a symbol represented by a cursor, to a predetermined target along the shortest path. [15] confirm these "Interaction ErrP" indicated by other research such as [9]. [15] suggest in order to take advantage of these aforementioned "Interaction ErrP" they must be detected in every single trial and using the feedback from a classifier in a BCI [15]. Report on achieving a recognition rate of 81.8 % for correct and 76.2 %, for erroneous single trials. [15] show the results of acquiring mental control in order to operate a brain-actuated device in addition to acquiring cognitive states such as error potentials for the brain-computer interaction. [4] used an inverse model (sLORETA), highlighting the main focus of activity at the occurrence of aforementioned ErrP are, as anticipated in the pre-supplementary motor area and in the anterior cingulate cortex.

2.4.3 Methods of corrective actions

Using the EEG signals the BCI will detect and decode an ErrP which indirectly identifies the participants' intention from and activity. Furthermore, if the BCI has information about the context of the activity. An inference can be made about the activity in addition to selecting the adaptation of the BCI and the assistive robot.

2.4.4 Methods of learning form mistakes

An ErrP once detected has the ability to adapt the ADS 5.2.3.5 in a type of self adjustment. Furthermore, other subsystems may adapt the operation of the BCI and external assistive robot and "learn from its mistakes" [94, 95].

A hybrid BMI / HCI using EEG brain signals such as motor imagery or stimulus recognition in addition used error-related potential as a means of correcting and an error [5, 43, 96, 97, 98, 99] the author [18] suggest that ErrP evoked during the interaction with the machine and the acquisition is implicit.

However, the authors [18] indicated the fact that ErrPs have been influenced by the participant's motivation in the activity [100].

2.4.5 Reinforcement learning

Iturrate et al [24] used EEG recorded brain activity as a reward signal obtained by observing a robot during a task. The authors suggest that an explicit model is not required for the reward signal and that it was possible to acquire subjective aspect for the individual. The authors designed a new protocol using the brain activity associated with the right or wrong action of the task by detecting the error signal which showed a statistical difference. Further, the authors applied reinforcement learning algorithms which used rewards from the brain activity. The authors choose AdaBoost classification algorithms [101] Moreover, the reinforcement learning was applied to a robot to learn about by using the detected ErrP. The robot used the signal from the ErrP to reward or penalise the behaviour of the robots. The authors suggest the framework

for reinforcement learning is Markov decision Process. The authors suggest that in the practice the classifier cannot be trained because each task has to be label. Furthermore, it would make the EEG signal ineffective.

2.4.6 Learning algorithms

Supervised learning The training processed will use a certain portion of the data for training and the other portion for testing the model or classifier moreover the training data is labelled. The model is configured with training data such that the model is able to make the correct classification. When the models' classification accuracy has reached a certain level after sufficient training. Example of supervised learning algorithms are logistic regression and backpropagation neural networks.[102]

Unsupervised learning: In the unsupervised learning the data is not labelled and there are no results. The model is configured by identifying patterns of structures in data. An example algorithm is the K-means.[102]

Semi-supervised learning: The input data is comprising of a mixture of labelled and unlabelled data. The model will organise the data to enable predictions. [102]

Regression algorithm: It is iterative by using the error in the prediction between variables to improve the models' prediction. Examples of a regression algorithm is linear regression.[102]

Instance-based algorithm: also known as winner take all methods adjust the model according to new training data this is determined to be significant such that the model is adjusted to accommodate the new training data. Example of this type of algorithm are K-Nearest Neighbour (KNN) and Self organising feature maps (SOM).[102]

Regularization algorithms This method is also known as Tikhonov a method of regularising ill-posed problems and reduces the problem of overfitting it has been developed in many different areas. In Integral equation, this method was developed by Andrey Tikhonov and David L. Phillips hence Tikhonov-Phillips regularization. In statistics regularization is known as ridge regression.[102]

Decision tree algorithms This method is used mainly for classification and regression the construction is of the tree is determined by the attributed in the training data.[102]

Bayesian algorithms These methods apply the Bayes theorem to data classification. Examples are Naive Bays, Gaussian Naive Bayes, Bayesian Belief Networks(BBN), Bayesian Networks(BN) [102]

Clustering algorithms These methods are generally used a modelling a process that uses a centroid to organise the data into groups or clusters. Examples are K-means, k-Median and Hierarchical clustering. [102]

Artificial neural networks algorithms These algorithms are inspired by the general basic function of a biological neuron such as each input is modified by a weighted value. The collection of inputs are summed such that the artificial neuron will generate and output depends on the input values. The actual biological neuron is far more complicated. Example of networks are , Perceptron, Back-propagation, Hopfield network, Radial Basis Function Network (RBFN) [102]

Deep learning algorithm These algorithms are a progression from the artificial neural network. The deep learning network has several layers and is generally more extensive, examples of deep learning algorithms are Deep Boltzmann Machine (DBM), Deep Belief Networks (DBN) and Convolution Neural Networks(CNN) [102]

Dimensionality reduction algorithm These methods are based on the data structure enables a process of identifying a separation of the data onto a number of classes. A method of using Eigenvalues and Eigenvectors can be used to identify the maximum separation within the data. Example of dimensionality reduction algorithms are Principle Component Analysis this method used unlabelled training data and Linear Discriminant Analysis this method uses labelled training data.[102]

2.5 Shared control strategy

2.5.1 Introduction

A shared control strategy is a high-level control to manage the interaction between the users and the assistive robotic device. A survey [7] was conducted which summarised the recent robot control using a BCI. The survey also identified the brain signals used and the classifiers.

The interactions could be between a very advanced robotics system will operate in a semi-autonomous and fully autonomous mode of operation or it could be a cursor on a computer screen. Montesano et al [103] proposed a shared control BCI such that the device decodes the commands and executes the shared-control BCI task of moving a cursor over a two-dimensional discrete grid of possible target destinations for the cursor the algorithms parameters are updated after detecting an ErrP.

Shared control BCI using error-related potentials the ErrP were used as a feedback or a supervision signal as the cursor is moved to the target destination by the device. The authors have indicated that the error-related potentials have been used to train a device in terms of reinforcement learning [24]. [96]

Shared control BCI correct potential decoding errors the results showed that the participants reached both a predefined and self-selected target destinations in 23 movements on the discrete grid in approximately 19 seconds of EEG signal data. The main objective of the user was to observe the movement of the device such that the users can assess the movement of the device as being correct or wrong. [103]

The authors suggest that within a shared control strategy it is possible to use other brain signals to correct any decoding errors in order to recover the navigation error by the device.[103]

2.5.2 Interactive control strategies BCI

Li et al [104] proposed a 'human-machine shared control strategy' for the application of controlling a wheelchair device. The authors used three different modes of operation:

Brain-computer interface control mode: Comprising of a brain-machine interface which uses a steady-state visual evoked potentials. The authors used two brain signals in order to identify a polar trajectory maintain a continuous curvature within the steering limits of the wheelchair.[104] Autonomous control mode: The authors proposed a 'synthesis of angle-based potential field and vision-based simultaneous localization and mapping technique', such that collision avoidance is maintained. The authors conducted experiments to evaluate a shard

control wheelchair.[104] The results verified the effectiveness of the authors proposed methodology of a shared control a scheme.[104]

2.5.3 Prosthetic hand restoring the sense of touch

Oddo et al [105] restored a sense of touch as an improvement over conventional upper arm prosthetics. They showed that textural discrimination can be implemented by neuromorphic real-time mechano-neuro-transduction (MNT). In addition MNT-evoked EEG activity showed physiologically plausible responses.

2.5.4 Prosthetic hand deep learning image recognition

Ghazaei [106] in Newcastle University used Deep learning-based artificial vision for grasp classification such that object could be automatically grabbed by configuring the hand to match the shape and size of the object.

2.5.5 Adaptive hierarchical brain-computer interface

Bryan et al[16] suggested that previous brain-computer interface systems that provided control for a humanoid robot required a higher-level control with non-adaptive pre-determine behaviour. The authors have suggested an adaptive hierarchical brain-computer interface, such that new skills are acquired by the BCI. In order that high-level commands may be invoked later. This approach reduces the workload of controlling the device. The authors explore hierarchical BCI to certain tasks. In addition, the authors investigate explicitly-defined command sequences for more complex tasks which will require 'multiple state spaces'. shared control [107]

2.6 Summary

In this literature review, we looked at the consequences of insufficient training data for the classifiers [15]. In addition the Limitations of BCI the amplitude of the EEG signals and the effects of age [22]. The non-stationary properties of the EEG signal were highlighted [25] was presented. A brief explanation of the areas of the brain [32] and the EEG electrode positions were shown [33].

The discussion proceeded with a definition of the BCI in section 2.2, followed by a definition of the different types of BCI 2.2.3 such as P300, Steady-state evoked potential, Motor Imagery BCI which will be the main focus. The discussion investigates the BCI components in section 2.2.4 such as BCI pre-processing, feature extraction, Common Spatial Patterns (CSP), the different types of BCI Classifiers. There are three main components Firstly discussing about 'Error detection using brain signals'.2.3 Secondly 'Error Driven Learning' 2.4 and finally a 'shared control strategy' 2.5

Chapter 3

Deficiencies of state of the art error detection algorithms between young and elderly

3.1 Introduction

The brain will evoke an ErrP in the EEG waveform recordings, which we could associate with the person's awareness of the situation or 'error', caused by themselves or by observation which is detectable in an EEG recordings. The ability to detect these ErrPs would enhance a future brain-computer interface.

Many research studies have investigated BCI with assistive robotics [7]. The effect of age in terms of working memory, attention, research has shown the changes during ageing [108]. Further the P300 development in a lifespan [36] has shown the effect of age. Therefore it would be prudent and beneficial to design a BCI that accommodate the abilities, of the different ages groups and in addition to adapting, to the individuals needs.

The present BCI systems do not necessarily take into account the differences in brain signals between the age groups young and elderly, therefore inadvertently diminishing the potential performance gains. As we age, specific brain signals inevitably become attenuated. It is, therefore, necessary to develop a BCI to maintain the performance of the BCI independent of the user's age.

To evaluate the performance of detecting error-related potentials (ErrP) using a Linear Discriminant Analysis LDA-based algorithm, please see appendix J for more information. The analysis of the ErrP in EEG recordings has shown the present performance in the young Fig 3.7 and elderly Fig 3.9 in terms of ErrP detection.

The ErrP occurred during an 'Error Awareness Task' experiment in which the participant will make a mistake and subsequently evoke the ErrP. Consequently, it would be a significant advantage, for a future brain-computer interface, to be able to detect the occurrence of these ErrPs in different age groups. In this chapter, the results have shown differences in performance between young Fig 3.6 and elderly participants Fig 3.8.

The experiments have shown the need to improve error detection in the elderly. Moreover, I anticipated that the primary users of the BCI would be the elderly. Improving error detection performance of a classifier would enable an assistive robotic device to function efficiently by improving the accuracy of detecting errors and subsequently correcting any mistakes or misinterpretations online.

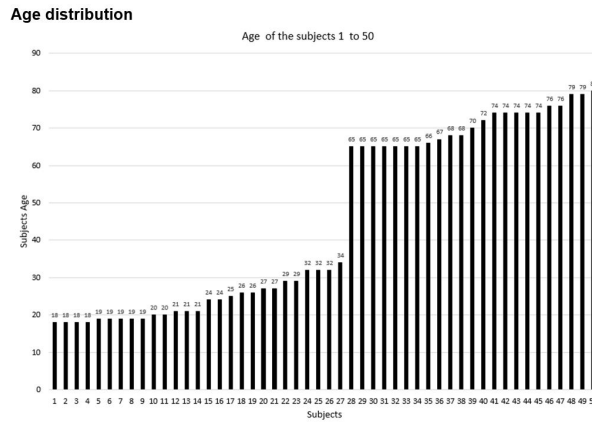


Figure 3.1: Participants age distribution

3.2 Methodology

The EEG feature extraction method has the following stages:

3.2.1 Sampling

EEG signals were acquired using a 64 channel cap with the electrodes located by the 10/20 international system. The sampling rate was in the range of between 512 to 1024 Hz.

3.2.2 Pre-processing

Time intervals were 350 ms window starting 50 ms after the correct/erroneous response (click) and ending 400 ms after. A band pass zero-phase filter of 1 to 10 Hz was applied to isolate the frequencies of interest. The baseline correction interval was -200 ms to 0. The criteria for artefact rejection was approximately 30 microvolts.

Table 3.1: Experiment preprocessing configuration

Parameter	Settings
Channel	Test 1 Fz,Cz Test 2 CPz,Fz,Cz
Re-sampling:	64 Hz
Electrode location:	10/20 international system
Sampling rate:	512 - 1024 Hz
Time intervals:	350 ms window
Starting:	50 ms after onset response
Ending:	400 ms after onset response
Baseline interval:	-200 to 0
Referencing:	Common average
Band pass filtering:	1 to 10 Hz
Criteria for artefact rejection:	30 micro volts

3.2.3 Feature extraction and classifications

Specific features are acquired from the EEG data and labelled with specific events and sent to the classifiers. During the session with the each subject, generally 8 blocks of data were captured. Generally blocks 1 to 4 were sent to the classifier for training and the other blocks of data from 5 to 8 were be sent to the classifiers for testing.

The evaluation is checked in terms of 'Error Accuracy' EA and 'Correct Accuracy' CA. The Brain Computer Interface BCI limitations are the potential for an incorrect classification of the

users intentions. The classifier method used was based on the Linear discriminant analysis (LDA) originating from the Fisher's linear discriminant.

Ronald A. Fisher formulated the Linear Discriminant in 1936. Linear Discriminant Analysis (LDA) is most commonly used as dimensionality reduction technique in the pre-processing step for pattern-classification and machine learning applications. The objective is to project a dataset onto a lower-dimensional space with good class-separability in order to avoid over-fitting 'curse of dimensionality' and also reduce computational costs. [72]

3.3 Experiment

3.3.1 Participants

The data [23] used was from fifty-three participants volunteered to participate in the study. 53 participants; 29 young (19 female) and 24 older (15 female). The younger participants had a mean age of 24.14 years (SD = 5.4), and the older participants mean age was 70.67 years (SD = 4.9). Please see Fig 3.1 for the age distributing.

All participants were healthy individuals. Procedures were approved by the Trinity College Dublin ethics committee and in accordance with the Declaration of Helsinki [109].

3.3.2 Experiment Protocol

3.3.2.1 Methodology

Each participant was presented with a series of visual information displayed in a certain sequence in order to capture the subjects awareness of an error.

The experiment protocol was develop by Lacey and his colleagues [23] in Trinity College Institute of Neuroscience based on a novel variant of Hester et al 'Error awareness task [110]. The participants carried out a practice session in order to establish the suitability for taking part in the experiments:

An outline of the experiment is provided as follows: The visual stimulus has a grey background upon which a centrally positioned white fixation. This fixation cross remains in place while the colour dots change during the experiment sessions.

- (A) The systems generates a random sequence of coloured dots from a given predefined set. Fig 3.5.

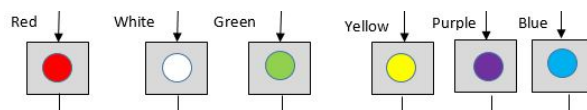


Figure 3.2: One colour is displayed at any given time.

- (B) The Participant receives the visual stimuli to their visual cortex
- (C) The systems waits for the participants response
- (D) The participant should withhold from responding in the following two cases:
- **Colour condition.** If the colour of the dot is blue.
 - **Repeat condition.** If colour of the dot was identical to the previous dot
- (E) The participant will actuate their response from their motor cortex to click a button in the following way:

(E.1) A “go target” indicated by ‘Event 3’ in the EEG recordings requires the participant press the left mouse button with their right hand. This indicate an “Acknowledgement” indicated by Event 1 from the participant.

(E.2) A “no go target” indicated by the following:

△ Colour condition, indicated by ‘Event 4’

△ Repeat condition, indicated by ‘Event 5’

(F) • An ‘Event 1’ after by ‘Event 4’ or ‘Event 5’ is an error made by the participant

(G) • An Error Correction Response ECR indicated by ‘Event 2’ in EEG recording is generated by participant pressing the right mouse button using their index finger of their right hand. This indicates that the participant is “aware” of the error.

Please see the schematic diagram Fig 3.3 and flowchart Fig 3.4

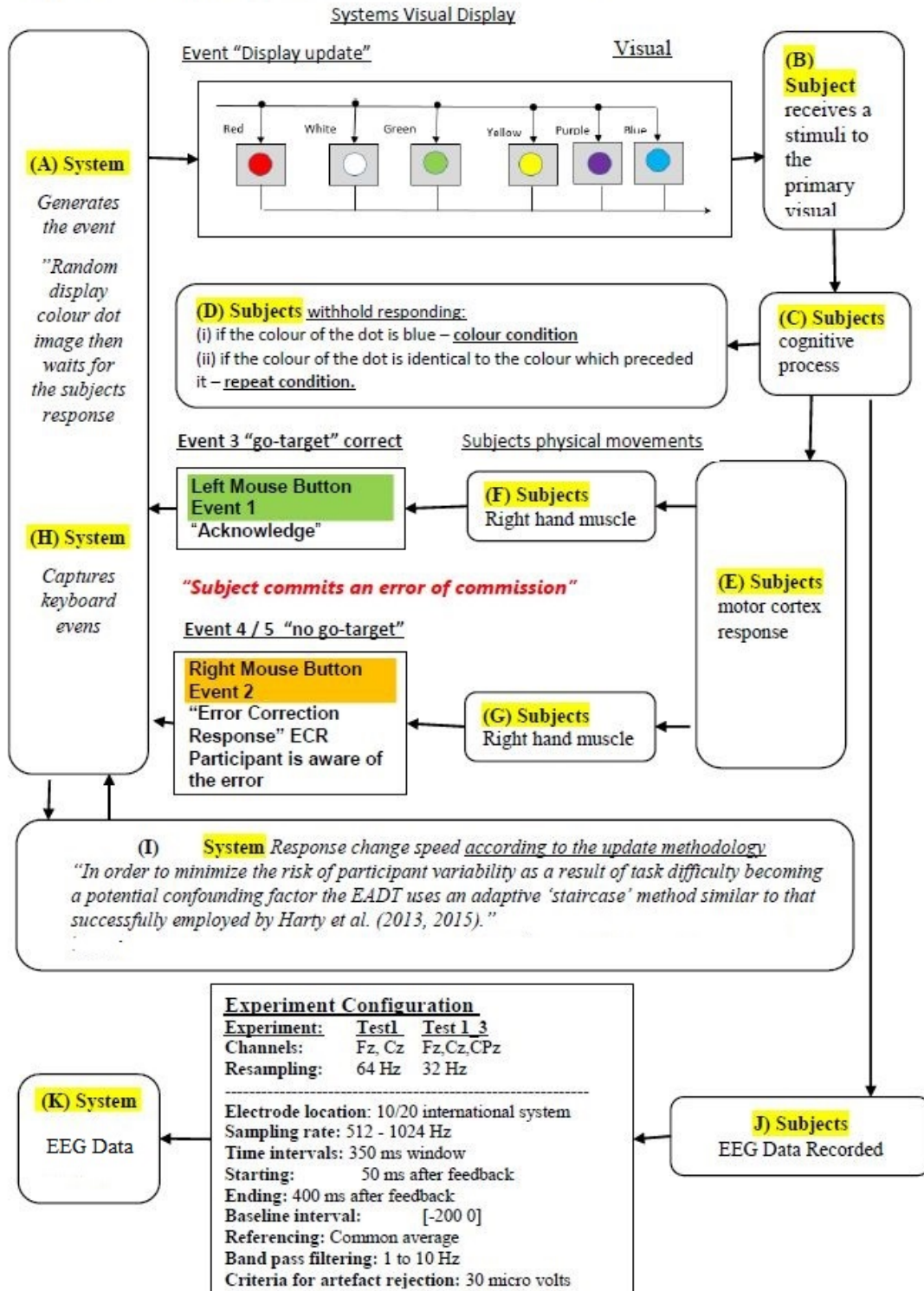


Figure 3.3: Error Awareness Task [23]

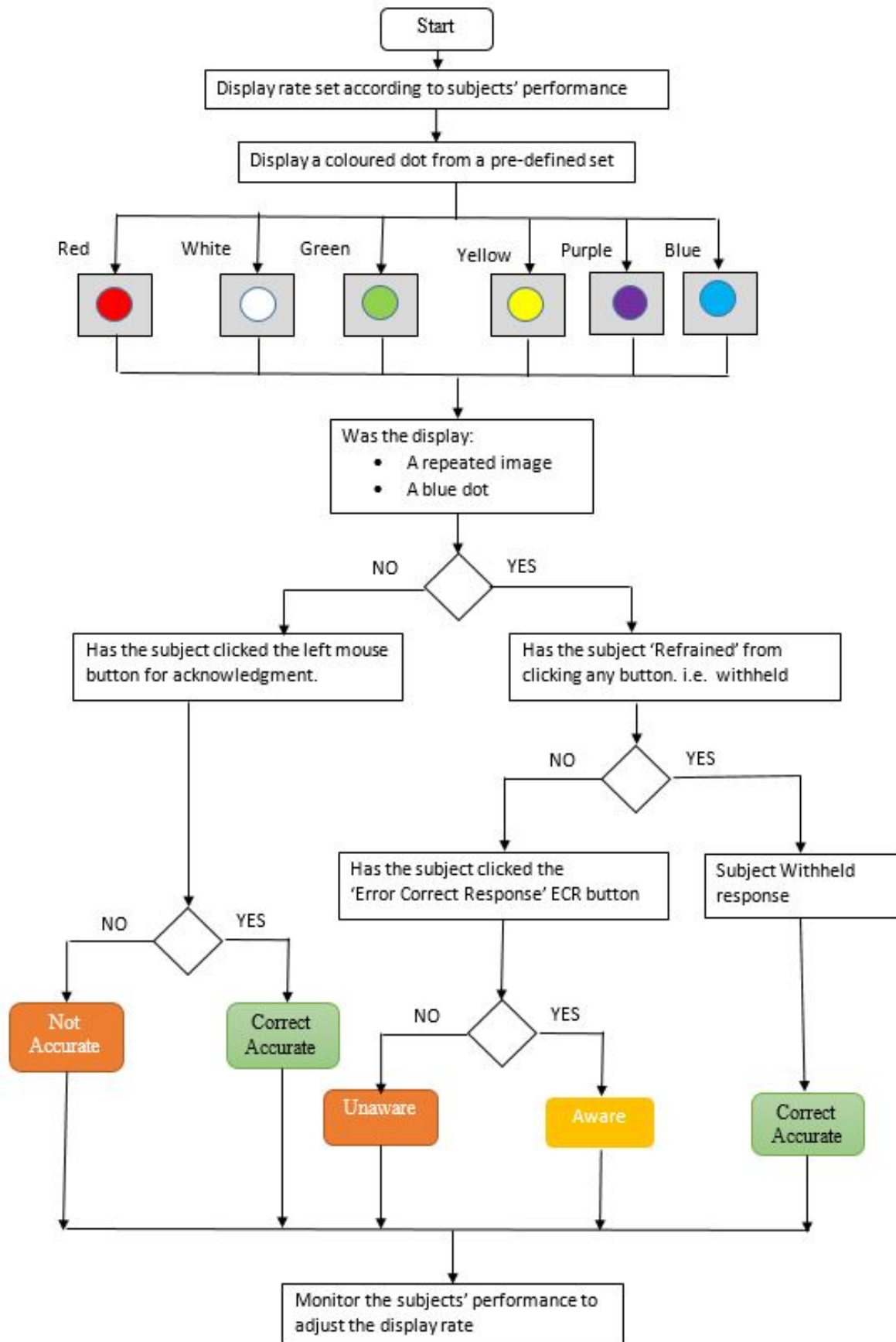
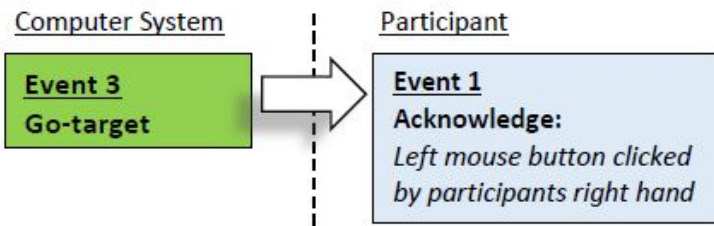


Figure 3.4: Error Awareness Task [23]

The diagram below illustrates the sequence of events during the experiment:

Sequence of events during the experiment

The participant responded correctly.



The participant makes an error in their response.

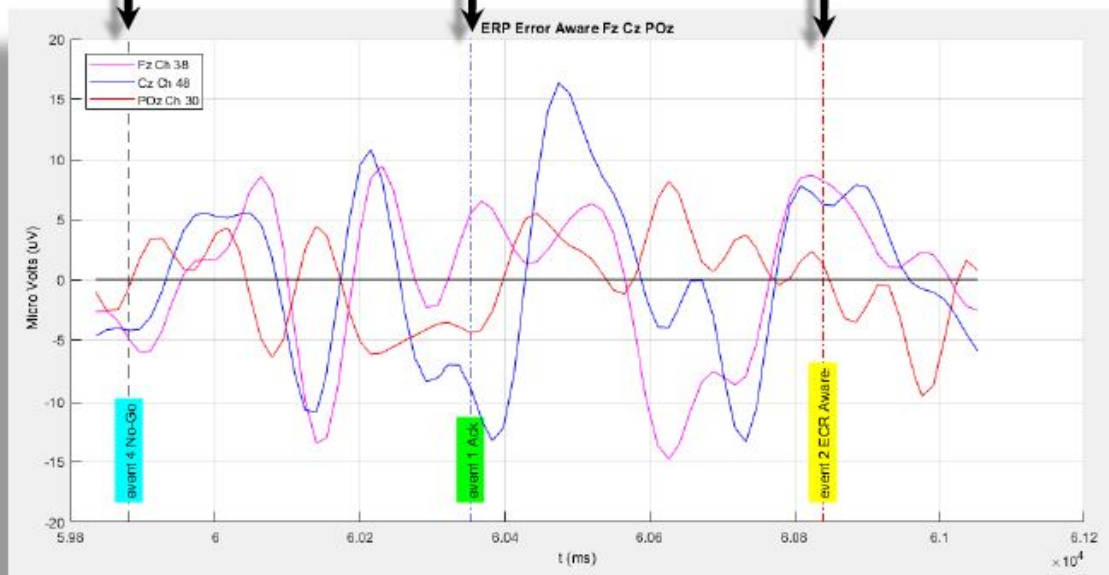
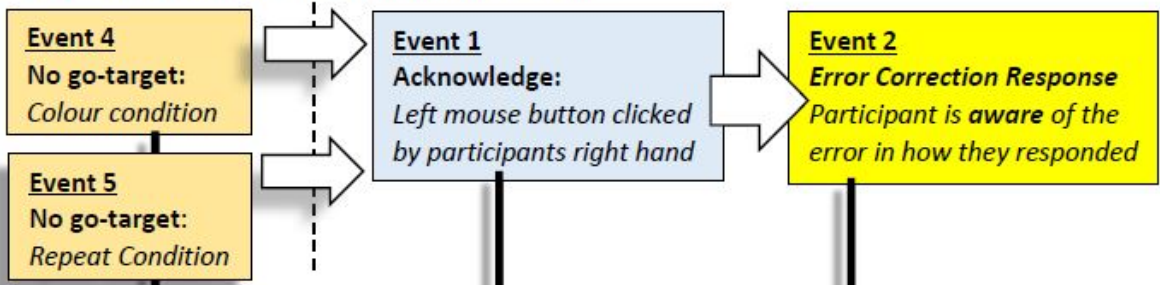


Figure 3.5: Sequence of events [23]

3.4 Results

3.4.1 Results of the statistical analysis

Experiment 1: Post-hoc comparisons using Tukey HSD test indicated the mean score for young subjects:18-34; was significantly different from elderly subjects: 65-80 years;
 $\rho = .021$

Descriptive statistics: Error Accuracy 60.8 % in comparison to the Correct Accuracy of 81.6 %

Experiment 2: Post-hoc comparisons using Tukey HSD test indicated the mean score for young subject:18-34; was significantly different from elderly subjects: 65-80 years;
 $\rho = .003$

Descriptive statistics: Error Accuracy 59.2 % in comparison to the Correct Accuracy of 82.9 %

3.4.2 Young / Old participants, Experiment 1

The repeated measure ANOVA revealed that the main effect on detection accuracies p -value is less than 0.001 $F(1,48) = 93.282$. It means that the accuracies obtained for detecting errors are significantly different from accuracies obtained for detecting correct trials

Table 3.2: Tests of Within-Subjects Contrasts Experiment 1

Source	Experiment 1	df	F	Sig.
Test_1	Err Acc. vs Corr Acc.	1	93.282	.000
Test_1 * Age Group	Err Acc. vs Corr Acc.	1	5.725	.021
Error(Test_1)	Err Acc. vs Corr Acc.	48		

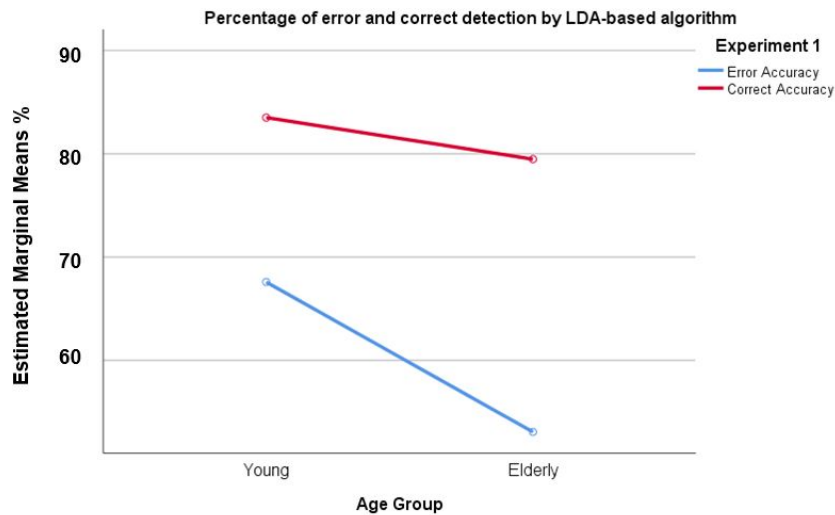


Figure 3.6: Participants age distribution

The interpretation of Tables 3.2 is that the error accuracy and the correct accuracy are significantly different regardless of their groups. The descriptive statistics shows that Error Accuracy 60.8 % in comparison to the Correct Accuracy of 81.6 % regardless of their groups are significantly different. Please see Fig 3.7

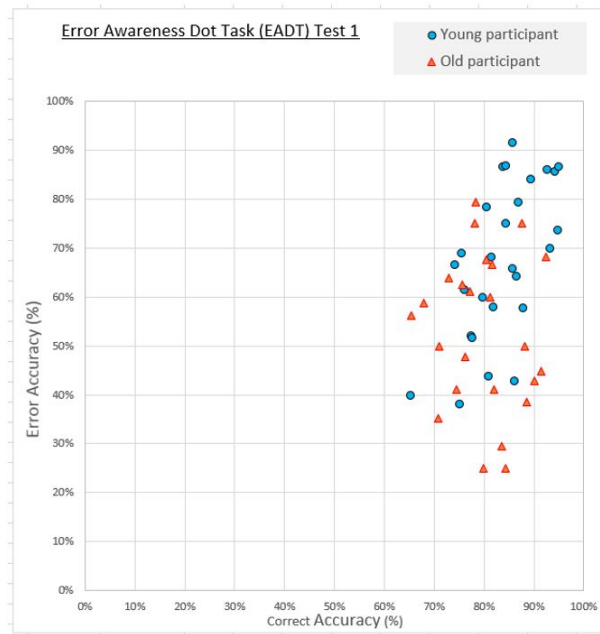


Figure 3.7: Error Awareness Dot Task(EADT) Experiment 1

3.4.3 Young / Elderly participants, Experiment 2

The repeated measure ANOVA revealed that the main effect on detection accuracies p -value is less than 0.001 $F(1,48) = 95.083$.

Table 3.3: Tests of Within-Subjects Contrasts Test Experiment 2

Source	Experiment 2	df	F	Sig.
Test_2	Err Acc. vs Corr Acc.	1	95.083	.000
Test_2 * Age Group	Err Acc. vs Corr Acc.	1	9.706	.003
Error(Test_1_3)	Err Acc. vs Corr Acc.	48		

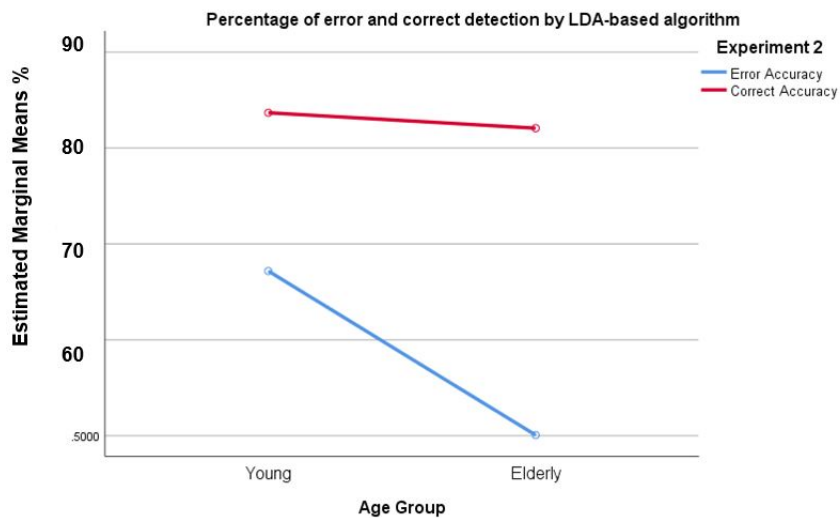


Figure 3.8: Participants age distribution

The descriptive statistics shows that Error Accuracy 59.2 % in comparison to the Correct Accuracy of 82.9 % Please see Fig 3.9

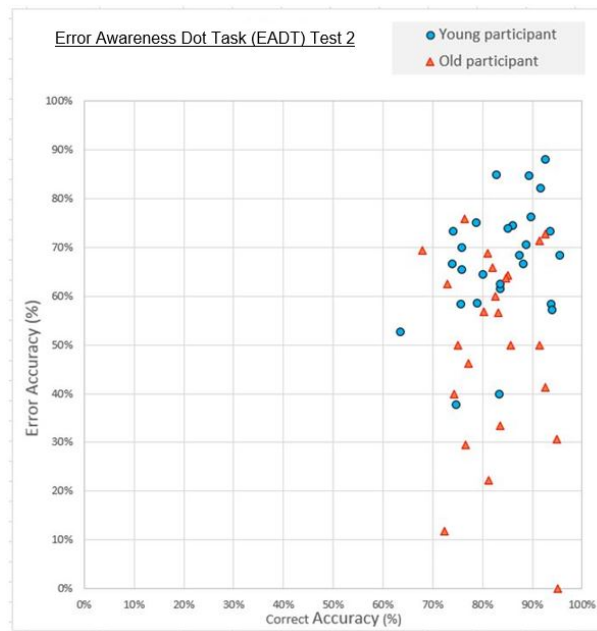


Figure 3.9: Error Awareness Dot Task(EADT) Experiment 2

3.5 Summary

This chapter shows that there is a significant difference between the young and old participants. The implications are that new BCI systems could take this finding into account and hence improve their performance by taking into account the age group of the user.

3.6 Discussion

There is a significant difference between the young and old participants in terms of error accuracy. Previous research has mainly used BCI with young adults and not with the elderly [15]. The implication is that any new BCI systems could take these findings into account and hence improve their performance by taking into account the age group of the user. Moreover, there is the potential for a future BCI system to infer the age group of the user by means of the initial interactions such that self-tuning or adaptability of the BCI settings could be a distinct advantage in terms of BCI performance. The limitation of the results is caused by using a limited number of participants.

Chapter 4

Improving EEG-based error detection using proposed novel relative peak features (RPF)

A brain-computer interface (BCI) potentially enables a person with a severe disability to communicate using brain signals. Automatic detection of error-related potentials (ErrPs) in electroencephalograph (EEG) could improve BCI performance by allowing to correct the erroneous action made by the machine. However, the current low accuracy in detecting ErrPs, particularly in some users, can reduce its potential benefits. This chapter addresses this problem by proposing a novel relative peak feature (RPF) selection method to improve performance and accuracy for recognising an ErrP in the EEG. Using data collected from 29 participants with a mean age of 24.14 years the relative peak features yielded an average across all classifiers of 81.63% accuracy in detecting the erroneous events and an average 78.87 % accuracy in detecting the correct events, using KNN, SVM and LDA classifiers. In comparison to the temporal feature selection, there was a gain in performance in all classifiers of 17.85% for error accuracy and a reduction of -6.16% for correct accuracy. Specifically, our proposed RPF significantly reduced the number of features by 91.7% when compared with the state of the art temporal features. In the future, this work will improve the human-robot interactions by improving the accuracy of detecting errors that enable the BCI to correct any mistakes.

4.1 Introduction

The brain will evoke an ErrP in the EEG waveform recordings, which we could associate with the person's awareness of an 'error', caused by themselves or by observation.

The ErrP has been characterised as a negative potential deflection called an error-related negative (ERN) in the frontal-central location electrodes within 100ms [22] of the participant's erroneous response. After the ERN in the frontal-central location electrodes, there may follow an error-related positive event (Pe). The participants' awareness of an error is associated with this positive deflection (Pe) located in the centro-parietal region of the scalp. The authors in [80, 81, 82] highlight a correlation between ERN events and behavioural adjustments. The authors further suggested that the ERN signal is associated with participant monitoring activity and hence reacting to an erroneous situation, this assertion supported by [83]. The authors suggest that the context of the task moreover what it means to the participants in terms of the importance has the effect of modulating the amplitude of the ERN ([81, 84]).

EEG data recordings have shown a rapid transient during an error-related potential. Generally, the Central ('Cz') Electrode detects a high transient dynamic in comparison to the other EEG electrodes. Error detection algorithms [78, 79] have shown some of the first EEG event-related potentials (ERP) during an experiment designed to evoke an (ERP) in a speed response task.

Previous research studies showed that ErrPs could be detectable in EEG on a single-trial basis [9],[91] The ability to detect these ErrPs would enhance a future BCI [9]. Many research studies have investigated BCI with assistive robotics presented in the following survey [7]. [9] has shown the presence of ErrP and investigated a simple human-robot interaction which satisfactorily recognised single trials using a Gaussian classifier.

This research aims to improve single-trial ErrP detection by proposing a new set of features, called relative peak features (RPF). The principle behind relative peak features is the selection of the essential characteristics of an ErrP event, specifically the latency between ERN and Pe and also the amplitude between ERN and Pe. We believe these sorts of important information are missing in the state of the art temporal features where only amplitude of the EEG samples are used as features. Moreover, compared to the state of the art temporal features, our proposed RPF considerably reduce the features to only two per channel. Using EEG data collected from 29 participants performing a decision making task, we evaluated the performance of the proposed features and compared with the temporal features across a range of different classifiers.

4.2 Methodology

4.2.1 Participants and the task

The 29 young participants (19 female) had a mean age of 24.14 years. All participants were healthy individuals. The study has been approved by the Trinity College Dublin ethics committee by following the Declaration of Helsinki [109]. All participants provided informed written consent to attend this study.

EEG data has been collected in Trinity College Institute of Neuroscience, Ireland [23] where participants performed a novel variant of Hester et al.' Error awareness task [110]. A 64 channel EEG Biosemi device was used for data collection. The data generated from the awareness of 'sustained attention' tasks. Generally, the cause of the 'Error related potential' ErrP is that the subject is aware of their mistake during the sustained attention task. The configuration of the experiment is such that the subjects will evoke ErrP events during the experiment.

Thus a performance comparison in terms of accuracy can be established between the features selection methods and the well-established KNN, SVM, LDA classifiers.

4.2.2 Preprocessing

A comparison between temporal feature selection and the proposed relative peak feature selection methods were investigated. In order to achieve this, we propose using well-established classifiers in Matlab®.

The preprocessing configuration details presented in, Table 4.1 indicating the setting for EEG segmenting, downsampling and filtering.

Table 4.1: Experiment preprocessing configuration

Parameter	Settings
Channel	Cz
Re-sampling:	64 Hz
Electrode location:	10/20 international system
Sampling rate:	512 - 1024 Hz
Time intervals:	350 ms window
Starting:	50 ms after onset response
Ending:	400 ms after onset response
Baseline interval:	-200 to 0
Referencing:	Common average
Band pass filtering:	1 to 10 Hz
Criteria for artefact rejection:	30 micro volts

4.2.3 Feature extraction and classifications

Specific features are acquired from the EEG data and labelled with specific events and sent to the classifiers. During the session with each subject; generally, eight blocks of data were captured. Moreover blocks 1 to 4 were sent to the classifier for training and the other blocks of data from 5 to 8 were sent later to the classifiers for testing.

The classifier's performance assessed in terms of 'Error Accuracy' (EA) and 'Correct Accuracy' (CA).

4.3 Relative peak features (RPF)

The objective of the relative peak feature is to improve the performance of a classifier. The advantages of the proposed RPF are neurophysiological characteristics of the signals that are captured. Information data loss can occur depending on exactly when the data is sampled from the EEG data recording. The time usually is locked; however, to improve the capture of the Cz signal the capture windows is dynamically adjusted to synchronise to the minimum value of Cz. Hence this will capture the dynamics of the Cz signal as Fig 5.9 shows.

Observing the error-related potentials appear to have a high amplitude between the (ERN) graph trough and (Pe) graph peaks. It was particularly noticeable in the 'Cz' EEG channel recordings.

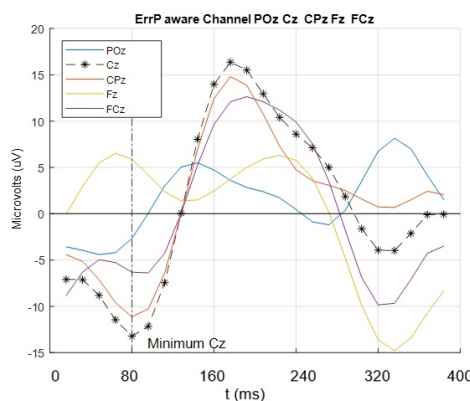


Figure 4.1: Synchronising the ErrP to Cz minimum

C_z is the EEG channel

$$X = x(t_0), x(t_1), x(t_2), \dots, x(t_n) \quad (4.1)$$

where
 t is the sample time as Fig 5.9 shows

$$x(t_{min}) = \min_{0 \leq i \leq n} x(t_i) \quad (4.2)$$

$$x(t_{max}) = \max_{0 \leq i \leq n} x(t_i) \quad (4.3)$$

$$\vec{a} = (t_{min}, x(t_{min})) \quad (4.4)$$

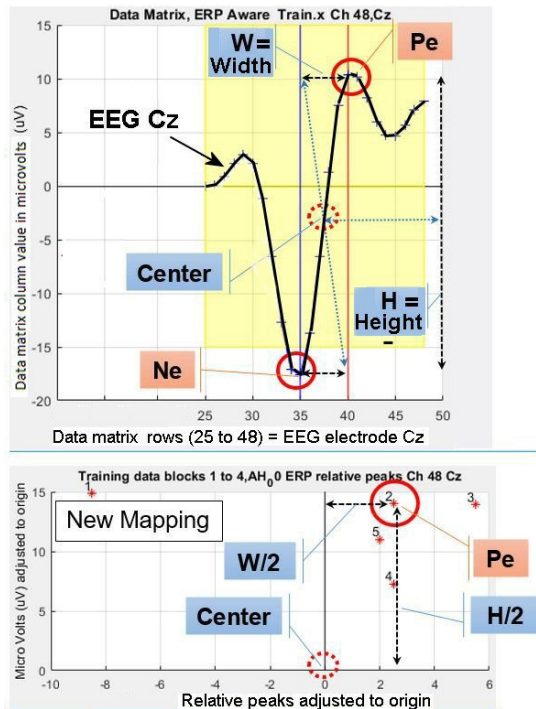
$$\vec{b} = (t_{max}, x(t_{max})) \quad (4.5)$$

$$\vec{v} = \frac{1}{2}(\vec{b} - \vec{a}) \quad (4.6)$$

$$V_{Train} = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\} \quad (4.7)$$

Please see the following diagrams: Fig 5.9 shows EEG electrode signals and Fig 4.2 shows the signal from the Cz electrode from the EEG in a data matrix. The vertical axis shows the amplitude in microvolts; the horizontal axis shows the EEG samples 25 to 48. We are interested in the amplitude height and the latency width between Pe and Ne to plot the relative peak the height and width are divided by 2 in order to generate the new mapping. The relative peak training data is presented in Fig 4.3 which shows the relative peak training data plot of relative amplitudes and latencies which significantly reduced the number of features by 91.7% when compared with using temporal features using all the 24 data points in equation 4.1 compared to only 2 data points used by RPF.

Locate the largest transient in the electrode channel Cz



New graph peak (Pe) mapping around a new graph origin

Figure 4.2: Extracting propose RPF from EEG waveform Cz.

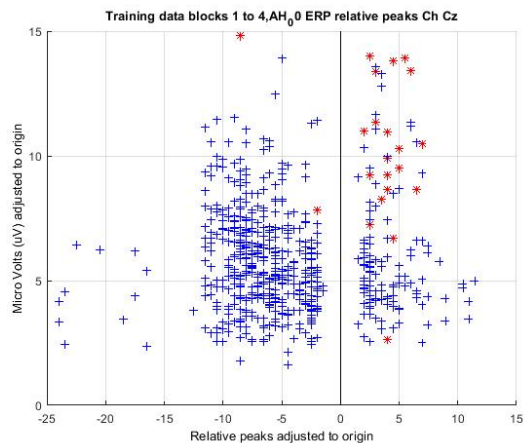


Figure 4.3: Distribution of RPF, Cz electrode ErrP red '*', blue '+' non-ErrP

4.4 Results

4.4.1 Classifier comparison results using temporal features

The table of classifiers in fig 4.5 shows using temporal features exhibited a relatively good performance from the SVM Linear and Linear discriminant analysis classifiers compared to the other classifiers. The temporal feature performance is further presented on the left side of the bar chart Fig 4.6 to visually clarify the comparison between the different classifiers using the temporal feature method.

4.4.2 Classifier comparison results using relative peak features

The fig 4.5 shows the relative peak feature performance presented on the right side Fig 4.6 visually clarifies the comparison between the different classifiers using the relative peak features, in addition to presenting the comparison between the different feature extraction methods. The relative peak features yielded an average across all classifiers of 81.63% accuracy in detecting the erroneous events and an average 78.87 % accuracy in detecting the correct events. In comparison to the relative peak feature selection method, there was an average increase in performance in all classifiers of 17.85% for error accuracy and a reduction of -6.16% for correct accuracy.

4.5 Summary

This chapter shows a method of feature selection that can improve the classification performance of well-established classifiers. It is beneficial to improve the accuracy of detecting errors for a BCI system. In the future, this work will improve the human-robot interaction by improving the accuracy of detecting errors, for example, when the prosthetic hand operates contrary to the users' intentions as shown in Fig 4.4



Figure 4.4: Assistive robotic prosthetic hand stacking blocks.

CLASSIFIER COMPARISON RESULTS USING TEMPORAL FEATURES AND RELATIVE PEAK FEATURES.

	Classifiers	Classifiers KNN LDA SVM using temporal features		Classifiers KNN LDA SVM using relative peak features	
		<i>Error Accuracy</i> [%]	<i>Correct Accuracy</i> [%]	<i>Error Accuracy</i> [%]	<i>Correct Accuracy</i> [%]
1	KNN Coarse	64.98	88.16	80.85	78.83
2	LDA	69.06	79.97	82.68	78.10
3	SVM Coarse Gaussian	50.33	92.68	80.85	80.71
4	SVM Linear	70.75	79.28	82.16	77.86
5	Average results among all classifiers	63.78	85.02	81.63	78.87

Figure 4.5: Classifier comparison results using temporal features and relative peak features

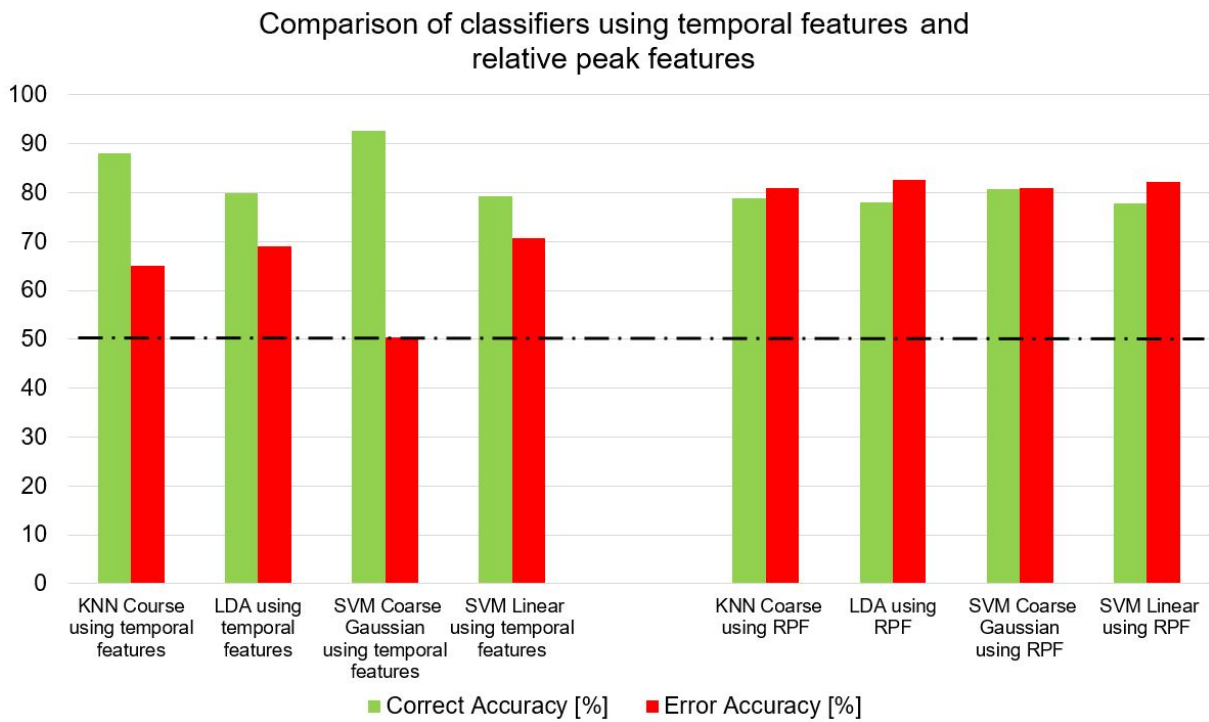


Figure 4.6: Classifier comparison results

4.6 Results of classifier comparisons using a novel feature extraction method

4.7 Introduction

4.7.1 Statistical Analysis Comparisons

In this research, a series of statistical tests were applied to the data from a selection of well-established classifiers. Firstly, a grand average comparison between a set of classifiers using temporal features and a set of classifiers using relative peak features. Secondly, the grand average comparison was further compared between different age groups young and elderly. Thirdly, a statistical comparison was completed to determine the statistical significance between the different extraction methods temporal features extraction and relative peak feature extraction. Finally, a summary of the completed statistical analysis will conclude this chapter.

4.7.2 Elderly participants classifier comparison results using temporal features

The table of classifiers 4.7 using temporal features and elderly participants showed an even performance for error accuracy 67.21% and correct accuracy 66.63% with the subspace discriminant classifier. The classifier comparison is listed in order from lowest to highest performance in terms of error accuracy.

	Classifiers	Error Accuracy [%]	Correct Accuracy [%]
1	Linear Discriminant (LDA)	50.61%	77.09%
2	Logistic Regression	55.15%	76.34%
3	KNN Coarse	56.57%	80.56%
4	SVM Linear	57.39%	74.98%
5	Subspace discriminant	67.21%	66.63%

Figure 4.7: Elderly participants classifier comparison results using temporal features

The table of classifiers 4.7 using temporal features were further presented in the following horizontal bar chart 4.8 to visually clarify the comparison between the different classifiers using the temporal feature extraction method with elderly participants.

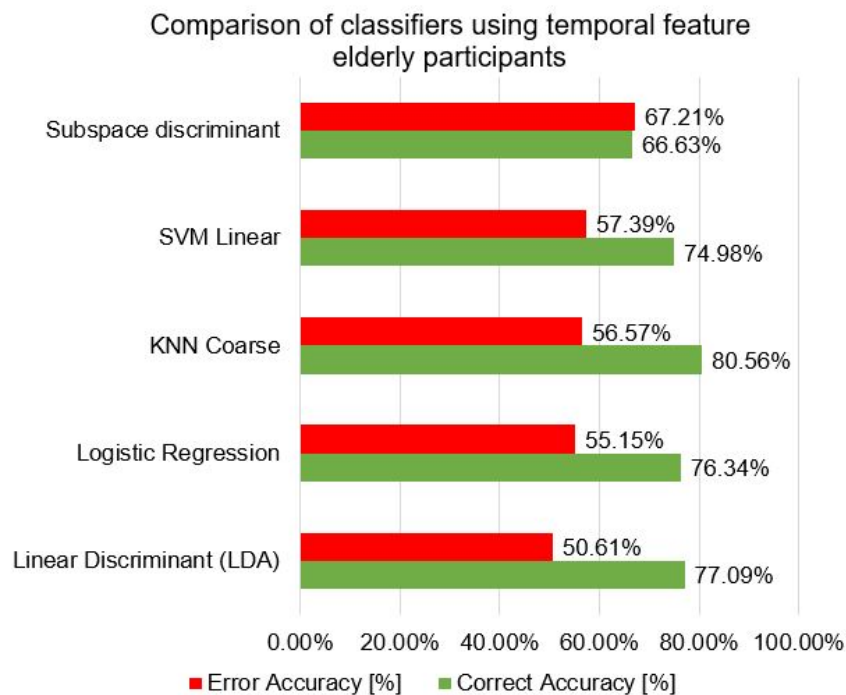


Figure 4.8: Elderly participants classifier comparison results using temporal features

4.7.3 Elderly participants classifier comparison results using relative peak features

The table of classifiers 4.9 using relative peak features and elderly participants showed a higher performance generally for both error accuracy and correct accuracy for a wider range of classifiers compared to table 4.7. The classifier comparison is listed in order from lowest to highest performance in terms of error accuracy.

	Classifiers	Error Accuracy [%]	Correct Accuracy [%]
1	ADS Prototype Classifier	63.63%	82.63%
2	SVM Medium Gaussian	64.99%	77.07%
3	Decision Tree Simple	69.10%	75.78%
4	Quadratic Discriminant(QDA)	75.29%	75.87%
5	Logistic Regression	76.47%	74.91%
6	KNN Coarse	77.03%	70.58%
7	Linear Discriminant (LDA)	77.36%	74.59%
8	Subspace discriminant	77.41%	75.02%
9	SVM Linear	78.83%	72.68%
10	SVM Coarse Gaussian	79.11%	73.63%

Figure 4.9: Elderly participants classifier comparison results using relative peak features

The table of classifiers 4.9 using temporal features with elderly participants were further presented in the following horizontal bar chart 4.10

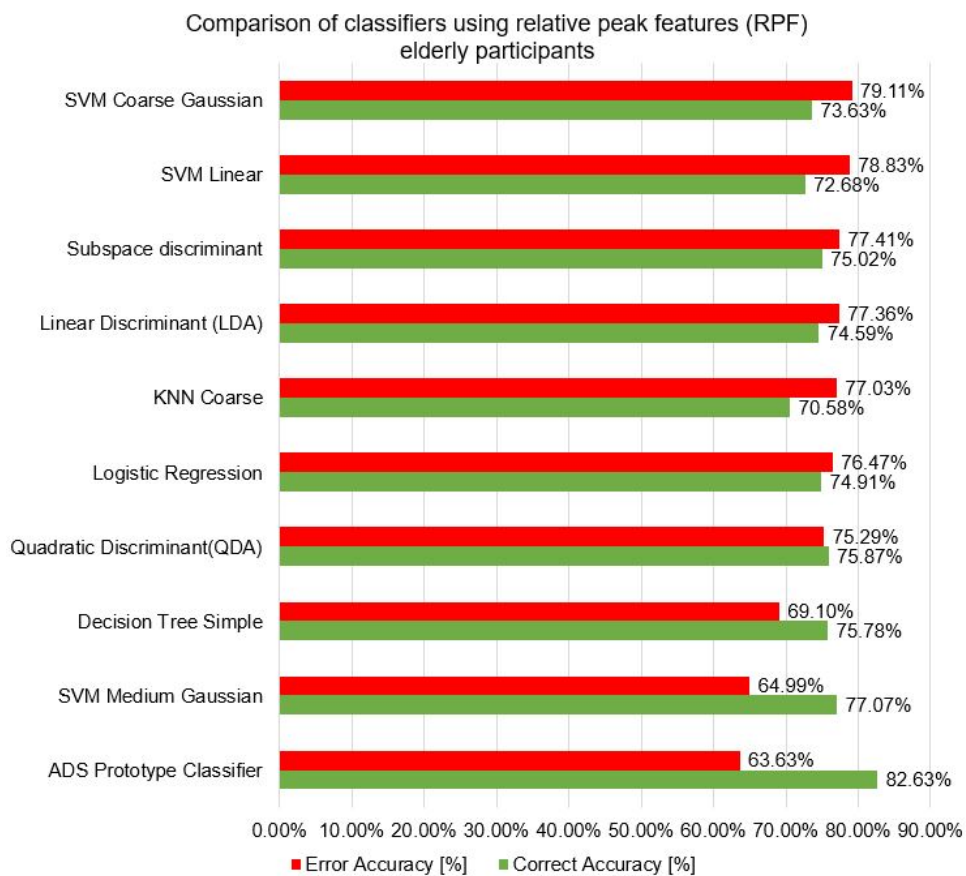


Figure 4.10: Elderly participants classifier comparison results using relative peak features

4.7.4 Young participants classifier comparison results using temporal features

The table of classifiers 4.11 using temporal features and young participants showed a relatively good performance from the subspace discriminant, SVM Linear and Linear discriminant (LDA) classifiers compared to the other classifiers. The classifier comparison is listed in order from lowest to highest performance in terms of error accuracy.

	Classifiers	Error Accuracy [%]	Correct Accuracy [%]
1	SVM Coarse Gaussian	50.33	92.68
2	Decision Tree Simple	60.80	77.71
3	KNN Coarse	64.98	88.16
4	Logistic Regression	68.44	80.34
5	Linear Discriminant (LDA)	69.06	79.97
6	SVM Linear	70.75	79.28
7	Subspace discriminant	77.73	70.57

Figure 4.11: Young participants classifier comparison results using temporal features

The table of classifiers 4.11 using temporal features is further presented in the following horizontal bar chart 4.12 to visually clarify the comparison between the different classifiers using the temporal feature extraction method with young participants.

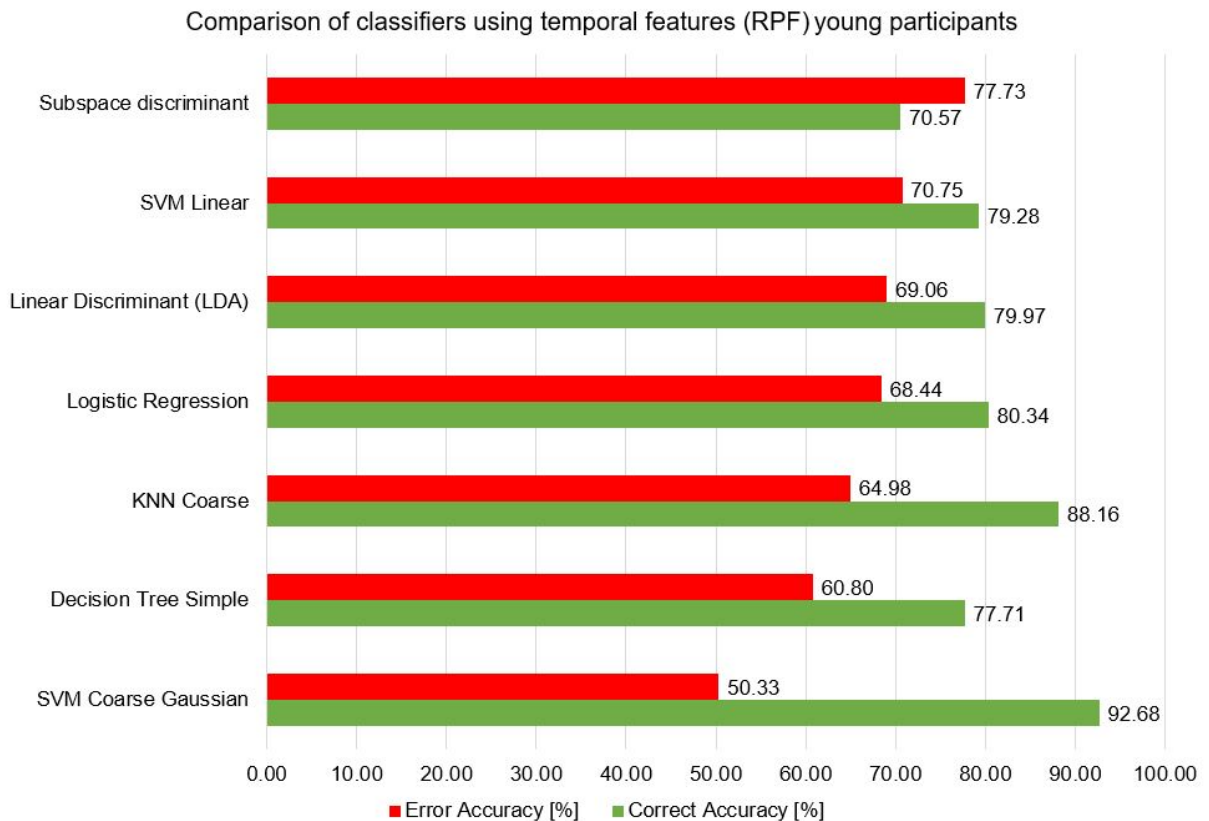


Figure 4.12: Young participants classifier comparison results using temporal features

4.7.5 Young participants classifier comparison results using relative peak features

The table of classifiers 4.11 using temporal features and young participants showed all classifier generally above 80% for error accuracy and generally above 78% for correct accuracy noticeably there was a good performance from the proposed novel ADS Prototype classifier compared to the other classifiers. The classifier comparison is listed in order from lowest to highest performance in terms of error accuracy. Please see the classifier performance bar chart Fig 4.14 for a visual comparison of the classifier performance.

	Classifiers	Error Accuracy [%]	Correct Accuracy [%]
1	Decision Tree Simple	74.28	82.38
2	SVM Medium Gaussian	75.37	81.86
3	Quadratic Discriminant(QDA)	80.33	80.78
4	KNN Coarse	80.85	78.83
5	SVM Coarse Gaussian	80.85	80.71
6	Logistic Regression	81.45	78.81
7	ADS Prototype Classifier	81.85	83.88
8	SVM Linear	82.16	77.86
9	Linear Discriminant (LDA)	82.68	78.10
10	Subspace discriminant	83.38	78.44

Figure 4.13: Young participants classifier comparison results using relative peak features

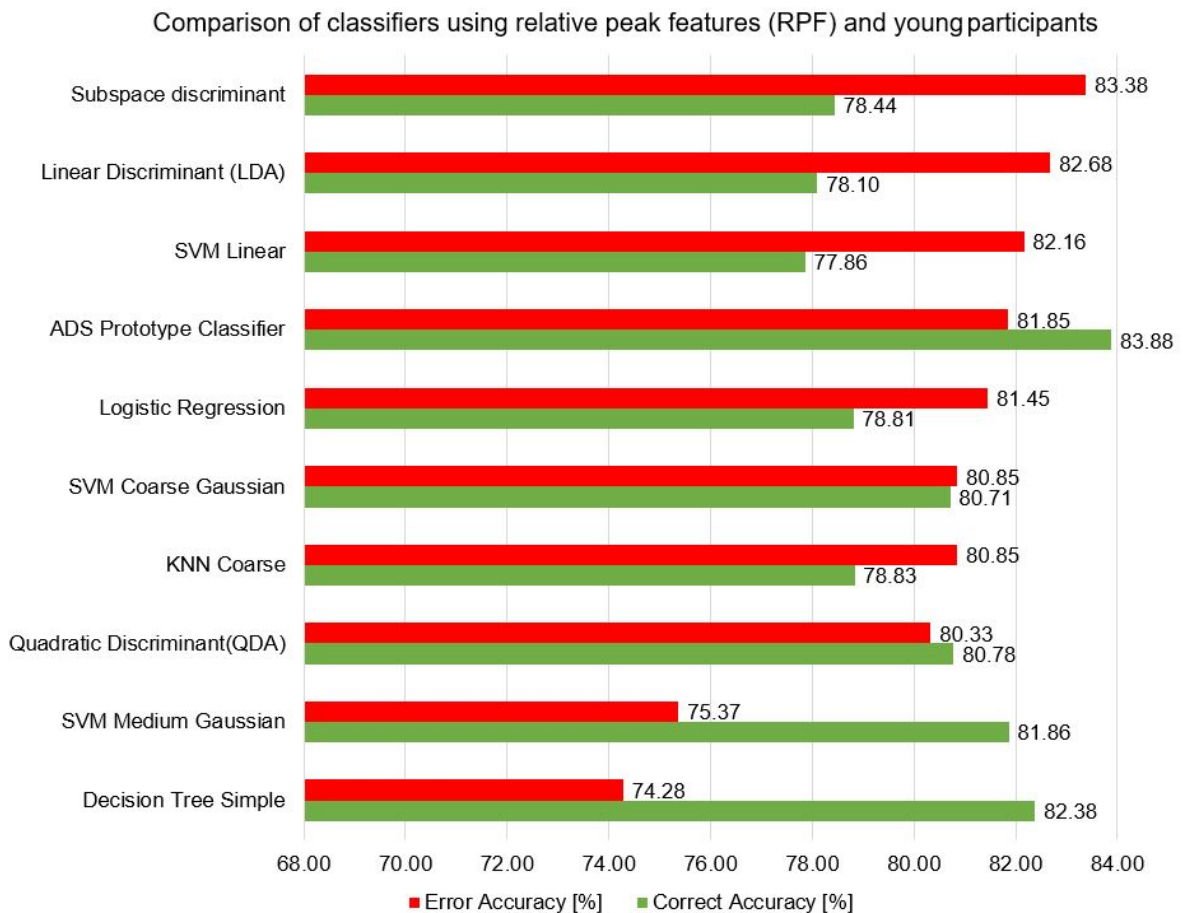


Figure 4.14: Young participants classifier comparison results using relative peak features

4.8 Results of classifier comparisons using a novel feature extraction method

The results of classifier comparisons using a novel feature extraction method are presented in Appendix F .

4.9 Conclusion

The results show that the relative peak feature selection provides a higher performance with a wider range of established classification methods including the adaptive decision surface ADS prototype classifier. The paired T-Test independent samples tests show that using the relative peak feature selection removes the statistically significant difference between young and elderly compared to temporal features method which shows a significant difference in the error accuracy between young and elderly groups. The summary of the T-Test further shows that the following classifiers: **SVM, LDA, Logistic regression** and **subspace discriminant** using a relative peak feature extraction remove the statistically significant difference between young and elderly however the **KNN coarse classifier** using the relative peak feature extraction does not remove the significant difference for correct accuracy. The results show a distinct advantage in using the relative peak feature extraction method please see table 4.15 Verification and support guidance for using the SPSS software, in particular, ANOVA the paired T-test analysis [1] advice was provided by Dr. Basile Marquier and Ellen Marshall, Mathematics and statistic tutor at Maths and Statistics Help (MASH).

Participants	Compare significances of different classifiers and different features extraction methods using independent T-Tests		SVM Linear classifier	LDA classifier	Logistic Regression	KNN coarse classifier	Subspace discriminant classifier	Bar chart key
			ρ	ρ	ρ	ρ	ρ	
Young Elderly	EA TF Error Accuracy Temporal Features	Equal variance not assumed	0.011	0.001	0.009	0.0954	0.014	Series 1
	CA TF Correct Accuracy Temporal Features	Equal variance assumed	0.143	0.254	0.126	0.017	0.377	Series 2
Young Elderly	EA RPF Error Accuracy Relative Peak Features	Equal variance not assumed	0.366	0.158	0.194	0.332	0.101	Series 3
	CA RPF Correct Accuracy Relative Peak Features	Equal variance assumed	0.08	0.146	0.1	0.012	0.152	Series 4

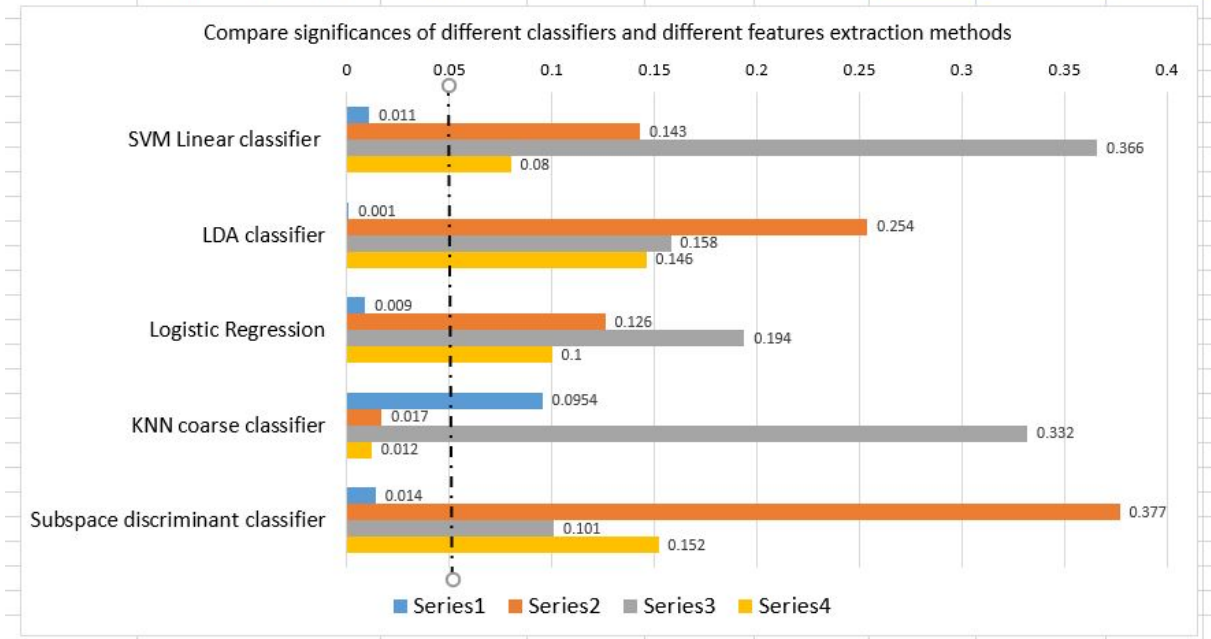


Figure 4.15: Compare significances of different classifiers and different features extraction methods using independent T-Test

4.10 Results of the performance gains summary

The results of the performance gains summary are presented in Appendix G .

4.11 Summary

The novel **relative peak feature** (RPF) extraction method improved the error detection performance and removed the statistical difference between the age groups young and elderly for a number of established classifiers. In addition, the relative peak feature is part of a novel error detection algorithm **adaptive decision surface** (ADS). This improvement in performance was proven with repeated ANOVA and T-Tests for using the following classifiers **SVM Linear**, **LDA**, **Logistic Regression** and **Subspace discriminant**. Improving the error detection and classification has implication for assistive supportive care robotics and human-robot interactions in the future. **LDA** classifier, compared to the state of the art temporal features the proposed relative peak features increase the accuracy of detection of an error by an average of 26.75% and 13.62% in young and elderly respectively and a -1.87% and -2.5% decrease in performance for young and elderly for correct accuracy. **Logistic Regression** classifier, compared to the state of the art temporal features the proposed relative peak features increase the accuracy of detection of error by an average of 26.75% and 13.62% in young and elderly respectively and a

-1.53% and -1.43% decrease in performance for young and elderly for correct accuracy. **SVM Linear** classifier compared to the state of the art temporal features the proposed relative peak features increase the accuracy of detection of an error by an average of 11.41% and 21.44% in young and elderly respectively and a -1.42% and -2.30% decrease in performance for young and elderly for correct accuracy. **Subspace discriminant** classifier compared to the state of the art temporal features the proposed relative peak features increase the accuracy of detection of an error by an average of 5.65% and 10.20% in young and elderly respectively. In addition, there was an increase in the accuracy of detection of correct by an average of 7.87% and 8.39% in young and elderly respectively for correct accuracy. The average performance of the novel adaptive decision surface ADS was 81.85% and 83.88 % for the error accuracy and correct accuracy for the young and 63 % and 82% for the error accuracy and correct accuracy for the elderly.

Chapter 5

Proposed novel adaptive decision surface (ADS) classifier

5.1 Introduction

The aims and objectives of the proposed ADS classifier will be to enhance the performance and efficiency of detection error-related potentials (ErrP) in the human brain. An ErrP may occur in the brain when the person experiences an unanticipated event. The ErrP has a negative potential deflection known as an error-related negative (ERN) within 100ms [22] of the unanticipated event usually followed an error-related positive event (Pe). The location of the ErrP is in the frontal-central position.

This contribution has improved the performance of detection ErrP events in addition to reducing the data required by 91.7% compared to the linear discriminant classifier LDA furthermore, this efficiency is advantageous to an embedded system such as the brain-computer interface BCI which may be wearable and operating with less power consumption.

EEG will capture a significant amount of data please see Fig 5.15, which can be analysed in the following domains: EEG Time-Frequency

- EEG Temporal Analysis.
- EEG Frequency Analysis.

What is the most reliable, fast and accurate methodology for recognising a particular brain-wave feature and accurately classifying this feature for the brain-computer interface to evaluate real time?

Electroencephalograph EEG data recordings have shown a rapid transient during an error-related potential. Specific electrodes, such as the 'Cz' located on top of the participants head. Generally, the 'Cz', 'CPz', 'FCz' electrodes have detected a high transient dynamic in comparison to the other EEG electrodes. Please see Fig 5.1.

Moreover, by using observations in the frequency domain and applying 'Morlet Wavelets' to the EEG data has highlighted these transients during an ErrP occurrence. Please see Fig 5.2 the graph shows an ErrP event characterised by the trough and peak of the Fz and Cz electrode EEG signals. The 'Morelet Wavelet' and the Power Spectral Density(PSD) Welch Hamming was evaluated.

However, the initial observation would require further investigation of the data across all the participants to support the evidence that there is a difference in ErrP response which is influenced by:

- The type of error stimuli presented to the participant affect the amplitude
- The literature suggests the effect of modulating the amplitude of the ERN ([81, 84])

Moreover, it could be the person's emotional perception of the stimuli which influence the error-related potential response.

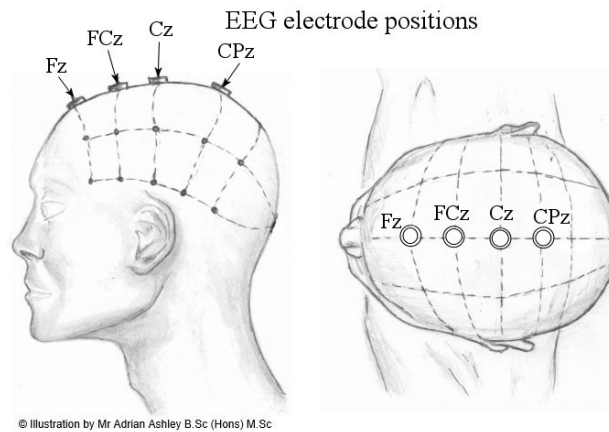


Figure 5.1: EEG Electrode positions Fz FCz Cz CPz

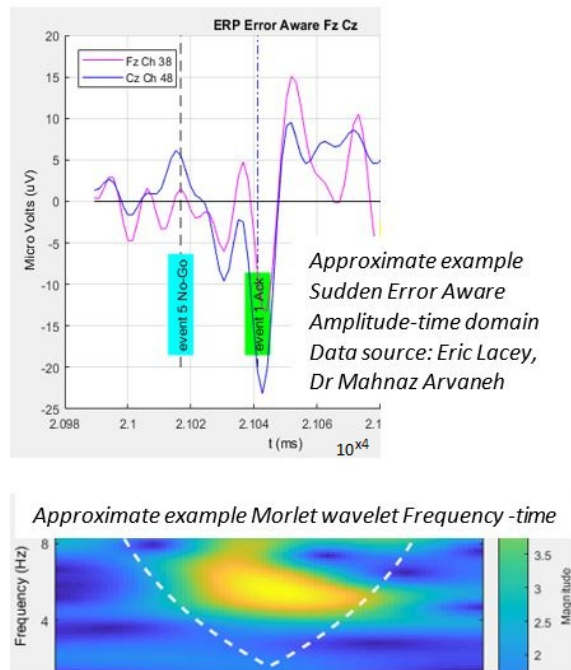


Figure 5.2: EEG waveforms Fz Cz Morlet Wavelet

Many research studies have investigated BCI with assistive robotics presented in the following survey [7]. [9] have shown the presence of ErrP

One of the challenges in feature extraction and classification can be identified as information data loss which can occur depending on exactly when the data is sampled from the EEG data recording. The time is normally locked there is an emergence of a characteristic waveform developing being clipped and hence the error-related potential waveform could be missed or clipped resulting in information loss. Please see Fig 5.3

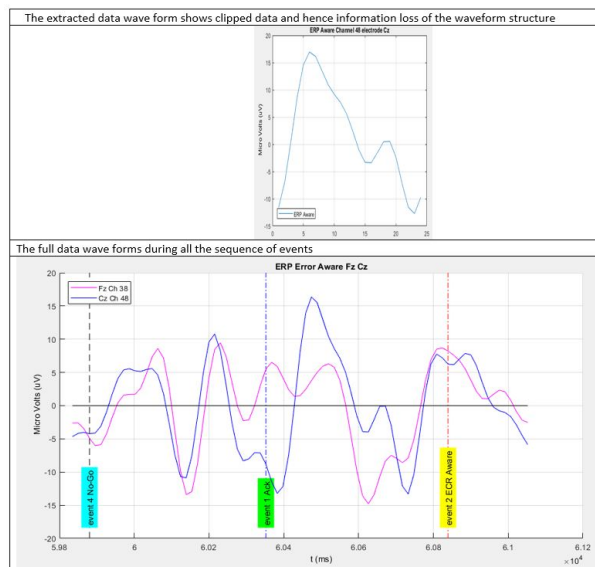


Figure 5.3: Error Awareness Dot Task(EADT) ERP waveform information loss

5.2 Methodology

The experiment procedural steps carried out during this experiment follow the previous investigations and are presented in the diagram Fig 5.4

The procedural processes are as follows:

Sampling the EEG brain waveforms from the selected EEG channels. Please see 3.2.1

Pre-processing re-sampling, Band-pass zero phase filter 1 to 10 Hz. Please see 3.2.2

EEG ERP extraction baseline correction, Artefact rejection.

Feature extraction using the relative peak features as shown in chapter 4

5.2.1 Electroencephalograph EEG feature classification process

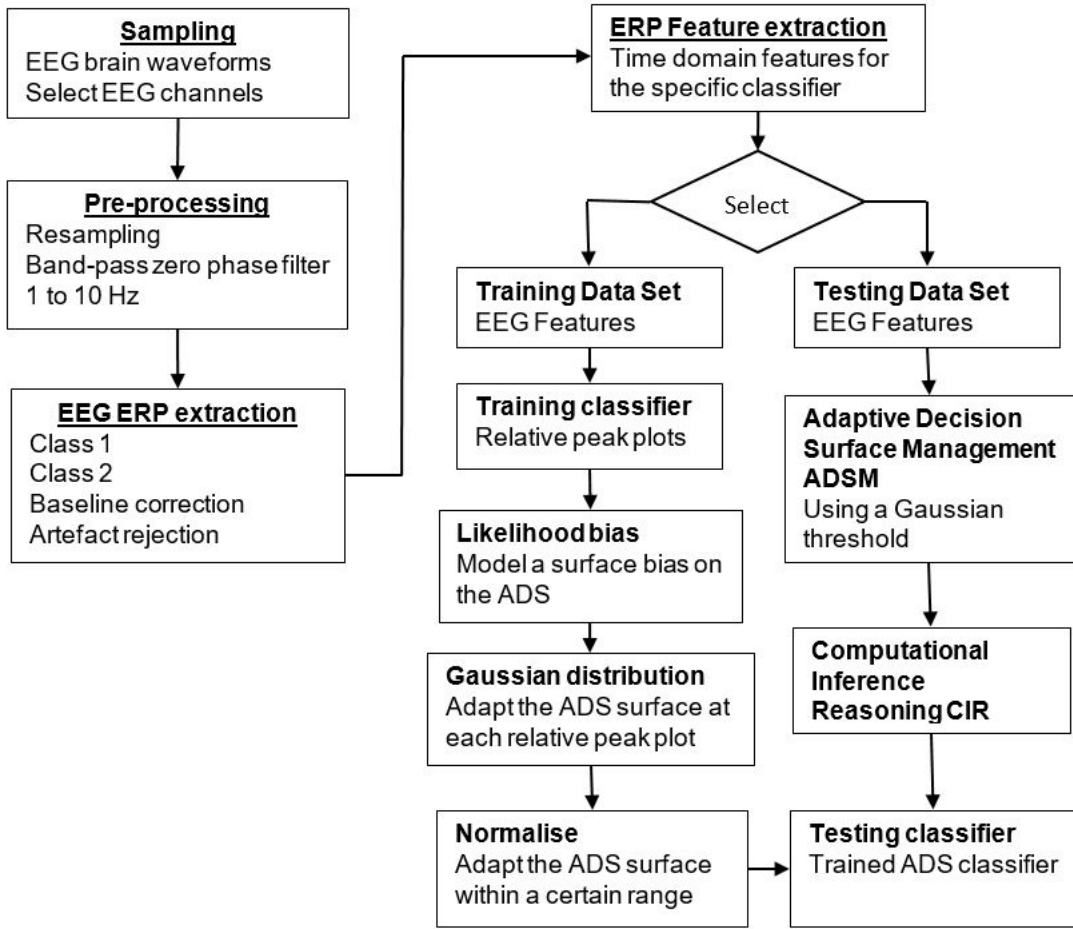


Figure 5.4: A schematic diagram of the envisaged adaptive decision surface ADS classifier data flow

5.2.2 Adaptive decision surface (ADS) mathematical definition

Using the relative peak features shown in chapter 4 whereby the *Training set of vertices*

$$V_{Train} = \{E_a, C_a\} \quad (5.1)$$

$$V_{Train} = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\} \quad (5.2)$$

The adaptive decision surface ADS represented by a 3 dimensional space

$$S \subseteq \mathbb{R}^3[1] \quad (5.3)$$

where,

$$S = [x_{min}, x_{max}] \times [y_{min}, y_{max}] \times [z_{min}, z_{max}] \quad (5.4)$$

$$V \in S \quad (5.5)$$

A Bivariate Gaussian distribution data structure is projected on to S around each training vector $\vec{v} \in V_{Train}$.

Classification 1 Error Accuracy EA

$$\vec{v} \in E_a \quad (5.6)$$

$$\vec{v} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \quad (5.7)$$

where,
 $f_1 = \mu_X$, feature 1
 $f_2 = \mu_Y$, feature 2

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} \right] \right) \quad (5.8)$$

where

μ is the mean in each dimension, $\mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}$

σ_x^2 is the variance

σ_y^2 is the variance

ρ is the correlation between X and Y such that $\sigma_X > 0$ and $\sigma_Y > 0$

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\rho\sigma_Y \\ \rho\sigma_X\rho\sigma_Y & \sigma_Y^2 \end{pmatrix} \quad (5.9)$$

Please see the approximate representation Figure 5.5

Classification 2 Correct Accuracy CA

$$\vec{v} \in C_a \quad (5.10)$$

$$\vec{v} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \quad (5.11)$$

where,
 $f_1 = \mu_X$, feature 1
 $f_2 = \mu_Y$, feature 2

$$f(x, y) = -\left\{ \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} \right] \right) \right\} \quad (5.12)$$

where

μ is the mean in each dimension, $\mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}$

σ_x^2 is the variance

σ_y^2 is the variance

ρ is the correlation between X and Y such that $\sigma_X > 0$ and $\sigma_Y > 0$

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\rho\sigma_Y \\ \rho\sigma_X\rho\sigma_Y & \sigma_Y^2 \end{pmatrix} \quad (5.13)$$

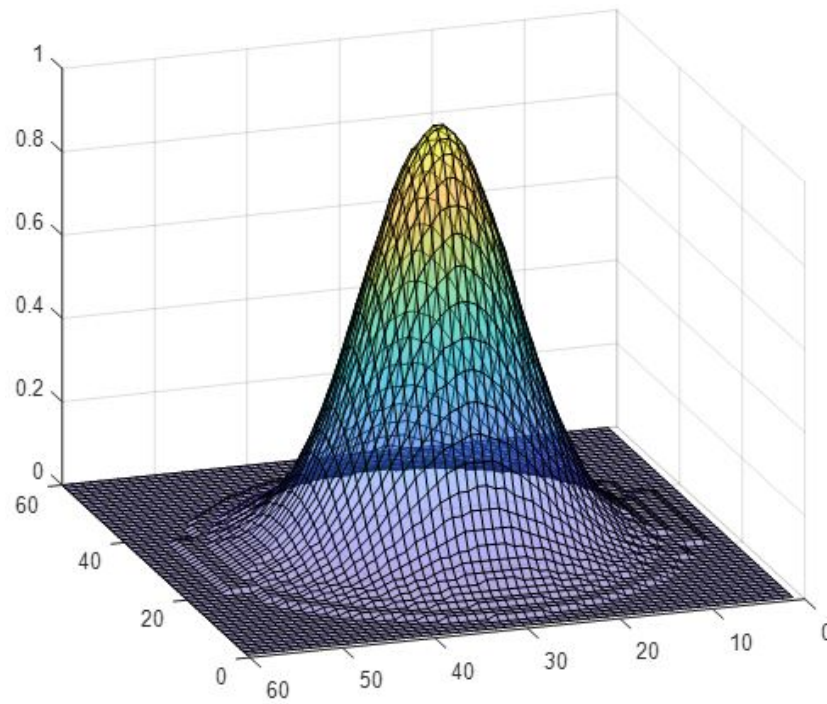


Figure 5.5: Shows the effect on the adaptive decision surface ADS after a training data point has been applied. The Gaussian distribution is with upwardly or inverted in its construction is determined by the class of the training data. The Gaussian distribution 3D view on the adaptive decision surface ADS approximation. ADS x,y,z axes 6.12

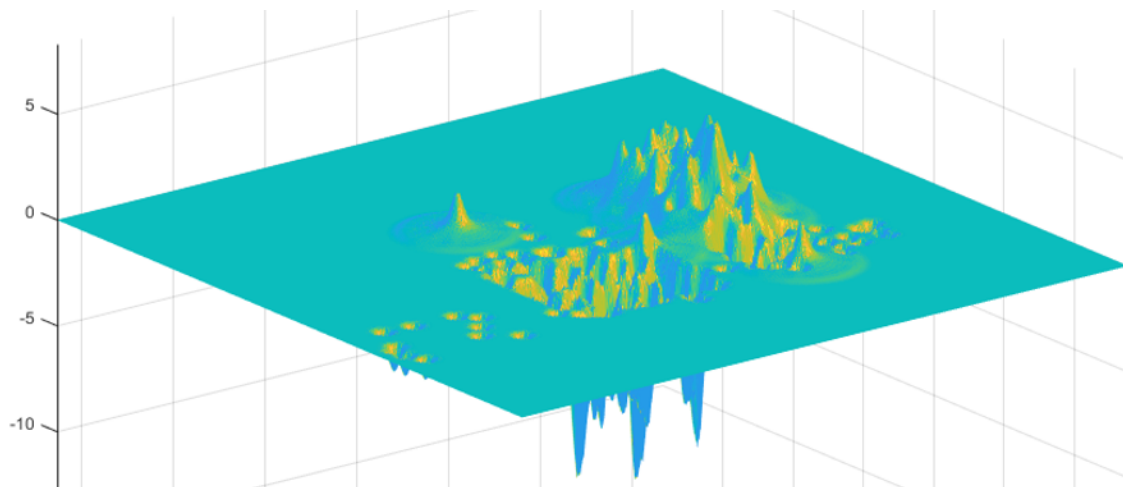


Figure 5.6: Shows the effect on the adaptive decision surface ADS after all training data points has been applied. The Gaussian distributions is are upwardly or inverted in its construction this is determined by the class of the training data. The ADS is now represents a data model of the training data in three dimensions. Top view of the adaptive decision surface ADS. x,y,z axes 6.12

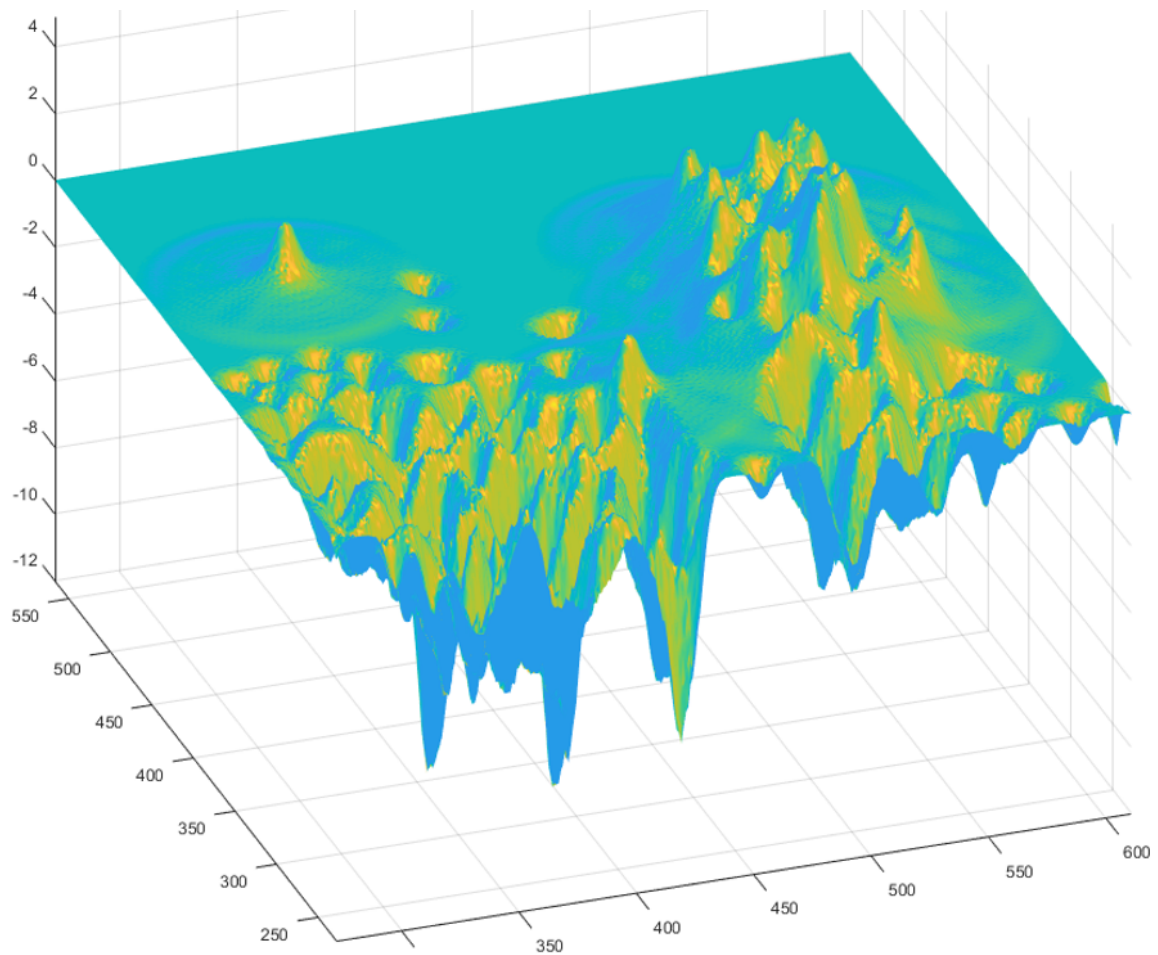
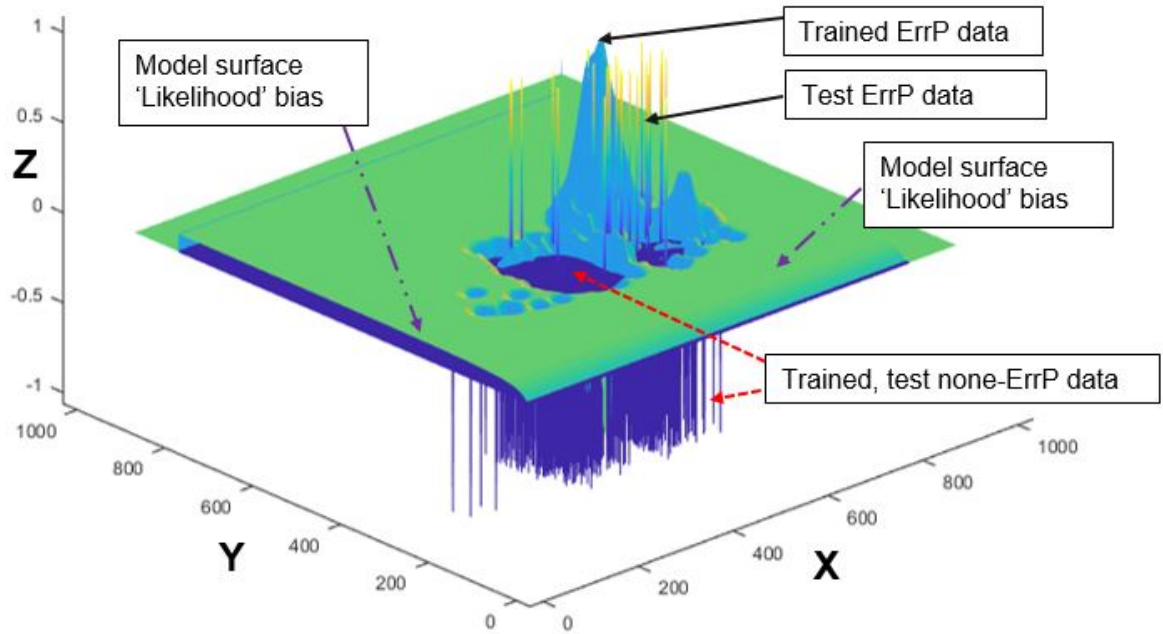


Figure 5.7: Shows the effect on the adaptive decision surface ADS after all training data points has been applied. The Gaussian distributions is are upwardly or inverted in its construction this is determined by the class of the training data. The ADS is now represents a data model of the training data in three dimensions. Top perspective view of the adaptive decision surface ADS. x,y,z axes 6.12



3D matrix of the Adaptive Decision Surface ADS, x,y,z axes

$$S \subseteq \mathbb{R}^3 \text{ Where } S = [x_{min}, x_{max}] \times [y_{min}, y_{max}] \times [z_{min}, z_{max}]$$

Figure 5.8: Shows the testing and evaluation phase of the ADS classifier. ADS x,y,z axes 6.12

5.2.2.1 Preprocessing

The proposed relative peak feature (RPF) selection method and an adaptive decision surface (ADS) classifier used research data from Ireland [23] generated from an awareness in a 'sustained attention' task. An LDA classifier has been applied to the aforementioned data to detect ErrP events. Thus a performance comparison can be established between the LDA and the proposed RPF selection method and the ADS classifier.

Generally, the cause of the 'Error related potential' ErrP is that the subject is aware of their mistake during the sustained attention task. The configuration of the experiment is such that the subjects will evoke ErrP events during the experiment. The test configuration details are as follows:

Table 5.1: Experiment Configuration

Parameter	Settings
Channel	Cz
Re-sampling:	64 Hz
Electrode location:	10/20 international system
Sampling rate:	512 - 1024 Hz
Time intervals:	350 ms window
Starting:	50 ms after feedback
Ending:	400 ms after feedback
Baseline interval:	-200 to 0
Referencing:	Common average
Band pass filtering:	1 to 10 Hz
Criteria for artefact rejection:	30 micro volts

5.2.2.2 Feature extraction and classifications

Information data loss can occur depending on exactly when the data is sampled from the EEG data recording. The time is normally time locked however to improve the capture of the Cz signal the capture windows is dynamically adjusted to synchronise to the minimum value of Cz. Hence this will capture the dynamics of the Cz signal Fig 5.9 , Please see algorithm 1.

Algorithm 1 *Adaptive Decision Surface (ADS) algorithm*

1: A set of coefficients were estimated heuristically
Class 1 Gaussian Scalar = $c1GS$
Class 2 inverted Gaussian Scalar = $c2GS$
Surface x Scale = xS
Surface y Scale = yS
time shift = tS

2: Select the EEG Cz channel

3: Select the 4 blocks of training data

for 1 : 4 do

- Import the EEG data from file
- Interpolate the bad channels
- Re-sampling
- Band-pass filtering
- Extract the time events
- Initialise variables
- for 1 : No go trials do**
 - if User is aware of an incorrect response then**
 - Adjust to the time shift tS
 - Extract the Cz waveform at time tS
 - find the minimum in the Cz waveform
 - Synchronise t at the minimum Cz
 - Extract the adjusted Cz waveform
 - Save the Cz 'ErrP' data
 - for 1 : Go trials do**
 - if User responded correctly then**
 - Extract the Cz waveform
 - Save the Cz 'None ErrP' data

4: Import the all the Cz waveforms

for 1 : All Cz waveforms do

- Find the maximum and minimum of Cz waveform
- Find the midpoint of Cz waveform
- Save the Cz 'relative peaks' data, Fig 5.10

5: Construct the Adaptive Decision Surface

6: Bias the Adaptive Decision Surface, Fig 5.12

for 1 : All Cz waveforms do

- if ErrP event then**
 - select the 'ErrP' feature coordinate
 - Plot a Gaussian scaled by $c1GS$
 - please see, Fig 5.11
- if None ErrP event then**
 - select the 'None ErrP' feature coordinate
 - Plot an inverted Gaussian scaled by $c2GS$
 - please see, Fig 5.11

6: Find the optimised level ADS classifier performance, Fig 5.14, Fig 5.13

7: Save the optimised Adaptive Decision Surface

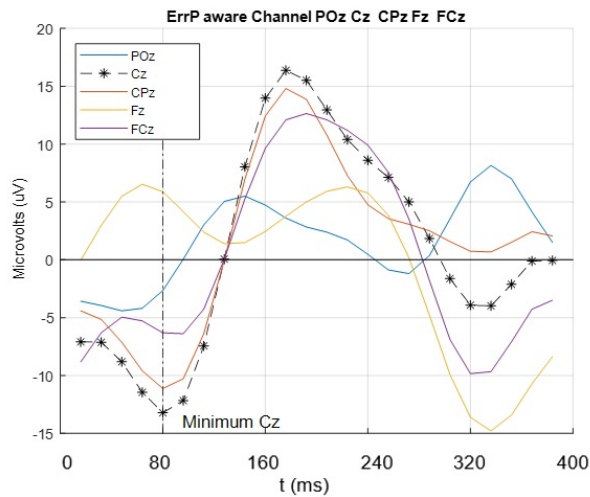


Figure 5.9: Shows an ErrP event in the temporal domain and indicates the process of synchronising the ErrP to Cz minimum.

5.2.3 Adaptive Decision Surface(ADS) Classifier

Observing the error-related potentials appear to have a high amplitude between the graph trough and graph peaks. It was particularly noticeable in the 'Cz' EEG channel recordings. Hence there will be a significant advantage in remapping the peaks of the graphs relative to each other. Moreover by remapping the peaks will not be affected by the lateral shifts of the EEG recording.

Within the capture window, the peak 'Pe' and the trough 'Ne' is located using a rectangle as a frame the central point is located. The data points are re-mapped to a new graph positioning the previous central points to the origin of the new graph.

This process is repeated for any subsequent event hence remapping the peaks relatively to each other waveforms. Please see Fig 5.10 for Cz Relative peak plot.

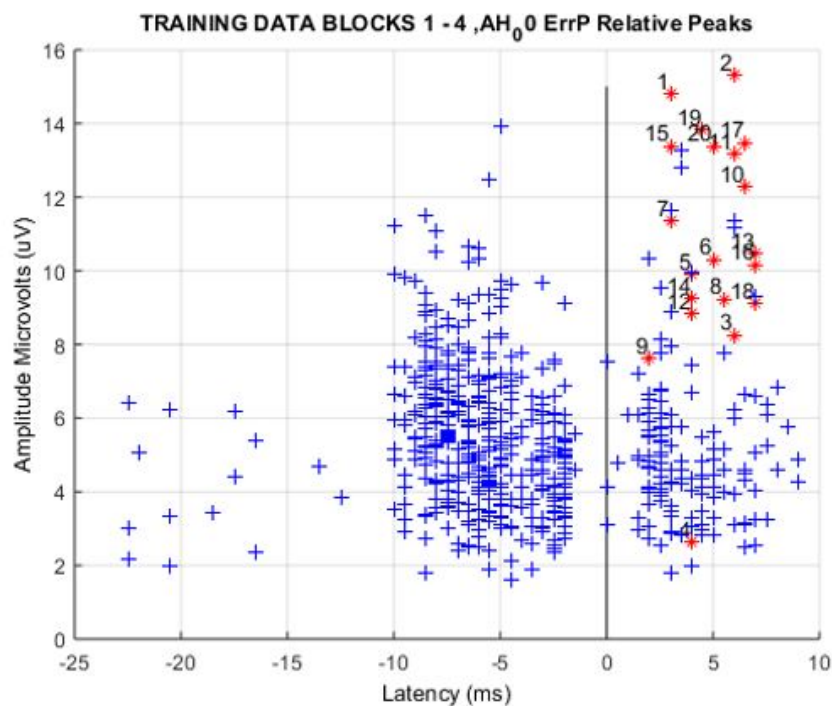


Figure 5.10: EEG waveform Cz Relative Peak Plot

5.2.3.1 Building the adaptive decision surface classifier

The adaptive decision surface is constructed in a three dimensional space equation 6.12. Information from the relative peak plot Fig 5.10, is transformed to the surface of the ADS. Such that the training data representing an ErrP event classified as class 1 'ErrP' positive and a Gaussian distribution is centralised on the transformed coordinate in the ADS Fig 5.11. Moreover the non-ErrP events are classified as class 2 'none ErrP' negative Fig 5.11 and an inverted Gaussian distribution are centralised on the transformed coordinate in the ADS.

$$S \subseteq \mathbb{R}^3 \quad (5.14)$$

where in the adaptive decision surface ADS [1].

$$S = [x_{min}, x_{max}] \times [y_{min}, y_{max}] \times [z_{min}, z_{max}] \quad (5.15)$$

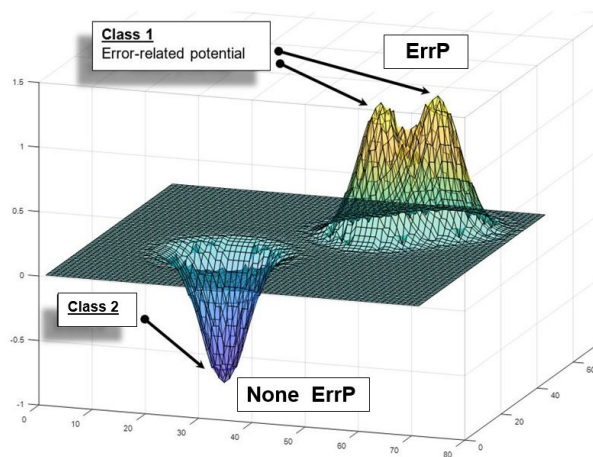


Figure 5.11: Adaptive Decision Surface ADS with three vector points. ADS x,y,z axes 6.12

5.2.3.2 Building 'Likelihood' bias into the adaptive decision surface ADS

The biased 'adaptive decision surface ADS' will support a likelihood for an event-related potential it will bias a region on the ADS where there is a higher likelihood of an ErrP Fig 5.12. The bias will be adaptable in a future decision processing system.

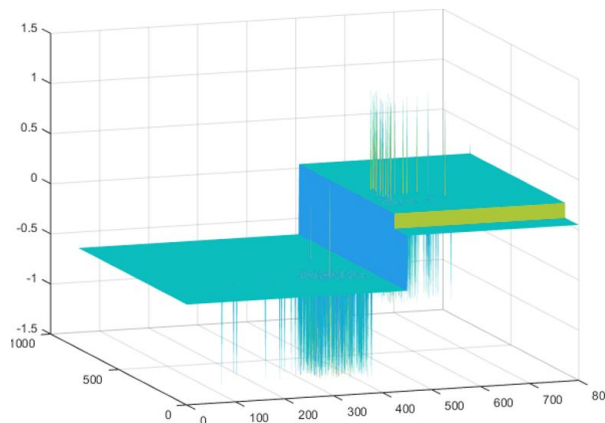


Figure 5.12: Adaptive Decision Surface ADS, incorporating a 'Likelihood' bias

5.2.3.3 Building the adaptive decision surface ADS

Using the real training data from data Block 1 to 4 from the relative peak plot Fig 5.10 and the likelihood bias Fig 5.12, it was possible to construct an initial 'adaptive decision surface' ADS. The following shows a cross-section through the structure to reveal the structure after the real data is applied to the surface. please see Fig 5.13

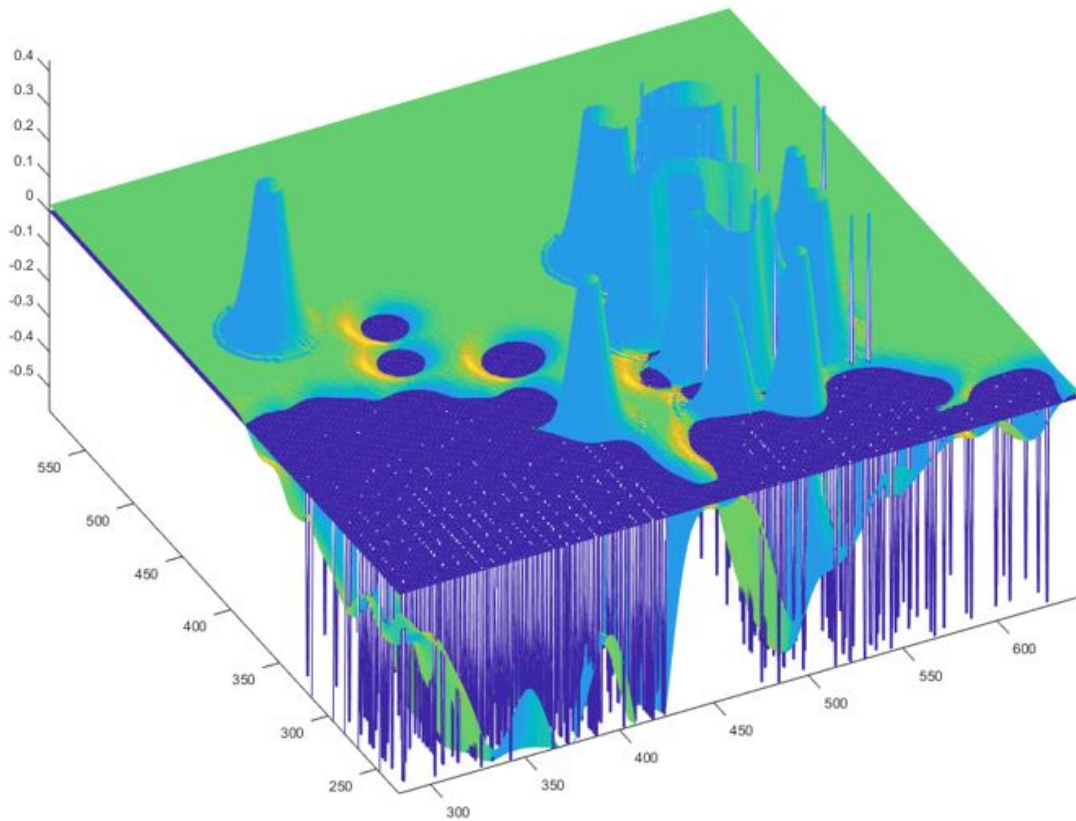


Figure 5.13: Adaptive Decision Surface ADS, cross-section training blocks 1 to 4 with 22 ErrP points. ADS x,y,z axes 6.12

5.2.3.4 Optimising and calibrating the ADS decision level

The ADS classifier decision level can be adjusted to achieve different performance characteristics Fig 5.14. The decision level is the cross-section horizontally through the ADS.

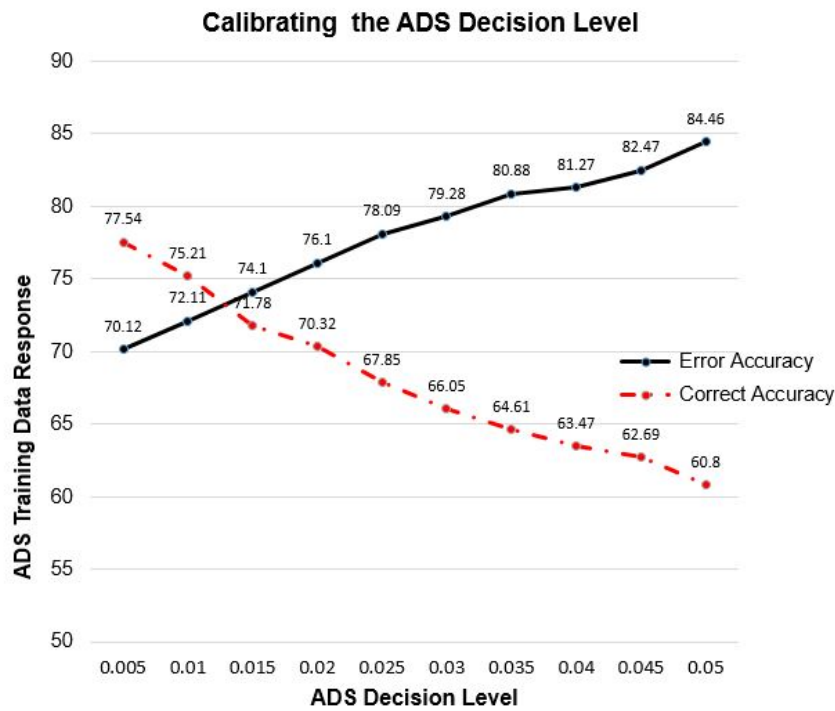


Figure 5.14: Calibrating the Adaptive Decision Surface

5.2.3.5 Adapting the surface of the ADS

The ADS classifier has the option to adapt the surface during classification. After the acquisition of the input vector feature, the ADS will classify the unknown vector. The adaptation will occur if the surface reading is above a certain threshold for the class. Please see algorithm 3

Algorithm 2 Adaptive Decision Surface (ADS) Adapter algorithm

```

1: A set of coefficients were estimated heuristically
   Class 1 threshold =  $th1$ 
   Class 2 threshold =  $th2$ 
2: Acquire the unknown relative peak plot feature coordinate
if AdaptiveDecisionsSurface > 0 then
   Classify as an ErrP
   if AdaptiveDecisionsSurface >  $threshold_{th1}$  then
     input feature coordinate
     Plot a Gaussian scaled by  $c1GS$ , Fig 5.11
if AdaptiveDecisionsSurface < 0 then
   Classify as an none ErrP
   if AdaptiveDecisionsSurface <  $threshold_{th2}$  then
     input feature coordinate
     Plot an inverted Gaussian scaled by  $c2GS$ , Fig 5.11

```

5.3 Experiment data

5.3.1 Participants

The data originated from experiments conducted by Dr. Eric Lacey [23]. All participants were healthy individuals comprising of the following: Fifty-three participants volunteered to participate in the study. 53 participants; 29 young (19 female) and 24 older (15 female). The younger

participants had a mean age of 24.14 years (SD = 5.4), and the older participants mean age was 70.67 years (SD = 4.9). During the study, procedures were approved by the Trinity College Dublin ethics committee and in accordance with the Declaration of Helsinki [109].

5.4 Results

5.4.1 Results from the adaptive decision surface ADS

The ADS classifier uses 25 times less training features. ADS used the relative peaks of the Cz electrode channel as a feature input. Interns of features quantities, Person 1 the ADS classifier used 606 Training features compared to LDA using 15,150 training features.

The average results for the classifiers for all the participants are presented in the following table.

Table 5.2: Young and elderly classifier results

Classifier	Participants	Error Accuracy [%]	Correct Accuracy [%]
ADS	27 Young	81.85	84.17
	24 Elderly	63.68	82.63
LDA	27 Young	69.13	81.46
	24 Elderly	55.18	76.95

The threshold setting for the 'Adaptive Decision System' ADS was: Any data reading off the surface of the ADS greater than 0 is defined as class 1. Conversely, any data reading off the surface of the ADS less than 0 is defined a class 2

Table 5.3: ADS Classifier with relative peak features performance

Table 5.4: Young Participants			Table 5.5: Elderly Participants		
Young Person	Error Accuracy [%]	Correct Accuracy [%]	Elderly Person	Error Accuracy [%]	Correct Accuracy [%]
1	100.00	94.51	1	60	79.93
2	91.67	78.20	2	68.75	67.9
3	79.49	81.10	3	38.46	82.68
4	89.29	97.34	4	50	84.33
5	76.92	80.86	5	40	88.61
6	64.00	82.66	6	34.29	85.58
7	78.95	90.51	7	71.43	82.35
8	78.13	75.70	8	62.5	86
9	84.78	82.16	9	42.86	84.95
10	66.67	73.52	10	56.25	81.54
11	70.00	85.69	11	60	86.67
12	87.5	92.95	12	71.88	80.55
13	90	91.30	13	64.71	82.98
14	83.33	80.15	14	89.29	73.81
15	77.42	91.73	15	56.25	81.73
16	78.95	80.82	16	70.37	85.92
17	73.91	82.75	17	100	87.96
18	89.66	74.39	18	80.95	86.01
19	95.24	86.79	19	58.82	81.96
20	84.62	72.93	20	66.67	79.26
21	73.33	82.95	21	70.59	74.68
22	76.19	85.32	22	100	85.02
23	83.33	86.38	23	62.96	87.27
24	85.71	74.37	24	50	85.43
25	90	92.02	mean	63.68	82.63
26	80	88.83			
27	80.77	86.67			
mean	81.85	84.17			

5.5 Summary

Using the Cz electrode and a relative peaks features selection with a trained adaptive decision surface provided an increase in performance for both young and elderly participants and opens the possibility for a new type of feature selection and classifier which uses fewer features.

5.6 Discussion future investigations

It may be beneficial to investigate image recognition using 'Morlet' Wavelet to the EEG research data please see Fig 5.15 there could be some advantages and possible innovative steps in pattern recognition in the time-frequency domain.

The pattern recognition could operate in parallel with these other subsystems please see section 5.2 to enhance accurate detection of patterns in the EEG data such as error-related potentials or motor imagery. The pattern recognition could be achieved by using a 'deep learning' neural networks such as a CNN. We propose to further investigate pattern recognition

with CNN.

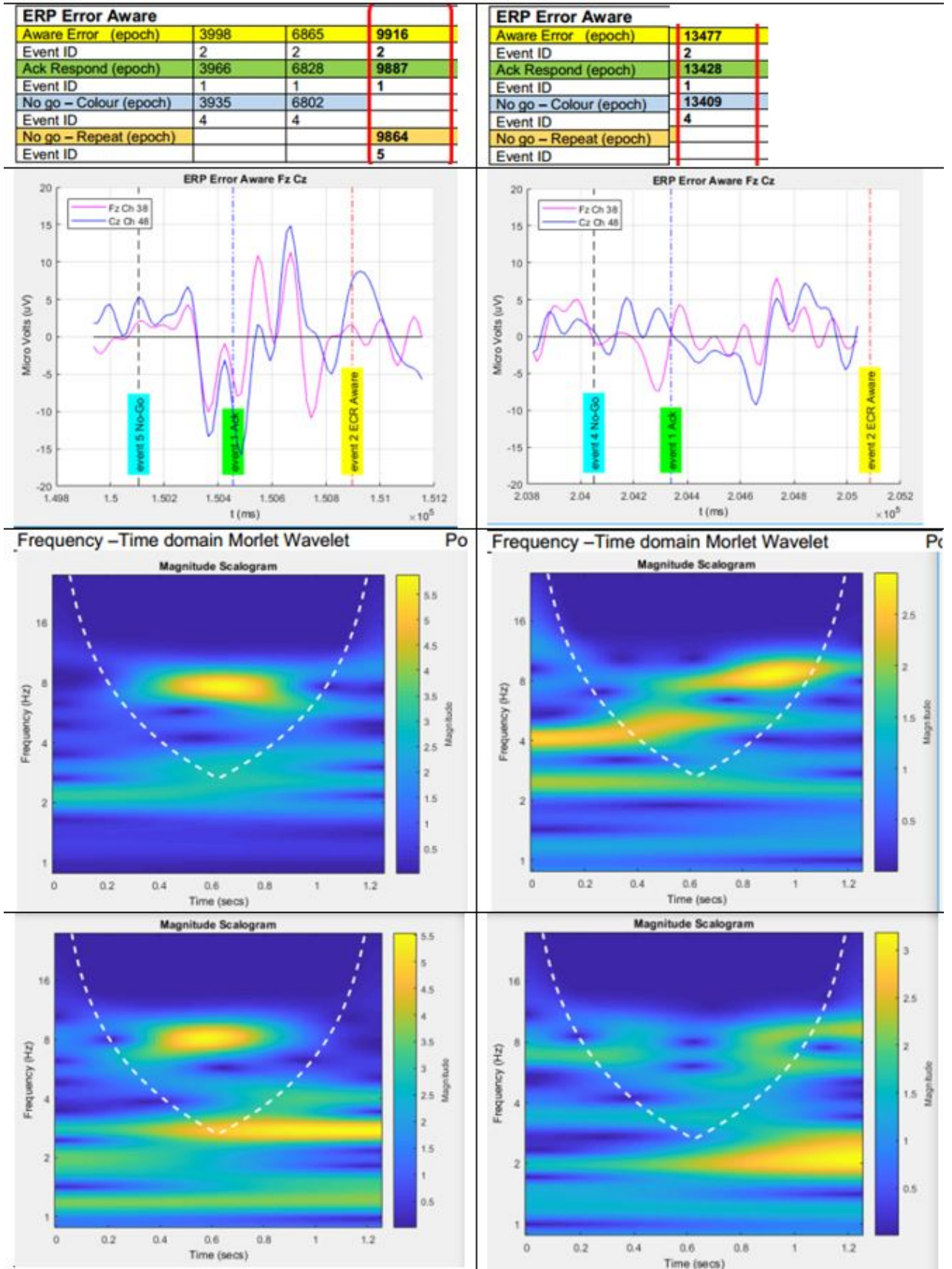


Figure 5.15: Shows the Fz and Cz EEG waveforms during an ErrP event and the associated Morlet Wavelets.

Chapter 6

Motor imagery based BCI and performance comparison of 23 classifiers

The purpose of motor imagery is to re-establish meaningful communications in terms of ascertaining the desired choice from a person with a disability. The methodology requires the persons to imagine moving parts of their body such that a motor imagery brain computer interface can determine the imagined movement hence it supports the needs of the person with a disability.

6.1 Introduction

A person with a disability unable to move their limbs such as their arms, hands, feet like they use to do in the past must be a debilitating condition. Previously the person may have been fully able bodied such that the person will have learned during the early stages of their life how to move their limbs. Having the prior skills of physical movement the person with a disability has the knowledge in their brains. Supporting the persons disability can be achieved in the following steps. Firstly, with this prior knowledge a person with a disability can imagine moving their left hand or right hand. Secondly a brain-computer interface using motor imagery will be able to determine the imagined movement. Finally an assistive device may support the person with a severe disability objectives.

6.1.1 Re-establish a sense of empowerment

How would a brain-computer interface controlling an assistive supportive device know about the person's desire to move their hands? This could be a scenario helping the person with a disability pick up any object such as a drinking cup. What sort of everyday daily tasks could be re-enabled for a person with a disability? What kind of assistance could make a positive supportive difference in that persons life, moreover it would be enabling hope and independence. There is a type of brain computer interface that can be placed over certain regions of the human brain (see Fig 7.1). Moreover when electrodes are placed at specific locations such as the motor cortex the person with a disability imagined movements can be detected (see Fig 6.2).

6.1.1.1 Functional regions of the human brain

The human brain can be divided into a number of regions according to the brain functionality.
Frontal lobe: behaviours coordinated information from other area movement concentration , thinking , personality

Parietal lobe: sensory skin information, muscle-skeletal system, taste buds, language , pressure, pain, temperature

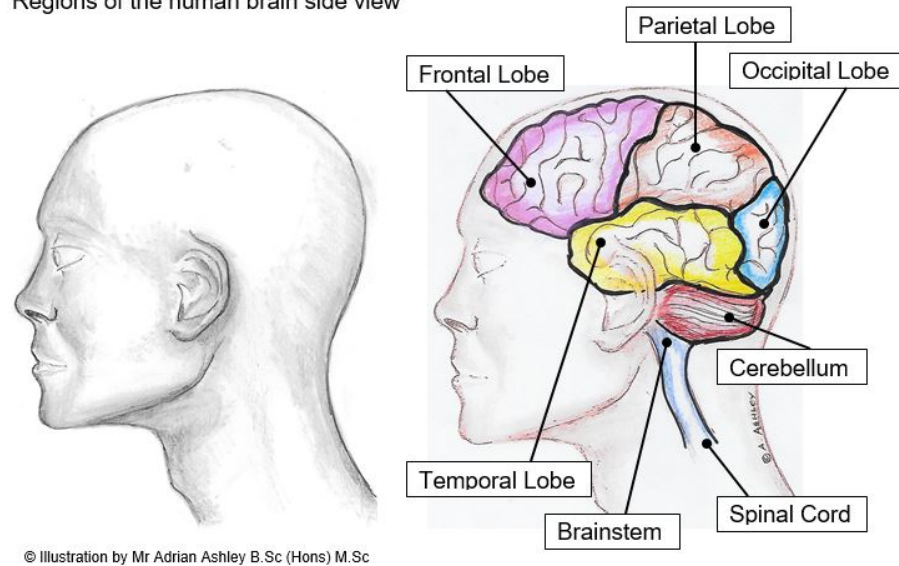
Occipital lobe: vision, perception

Brain stem: consciousness breathing , heart rate

Cerebellum: posture balance coordination of movements

Temporal lobe: hearing , language memory

Regions of the human brain side view



© Illustration by Mr Adrian Ashley B.Sc (Hons) M.Sc

Figure 6.1: Regions of the human brain

6.1.1.2 Locations for motor imagery in the human brain

The human brains functional regions can be further subdivided into specific functions for example the primary motor cortex (see Fig 6.2).

Motor imagery brain side view

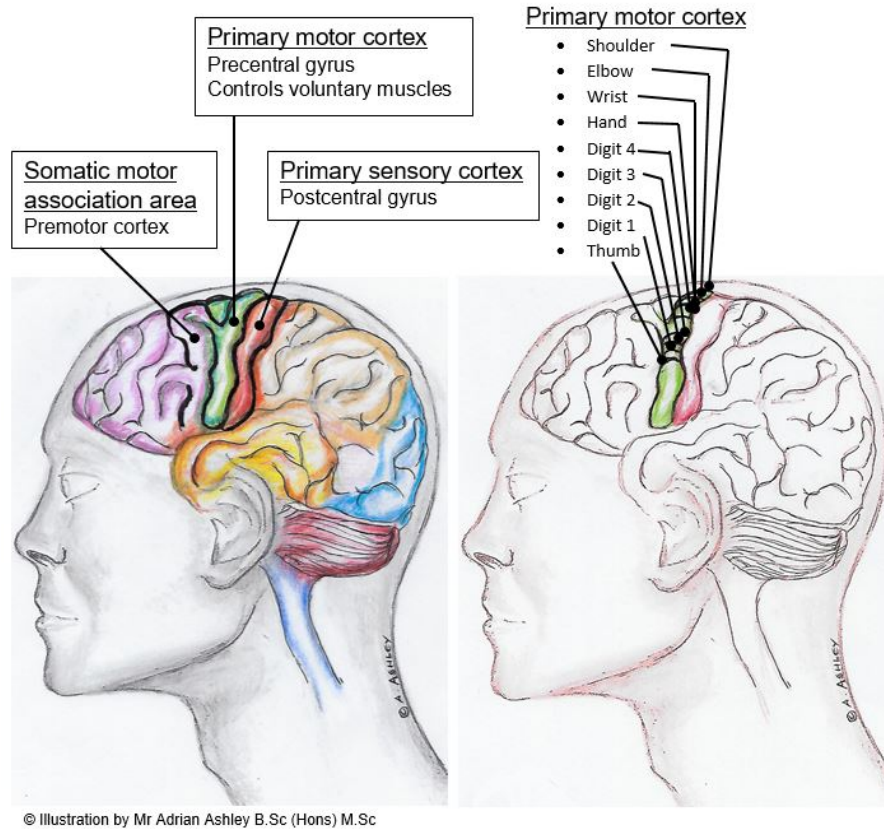


Figure 6.2: Regions of the human brain

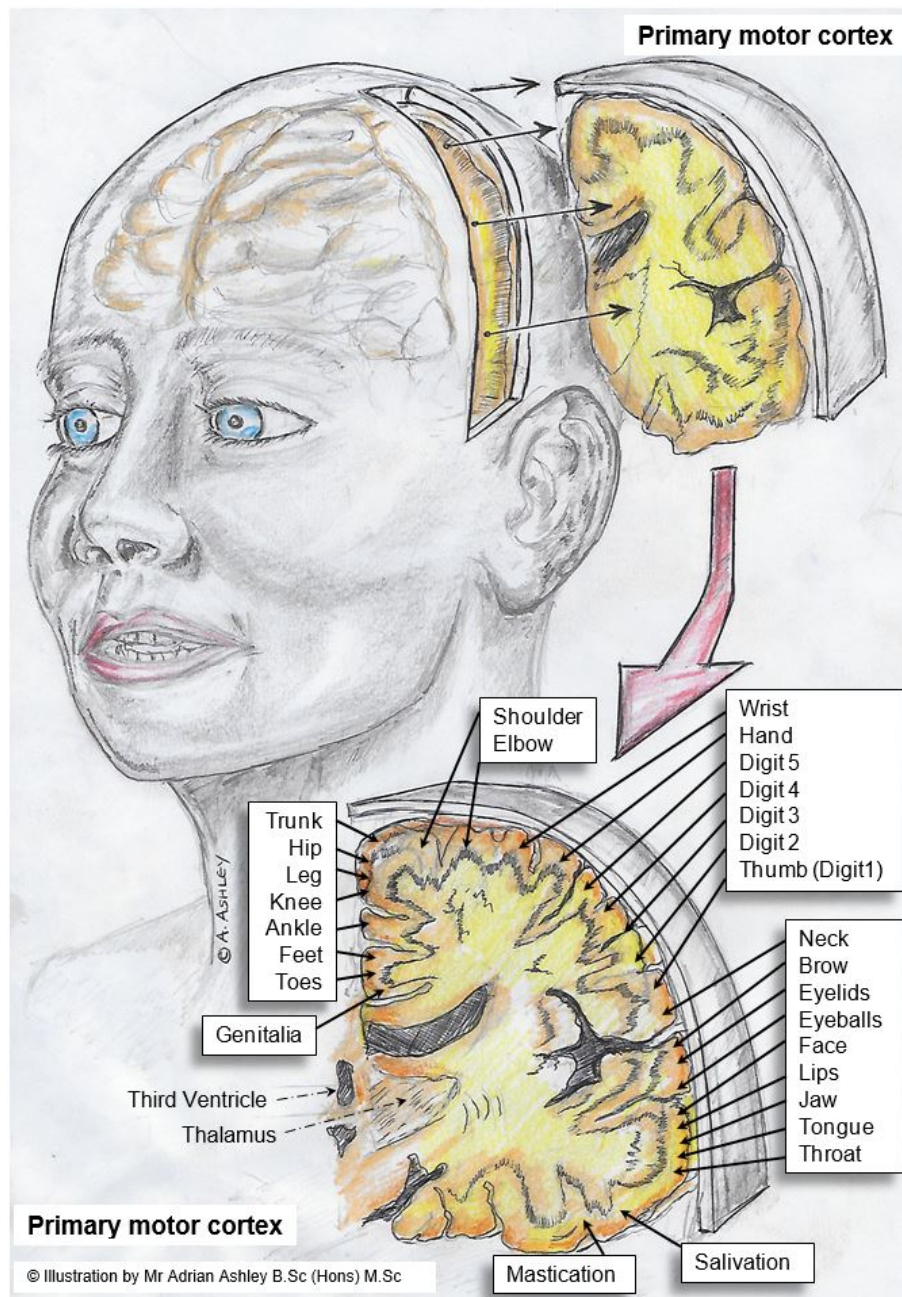


Figure 6.3: Human brain motor imagery labelled

6.1.2 Motor Imagery Brain Computer Interface (MI-BCI) Components

As an overview for motor imagery brain-computer interface comprises of a number of components used to establish the MI-BCI. Firstly there should be a method of acquiring information about the brains activity. One of the methods of acquisition is by using EEG which is relatively low cost and portable. Secondly, EEG signal processing which consists of at least three components which are preprocessing features extraction and classification. Finally the feedback to the user which can be any number of stimuli in terms of audio, visual or tactile. Moreover, the feedback is provided by systems, for example, the feedback could be the user observing a graphical user interface from a portable computer or the feedback could be as an assistive robotic device such as a prosthetic hand or a robotic arm or mobile robot. Furthermore, the feedback which is driven by the output from the MI-BCI classification. Please see Fig 6.4 for an overview of a MI-BCI system

EEG based motor Imagery BCI System

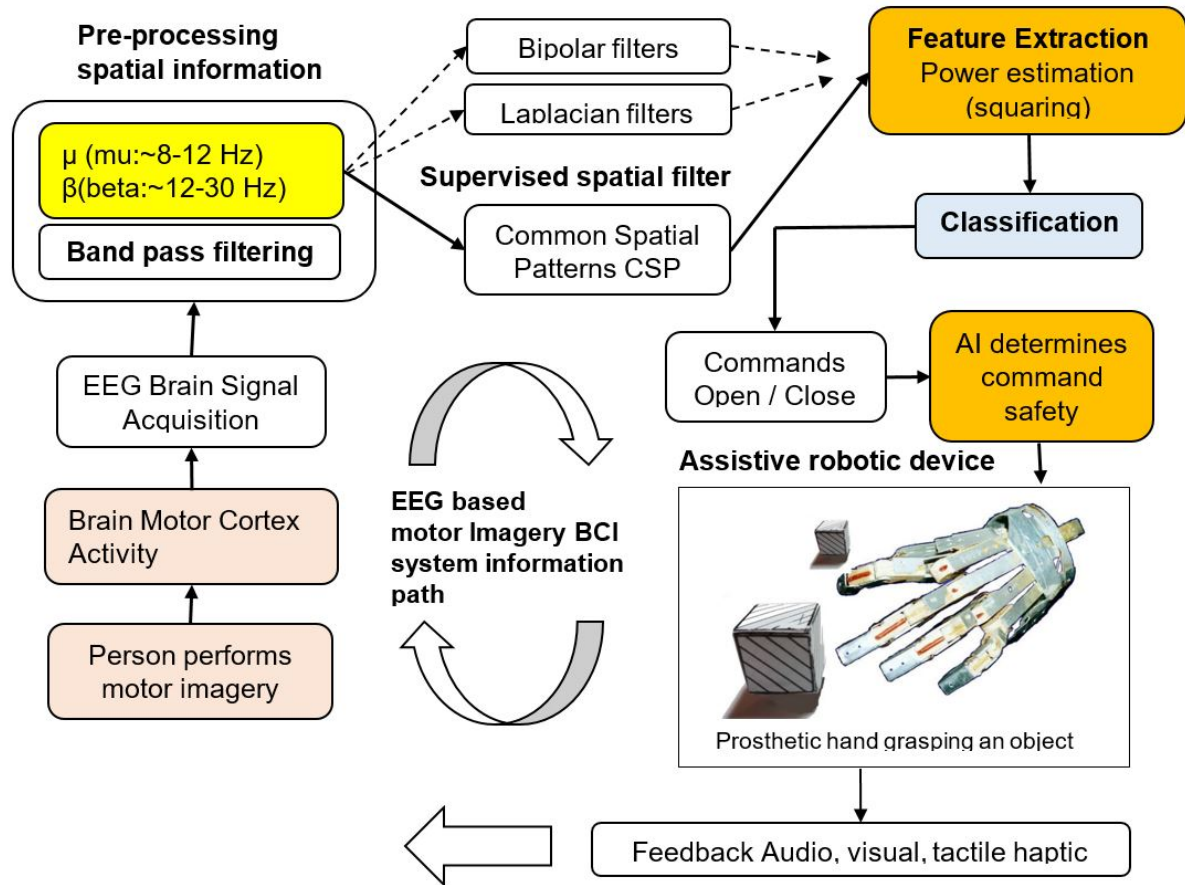


Figure 6.4: A schematic diagram of an EEG based MI-BCI System

6.1.2.1 MI-BCI Electrode channel location

The basic EEG channels are located at certain locations on top of the scalp in the area near to the motor cortex under electrodes C3 and C4, Fig 6.5 The right hand motor imagery can be detected by electrode C3 and left hand motor imagery can be detected by electrode C4. Electrode Cz is associated the foot motor imagery.

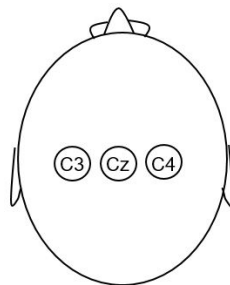


Figure 6.5: EEG Channel C3 and C4 location on top of the scalp

6.1.2.2 Feature extraction using band power

In the previous section we explained about the components of BCI 2.2.4 and methods of feature extraction 2.2.4.2 the **Power spectral density (PSD)**: PSD in fig 6.6 has been applied to the a data set [23] in the PSD using the Welch method with Hamming in order extract features. Jatupaiboon et al [111] proposed a real-time happiness detection BCI system was researched

using PSD. In [112] used PSD over 250 milliseconds of data with a temporal shift of 32.1 milliseconds in the bandpass of 4 to 40 Hz using 19 electrodes. The authors used the Hidden Markov Models (HMM) to discriminated between different motor imageries In [113] used the Welch periodogram algorithm to estimate the power spectrum density of each surface Laplacian transformed channel. Further, a neural classifiers were used in this asynchronous BCI.

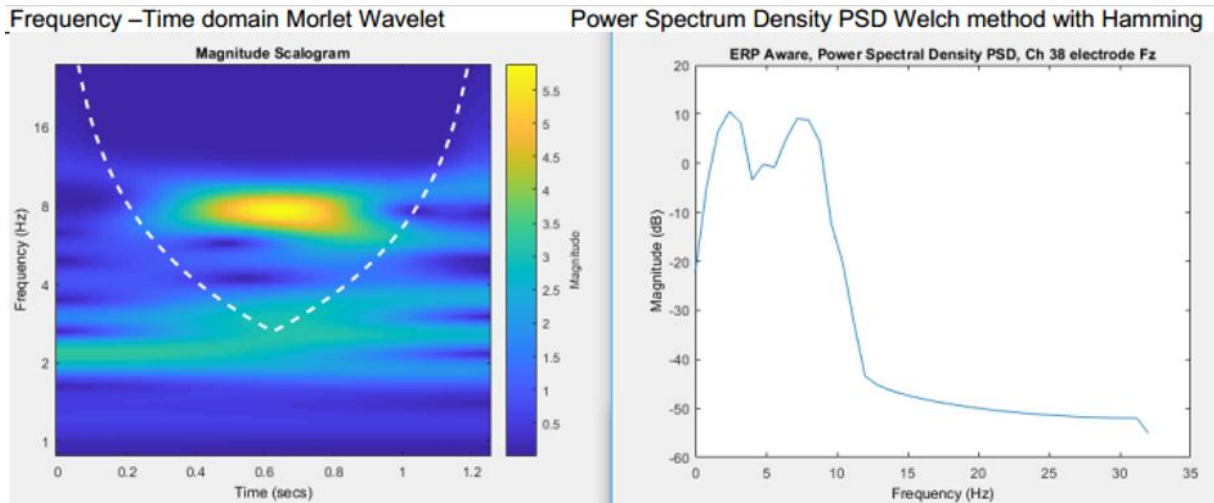


Figure 6.6: Morlet wavelet ErrP Fz Cz by A Ashley

6.1.2.3 MI-BCI left and right hand detection performance

To establish the very basic motor imagery BCI will require the following components Calculate the band power features P in the specific frequency bands μ (7.5 - 12.5) and β (12 - 30 Hz). We are looking a specific EEG Channels such as C3 for right hand motor imagery and C4 for left hand motor imagery. A feature vector can be constructed from the power such as:

$$V = [P_{c3-\mu}, P_{c4-\mu}, P_{c3-\beta}, P_{c4-\beta}] \quad (6.1)$$

A classifier such as linear discriminant analysis LDA can be used to determine the classification based upon the input feature vector V equation 6.1 comprising of the power from the specifies electrodes and frequency range.

This basic configuration for motor imagery does not generate a high performance and it is not really optimal for the following reasons. EEG information maybe missing because only C3 and C4 channels are used. There is a risk of being subject dependent by using only C3 and C4 channels and also using fixed frequency bands frequency bands μ (7.5 - 12.5) and β (12 - 30 Hz). The above may work for certain people but becomes less universal for a MI-BCI.

The basic MI-BCI performance with the following configuration has been shown to have and accuracy of about 61% [114]

- ◇ **BCI competition IV** [115] data set 2a
- ◇ **9 subjects** in a motor imagery task using left and right hand motor imagery
- ◇ **72 trials** per class for both training and testing
- ◇ **LDA classification algorithm** [116] [117]

What can be done to improves the performance beyond the 61% for detecting motor imagery events? Perhaps using more features from other EEG channels near to C3 and C4? The possible solution of using more channels will inevitably require more training and testing data.

One possible solution would be to use a well established method called spatial filtering [114] [69] [20].

In a method called linear spatial filtering a number of additional channels are defined as a linear combination (see Equation 6.2) of other electrodes such as C3 and C4 as explained in Lottes lectures [114]. The EEG signals emanating exactly from the expected location, the effect of the brain and the skull bone affects the exact location of the EEG signals the method of linear spatial filtering enables the ability to acquire the signals of interest such as C3 and C4.

$$X' = \sum_i w_i x_i = wX \quad (6.2)$$

6.1.2.4 MI-BCI Improving performance with Spatial filters

The purpose of spatial filters is to reduce unnecessary spatial EEG electrical activity and highlight a particular location of interest. In addition, the spatial filter will maximize the signal to noise ratio such that accuracy of EEG-based communication will be improved shown in [20]. The classification process will benefit from the improved EEG signal with a more accurate classification. There are a number of well know spatial filters such as the basic 'Bipolar' filters and the 'Laplacian' filters. In addition, there is a supervised spatial filter and the common spatial filter (CSP) which provides an increase in performance as shown in [118].

6.1.2.5 MI-BCI basic spatial filters

Bipolar filters use the EEG information from a pair of electrodes for example if we are interested in capturing the EEG signal from C3 and C4 were acquired by the following:

$$\begin{aligned} C3 &= FC3 - CP3 \\ C4 &= FC4 - CP4 \end{aligned}$$

Laplacian filters use more EEG information essentially the filters enhance the local electrical activity detected by the EEG system and reduce defuse spatial activity as explained in [114].

$$C3 = 4 * C3 - FC3 - C5 - C1 - CP3$$

There is also other variations 'Laplacian' filter such as 'Large Laplacian' which use electrodes from further away as shown in [20].

6.1.2.6 MI-BCI supervised spatial filters

Training a classifier will require the training data that is labelled. The labelled training data is used during the supervised training phase to adjust the weights of the common spatial filter. This process finds the spatial filter weights W which provided a maximum for class 1 and a minimum for class 2 of a two class system in order to achieve a maximum class separation. For example A CSP filter using band-power features will use the variance of the band passed signal such the spatial filters are adjusted as the CSP changes the weights W so that there is an optimised discrimination between the classes based on band-power features as explained in [20].

Let X_1 be of size (n, t_1) of a sample window frame of multivariate signals comprising of n EEG channels in parallel for a duration of t_1 discrete samples.

Let X_2 be of size (n, t_2) of a sample window frame of multivariate signals comprising of n EEG channels in parallel for a duration of t_2 discrete samples.

The objective of the common spatial pattern(CSP) algorithm will be to find the weights represented by \mathbf{w}^T such that the variance of X_1 and X_2 is at the maximum separation between the two classes

Common Spatial Patterns CSP Channels 1 to 22

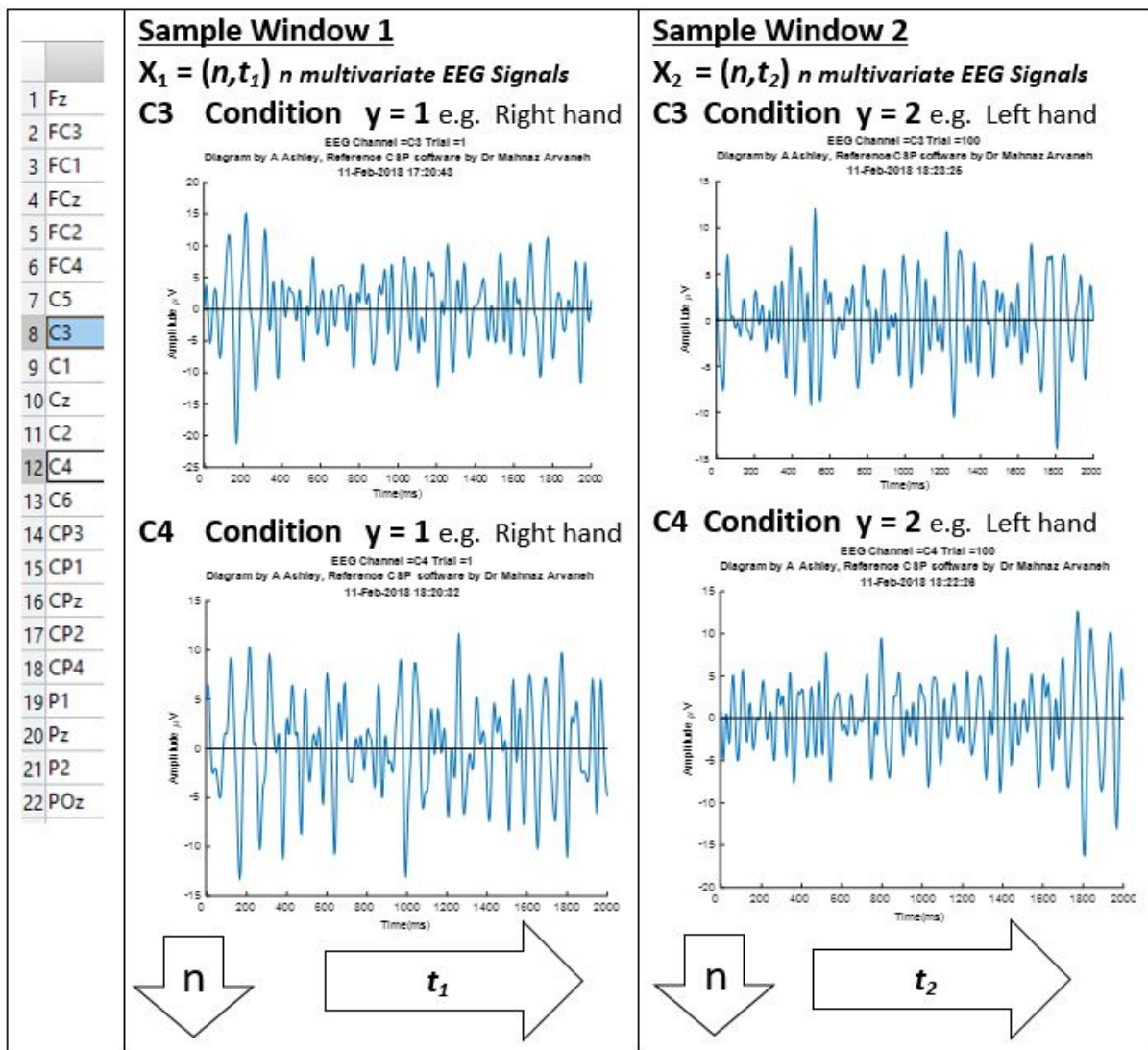


Figure 6.7: Common Spatial Patterns

6.1.2.7 MI-BCI When classification goes wrong

In 6.1.2 shows a MI-BCI with the assumption that motor imagery classification operates accurately and never generating the wrong classification. Due to the none stationarity and the myriad of problems that can occur during EEG acquisition, there is a risk of a misclassification. The consequences of the systems executing a misclassification will be the wrong feedback in the form of an audio, visual or tactile is presented to the user. For example, the wrong choice is displayed to the user or the assistive robotic device moves in the wrong direction moreover it maybe an assistive prosthetic hand which opens the robotic hand and drops the beverage on the table spilling the drink. There is a possible solution to the problem of MI-BCI misclassification which will use the error signal from the user's brain as a means of abating the error and hence mitigating the mistake and reduce the consequences of the error cascading across the system. Furthermore, it would be prudent to adapt the MI-BCI classifiers to reduce the likelihood of a future misclassification. Such that the MI-BCI will be more robust and less frustrating for the users.

6.2 Methodology

The possibility of detecting the changes in brain activity following muscle movements such as moving an arm or leg is well known in research as shown in [3]. Changes in the cortex area of the brain occur when a person moves their limbs can be detected with EEG [30]. Sensorimotor rhythms associated with oscillations in brain activity involving both sensory and motor functions comprises:

μ (7.5 - 12.5) Located over the motor cortex of the brain are synchronized patterns of electrical activity associated with a person's voluntary movement such as opening or closing your right hand.

β Range in frequency from (12 - 30 Hz) discovered by Hans Berger he also invented EEG in 1924. During early experiments it was noticed that when a person's eyes were closed the alpha waves with neural oscillation in 7.5 - 12.5 Hz reduced with movement or imagined movement and open eyes. Moreover the alpha wave is replaced by the beta wave with a reduced amplitude and higher frequency was observed when the person opened their eyes. Beta waves are associated with muscle movement furthermore beta waves increase when a person is voluntarily suppressing or resisting movement

β Low Beta Waves (12.5–16 Hz) associated with various levels of consciousness.

β Beta Waves (16.5–20 Hz) associated with various levels of consciousness.

β High Beta Waves (20.5–28 Hz) associated with various levels of consciousness.

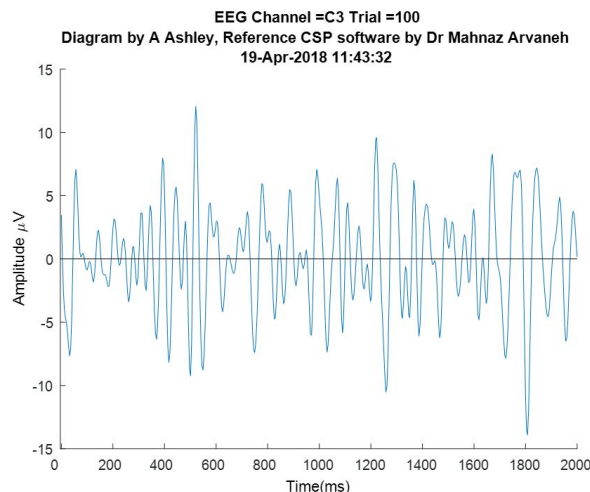


Figure 6.8: Motor Imagery Sensorimotor rhythm

These Sensorimotor rhythms can be detected in the EEG during physical movement or imagined movement [119]. Before a person moves there is a decrease in μ (7.5 - 12.5) and β (12 - 30 Hz) rhythms in the cortical area Fig 2.1 show the cortical region. The decrease is labelled as event-related desynchronization (ERD). After the movement followed by relaxation the rhythm increases and is known as event-related synchronization (ERS) [119]. In addition, ERD and ERS can occur by imagining the physical movement [3], [120]. Hence an application in BCI can enable the detection of a person's intentions and therefore restore physical movement via the BCI and assistive robotic device. Furthermore, a BCI is able to detect an ErrP when those intentions are not interpreted.

Ang et al [42] showed a method for using a filter bank common spatial patterns (FBCSP) algorithm using 4 progressive stages that incorporated signal processing and machine learning using EEG data from the BCI competition 2008 Datasets 2a and 2b Fig 6.9 and Fig 6.10. The filter bank [42] comprising Chebyshev Type II bandpass filters, spatial filtering used a CSP algorithm, CSP feature selection. A CSP projection matrix for each filter band, the discriminative

CSP features and the classifier model labelling the training data according to the motor imagery. The acquired parameters during the training phase and used the evaluation phase [42]. There are other motor imagery data sets available, however, the Graz dataset A is well known in the BCI research community.

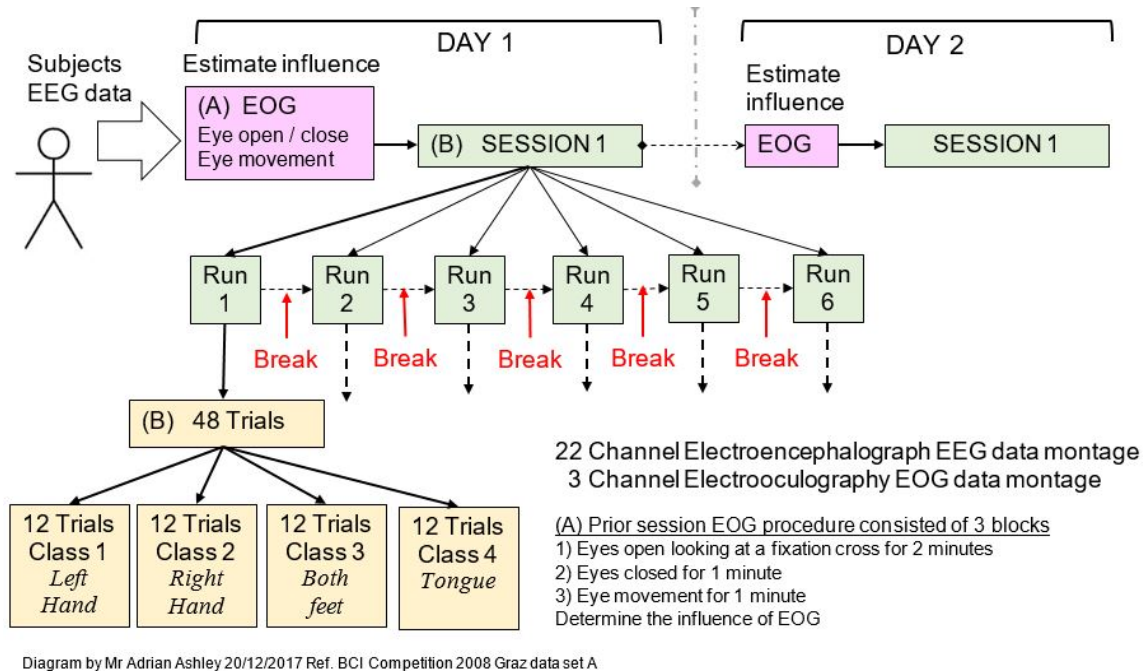


Figure 6.9: BCI competition 2008 Graz dataset A

The subject imagines moving parts of their body such as their; left-hand class1, right-hand class 2, both feet class 3 or tongue class 4. The movements are synchronised according to a predetermined timing scheme paradigm per trial per body movement class Fig 6.10

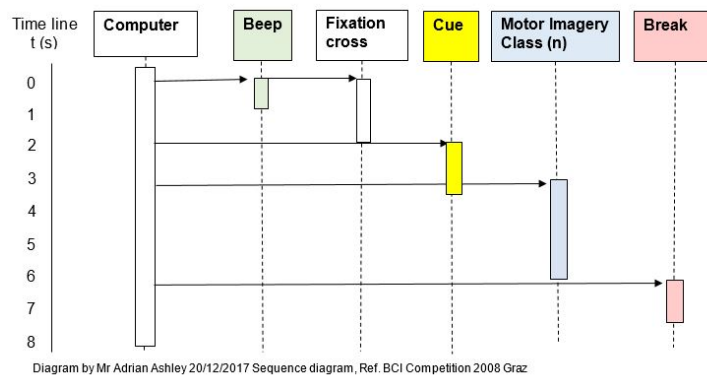


Figure 6.10: BCI competition 2008 Graz dataset A

6.3 Classifier performance experiments using 23 motor imagery classifiers

During the research into the performance of motor imagery classifiers, an experiment was carried out using MATLAB® classifiers to determine the performance of each classifier. The data used is from the 'BCI competition 2008 Graz dataset (A) Motor imagery'.

6.4 Results from the 23 motor imagery classifiers

The results from the 23 motor imagery classifiers are presented in Appendix E .

6.5 Summary of the 23 motor imagery classifier comparisons

Results show SVM Linear and SVM Quadratic are approximately balanced between the Class 1 and Class 2 at approximately 80% performance for both classes. In section 6.6, a novel three dimensional probability-based classifier is proposed for improving motor imagery-based BCI.

6.6 A novel three dimensional probability-based classifier for improving motor imagery-based BCI

Motor imagery BCI based assistive robotics solution has the potential to empower the upper mobility independence of a person with a disability. The objective of this work was to compare the classification performance of well-established classifiers with a novel prototype classifier. An adaptive decision surface ADS classifier with the future objective to augment an assistive robotic prosthetic hand to open and close to grasp an object in cooperation with LIDAR sensors is proposed. The ADS was trained with a training data set from the BCI competition IV dataset 2a from Graz University of Technology. *Main results:* The classification accuracy in the offline tests reached 76.06 % class 1 and 81.50 % class 2 using a non-adaptive ADS and 79.55 % class 1 and 99.69 % class 2 using an adaptive ADS classifiers. We show a prototype adaptive decision classifier used with motor imagery datasets.

6.7 Introduction

A person with a disability unable to move their limbs such as their arms, hands, feet like they used to do in the past must be a debilitating condition. Previously the person may have been fully able-bodied such that the person will have learned during the early stages of their life how to move their limbs. Having the prior skills of physical movement the person with a disability has the knowledge in their brains. Supporting a person's disability can be achieved in the following steps. Firstly, with this prior knowledge, a person with a disability can imagine opening or closing their left or right hand. Secondly, a brain-computer interface using motor imagery will be able to determine the imagined movement. Finally, an assistive device may support the person's objectives.

6.8 Methodology

6.8.1 MI-BCI Improving performance with Spatial filters

The purpose of spatial filters is to reduce unnecessary spatial EEG electrical activity and highlight a particular location of interest. In addition, the spatial filter will maximize the signal to noise ratio such that accuracy of EEG-based communication will be improved shown in [20]. The classification process will benefit from the improved EEG signal with a more accurate classification.

6.8.2 Motor Imagery Adaptive decision surface(MI-ADS) mathematical definition

Training set of vertices

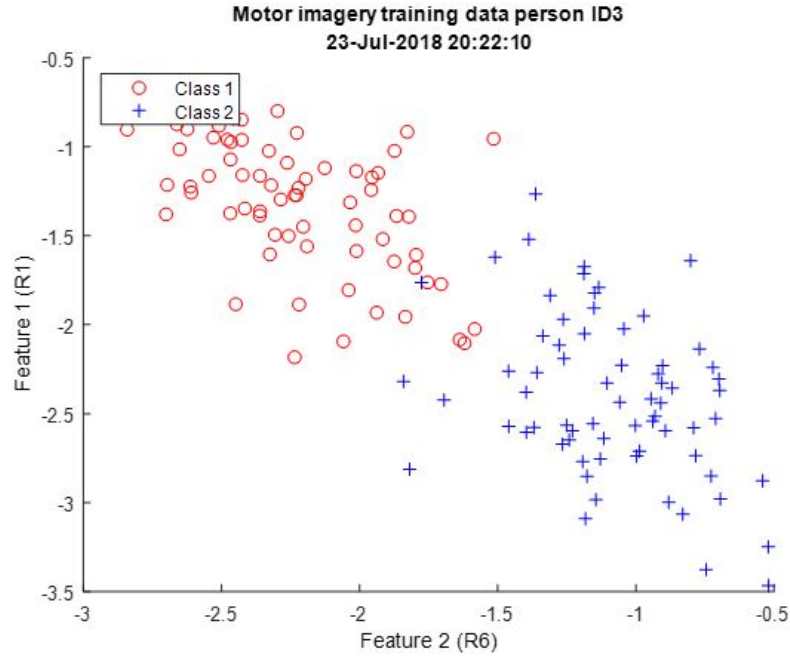


Figure 6.11: Training data Class 1 and Class 2

$$V_{Train} = \{C_1, C_2\} \quad (6.3)$$

where,

C_1 = Class 1, Right hand motor imagery

C_2 = Class 2, Left hand motor imagery

$$C_1 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\} \quad (6.4)$$

$$C_2 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\} \quad (6.5)$$

Classification 1, Right hand motor imagery

$$\vec{\mu}_{C_1} \in C_1 \quad (6.6)$$

$$\vec{\mu}_{C_1} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \quad (6.7)$$

where,

$f_1 = \mu_{feature1}$

$f_2 = \mu_{feature2}$

$$\vec{\mu}_{C_1} = \frac{1}{n} \sum_{i=1}^n \vec{v}_i \quad (6.8)$$

Classification 2, Left hand motor imagery

$$\vec{\mu}_{C_2} \in C_2 \quad (6.9)$$

$$\vec{\mu}_{C_2} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \quad (6.10)$$

where,
 $f_1 = \mu_{feature1}$
 $f_2 = \mu_{feature2}$

$$\vec{\mu}_{C_2} = \frac{1}{n} \sum_{i=1}^n \vec{v}_i \quad (6.11)$$

The adaptive decision surface ADS represented by a 3 dimensional space

$$S \subseteq \mathbb{R}^3[1] \quad (6.12)$$

where,

$$S = [x_{min}, x_{max}] \times [y_{min}, y_{max}] \times [z_{min}, z_{max}] \quad (6.13)$$

$$V \in S \quad (6.14)$$

A Bivariate Gaussian distribution data structure is projected on to S around $\vec{\mu}_{C_1} \in C_1$ and $\vec{\mu}_{C_2} \in C_2$

6.8.3 Building 'Likelihood' bias into the ADS model for motor imagery

When the ADS is used for motor imagery the likelihood is shown in Fig 6.12 The bias will be adaptable in a future decision processing system.

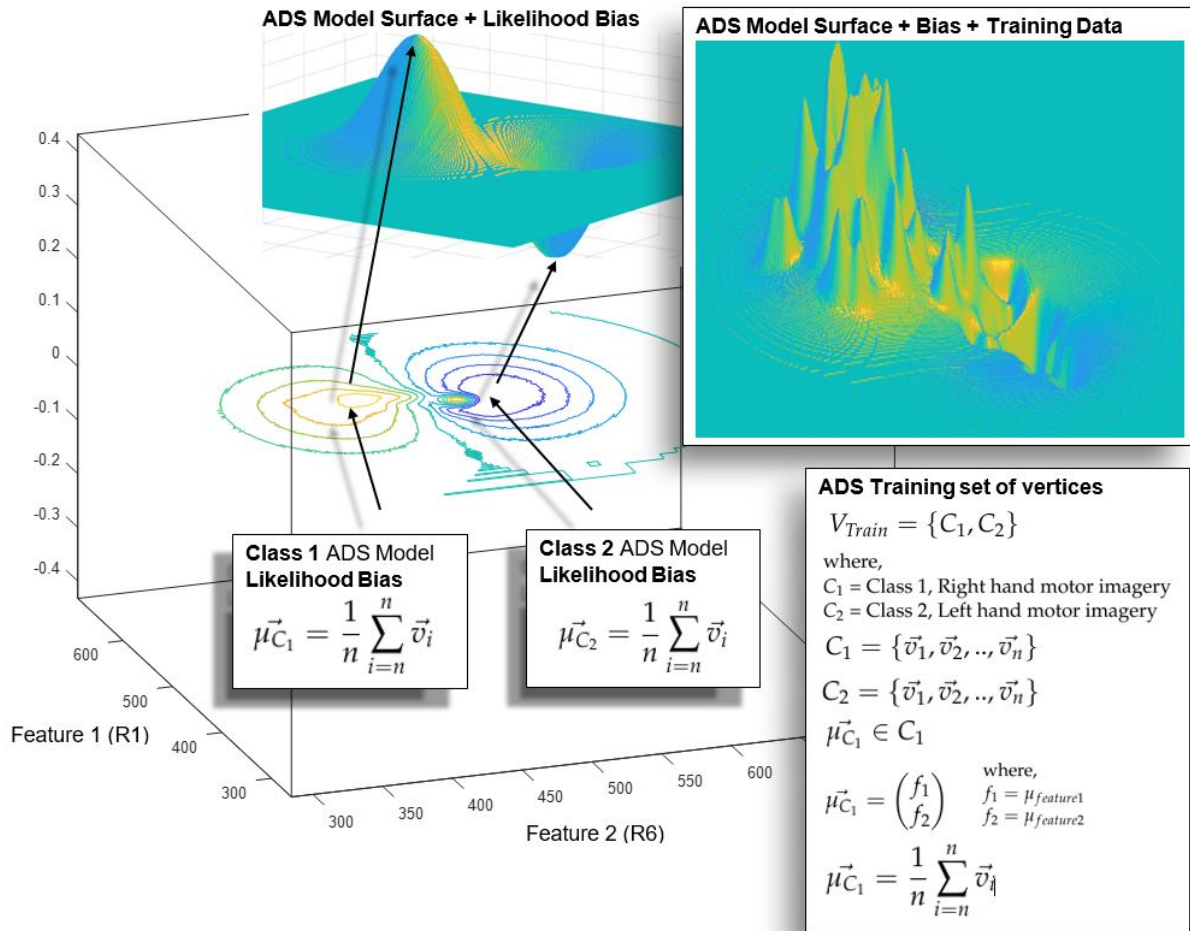


Figure 6.12: ADS Likelihood Bias

6.8.4 A constructed motor imagery adaptive decision surface model

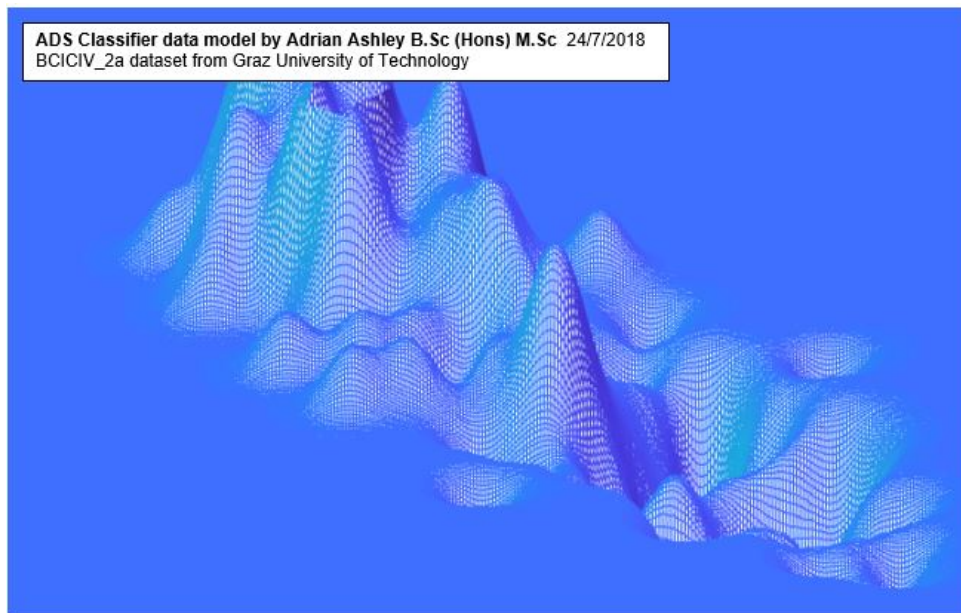


Figure 6.13: Trained motor-imagery adaptive decision surface

6.8.5 Motor imagery test data

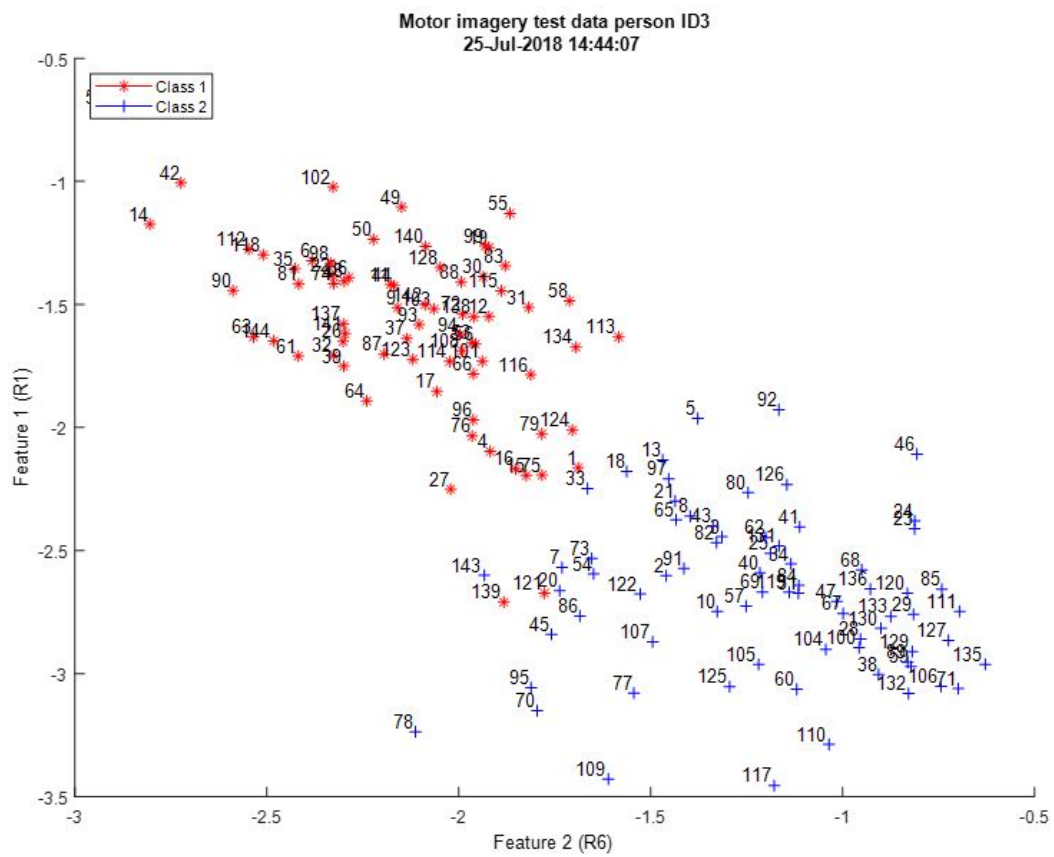


Figure 6.14: ADS Test data

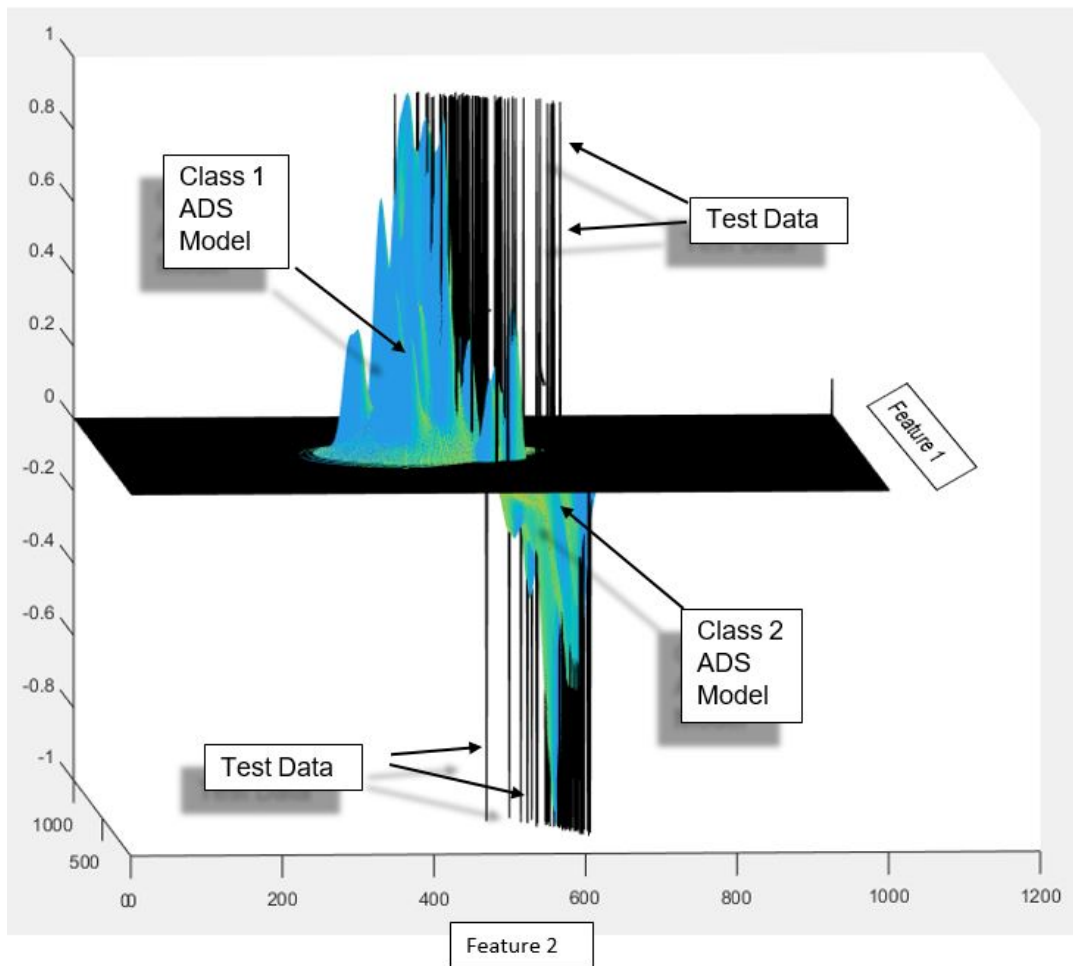


Figure 6.15: Trained motor-imagery adaptive decision surface with test data

6.8.6 Adapting the surface of the ADS

The ADS classifier has the option to adapt the surface during classification. After the acquisition of the input vector feature, the ADS will classify the unknown vector. The adaptation will occur if the surface reading is above a certain threshold for the particular class.

Algorithm 3 Adaptive Decision Surface (ADS) Adapter algorithm

```

1: A set of coefficients were estimated heuristically
   Class 1 threshold =  $th1$ 
   Class 2 threshold =  $th2$ 
2: Acquire the unknown feature coordinate
if AdaptiveDecisionsSurface > 0 then
   Classify as class 1
   if AdaptiveDecisionsSurface > threshold $th1$  then
     input feature coordinate
     Plot a Gaussian scaled by  $c1GS$  on the ADS ,Fig 6.13
if AdaptiveDecisionsSurface < 0 then
   Classify as class 2
   if AdaptiveDecisionsSurface < threshold $th2$  then
     input feature coordinate
     Plot an inverted Gaussian scaled by  $c2GS$  on the ADS ,Fig 6.13

```

6.9 Summary

6.9.1 Results from the motor imagery adaptive decision surface MI ADS

The classification accuracy in the offline tests reached 76.06 % class 1 and 81.50 % using a non adaptive ADS and 79.55 % class 1 and 99.69 % using an adaptive ADS classifiers.

	Classifier	Class 1	Class 2
1	SVM Coarse Gaussian	83.49	77.16
2	Subspace discriminant	82.41	77.62
3	SVM Medium Gaussian	81.94	77.78
4	SVM Cubic	81.17	77.78
5	SVM Linear	80.86	79.17
6	SVM Quadratic	80.56	79.32
7	SVM Fine Gaussian	80.55	74.69
8	KNN Weighted	79.78	76.23
9	Linear Discriminant (LDA)	79.63	79.01
10	Quadratic Discriminant(QDA)	79.63	79.47
11	ADS Adaptive v3	79.55	99.69
12	Logistic Regression	79.17	79.94
13	KNN Cubic	78.24	83.95
14	KNN Coarse	78.08	78.55
15	KNN Medium	77.93	84.26
16	KNN Fine	77.62	72.84
17	KNN Cosine	76.85	81.95
18	Ensemble Bagged Tree	76.70	77.78
19	Decision Tree Coarse	76.24	77.78
20	ADS None adaptive	76.07	81.50
21	Decision Tree Simple	75.92	80.71
22	Decision Tree Medium	75.77	78.55
23	Subspace KNN	67.44	76.08
24	Ensemble Boosted Tree	59.41	64.97
25	RUSBoosted Trees	58.49	67.44

Figure 6.16: BCI competition IV dataset 2a from Graz University of Technology

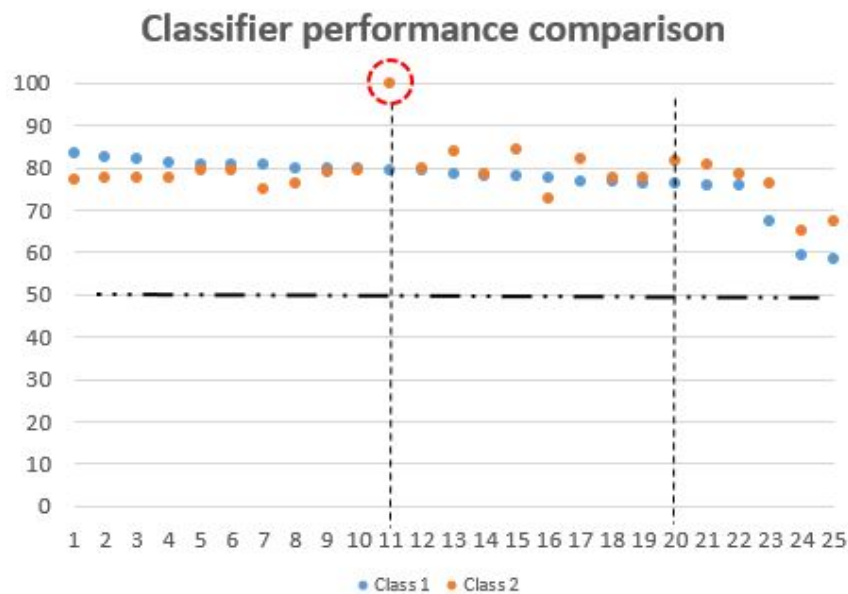


Figure 6.17: BCI competition IV dataset 2a from Graz University of Technology

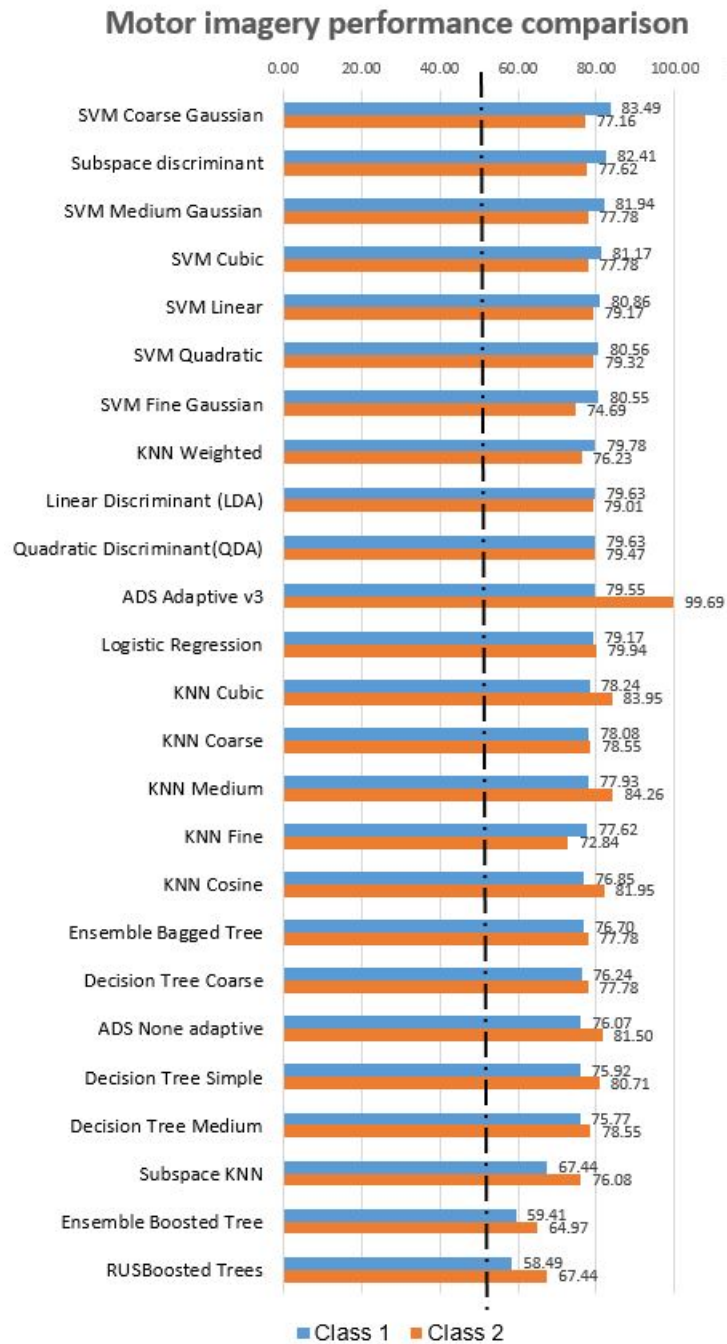


Figure 6.18: 25 Classifier MI performance comparison BCI competition IV dataset 2a from Graz University of Technology

6.10 Discussion

The chapter showed a probability-based classifier for improving motor imagery based BCI performance compared with other classifiers. The classification accuracy in the offline tests reached 76.06 % class 1 and 81.50 % class 2 using a non-adaptive ADS and 79.55 % class 1 and 99.69 % class 2 using an adaptive ADS classifier. Finally, a direction for future development an artificial intelligent controller and this classifier is shown in Fig 6.20. This AI controller could decide to combine proximity sensor data to augment certain features to close the prosthetic hand on a nearby object Fig 6.19. Other inputs such as electromyogram could be part of a multi-modal input used by an AI controller. In addition AI control may update the trained ADS classifier in the event of error correction during usage.

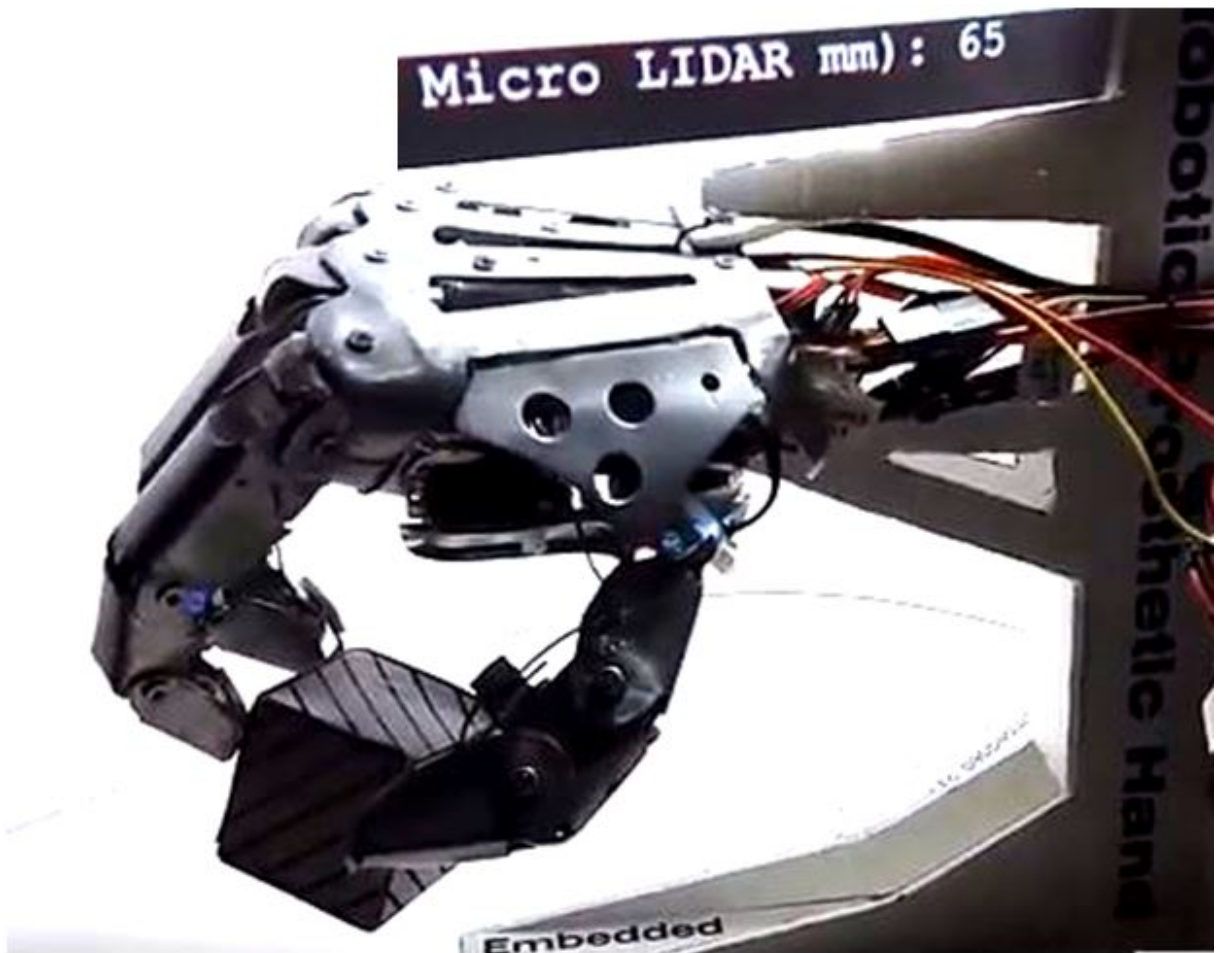


Figure 6.19: Prosthetic hand with a LIDAR sensor.

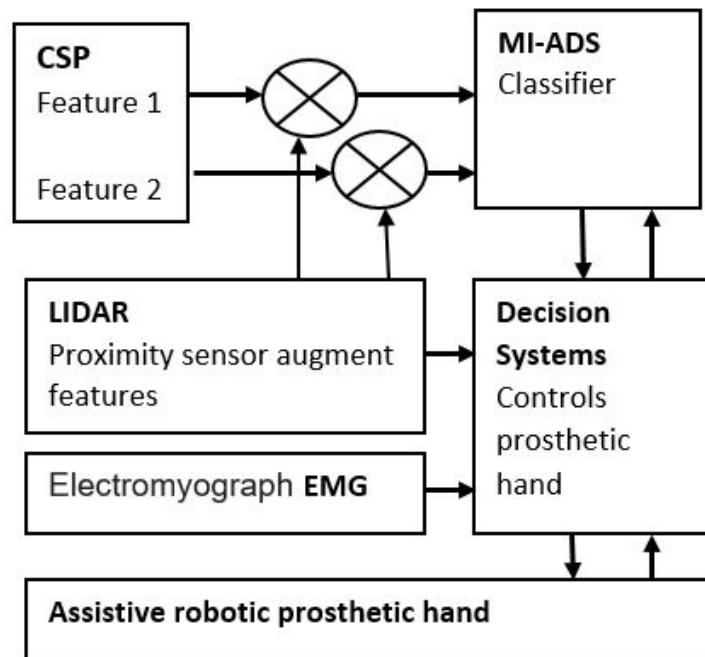


Figure 6.20: BCI intelligent controller system.

6.11 MI ADS Classifier Results

Person ID	Class 1 Results [%]	Class 2 Results [%]	Class 1 Results [%]	Class 2 Results [%]	Class 1 Results [%]	Class 2 Results [%]	Class 1 Results [%]	Class 2 Results [%]
1	69.44	98.61	77.78	98.61	79.17	98.61	69.44	98.61
2	61.11	62.5	69.44	48.61	63.89	48.61	63.89	61.11
3	100	97.22	100	90.28	100	90.28	100	97.22
4	70.83	77.78	81.94	66.67	83.33	62.5	69.44	79.17
5	81.94	68.06	84.72	63.89	84.72	62.5	81.94	68.06
6	51.39	79.17	73.61	62.5	66.67	70.83	52.78	79.17
7	76.39	90.28	73.61	86.11	75	90.28	76.39	88.89
8	93.06	100	93.06	100	90.28	100	93.06	100
9	97.22	84.72	97.22	77.78	98.61	75	97.22	83.33
Average	77.93	84.26	83.49	77.16	82.41	77.62	78.24	83.95
KNN Medium MATLAB® Ver.(2018a) Feature selection method CSP		SVM Coarse Gaussian MATLAB® Ver.(2018a) Feature selection method CSP		Subspace discriminant MATLAB® Ver.(2018a) Feature selection method CSP		KNN Cubic MATLAB® Ver.(2018a) Feature selection method CSP		

Figure 6.21: ADS Test data

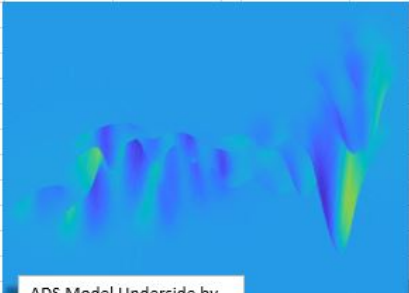
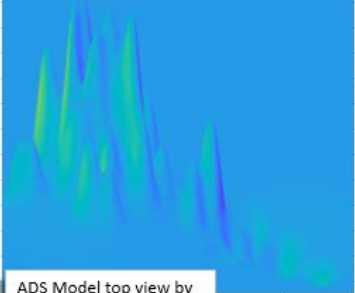
Class 1 Results [%]	Class 2 Results [%]	Class 1 Results [%]	Class 2 Results [%]	Class 1 Results [%]	Class 2 Results [%]	Class 1 Results [%]	Class 2 Results [%]	Person ID
95.83	73.61	95.83	100	95.83	100	98.61	100	1
50.72	73.61	50.72	93.06	52.17	94.44	52.17	100	2
91.55	97.22	91.55	100	90.14	100	90.14	100	3
67.14	81.43	67.14	100	67.14	100	68.57	100	4
60	85.71	61.43	100	58.57	100	62.87	100	5
62.86	54.17	68.57	97.22	78.57	100	78.57	100	6
84.72	77.46	86.11	100	88.89	100	86.11	100	7
100	91.67	100	98.61	100	98.61	100	98.61	8
71.83	98.61	71.83	98.61	78.87	98.61	78.87	98.61	9
76.07	81.50	77.02	98.61	78.91	99.07	79.55	99.69	
ADS None adpative Feature selection method CSP 26/07/2018 by Adrian Ashley	ADS Adaptive Feature selection method CSP 26/07/2018 by Adrian Ashley	ADS Adaptive Feature selection method CSP 30/07/2018 by Adrian Ashley	ADS Adaptive Feature selection method CSP 31/07/2018 by Adrian Ashley					
								
ADS Model Underside by A Ashley 31/07/2018				ADS Model top view by A Ashley 31/07/2018				

Figure 6.22: ADS Test data

Chapter 7

Shared control image decoding and tactile LiDAR of kitchen utensils for an assistive prosthetic hand

This chapter describes image decoding for object identification, position and orientation based on the discovery made by David Hubel and Torsten Wiesel who research the primary visual cortex and receiving the Nobel prize for their work. This work is distinguished by three key contributions. The first is image feature selection called synthesise image decoding layers which enables a stylus pattern to select features with a predefined threshold for the image. The second is a vertical stylus pattern to select vertically orientated features enabling the functional process of orientation cells inspired by the contribution of []. Third signals from LiDAR sensors to enable the identification and verification of the utensil such the visual system is presented with a real physical three-dimensional object.

7.1 Introduction

Disability can occur to any person at any stage in life. Debilitating conditions spinal cord injury, brain-stem strokes, accidents or illness resulting in upper limb amputation will inhibit a person from interacting with their environment. Brain-computer interface (BCI) enables a direct interaction of the human brain with the external environment. Moreover an assistive robotic device with the capability to interact with the persons environment by recognising utensils using image processing and LiDAR Fig 7.4

Furthermore reduce the workload of manipulating the objects on behalf of the person with a disability will empower and support the people in the future.

Many research studies have investigated BCI with assistive robotics presented in the following survey [7]. Microsoft[®] Kinect[®] has the ability to acquire data and track the motion of objects and has been applied to many applications based on user physical interactions [121].

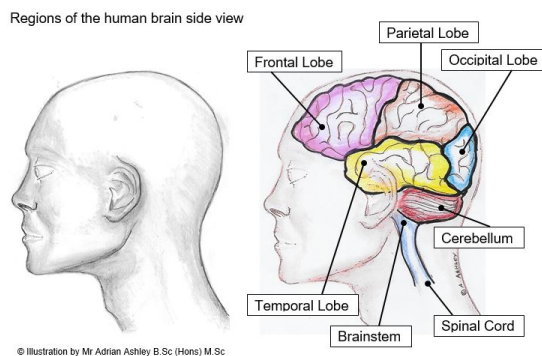


Figure 7.1: Regions of the human brain

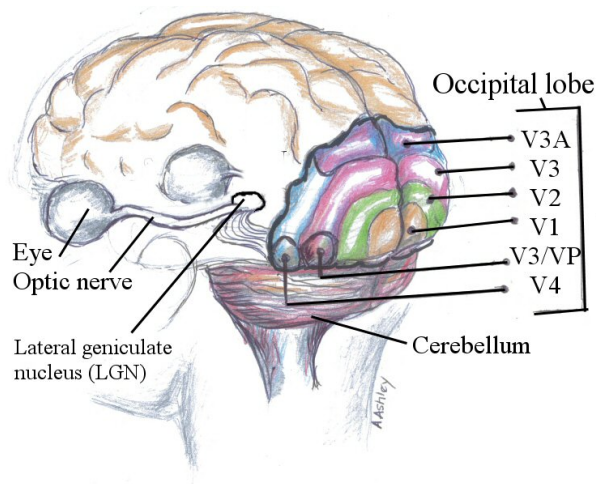


Figure 7.2: Regions of the human brain

Object detection and classification is well established moreover convolution neural network has the ability to recognise objects in images and has been used in many research fields and in industrial patents [122] in [123]. Matrix pattern recognition, decision making and adaptive learning process has also been applied to identify objects and dynamically update its own model concepts are updated with previously unknown input.

The stages of visual processing has been researched for many years. Visual processing consists of the following: Light entering the eye through the lens is focus on the back of the eye on the retina. Visible light is part of the electromagnetic spectrum and generally the human eye can detects wavelengths between 480 to 740 nanometres. Photoreceptor cells in the retina convert visible electromagnet radiation into signals such as action potentials for the biological system. Action potential signals go down axons of the optic nerve to the lateral genicular nucleolus (LGN) of the thalamus. The axons from the thalamus are connected to different areas of the primary visual cortex which are known as the occipital lobe of the brain the primary visual cortex is comprising of different regions known as v1 primary input , v2, v3, v4 Fig 7.2 collectively known as the functional subdivisions of the visual cortex, Hubel Wiesel made a discovery of cells in the visual cortex is sensitive to edges and the orientation of a visual stimulation shown in [21]

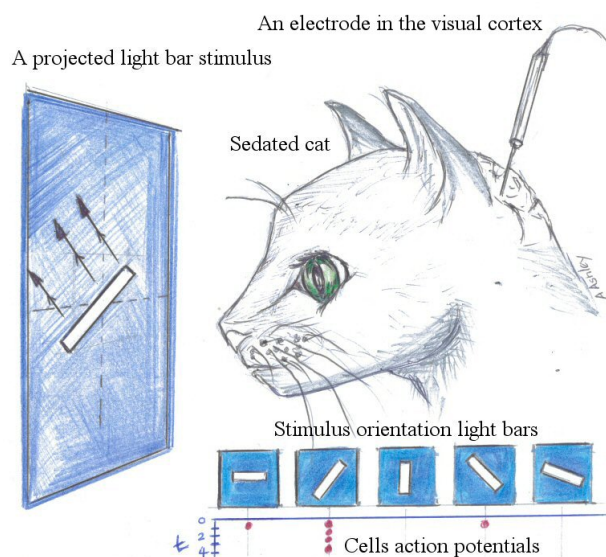


Figure 7.3: Regions of the human brain, Hubel Wiesel [21]

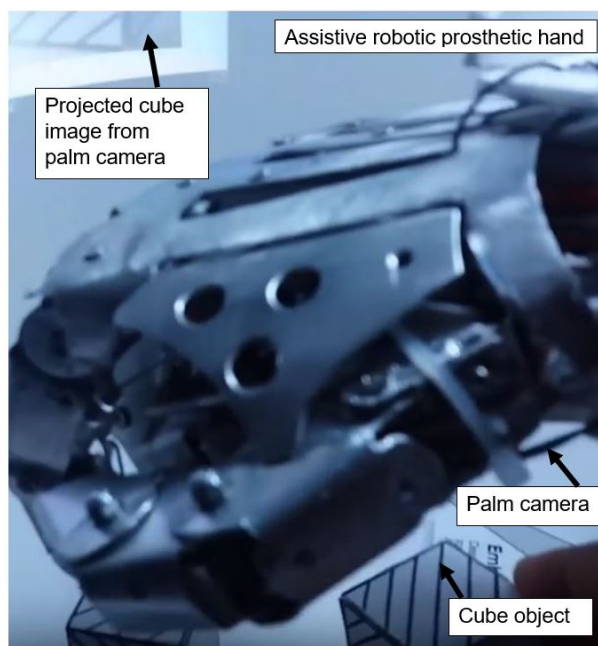


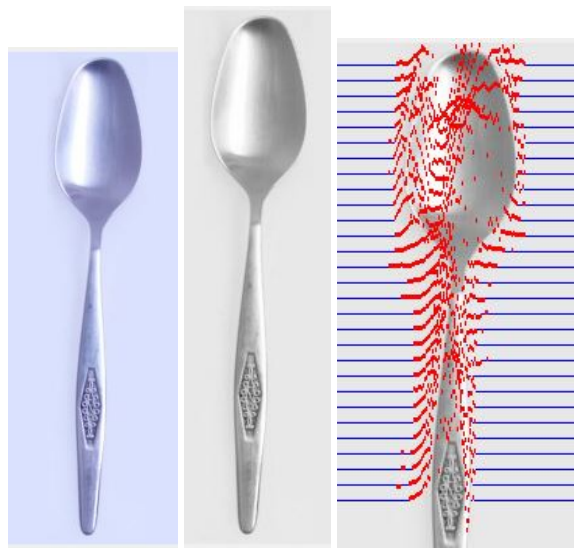
Figure 7.4: Assistive prosthetic hand palm camera

The remainder of this paper is organized as follows. Section II describes the method. Section III discussion Section IV. finally Section V concludes this paper

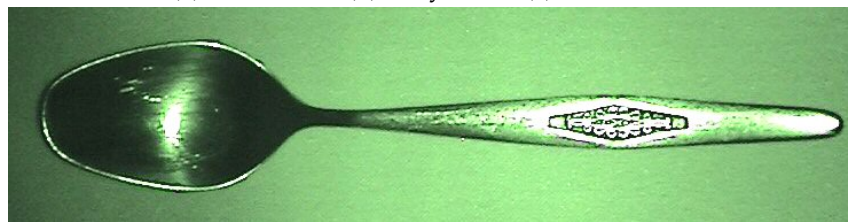
7.2 Methodology

7.2.1 Initial image processing stage

Initial image processing will remove the shadows to reduce the detrimental effect on subsequent image processing stages, such as determine the boundary of the object. The kitchen utensil is below the prosthetic hands camera to identify the object and ascertain the orientation relative to the hand. Fig 7.6 (a). The range of the grey scale is adjusted and a threshold determined Fig 7.6 (b). Contrast adjustment removed shadows in the image sample Fig 7.6 (c). A basic image segmentation by setting a threshold between a certain range in comparison to the background Fig 7.6 (d)



(a) Colour (b) Grey (c) Shadows

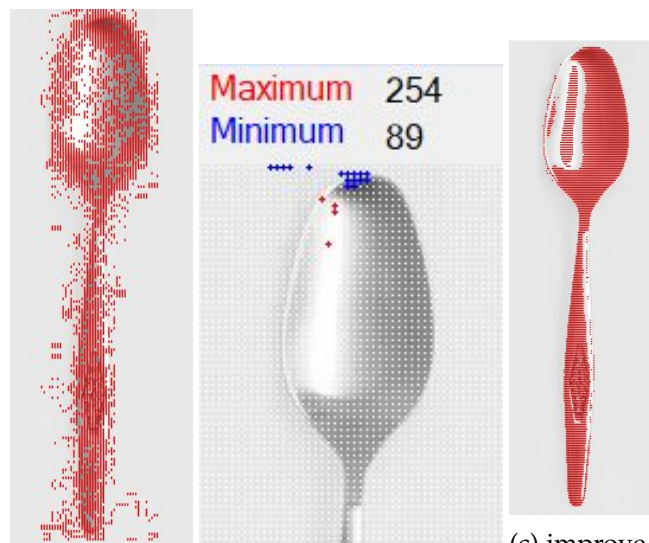


(d) Infra-red Image



(e) Infra-red manually remove reflections

Figure 7.5



(a) interference (b) Image analysis (c) improvement

Figure 7.6

Algorithm 4 *Adjust the image contrast phase 1*

Bitmap Picture = *BitMap*
Boolean switch = *sw1*
pixel x position = *x*
pixel y position = *y*
1: A set of coefficients were estimated heuristically
contrast Threshold = *Theshold*
2: Adjust the contrast horizontally from left to right.
for *y* = 1 : *BitMap.Height*-1 increment *y* by 1 **do**
 sw1 = true
 for *x* = 1 : *BitMap.Width* increment *x* by 2 **do**
 BitMap get pixel value $P_0(x-1, y)$
 BitMap get pixel value $P_1(x, y)$
 $n = P_0 - P_1$
 if ($n < 0$) **then**
 $n = n$ multiplied by (-1)
 if ($n > Theshold$) **then**
 if (*sw1* = true) **then**
 BitMap set the pixel(*x*, *y*) P_1 to white
 sw1 = false
 else
 if (*sw1* = true) **then**
 BitMap set pixel(*x* - 1, *y*) P_0 white
 BitMap set pixel(*x*, *y*) P_1 white
2: Adjust the contrast horizontally from right to left.
for *y* = 1 : *BitMap.Height*-1 increment *y* by 1 **do**
 sw1 = true
 for *x* = *BitMap.Width* - 1 : 1 decrement *x* by (-2) **do**
 BitMap get pixel value $P_0(x-1, y)$
 BitMap get pixel value $P_1(x, y)$
 $n = P_0 - P_1$
 if ($n < 0$) **then**
 $n = n$ multiplied by (-1)
 if ($n > Theshold$) **then**
 if (*sw1* = true) **then**
 BitMap set the pixel(*x*, *y*) P_1 to white
 sw1 = false
 else
 if (*sw1* = true) **then**
 BitMap set pixel(*x* - 1, *y*) P_0 white
 BitMap set pixel(*x*, *y*) P_1 white

$P_1(x-1, y-1)$	$P_2(x, y-1)$	$P_3(x+1, y-1)$
$P_4(x-1, y)$	$P_5(x, y)$	$P_6(x+1, y)$
$P_7(x-1, y+1)$	$P_8(x, y+1)$	$P_9(x+1, y+1)$

Figure 7.7: Patch 3x3 array

Algorithm 5 *Adjust the image contrast by scaling phase 2*

```
Bitmap picture = BitMap
Bitmap picture Height = BMapH
Bitmap picture Width = BMapW
Maximum contrast value = Cmax set to 0
Minimum contrast value = Cmin set to 255
Span contrast = Cspan
Scale contrast = Cscale
New contrast value = vn
Evaluation patch size = Ps set to 3 pixels
1: Find the average patch maximum and minimum.
for y = ps : BMapH - Ps increment y by Ps do
  for x = ps : BMapW - Ps increment x by Ps do
    BitMap get pixel value P1(x - 1, y - 1)
    BitMap get pixel value P2(x, y - 1)
    BitMap get pixel value P3(x + 1, y - 1)
    BitMap get pixel value P4(x - 1, y)
    BitMap get pixel value P5(x, y);
    BitMap get pixel value P6(x + 1, y)
    BitMap get pixel value P7(x - 1, y + 1)
    BitMap get pixel value P8(x, y + 1)
    BitMap get pixel value P9(x + 1, y + 1)
    patch average = (P1 + P2 + P3 + P4 + P5 + P6 + P7 + P8 + P9) / 9
    if (Cmax < patch average) then
      Cmax = patch average
    if (Cmin > patch average) then
      Cmin = patch average
2: Calculate the span and scale.
Cspan = Cmax - Cmin
Cscale = 255 / Cspan
3: Adjust the contrast.
for y = ps : BMapH increment y by 1 do
  for x = ps : BMapW increment x by 1 do
    BitMap get pixel value P1(x, y);
    P1 = P1 - Cmin
    vn = P1 multiply by Cscale
    if ((vn > 0) AND (vn < 255)) then
      BitMap set pixel value P1(x, y) = vn
```



(a) Grey image before enhancement.



(b) After contrast enhancement.

Figure 7.8: Contrast adjustment

Algorithm 6 Image feature selection by orientation lines.

Bitmap picture = *BitMap*
Bitmap picture Height = $BMap_H$
Bitmap picture Width = $BMap_W$
pixel x position = x
pixel y position = y
Bitmap picture L1 = $BitMap_{L1}$
Bitmap picture L2 = $BitMap_{L2}$
Bitmap picture L3 = $BitMap_{L3}$
Bitmap picture L4 = $BitMap_{L4}$
Bitmap picture Ls = $BitMap_{Ls}$
Bitmap picture SynthNet = $BitMap_{SynthNet}$
Boolean switch $sw1$ = false
Boolean switch $sw2$ = false
Boolean switch $sw3$ = false
Boolean switch $sw4$ = false
Evaluation patch size = P_s set to 3 pixels
Upper limit threshold P_{5u}
Lower limit threshold P_{5l}
1: Find the pixel values in the 3x3 patch.
for $y = 1 : BMap_H - P_s$ increment y by 1 **do**
 for $x = 1 : BMap_W - P_s$ increment x by 1 **do**
 $BitMap$ get pixel value $P_1(x - 1, y - 1)$
 $BitMap$ get pixel value $P_2(x, y - 1)$
 $BitMap$ get pixel value $P_3(x + 1, y - 1)$
 $BitMap$ get pixel value $P_4(x - 1, y)$
 $BitMap$ get pixel value $P_5(x, y)$;
 $BitMap$ get pixel value $P_6(x + 1, y)$
 $BitMap$ get pixel value $P_7(x - 1, y + 1)$
 $BitMap$ get pixel value $P_8(x, y + 1)$
 $BitMap$ get pixel value $P_9(x + 1, y + 1)$
2: Thresholds for feature detection.
 $P_{5u} = P_5 + (P_5 \times 0.3)$
 $P_{5l} = P_5 - (P_5 \times 0.3)$
3: Horizontal evaluation
 if $((P_4 <> P_5) \text{ AND } (P_5 <> P_6))$ **then**
 if $((P_4 < P_{5u}) \text{ AND } (P_4 > P_{5l})) \text{ AND } ((P_6 < P_{5u}) \text{ AND } (P_6 > P_{5l}))$ **then**
 $BitMap_{L1}$ Set Pixel($x - 1, y$) P_4 Black
 $BitMap_{L1}$ Set Pixel(x, y) P_5 Black
 $BitMap_{L1}$ Set Pixel($x + 1, y$) P_6 Black
 $sw_{L1} = \text{true}$
 $n1 = (n1 + 1)$
4: Vertical evaluation
 if $((P_2 <> P_5) \text{ AND } (P_5 <> P_8))$ **then**
 if $((P_2 < P_{5u}) \text{ AND } (P_2 > P_{5l})) \text{ AND } ((P_8 < P_{5u}) \text{ AND } (P_8 > P_{5l}))$ **then**
 $BitMap_{L2}$ Set Pixel($x, y - 1$) P_2 Black
 $BitMap_{L2}$ Set Pixel(x, y) P_5 Black
 $BitMap_{L2}$ Set Pixel($x, y + 1$) P_8 Black
 $sw_{L2} = \text{true}$
 $n2 = (n2 + 1)$

5: Cross top right, bottom left

if $((P_3 <> P_5) \text{ AND } (P_5 <> P_7))$ **then**
 if $((P_3 < P_{5u}) \text{ AND } (P_3 > P_{5l})) \text{ AND } ((P_7 < P_{5u}) \text{ AND } (P_7 > P_{5l}))$ **then**
 BitMap_{L3} Set Pixel(x+1, y-1) *P₃ Black*
 BitMap_{L3} Set Pixel(x, y) *P₅ Black*
 BitMap_{L3} Set Pixel(x-1, y+1) *P₇ Black*
 sw_{L3} = true
 n3 = (n3 + 1)

6: Cross top left, bottom right

if $((P_1 <> P_5) \text{ AND } (P_5 <> P_9))$ **then**
 if $((P_1 < P_{5u}) \text{ AND } (P_1 > P_{5l})) \text{ AND } ((P_9 < P_{5u}) \text{ AND } (P_9 > P_{5l}))$ **then**
 BitMap_{L4} Set Pixel(x-1, y-1) *P₁ Black*
 BitMap_{L4} Set Pixel(x, y) *P₅ Black*
 BitMap_{L4} Set Pixel(x+1, y+1) *P₉ Black*
 sw_{L4} = true
 n4 = (n4 + 1)

if (*sw₁* = true) **then**
 BitMap_{LS} Set Pixel(x, y) *Black*

if (*sw₂* = true) **then**
 BitMap_{LS} Set Pixel(x, y) *Black*

if (*sw₃* = true) **then**
 BitMap_{LS} Set Pixel(x, y) *Black*

if (*sw₄* = true) **then**
 BitMap_{LS} Set Pixel(x, y) *Black*

sw₁ = false
sw₂ = false
sw₃ = false
sw₄ = false

7: Evaluate bitmap *BitMap_{LS}* 3x3 patch.

for $y = 1 : BMap_H - P_s$ increment y by P_s **do**
 for $x = 1 : BMap_W - P_s$ increment x by P_s **do**
 BitMap_{LS} get pixel value $P_1(x - 1, y - 1)$
 BitMap_{LS} get pixel value $P_2(x, y - 1)$
 BitMap_{LS} get pixel value $P_3(x + 1, y - 1)$
 BitMap_{LS} get pixel value $P_4(x - 1, y)$
 BitMap_{LS} get pixel value $P_5(x, y)$
 BitMap_{LS} get pixel value $P_6(x + 1, y)$
 BitMap_{LS} get pixel value $P_7(x - 1, y + 1)$
 BitMap_{LS} get pixel value $P_8(x, y + 1)$
 BitMap_{LS} get pixel value $P_9(x + 1, y + 1)$
 if $((p_1 = 255) \text{ AND } (p_5 = 255))$ **then**
 c1 = c1 + 1
 BitMap_{SynthNet} Set Pixel(x, y) *Black*
 if $((p_2 = 255) \text{ AND } (p_5 = 255))$ **then**
 c2 = c2 + 1
 BitMap_{SynthNet} Set Pixel(x, y) *Black*
 if $((p_3 = 255) \text{ AND } (p_5 = 255))$ **then**
 c2 = c2 + 1
 BitMap_{SynthNet} Set Pixel(x, y) *Black*
 if $((p_4 = 255) \text{ AND } (p_5 = 255))$ **then**
 c4 = c4 + 1
 BitMap_{SynthNet} Set Pixel(x, y) *Black*

```

if ((p5 = 255) AND (p5 = 255)) then
    c5 = c5 + 1
    BitMapSythNet Set Pixel(x, y) Reference
if ((p6 = 255) AND (p5 = 255)) then
    c6 = c6 + 1
    BitMapSythNet Set Pixel(x, y) Black
if ((p7 = 255) AND (p5 = 255)) then
    c7 = c7 + 1
    BitMapSythNet Set Pixel(x, y) Black
if ((p8 = 255) AND (p5 = 255)) then
    c8 = c8 + 1
    BitMapSythNet Set Pixel(x, y) Black
if ((p9 = 255) AND (p5 = 255)) then
    c9 = c9 + 1
    BitMapSythNet Set Pixel(x, y) Black

```

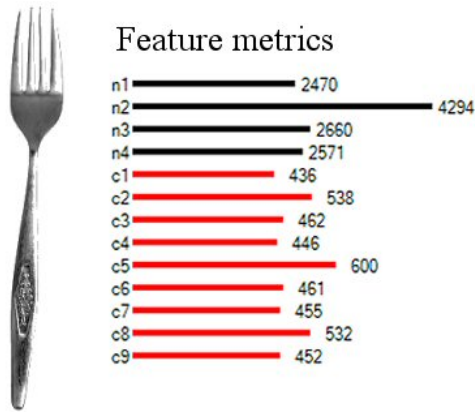


Figure 7.9: fork feature metrics

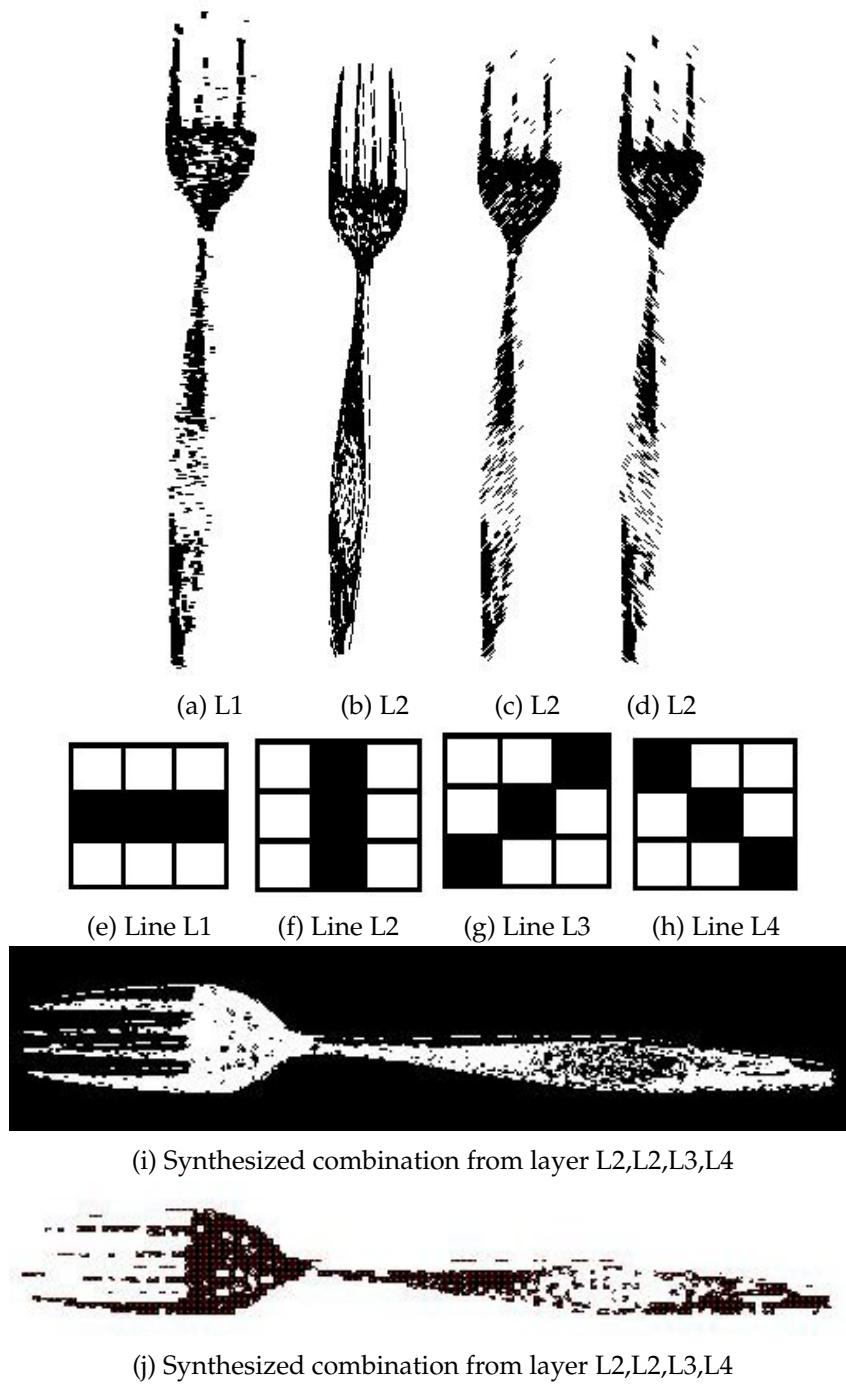


Figure 7.10: Filter orientation line processing

Algorithm 7 *Image processing*

```
Bitmap picture = BitMapsynth
Bitmap picture Height = BMapH
Bitmap picture Width = BMapW
pixel x position = x
pixel y position = y
Patch size =  $P_s$ 
1: Horizontal scan
  n = 0
  last0 = 0
  for  $y = 1 : \text{BitMap}_{synth}.\text{Height} - P_s$  increment  $y$  by 1 do
    sw1 = false
    for  $x = 1 : \text{BitMap}_{synth}.\text{Width} - P_s$  increment  $x$  by 1 do
      BitMapsynth get pixel value  $V_{scan}(x,y)$ 
      if ( $V_{scan} = 255$ ) then
        n = n + 1
        sw1 = true
        lastx =  $x$ 
      else
        if (sw1 = true) then
           $n_{mid} = (n / 2)$ 
          if ( $(n_{mid} > 1)$  AND  $((last_x - n_{mid}) > 1)$ ) then
            BitMapsynth set pixel  $P_{synth}((last_x - n_{mid}), y) = \text{Black}$ 
          n = 0
          sw1 = false
1: Vertical scan
  n = 0
  lasty = 0
  for  $x = 1 : \text{BitMap}_{synth}.\text{Height} - P_s$  increment  $x$  by 1 do
    sw1 = false
    for  $y = 1 : \text{BitMap}_{synth}.\text{Width} - P_s$  increment  $y$  by 1 do
      BitMapsynth get pixel value  $V_{scan}(x,y)$ 
      if ( $V_{scan} = 255$ ) then
        n = n + 1
        sw1 = true
        lasty =  $y$ 
      else
        if (sw1 = true) then
           $n_{mid} = (n / 2)$ 
          if ( $(n_{mid} > 1)$  AND  $((last_y - n_{mid}) > 1)$ ) then
            BitMapsynth set pixel  $P_{synth}(x, (last_y - n_{mid})) = \text{Black}$ 
          n = 0
          sw1 = false
1: Vertical cells orientation lines indicator
  for  $x = 1 : \text{pnlProcess}.\text{Height}$  increment  $x$  by 2 do
    Draw vertical lines to represent the orientation cells
```

2: Vertical orientation cells. Fig 7.11

$n = 0$

$last_0 = 0$

Vertical orientation cells activity = $n_{ln}[x]$

for $y = 1 : BitMap_{synth}.Width - 10$ inc.. y by 4 **do**

for $x = 1 : BitMap_{synth}.Height - 2$ inc.. x by 1 **do**

$BitMap_{synth}$ get pixel value $C_0(x, y)$

$BitMap_{synth}$ get pixel value $C_1(x, y + 1)$

$BitMap_{synth}$ get pixel value $C_2(x, y + 2)$

$BitMap_{synth}$ get pixel value $C_3(x, y + 3)$

$BitMap_{synth}$ get pixel value $C_4(x + 1, y)$

$BitMap_{synth}$ get pixel value $C_5(x + 1, y + 1)$

$BitMap_{synth}$ get pixel value $C_6(x + 1, y + 2)$

$BitMap_{synth}$ get pixel value $C_7(x + 1, y + 3)$

if $((C_0 = 0) \text{ OR } (C_4 = 0)) \text{ AND } ((C_1 = 0) \text{ OR } (C_5 = 0)) \text{ AND } ((C_2 = 0) \text{ OR } (C_6 = 0))$

AND $((C_3 = 0) \text{ OR } (C_7 = 0))$ **then**

 3: Indicator Marker

if $((x > last_x - 3) \text{ AND } (x < last_x + 3))$ **then**

 Indicate vertical orientation cell firing

 Activity $n_{lv[x]} = n_{lv[x]} + 1$

$last_x = x$

Find the position of the maximum firing cells

Maximum vertical cell quantity = $max_{lvqty} = 0$

Location maximum firing cells = $mx_{Locate} = 0$

for $x = 1 : BitMap_{synth}.Width$ increment x by 1 **do**

if $(max_{lvqty} < last_x - 3)$ **then**

 maximum value $max_{lvqty} = n_{ln}[x]$

 location $mx_{Locate} = x$

$C_0(x, y)$	$C_4(x+1, y)$
$C_1(x, y+1)$	$C_5(x+1, y+1)$
$C_2(x, y+2)$	$C_6(x+1, y+2)$
$C_3(x, y+3)$	$C_7(x+1, y+3)$

Figure 7.11: Vertical orientation cells

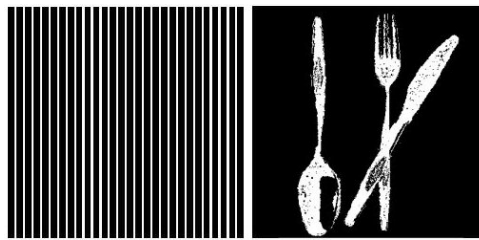


Figure 7.12: Vertical cells indicator

7.2.2 Tactile LiDAR signals

A signal profiles are acquired from the LIDAR sensor on the index finger tip Fig 7.13 transmitting the laser beam on to the object. The prosthetic hand will execute lateral movement such as longitudinally and laterally to acquire the surface profile properties of the kitchen utensil.

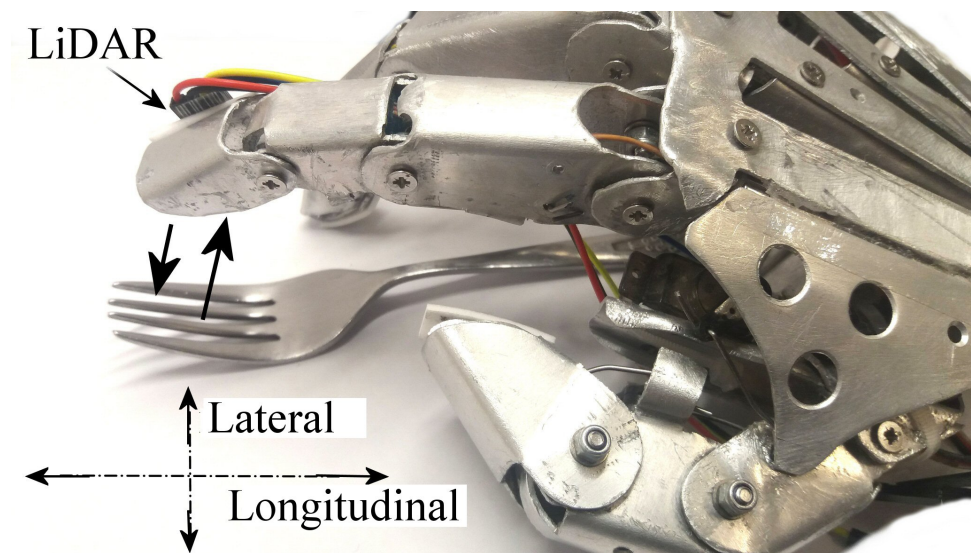


Figure 7.13: Tactile LIDAR from the finger tip

7.3 Summary

The assistive robotic device has the ability to retrieve an object for the disabled person. Moreover the prosthetic hand will be aware of how to handle the objects. Fig 7.14 shows a selection of kitchen utensils.

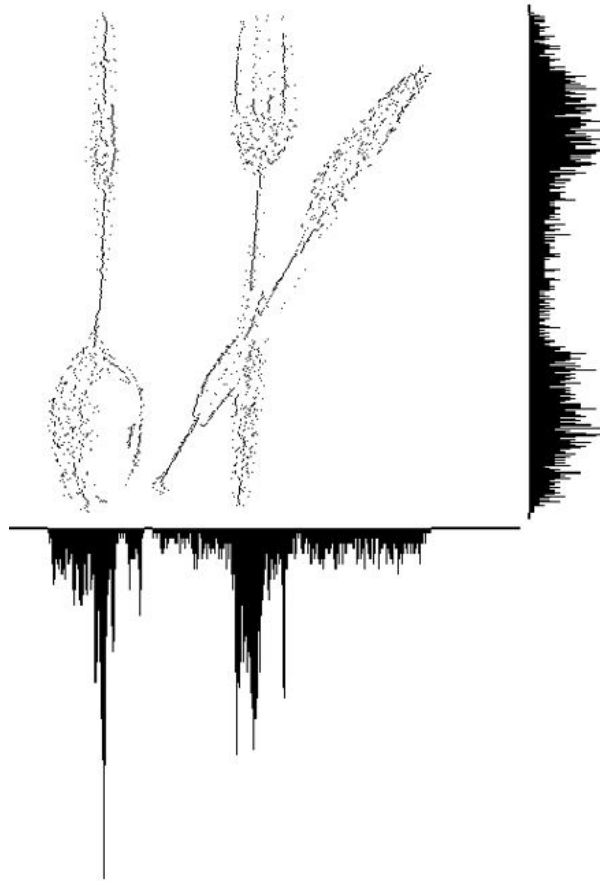


Figure 7.14: Cutlery



Figure 7.15: Vertical orientation cells firing with the cutlery image

7.3.1 Kitchen utensil tactile LiDAR scanning

A selection of kitchen utensils was scanned by a micro LiDAR. The signal could be used in the future development to support the identification of the utensil in cooperation with the visual recognition system to support the evidence that the object is an actual three-dimensional object and not just a picture of an object.

Fig 7.16 Shows the LiDAR output from spoon passing under the LiDAR three times it is characterised with twin peak link by a 'U' shape. Other characteristics may be obtained from other utensils.

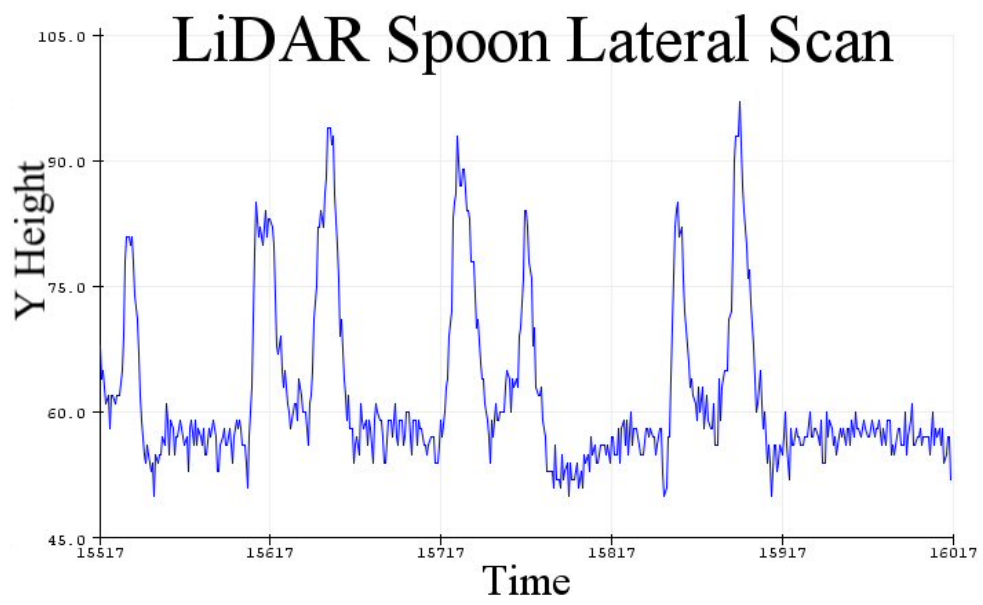


Figure 7.16: Tactile LiDAR from knife lateral scan

Chapter 8

Conclusions and Future Work

8.1 Conclusions

The aim of this research was to improve the classification accuracy with the minimum features from EEG data for error detection in the elderly and young participants using the established classifiers in addition to developing a novel adaptive decision surface classifier(ADS). The aims and objectives of this research work presented in this MPhil thesis were to improve the accuracy in terms of the following objectives:

- 1) Improving the accuracy of EEG-based error detection Please see 1.3.1
- 2) Improving the accuracy of motor imagery BCI. Please see 1.3.3
- 3) Shared control strategy such that the prosthetic robotic hand may recognise kitchen utensils to support the user interaction with less workload. Please see 1.3.4

To address error detection, we propose improving EEG-based error detection using the proposed novel feature extraction method. Thus in chapter 4, we propose a novel method called the relative peak features (RPF). Further, we combine the aforementioned proposed feature extraction method with a proposed novel classification algorithm in chapter 5 called the adaptive decision surface ADS Classifier.

To address the accuracy in motor imagery, we propose to use common spatial filters combined with the adaptive decision surface ADS Classifier. Spatial filters are used to reduce unnecessary spatial EEG electrical activity and highlight a particular location of interest. In addition, the spatial filter will maximize the signal to noise ratio such that accuracy of EEG-based communication will be improved shown in [20]. The classification process will benefit from the improved EEG signal with a more accurate classification. Please see chapter 6.6 a novel three dimensional probability-based classifier for improving motor imagery-based BCI.

Chapter 2 presented different types of BCI systems to determine the merits of each type. An investigation into the state of art literature is further provided by chapter 2 was highlighted by the report's aims and objectives 1.3 such that the literature review will identify any gaps in features extraction methods and error detection classification algorithms.

In chapter3, we evaluated the **deficiencies of state of the art error detection algorithms between young and elderly**. The process of evaluating the performance of detecting error-related potentials (ErrP) using an 'Linear Discriminant Analysis' LDA-based algorithm to detect the ErrP in EEG recordings has shown the present performance in young Fig 3.7 and elderly Fig 3.9 in terms ErrP detection. The results showed that there is a significant difference between the young and old participants. The implications are that new BCI systems could take this finding into account and hence improve their performance by taking into account the age

group of the user.

We proposed in chapter4 a novel features selection method for **Improving EEG-based error detection using proposed novel relative peak features (RPF)**. The table of classifiers Fig 8.1 below and in Fig 4.5 shows the relative peak feature performance presented on the right side of the bar chart below in Fig 8.2 and in Fig 4.6 to visually clarify the comparison between the different classifiers using the relative peak features, in addition to presenting the comparison between the different feature extraction methods. The relative peak features yielded an average across all classifiers of 81.63% accuracy in detecting the erroneous events and an average 78.87 % accuracy in detecting the correct events. In comparison to the relative peak feature selection method, there was an average increase in performance in all classifiers of 17.85% for error accuracy and a reduction of -6.16% for correct accuracy.

CLASSIFIER COMPARISON RESULTS USING TEMPORAL FEATURES AND RELATIVE PEAK FEATURES.

Classifiers	Classifiers KNN LDA SVM using temporal features		Classifiers KNN LDA SVM using relative peak features	
	Error Accuracy [%]	Correct Accuracy [%]	Error Accuracy [%]	Correct Accuracy [%]
1 KNN Coarse	64.98	88.16	80.85	78.83
2 LDA	69.06	79.97	82.68	78.10
3 SVM Coarse Gaussian	50.33	92.68	80.85	80.71
4 SVM Linear	70.75	79.28	82.16	77.86
5 Average results among all classifiers	63.78	85.02	81.63	78.87

Figure 8.1: Classifier comparison results using temporal features and relative peak features

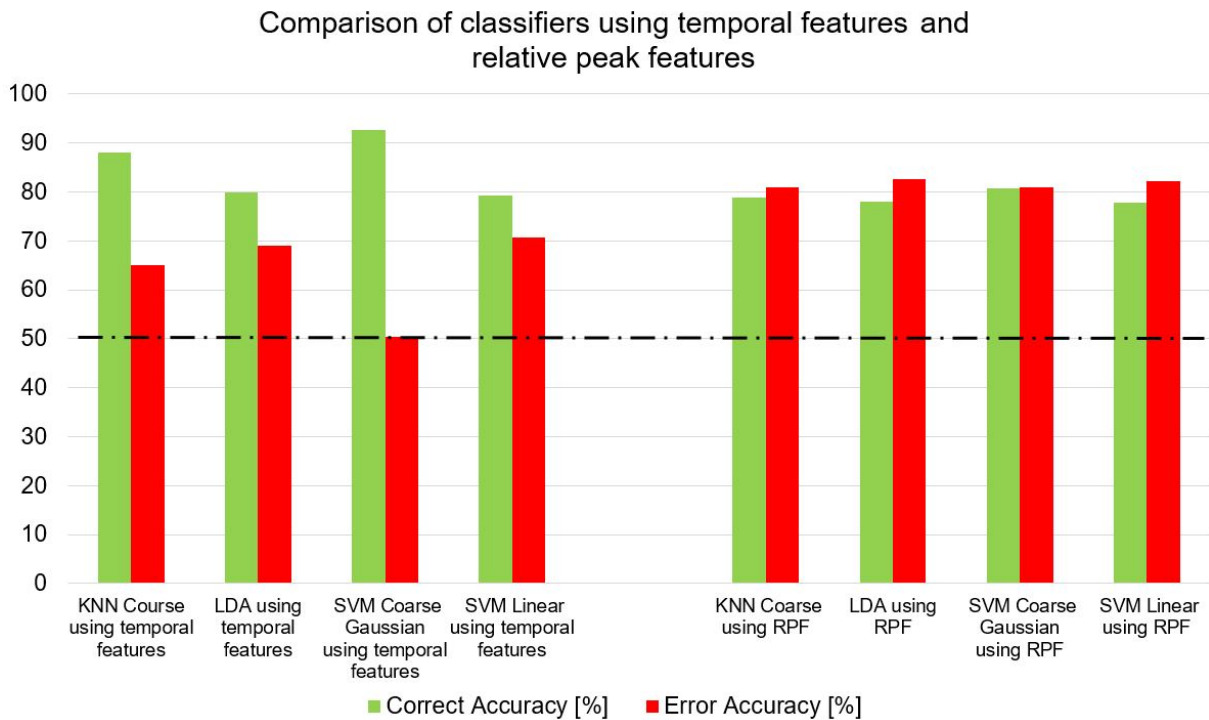


Figure 8.2: Classifier comparison results

This research is the second IEEE accepted conference paper publication with the title of **“Improving EEG-Based Error Detection Using Relative Peak Features”** 2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) in conjunction with the 43rd Annual Conference of the Canadian Medical and Biological Engineering Society Montréal, Canada, July 20-24, 2020

In chapter5 we proposed a novel error detection algorithm called an **adaptive decision surface (ADS) classifier**. The ADS classifier uses 25 times less training features. ADS used the relative peaks in chapter4 of the Cz electrode channel as a feature input. In terms of features quantities, Person 1 the ADS classifier used 606 Training features compared to LDA using 15,150 training features.

The average results for the classifiers for all the participants are presented in the following table.

Table 8.1: Young and elderly classifier results

Classifier	Participants	Error	Correct
		Accuracy [%]	Accuracy [%]
ADS	27 Young	81.85	84.17
	24 Elderly	63.68	82.63
LDA	27 Young	69.13	81.46
	24 Elderly	55.18	76.95

In chapter6.6 we proposed a novel motor combination of CSP and a motor imagery classifier algorithm called an **A novel three dimensional probability-based ADS classifier for improving motor imagery-based BCI**. The results from the motor imagery adaptive decision surface ADS classification accuracy in the off-line tests reached 76.06 % class 1 and 81.50 % using non adaptive ADS and 79.55 % class 1 and 99.69 % using an adaptive ADS classifiers. Please see the results in the following Fig 8.3, Fig 8.4, Fig 8.5.

	Classifier	Class 1	Class 2
1	SVM Coarse Gaussian	83.49	77.16
2	Subspace discriminant	82.41	77.62
3	SVM Medium Gaussian	81.94	77.78
4	SVM Cubic	81.17	77.78
5	SVM Linear	80.86	79.17
6	SVM Quadratic	80.56	79.32
7	SVM Fine Gaussian	80.55	74.69
8	KNN Weighted	79.78	76.23
9	Linear Discriminant (LDA)	79.63	79.01
10	Quadratic Discriminant(QDA)	79.63	79.47
11	ADS Adaptive v3	79.55	99.69
12	Logistic Regression	79.17	79.94
13	KNN Cubic	78.24	83.95
14	KNN Coarse	78.08	78.55
15	KNN Medium	77.93	84.26
16	KNN Fine	77.62	72.84
17	KNN Cosine	76.85	81.95
18	Ensemble Bagged Tree	76.70	77.78
19	Decision Tree Coarse	76.24	77.78
20	ADS None adaptive	76.07	81.50
21	Decision Tree Simple	75.92	80.71
22	Decision Tree Medium	75.77	78.55
23	Subspace KNN	67.44	76.08
24	Ensemble Boosted Tree	59.41	64.97
25	RUSBoosted Trees	58.49	67.44

Figure 8.3: BCI competition IV dataset 2a from Graz University of Technology

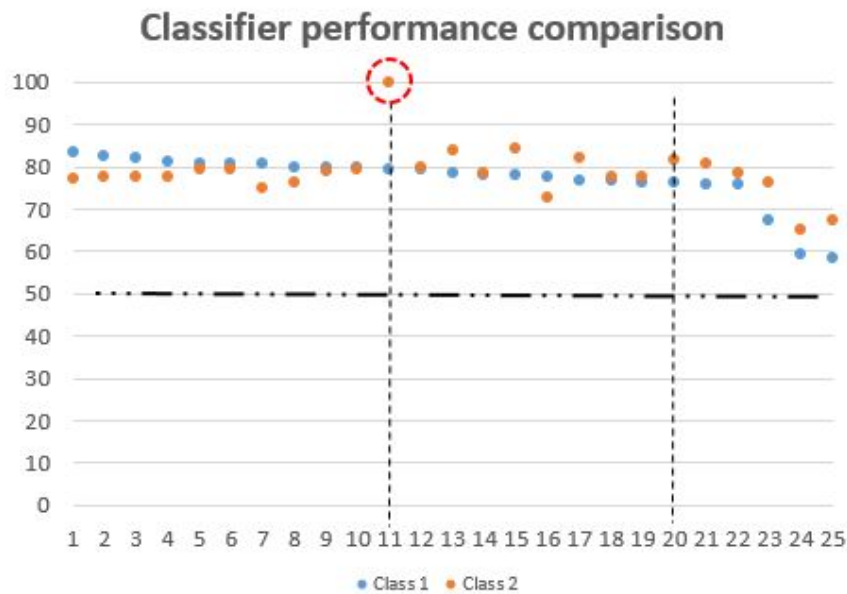


Figure 8.4: BCI competition IV dataset 2a from Graz University of Technology

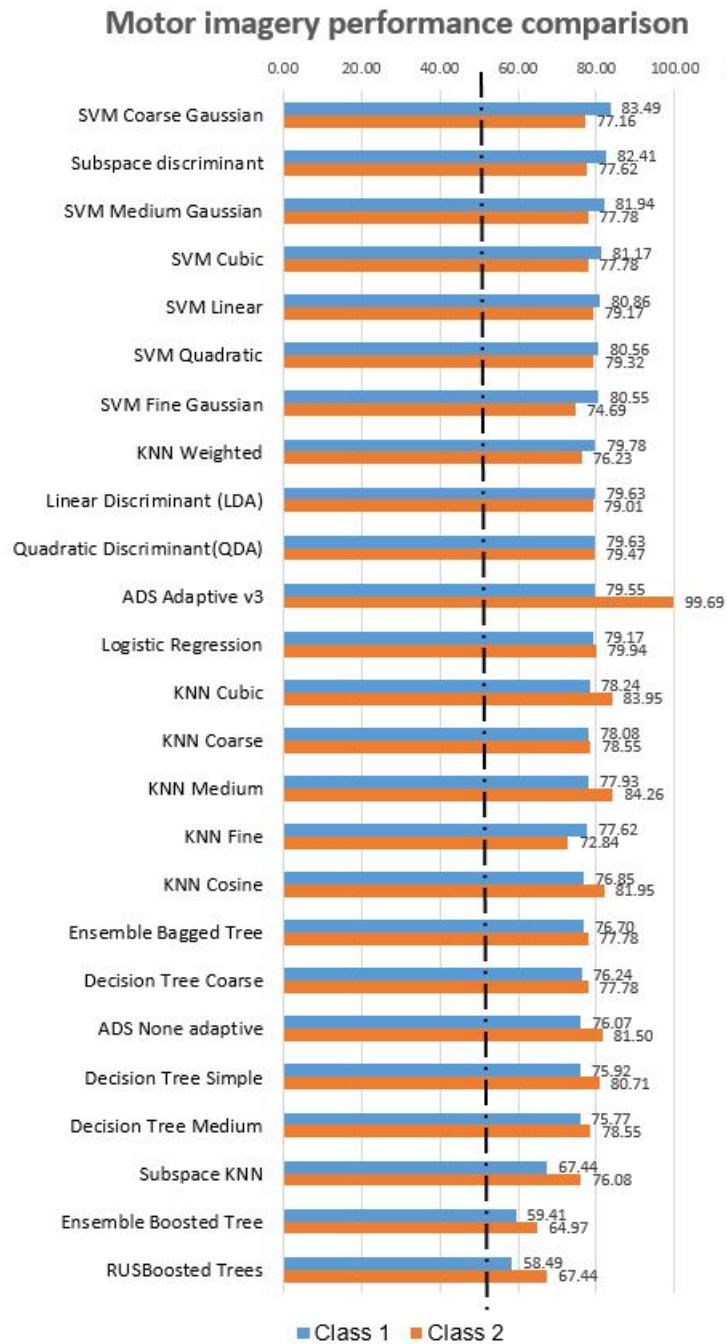


Figure 8.5: 25 Classifier MI performance comparison BCI competition IV dataset 2a from Graz University of Technology

This research is the first IEEE accepted conference paper publication with the title of “**A Novel Three Dimensional Probability-based Classifier for Improving Motor Imagery-based BCI**” published in: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) date of Conference: 12-17 May 2019 Date Added to IEEE Xplore: 17 April 2019 INSPEC accession Number: 18813183 DOI: 10.1109/ICASSP.2019.8683136

8.2 Future work

8.2.1 Future development the absorption adaptability function of the (ADS)

The adaptive decision surface(ADS) classifier could be improved by modifying the adaptability by using the ADS surface readings and depending on the threshold accommodate a wider distance from previously known data. Such that once the threshold is reached so that learning step is applied and the ADS surface is modified to accommodate the new ErrP data relative peak plot. This process of adaptation can be regarded as adsorption to a new state. The adaptive process has been in applied and tested resulting in a small percentage increase in performance however, there is a risk of having an out of control adaptation of the surface which may corrupt the integrity of the surface. This adaptability function would need to be carefully applied and tested after each adaptation for surface integrity.

8.2.2 Future development reverse adaptation function of the (ADS)

The adaptive decision surface(ADS) classifier could be further improved to record the adaptation process in section 8.2.1 for self-auditing and maintenance such that a reversal of the adaptation is possible if the subsequent adaptation was later found to be erroneous. There must be a controlled process of reversing the learning process in order to preserve and maintain the ADS. This could be an envisage function of the previously mentioned 'Adaptive Decision Surface Management', please see Section 8.2.5 .

8.2.3 Future development a computational inference reasoning CIR

In the situation where unknown data is acquired and, the system has no prior training data. Such that an 'Adaptive Decision Surface' ADS has no data, and the surface is flat with no prior values to support a decision system. An equivalent scenario is a blank neural network with all the weights set to zero. We propose additional functions to support the decision system either using the information from the 'likelihood bias', please see 5.2.3.2 or building a type of mathematical system model about the error-related potentials ErrPs dynamics in a peak plot. With the model, a full range of data tests can be evaluated so the outer limits can be explored, and some form of reasoning can be applied. However, developing a model and then applying every possible combination would eventually lead to a surface which would be the same as the aforementioned 'likelihood bias' which is every permutation of feasible location for an anticipated relative peak plot. The main objective is to remove the possibility of gaps or zero readings from the adaptive decision surface ADS such that the classification is unknown data as highlighted in a sparse adaptive decision surface SADS.

8.2.4 Future development of a method for prior computational inference

Develop a method for prior knowledge with computational inference using a model with every possible permutation or using a 'likelihood bias' would lead to the same surface. The classification of an unknown input should acquire a classification and not fall between the gaps however, there is no indication of certainty based on prior experience?

The advantage of using the 'Adaptive Decision Surface' ADS is having been exposed to a set of training data the surface is shaped and distorted according to the prior training data. Moreover, the decision system may indicate the certainty according to the prior training data, which shapes the surface.

8.2.5 Future work self-organising surface management

The proposed algorithm must incorporate a type of management process to maintain the surface of the 'Adaptive Decision Surface' ADS to remove the risk of zeros 'the indeterminate state' in terms of classification. Previously mentioned using the 'likelihood bias' please see Fig 5.12. would bias a particular region. However, the adaptive decision surface is subject to adaptation in the learning process, and changes during the initial application of the training data zero values may be on the surface due to the changes in the surface data. Please see Fig-8.6 for an example.

The risk of a zero on the surface would mean this is an ambiguous or indeterminate state in terms of classification. The decision system would have to determine the classification of the data by an alternative method.

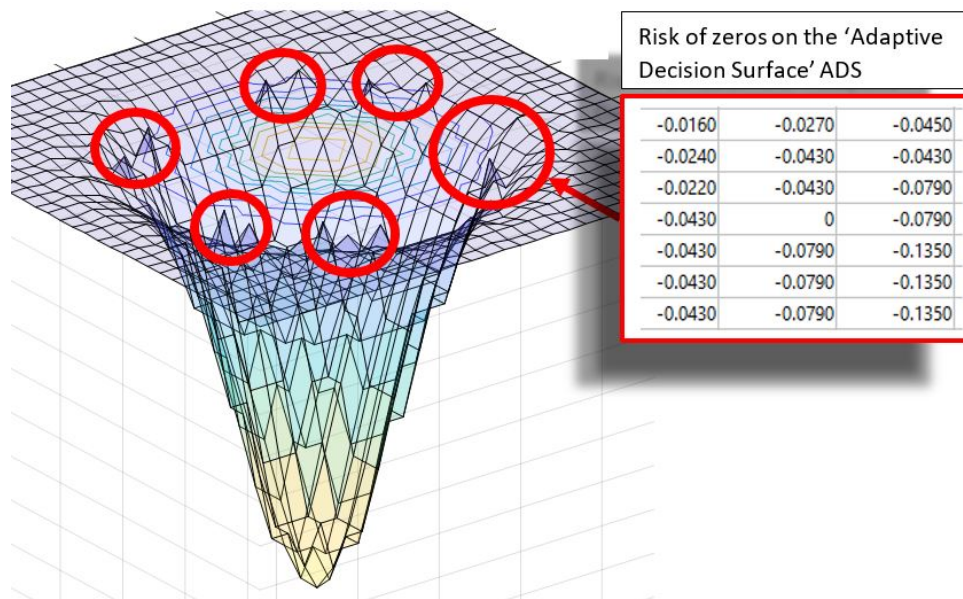


Figure 8.6: Adaptive Decision Surface ADS, risk of zeros on the surface

8.2.6 Future development an adaptive decision surface data threshold

ErrP classifier zone depends on a threshold on cross-section horizontal, please see section 5.2.3.4. The threshold level could be adaptable on the horizontal plane and tilted. Please see Fig:8.7 before the intersection. Please see an example of a planar threshold intersection. Fig 8.8

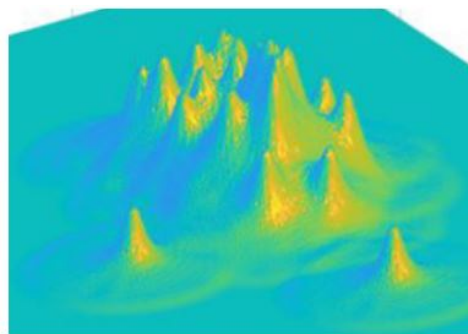


Figure 8.7: zone proximity structure threshold before intersection. ADS x,y,z axes 6.12

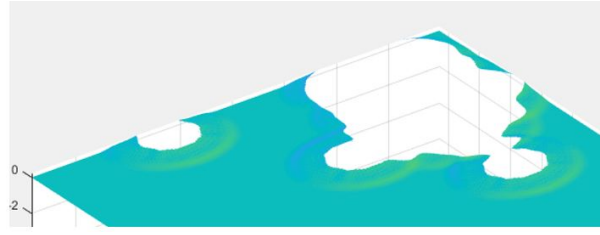


Figure 8.8: zone proximity structure threshold after intersection. ADS x,y,z axes 6.12

The threshold intersection can be either a planar or hyperbola surface. The intersecting surface can be pitched in any x,y,z rotation in addition to lateral repositioning in the x,y,z coordinated direction. Moreover, the intersection could be adjusted during an adaptation phase in addition to changes in the environment such as the persons condition such as temperature or hydration state which may affect the decision threshold.

Furthermore there could be a systematic parameter optimisation for the adaptive decision surface(ADS). In addition an error-driven learning such that previous errors are less likely to occur by adjust parameters adjusting the classification algorithm or feature extraction methods.

8.2.7 Future work on error driven learning algorithm

Herman et al [74] used a learning algorithm as part of their solution to handle uncertainty effects in BCI classification. Moreover, the authors used a type-2 fuzzy logic (T2FL) methodology for dealing with uncertainty. The uncertainties in the variability of brain dynamics are due to the non-stationary of brain signals. The authors' examined the applicability of the T2FL approach for EEG pattern recognition. The authors worked on a design methodology for the interval T2FL system (IT2FLS) to handle intersession in addition to within-session manifestations of non-stationary spectral EEG correlates of motor imagery. The authors examined the proposed fuzzy classifier in both off-line and on-line EEG classification case studies.

We propose to develop a learning algorithm such that when an error-related potential occurs the following parameters; feature extraction, feature selection, motor imagery classification parameters could be adjusted in order that the error is less likely occur in the future operation by learning from the previous mistakes.

8.2.8 Future work on a novel shared control strategy algorithm

Develop a shared control strategy algorithm such that reinforcement-learning methods should be incorporated into the BCI system and the robotic assistive device. In addition, the robotic interactions will support the person with a disability using the BCI system.

Moreover, the robotic device operating in the shared environment must have an array of sensors to maintain safety and security in the shared environment. Sensors may be of several forms such as vision, sound, inferred vision, ultrasound, LiDAR, thermal temperature sensing. Please see Chapter 7 for more information.

Collectively these sensors with a shared controlled strategy will enable a safe interaction with the robotic device. Furthermore, the operating environment will be a changing environment, therefore, a dynamic risk register could be a part of the control policy of the robotic device to ensure a safe interaction.

In a shared control strategy, there must be some type of communications protocol between the BCI and the assistive robotic device or prosthetic hand. For example, the participant may desire to pick up an object. The prosthetic hand will be able to assist the person in reducing the workload in retrieving the object. Several studies have explored the possibilities of different types of sensors attached to the supportive device such as the sense of touch in [105] or object

recognition as in [106].

A communications protocol that could function as follows: **Step 1)** Detect an object and recognise a particular object such as a box or cup Fig 8.10, Fig 8.11, Fig 8.12. **Step 2)** Wait for the BCI to signal an action such as open the hand following a successful motor imagery event Fig 8.9. **Step 3)** Then using local sensors such as proximity and vision to control the prosthetic hand to grasp the object. [106]**Step 4)** The prosthetic hand waits for a successful motor imagery event to release the object. Please see Chapter 7 for more information details regarding vision.

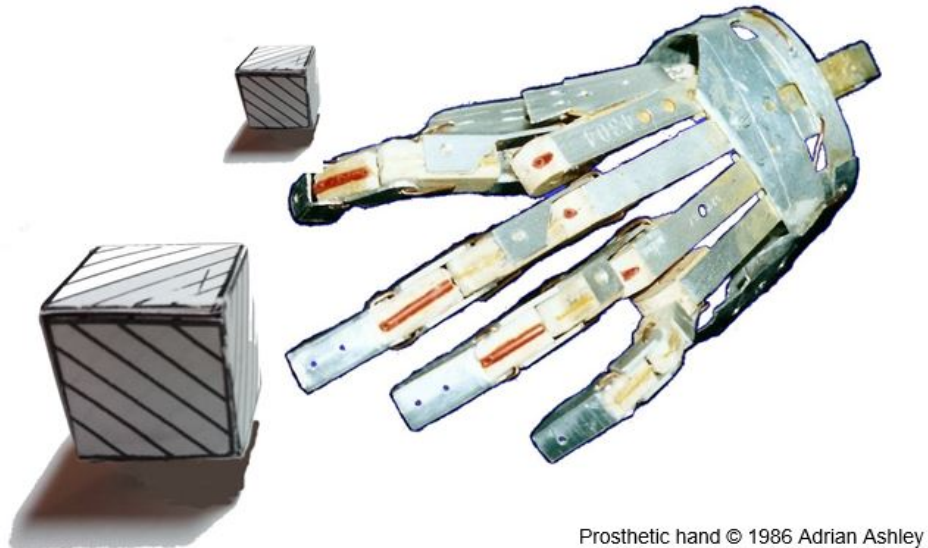


Figure 8.9: Prosthetic Hand

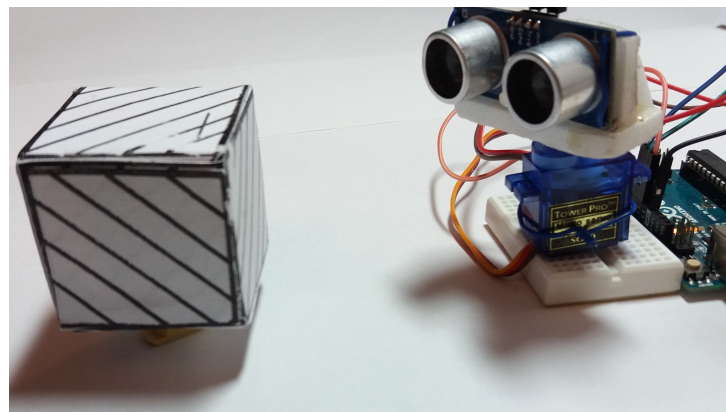


Figure 8.10: Ultra sonic range detection

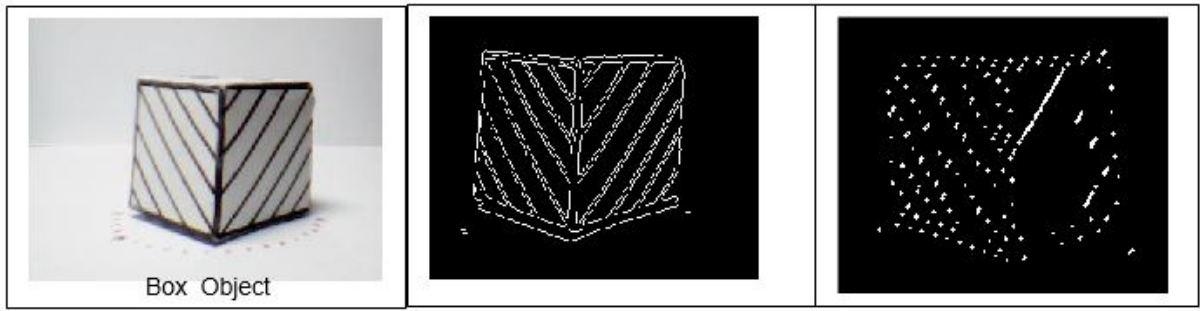


Figure 8.11: Physical box, edge detection, line interactions

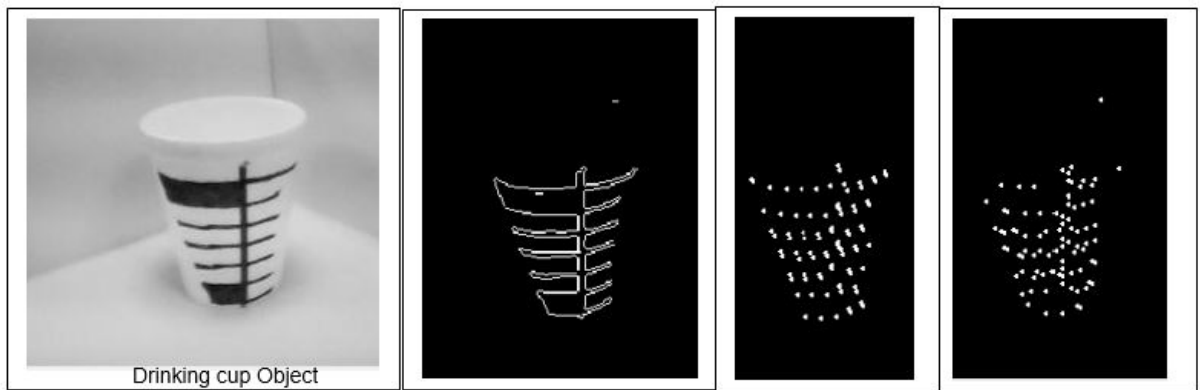


Figure 8.12: Physical Drinking cup, edge detection, line interactions

Appendix A

Participants Dataset

Table 1 A.1 : Young participants dataset

Table 2 A.2 : Elderly participants dataset

Bar chart A.3 : Participants age

A.1 Participants

A.1.1 Young participants data table

Person	Session ID	Age	Gender	Age Category
subject 1	AI_00	18	F	Young
Subject 2	DS_00	18	M	Young
subject 3	NL_00	18	F	Young
subject 4	NM_00	18	F	Young
subject 5	COC2_00	19	M	Young
subject 6	NIK_00	19	F	Young
subject 7	CJ_00	19	M	Young
subject 8	LS_00	19	F	Young
subject 9	SC_00	19	F	Young
subject 10	SH_00	20	F	Young
subject 11	EJD_00	20	M	Young
subject 12	CD_00	21	F	Young
subject 13	GG_00	21	F	Young
subject 14	KM_00	21	M	Young
subject 15	MB_00	24	F	Young
subject 16	IC_00	24	F	Young
subject 17	AU_00	25	M	Young
subject 18	HJ_00	26	F	Young
subject 19	PV_00	26	M	Young
subject 20	LF_00	27	F	Young
subject 21	KG_00	27	F	Young
subject 22	AM_00	29	M	Young
subject 23	BW_00	29	F	Young
subject 24	AH_00	32	F	Young
subject 25	JG_00	32	F	Young
subject 26	MA_00	32	F	Young
subject 27	NK_00	34	F	Young

Figure A.1: Young participants dataset table

A.1.2 Elderly participants data table

Person	Session ID	Age	Gender	Age Category
subject 1	ED_00	65	F	Elderly
subject 2	JM_00	65	F	Elderly
subject 3	KF_00	65	F	Elderly
subject 4	KB_00	65	F	Elderly
subject 5	BJC_00	65	M	Elderly
subject 6	EID_00	65	M	Elderly
subject 7	UL_00	65	F	Elderly
subject 8	CC2_00	66	F	Elderly
subject 9	BL_00	67	F	Elderly
subject 10	BL2_00	68	F	Elderly
subject 11	SM_00	68	F	Elderly
subject 12	CMD_00	70	M	Elderly
subject 13	DMCH_00	72	F	Elderly
subject 14	EOB_00	74	M	Elderly
subject 15	JD_00	74	M	Elderly
subject 16	BC_00	74	F	Elderly
subject 17	FMC_00	74	M	Elderly
subject 18	MM_00	74	F	Elderly
subject 19	FM_00	76	M	Elderly
subject 20	GD_00	76	F	Elderly
subject 21	JS_00	79	M	Elderly
subject 22	DD_00	79	M	Elderly
subject 23	MS_00	80	F	Elderly
subject 24	MM2_00	??	F	Elderly

Figure A.2: Young participants dataset table

A.1.3 Participants age bar chart

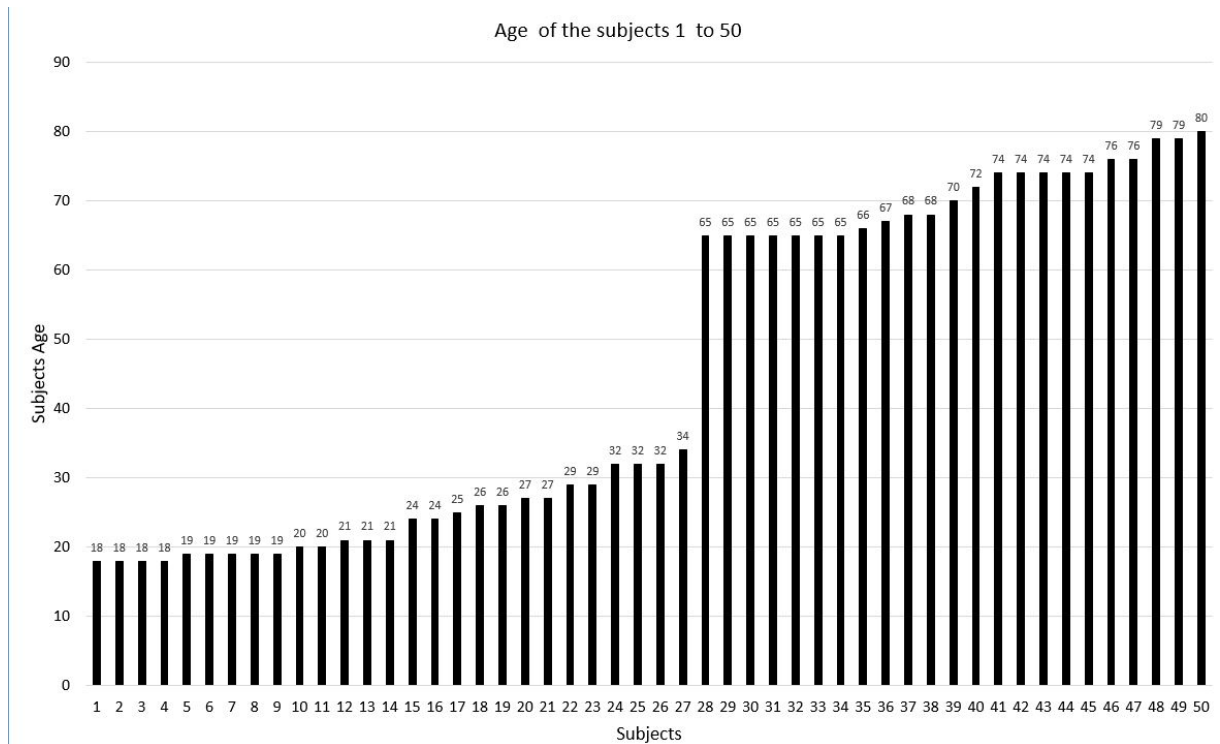


Figure A.3: Participants age bar chart

Appendix B

Participants relative peak feature(RPF) datasets

B.1 Dataset elderly participants relative peak features(RPF)

Dataset 1, 2 B.1 : Elderly participants 1 and 2, showing their relative peak features(RPF)

Dataset 3, 4 B.2 : Elderly participants 3 and 4, showing their relative peak features(RPF)

Dataset 5, 6 B.3 : Elderly participants 5 and 6, showing their relative peak features(RPF)

Dataset 7, 8 B.4 : Elderly participants 7 and 8, showing their relative peak features(RPF)

Dataset 9, 10 B.5 : Elderly participants 9 and 10, showing their relative peak features(RPF)

Dataset 11, 12 B.6 : Elderly participants 11 and 12, showing their relative peak features(RPF)

Dataset 13, 14 B.7 : Elderly participants 13 and 14, showing their relative peak features(RPF)

Dataset 15, 16 B.8 : Elderly participants 15 and 16, showing their relative peak features(RPF)

Dataset 17, 18 B.9 : Elderly participants 17 and 18, showing their relative peak features(RPF)

Dataset 19, 20 B.10 : Elderly participants 19 and 20, showing their relative peak features(RPF)

Dataset 21, 22 B.11 : Elderly participants 21 and 22, showing their relative peak features(RPF)

Dataset 23, 24 B.12 : Elderly participants 23 and 24, showing their relative peak features(RPF)

B.1.1 Elderly participants 1 and 2, showing their relative peak features(RPF)

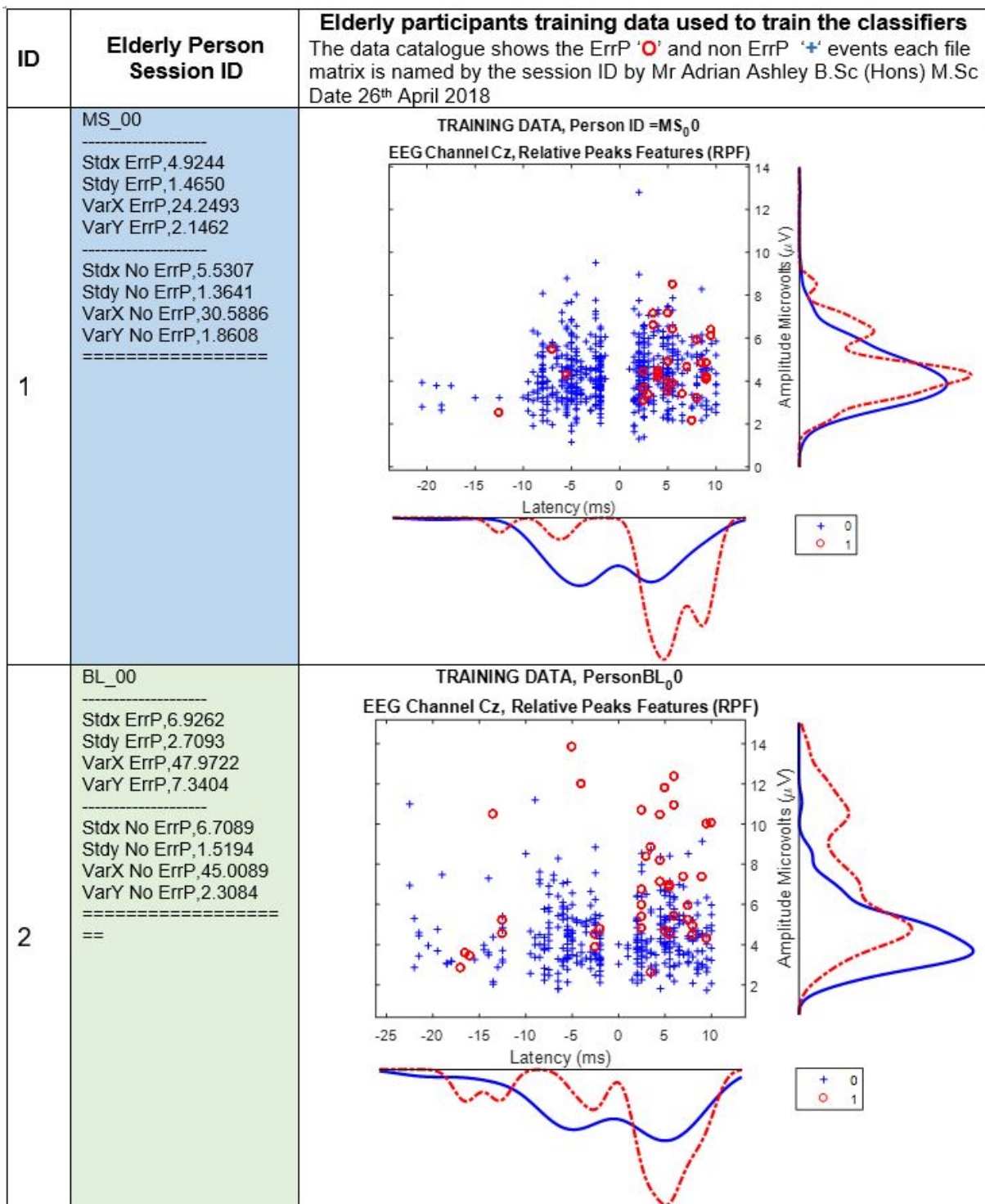


Figure B.1: Dataset elderly participants relative peak features ID 1,2

B.1.2 Elderly participants 3 and 4, showing their relative peak features(RPF)

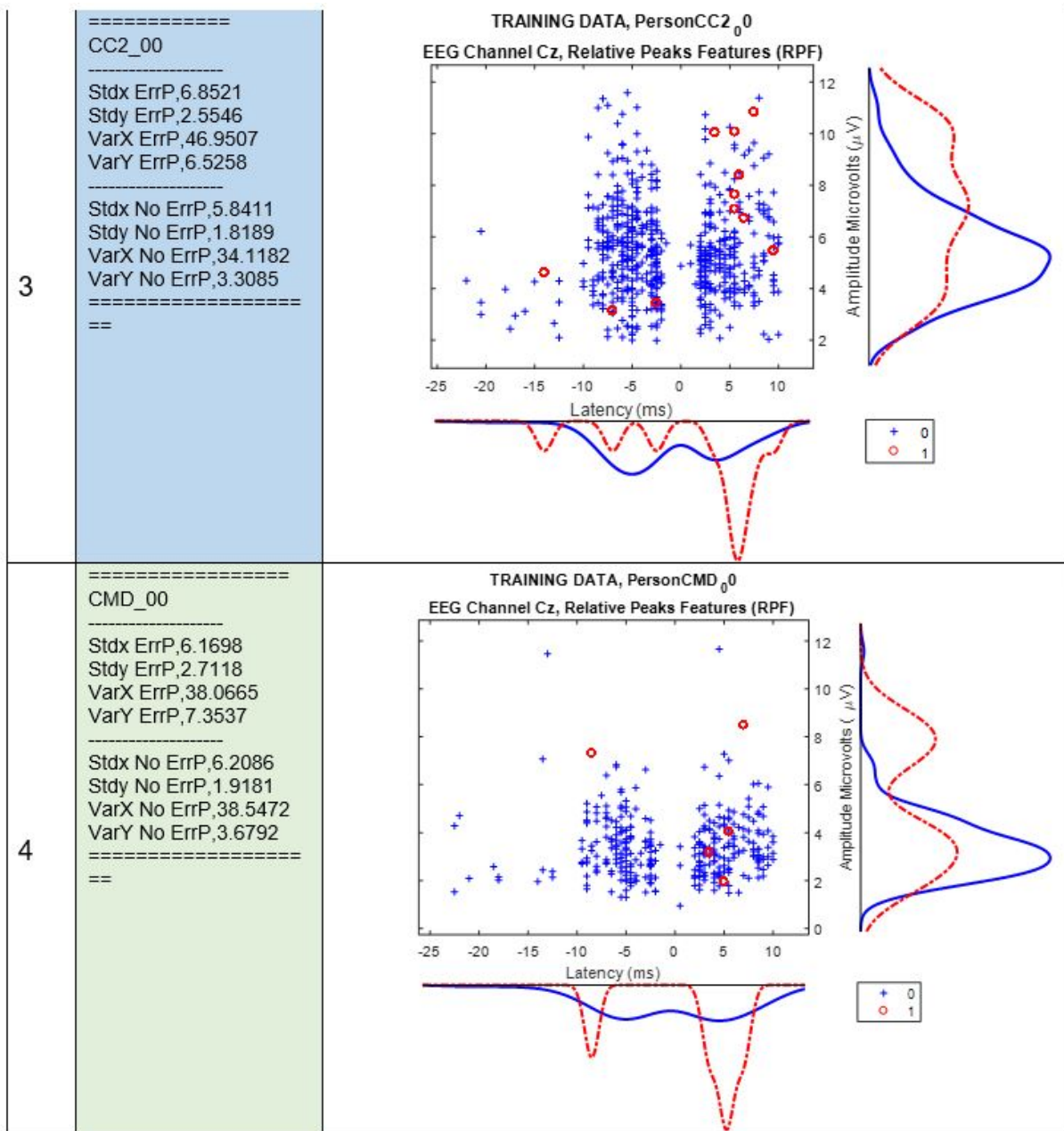


Figure B.2: Dataset elderly participants relative peak features ID 3,4

B.1.3 Elderly participants 5 and 6, showing their relative peak features(RPF)

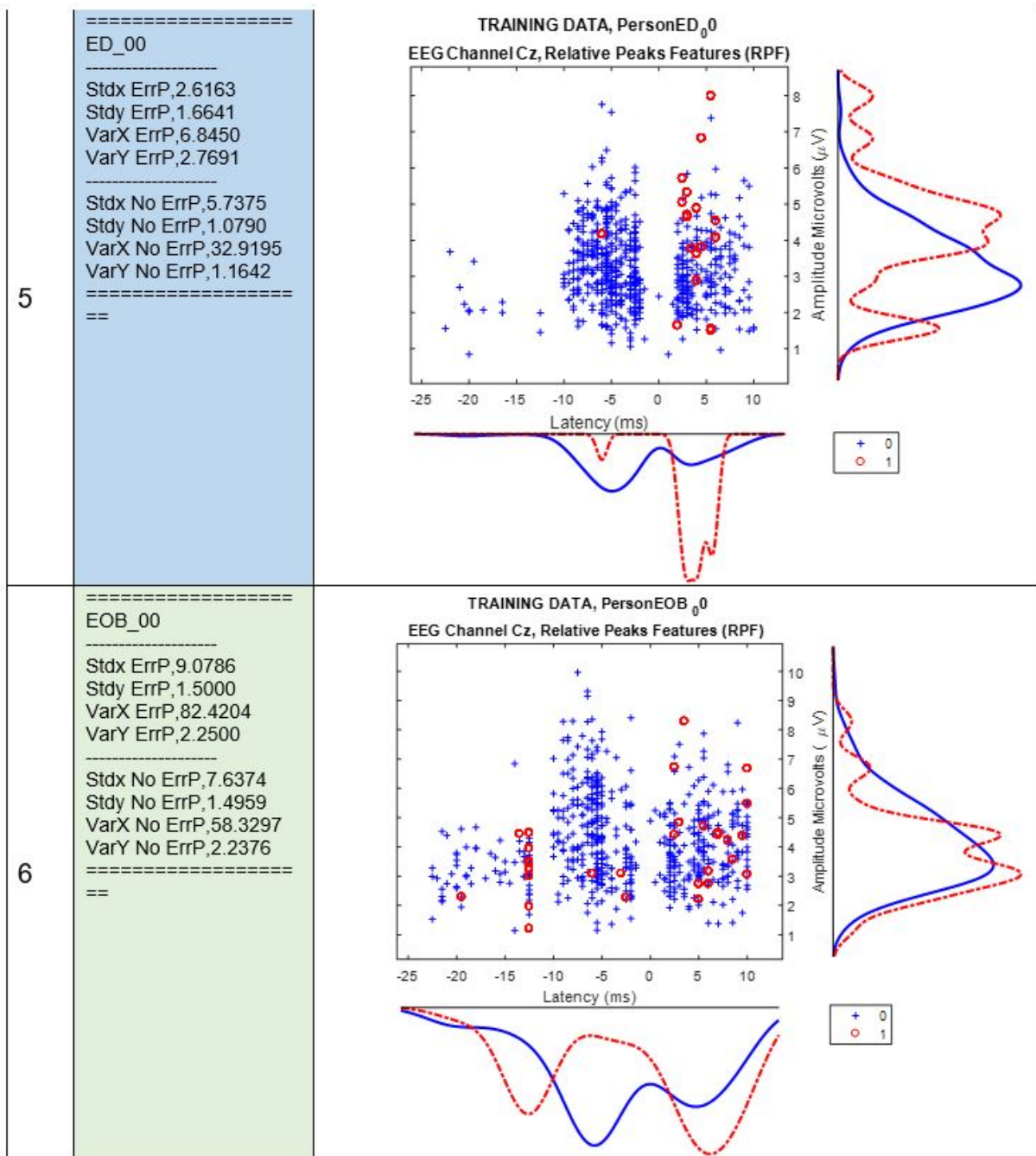


Figure B.3: Dataset elderly participants relative peak features ID 5,6

B.1.4 Elderly participants 7 and 8, showing their relative peak features(RPF)

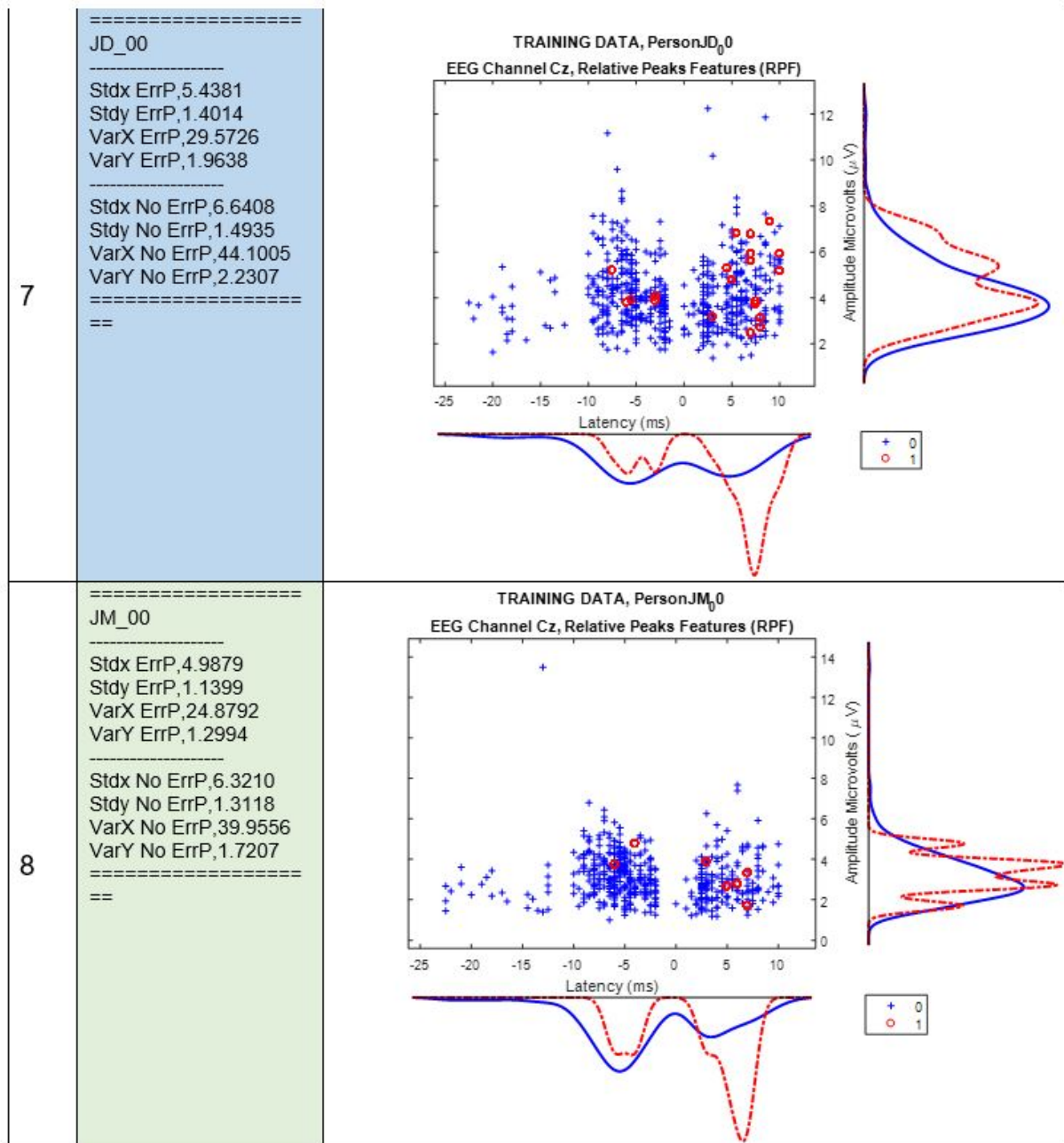


Figure B.4: Dataset elderly participants relative peak features ID 7,8

B.1.5 Elderly participants 9 and 10, showing their relative peak features(RPF)

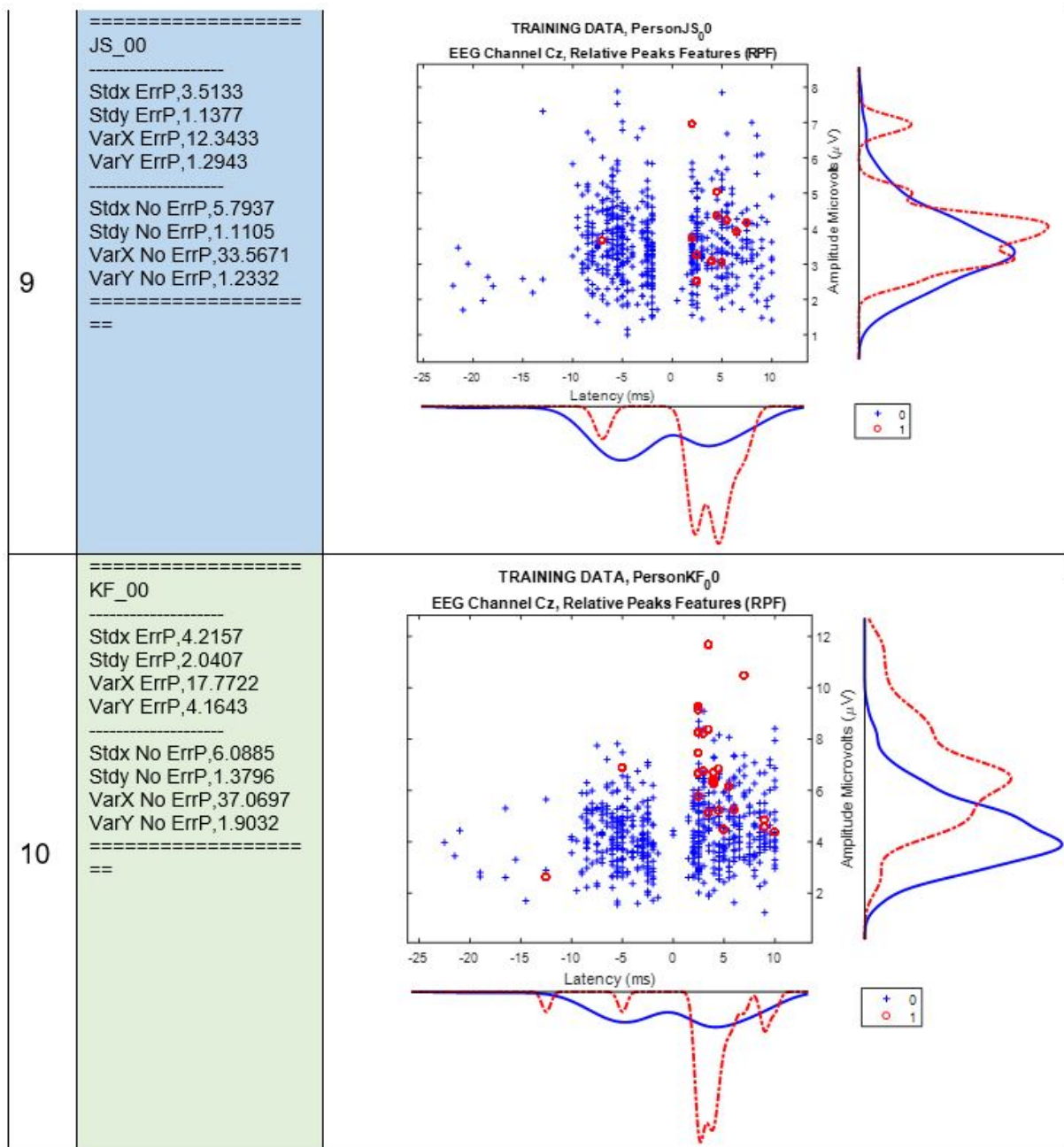


Figure B.5: Dataset elderly participants relative peak features ID 9,10

B.1.6 Elderly participants 11 and 12, showing their relative peak features(RPF)

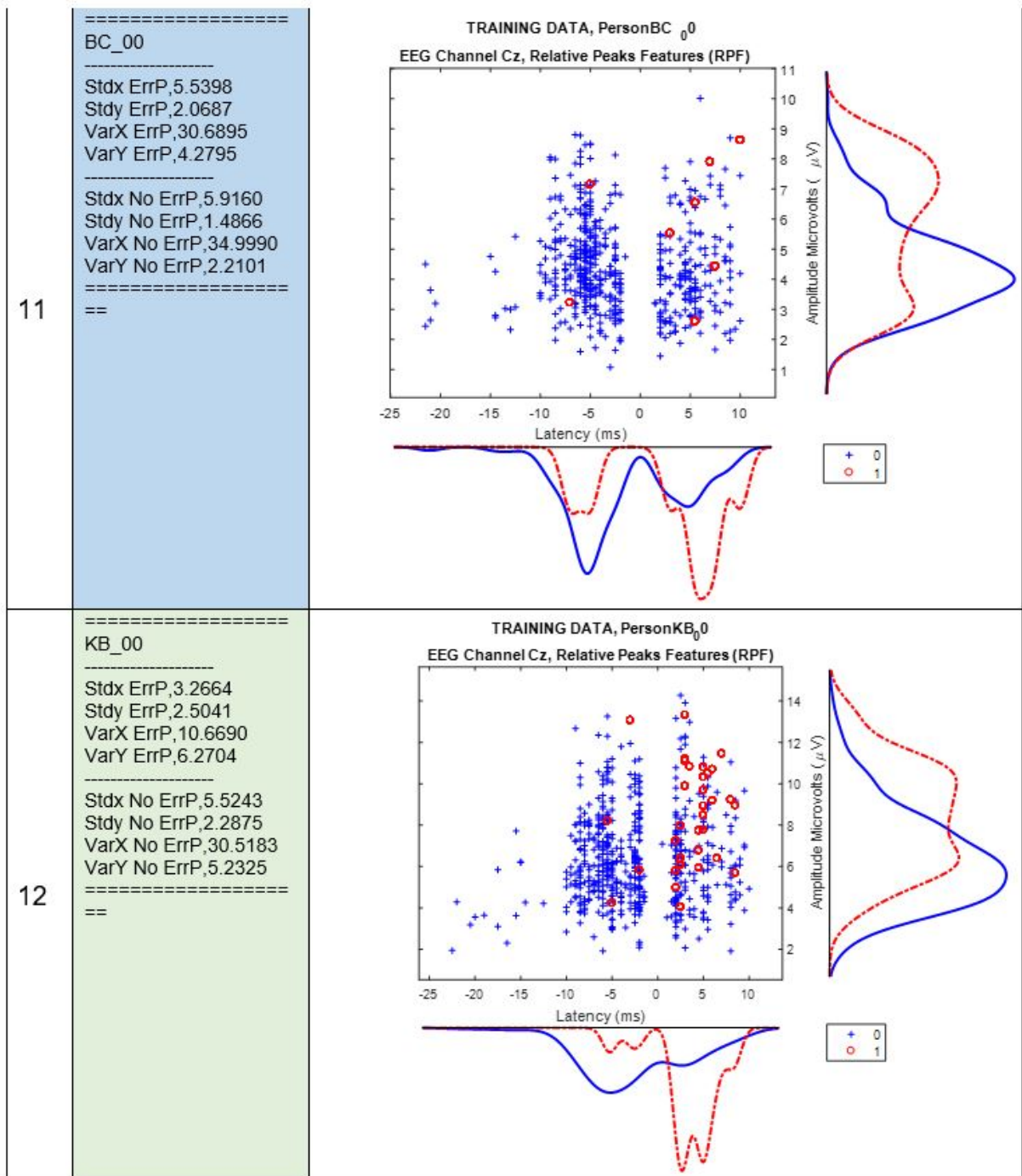


Figure B.6: Dataset elderly participants relative peak features ID 11,12

B.1.7 Elderly participants 13 and 14, showing their relative peak features(RPF)

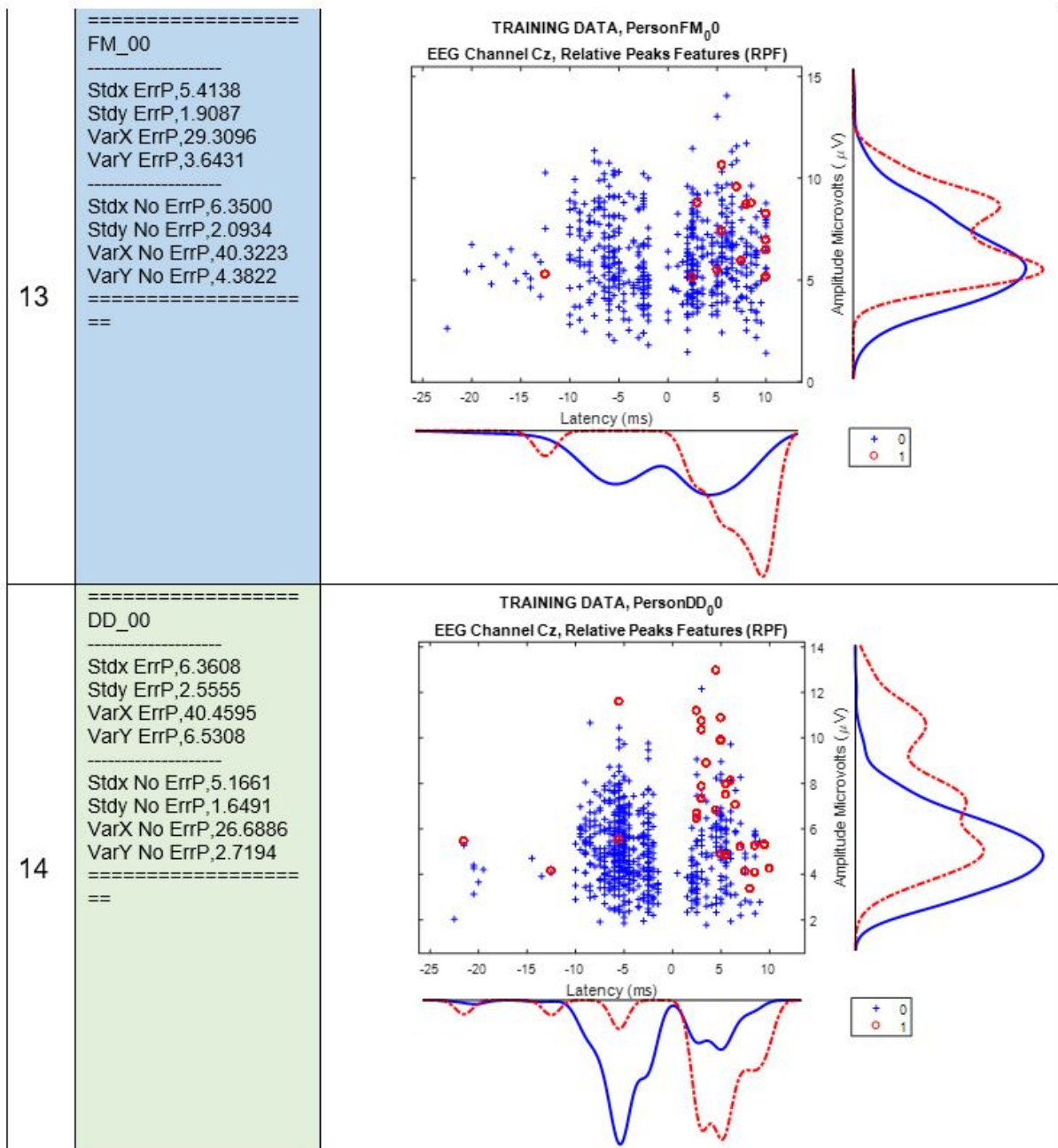


Figure B.7: Dataset elderly participants relative peak features ID 13,14

B.1.8 Elderly participants 15 and 16, showing their relative peak features(RPF)

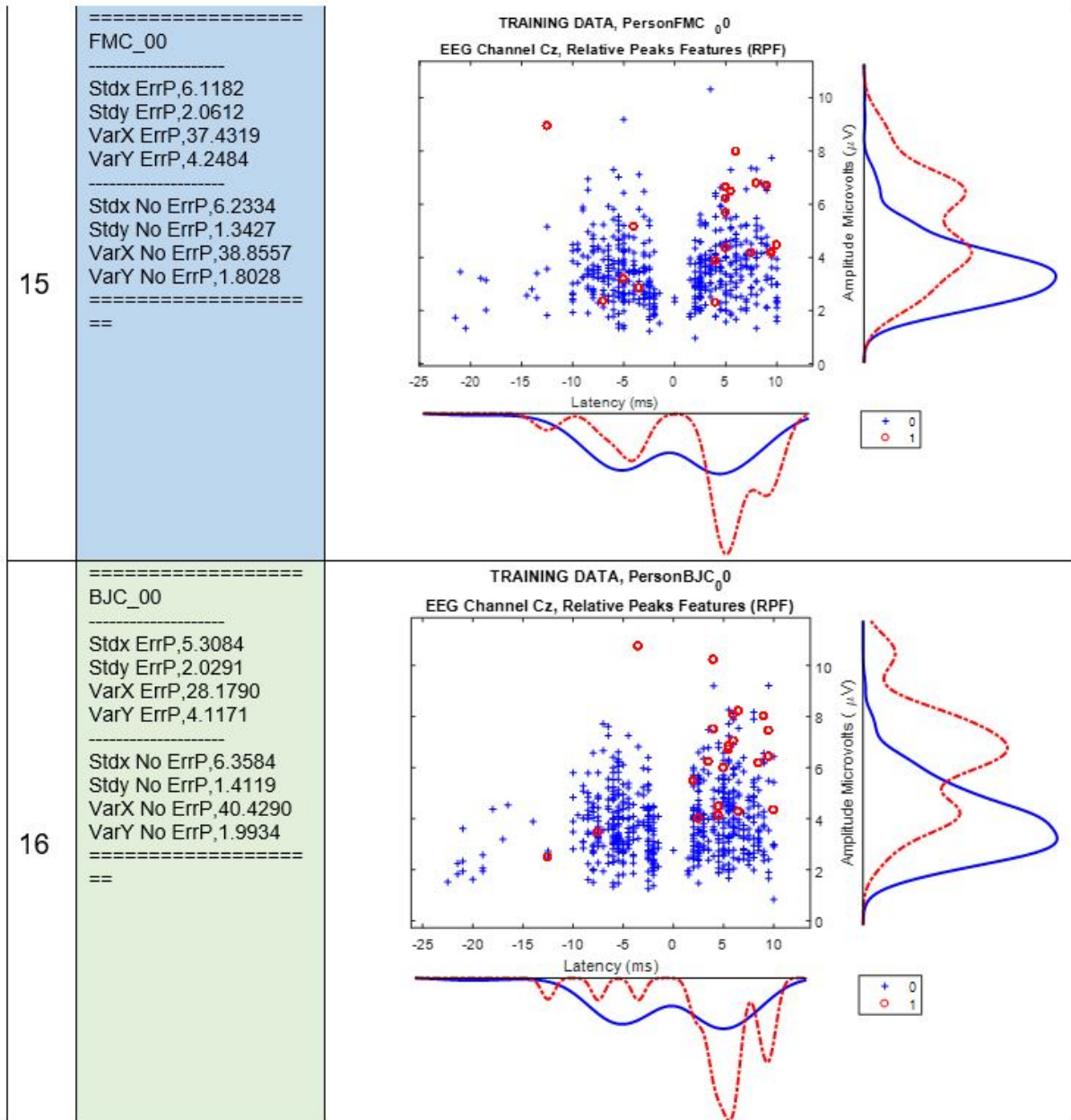


Figure B.8: Dataset elderly participants relative peak features ID 15,16

B.1.9 Elderly participants 9 and 10, showing their relative peak features(RPF)

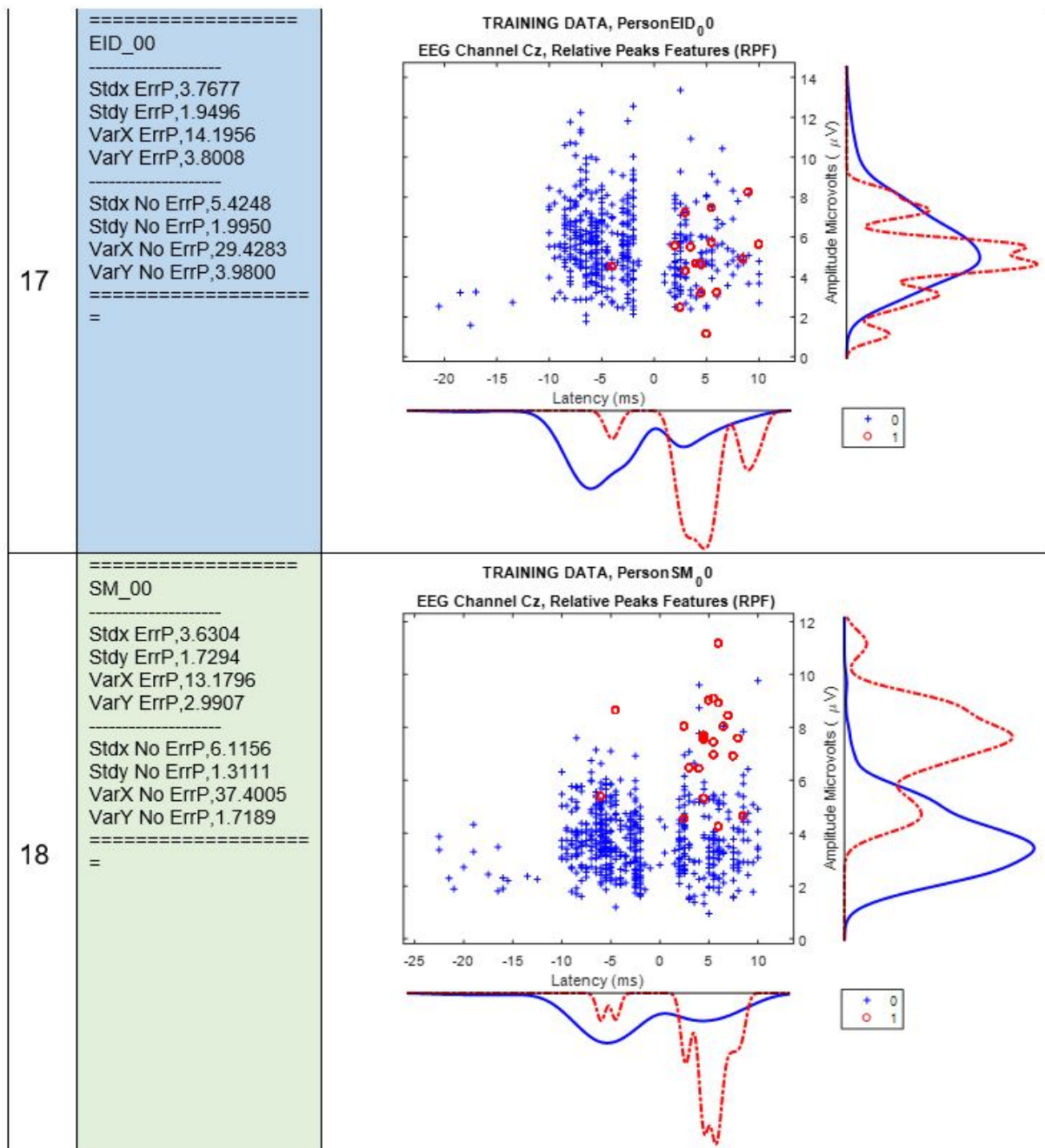


Figure B.9: Dataset elderly participants relative peak features ID 17,18

B.1.10 Elderly participants 11 and 12, showing their relative peak features(RPF)

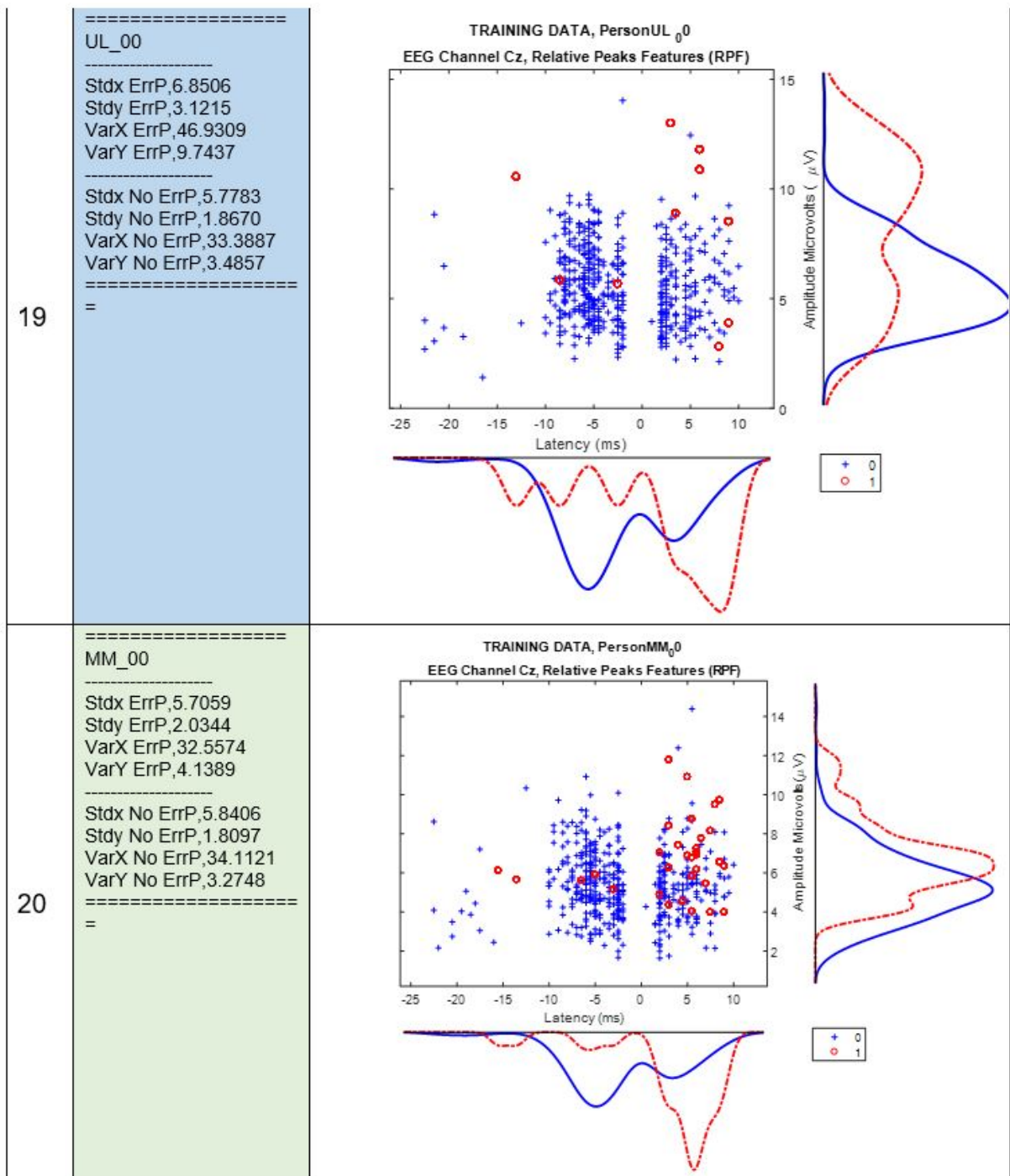


Figure B.10: Dataset elderly participants relative peak features ID 19,20

B.1.11 Elderly participants 13 and 14, showing their relative peak features(RPF)

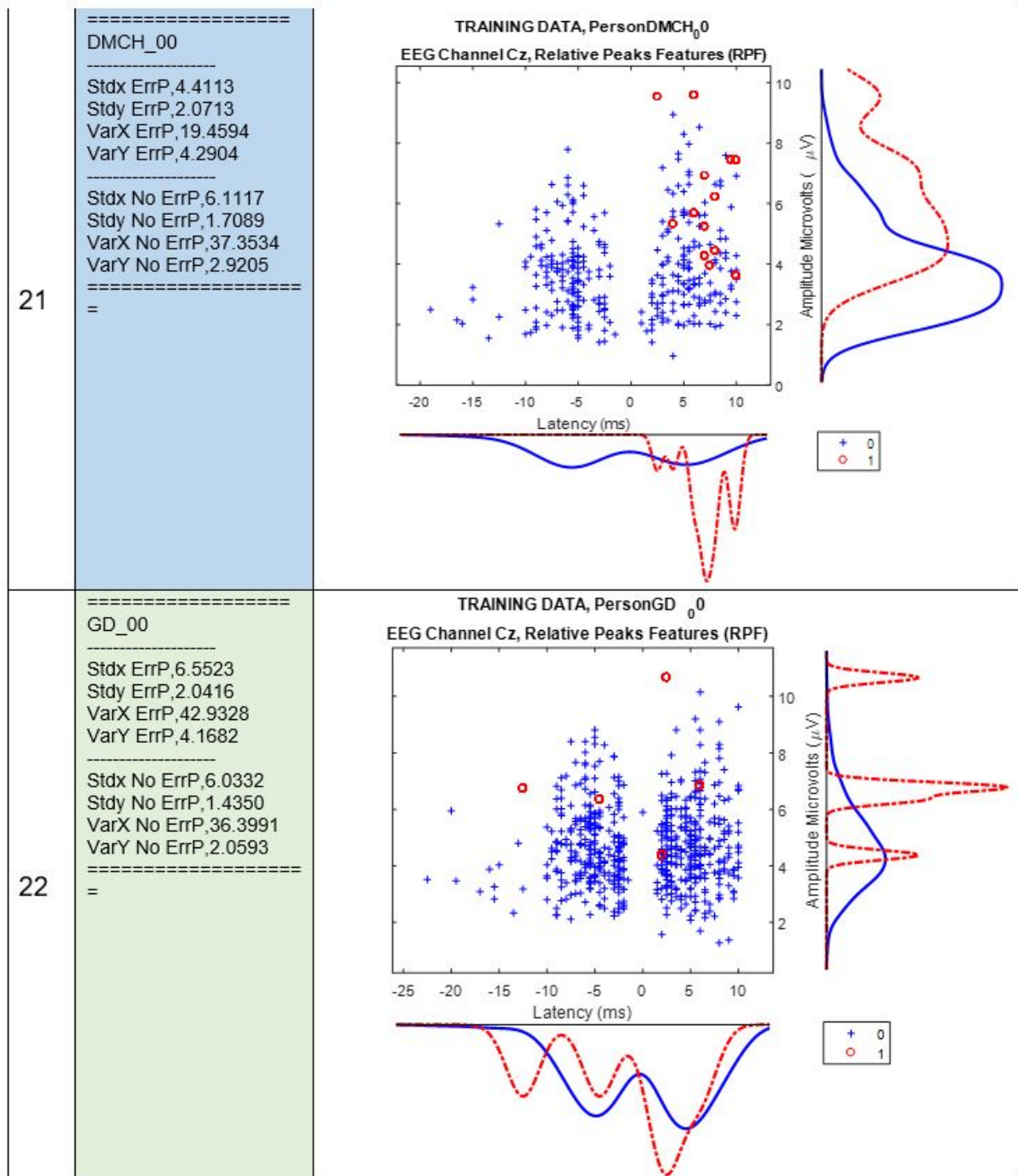


Figure B.11: Dataset elderly participants relative peak features ID 21,22

B.1.12 Elderly participants 15 and 16, showing their relative peak features(RPF)

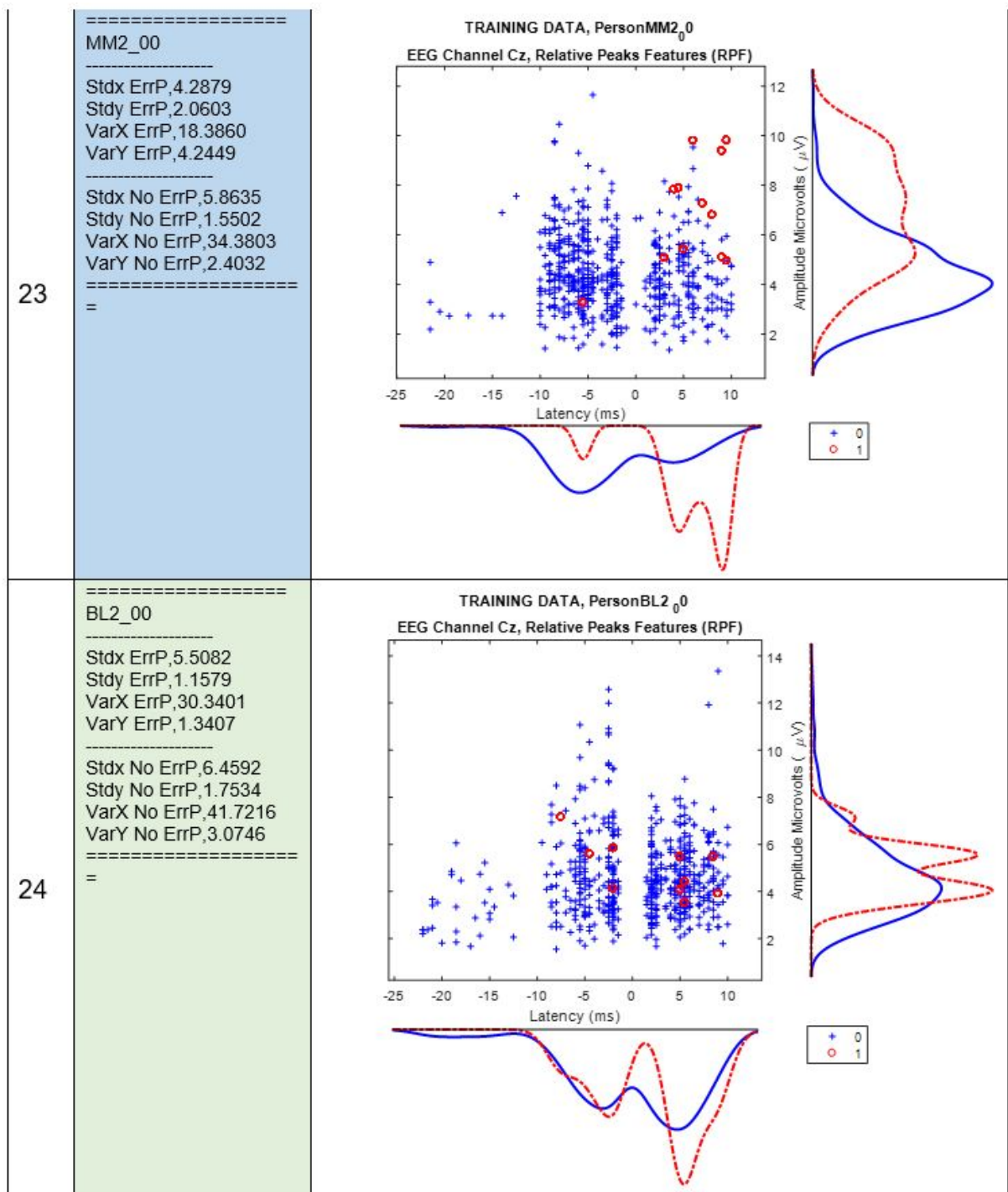


Figure B.12: Dataset elderly participants relative peak features ID 23,24

B.2 Dataset young participants relative peak features(RPF)

.
Dataset 1, 2 B.13 : Young participants 1 and 2, showing their relative peak features(RPF)

Dataset 3, 4 B.14 : Young participants 3 and 4, showing their relative peak features(RPF)

Dataset 5, 6 B.15 : Young participants 5 and 6, showing their relative peak features(RPF)

Dataset 7, 8 B.16 : Young participants 7 and 8, showing their relative peak features(RPF)

Dataset 9, 10 B.17 : Young participants 9 and 10, showing their relative peak features(RPF)

Dataset 11, 12 B.18 : Young participants 11 and 12, showing their relative peak features(RPF)

Dataset 13, 14 B.19 : Young participants 13 and 14, showing their relative peak features(RPF)

Dataset 15, 16 B.20 : Young participants 15 and 16, showing their relative peak features(RPF)

Dataset 17, 18 B.21 : Young participants 17 and 18, showing their relative peak features(RPF)

Dataset 19, 20 B.22 : Young participants 19 and 20, showing their relative peak features(RPF)

Dataset 21, 22 B.23 : Young participants 21 and 22, showing their relative peak features(RPF)

Dataset 23, 24 B.24 : Young participants 23 and 24, showing their relative peak features(RPF)

Dataset 25, 26 B.25 : Young participants 25 and 26, showing their relative peak features(RPF)

Dataset 27 ?? : Young participant 27, showing their relative peak features(RPF)

B.2.1 Young participants 1 and 2, showing their relative peak features(RPF)

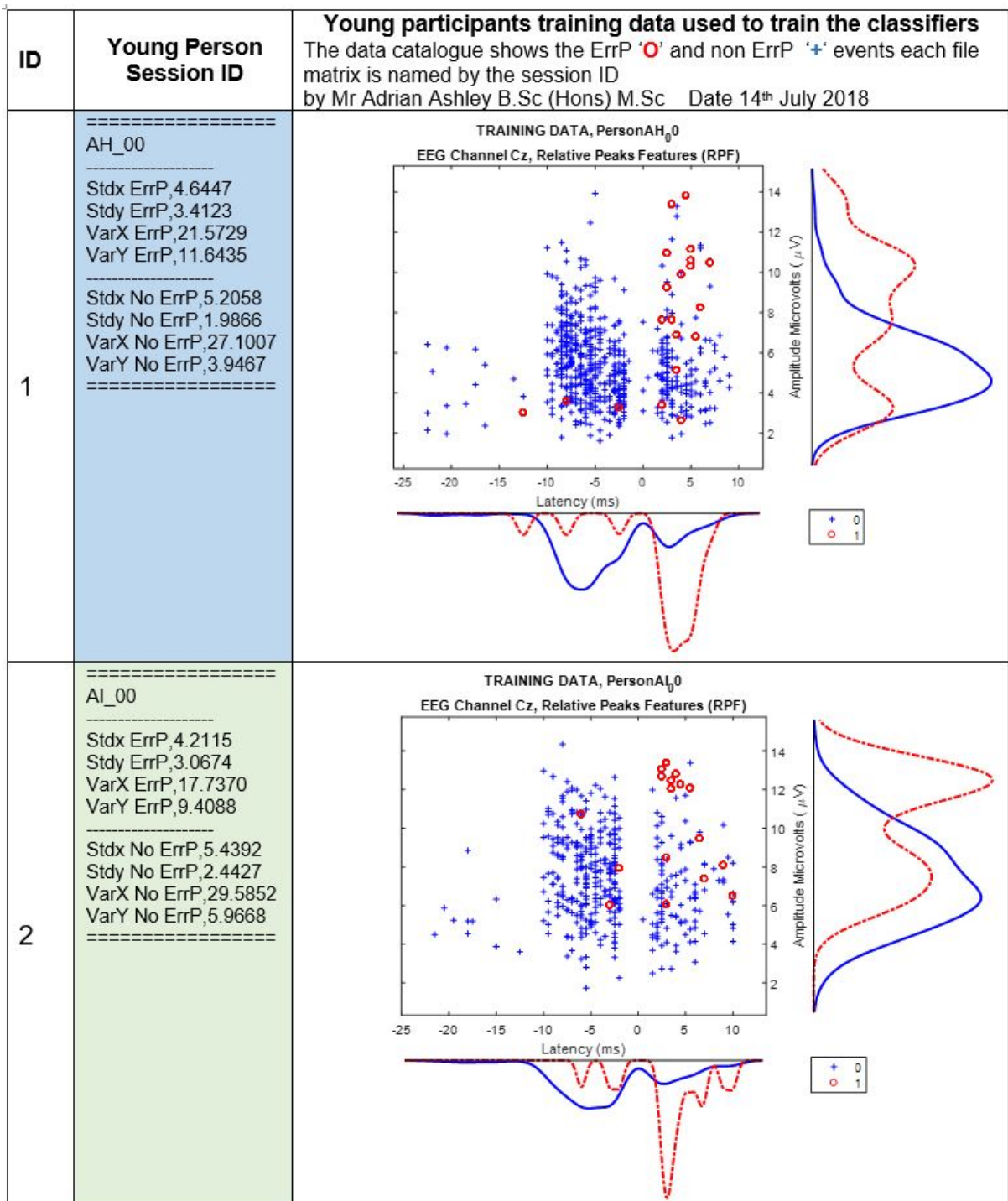


Figure B.13: Dataset young participants relative peak features ID 1,2

B.2.2 Young participants 3 and 4, showing their relative peak features(RPF)

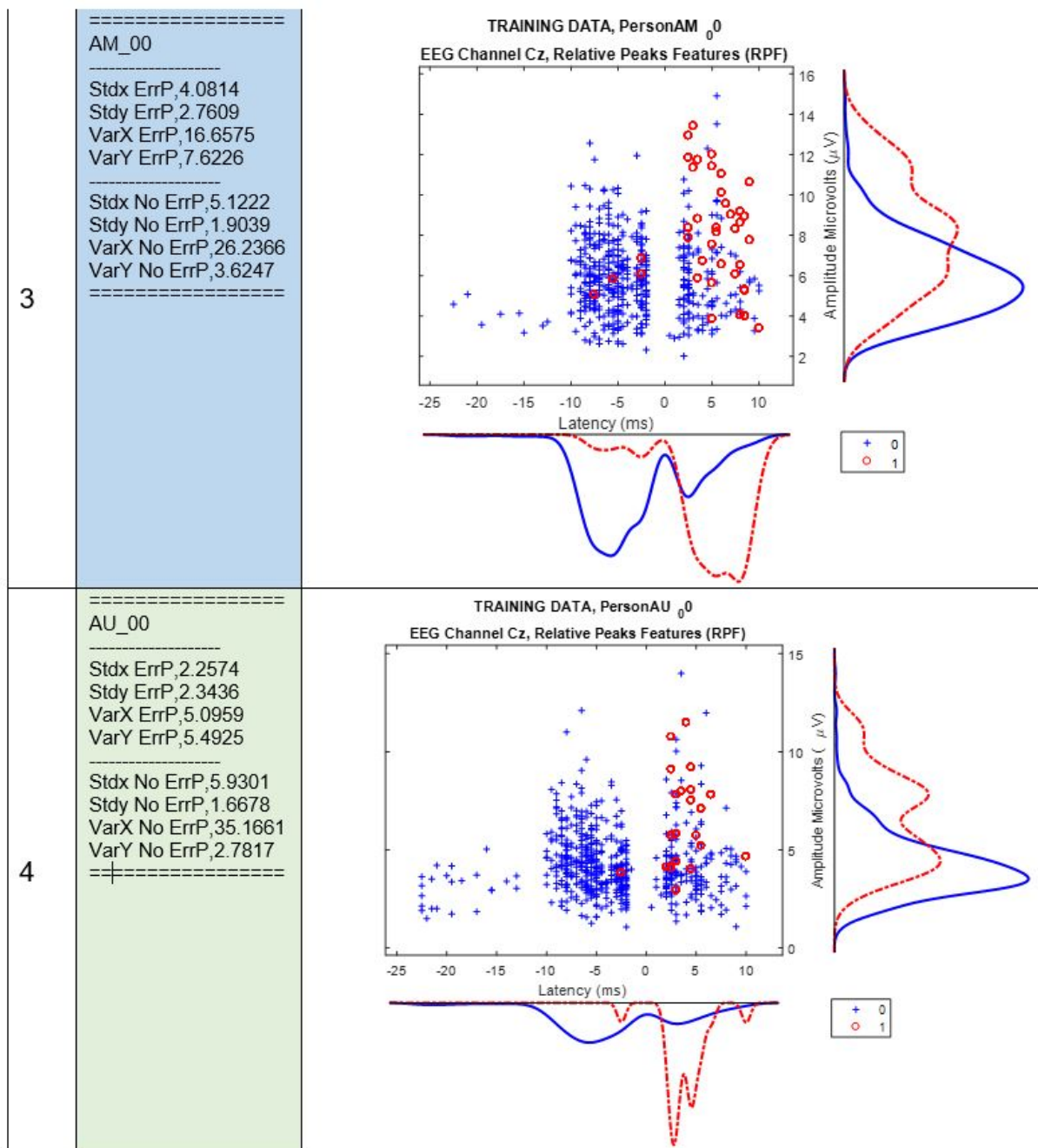


Figure B.14: Dataset young participants relative peak features ID 3,4

B.2.3 Young participants 5 and 6, showing their relative peak features(RPF)

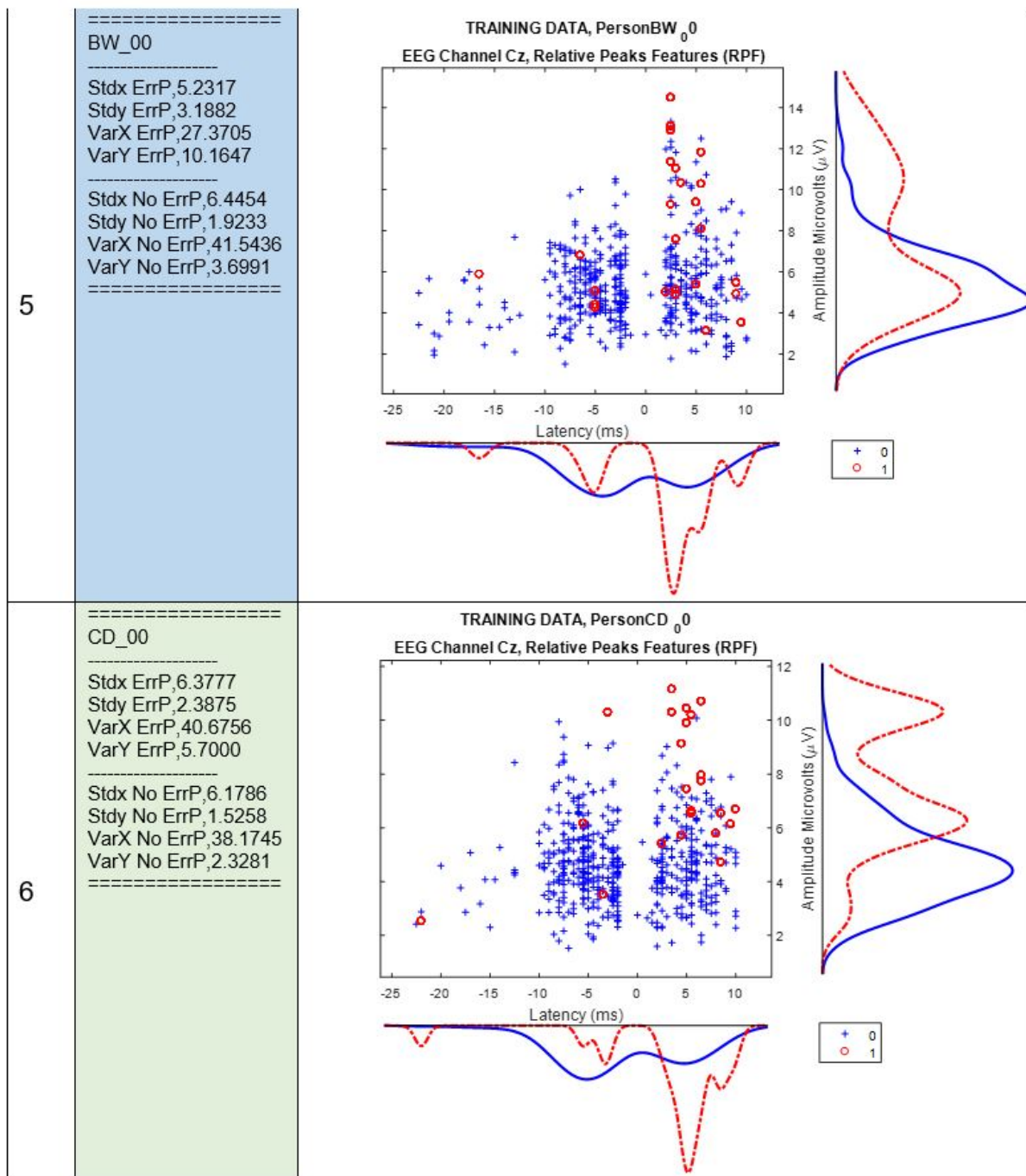


Figure B.15: Dataset young participants relative peak features ID 5,6

B.2.4 Young participants 7 and 8, showing their relative peak features(RPF)

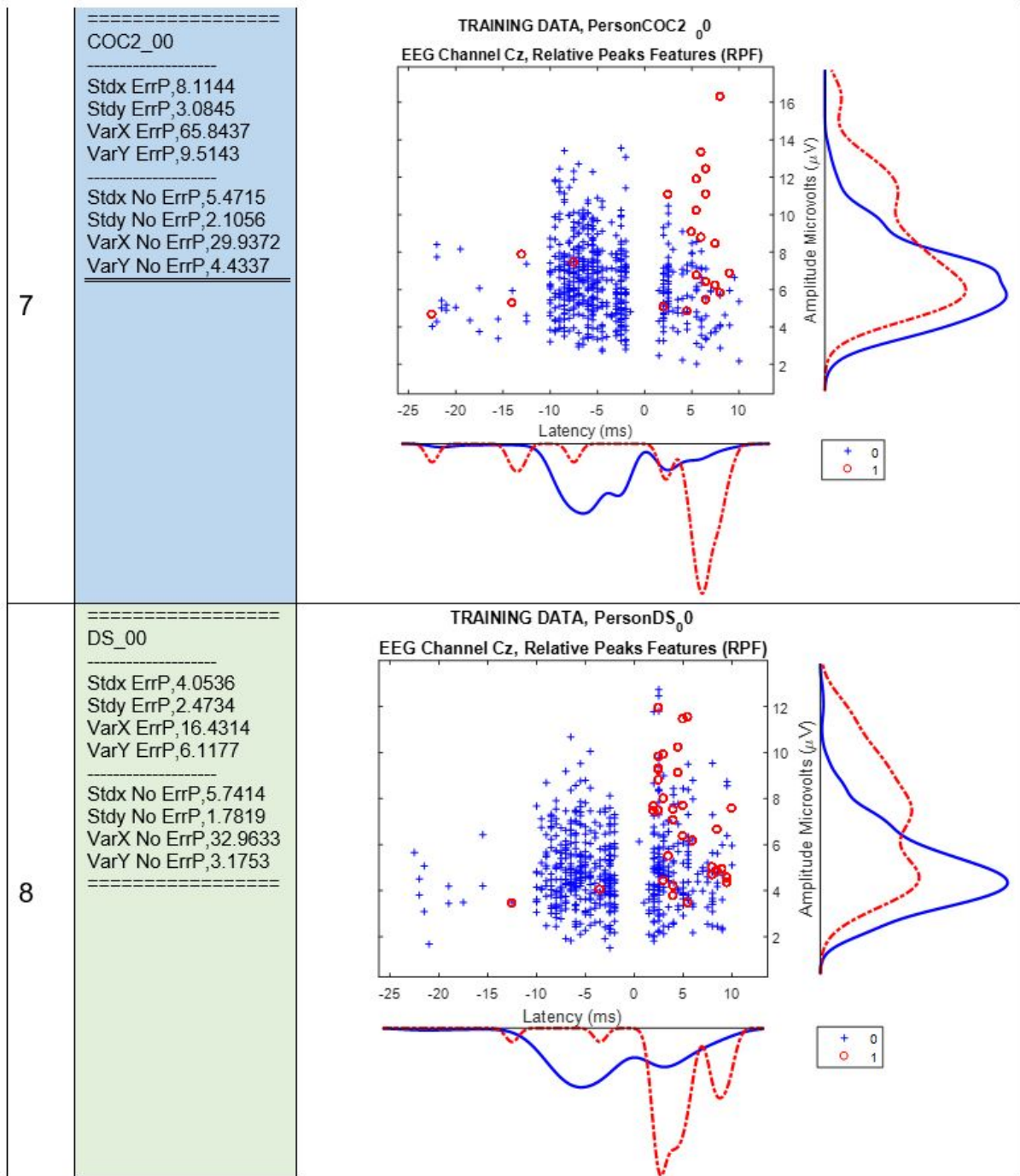


Figure B.16: Dataset young participants relative peak features ID 7,8

B.2.5 Young participants 9 and 10, showing their relative peak features(RPF)

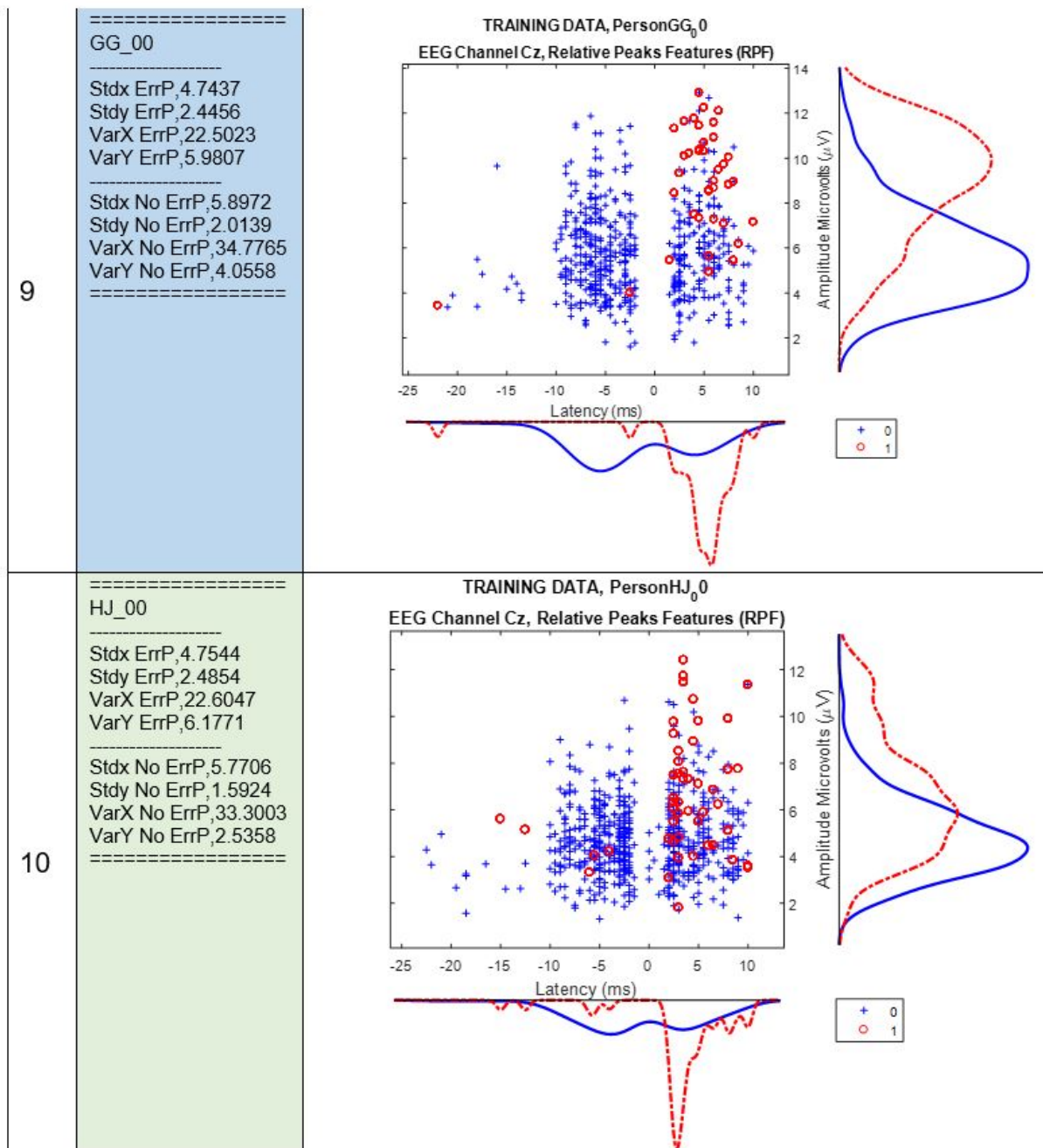


Figure B.17: Dataset young participants relative peak features ID 9,10

B.2.6 Young participants 11 and 12, showing their relative peak features(RPF)

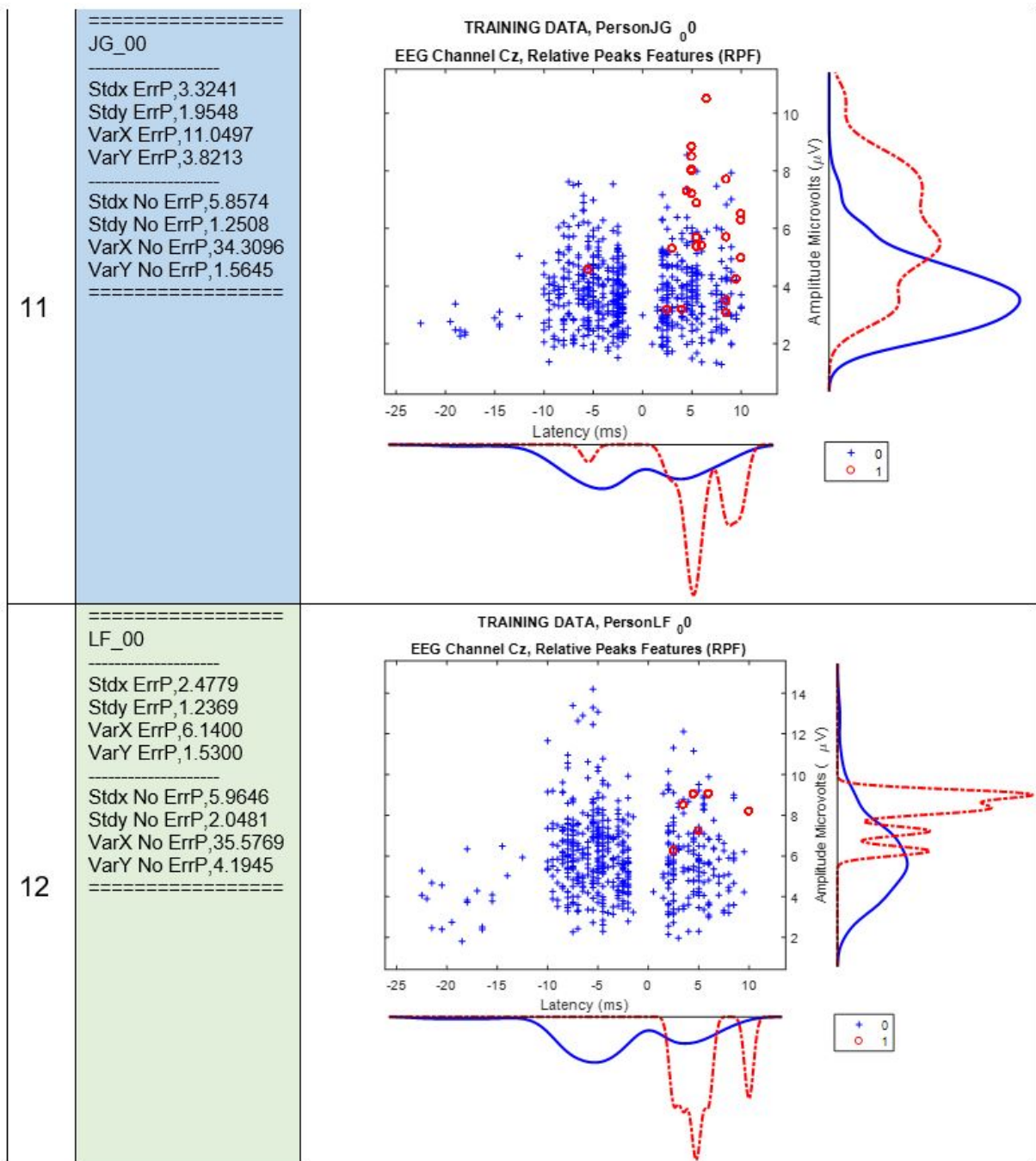


Figure B.18: Dataset young participants relative peak features ID 11,12

B.2.7 Young participants 13 and 14, showing their relative peak features(RPF)

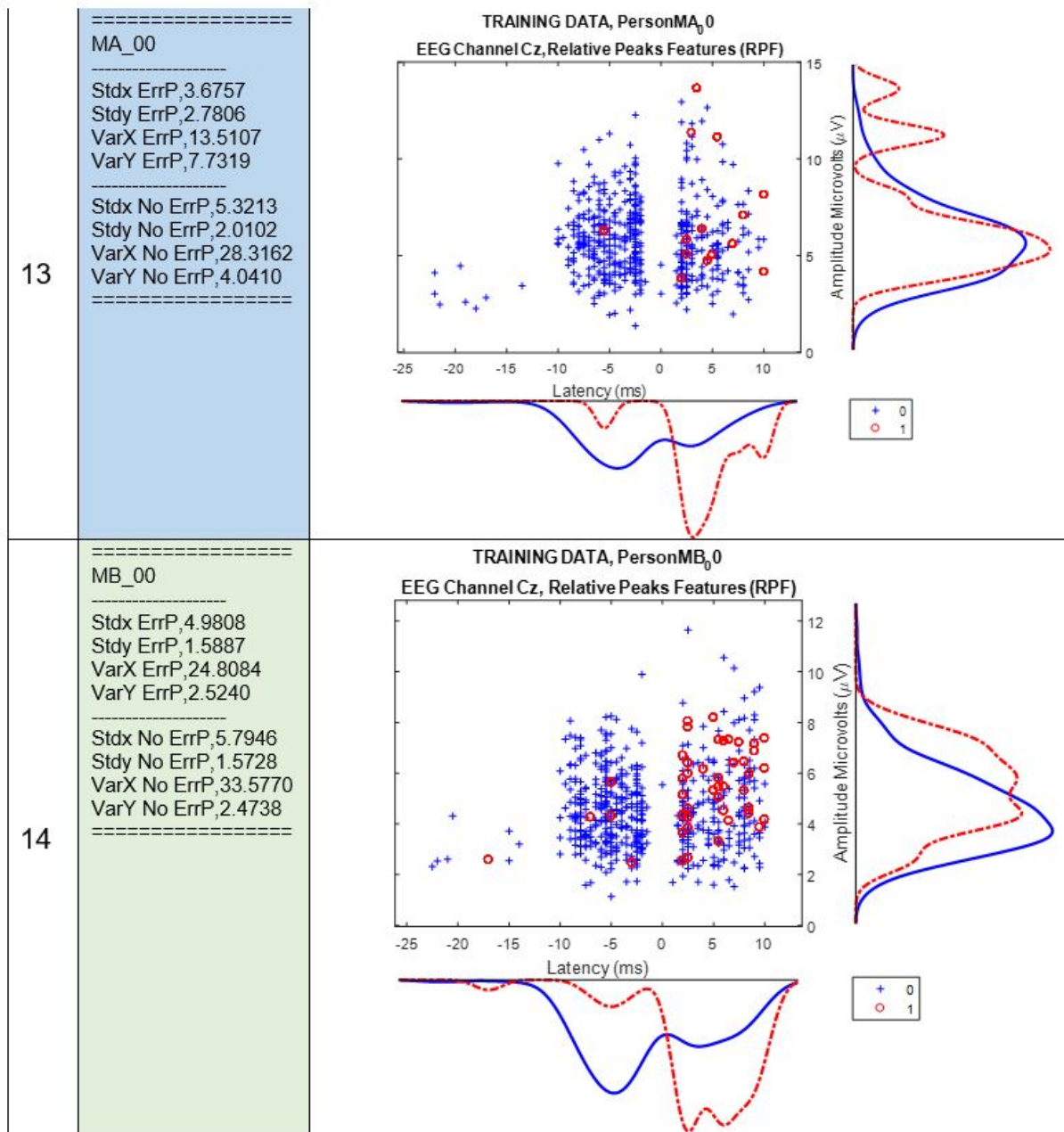


Figure B.19: Dataset young participants relative peak features ID 13,14

B.2.8 Young participants 15 and 16, showing their relative peak features(RPF)

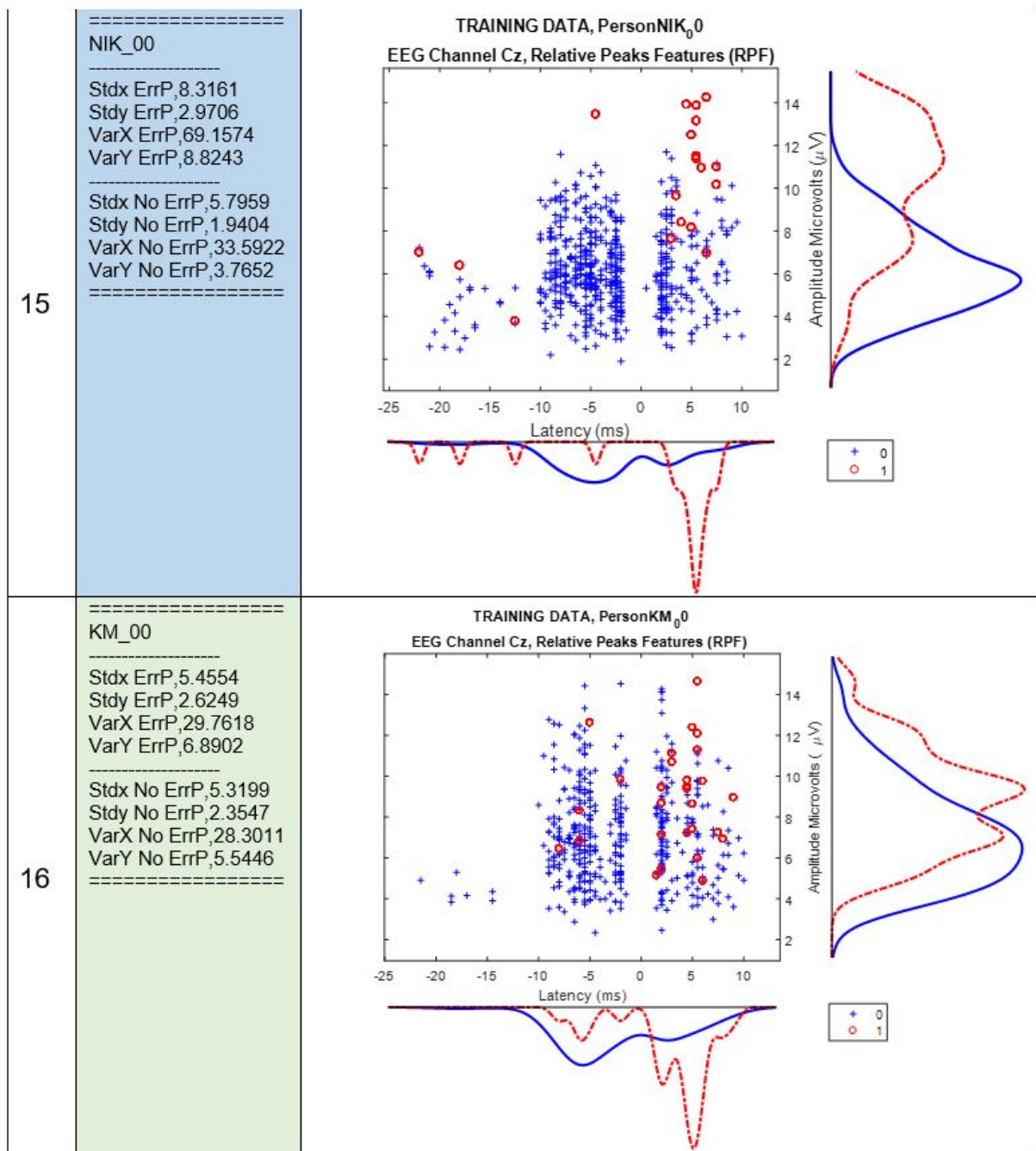


Figure B.20: Dataset young participants relative peak features ID 15,16

B.2.9 Young participants 17 and 18, showing their relative peak features(RPF)

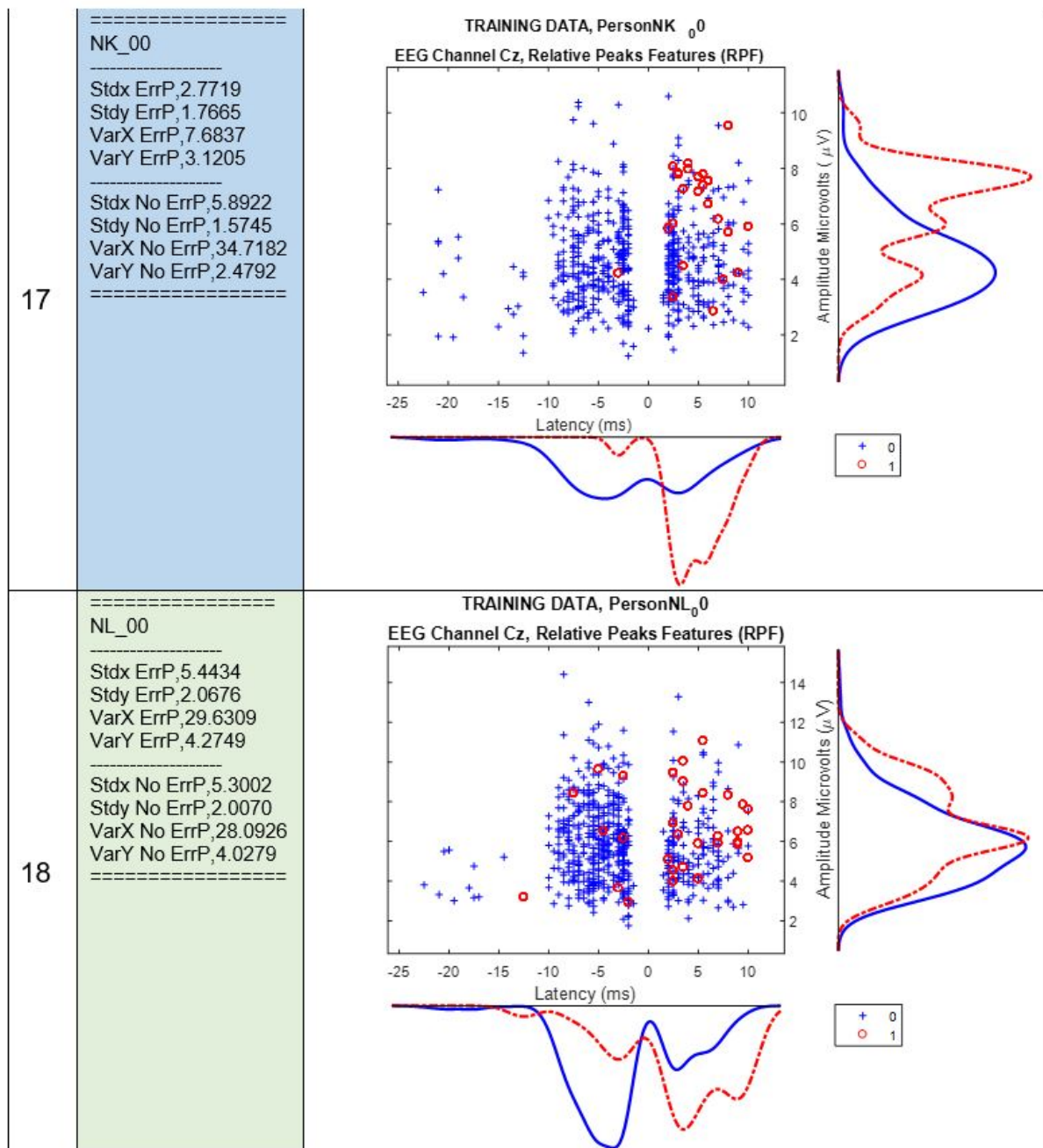


Figure B.21: Dataset young participants relative peak features ID 17,18

B.2.10 Young participants 19 and 20, showing their relative peak features(RPF)

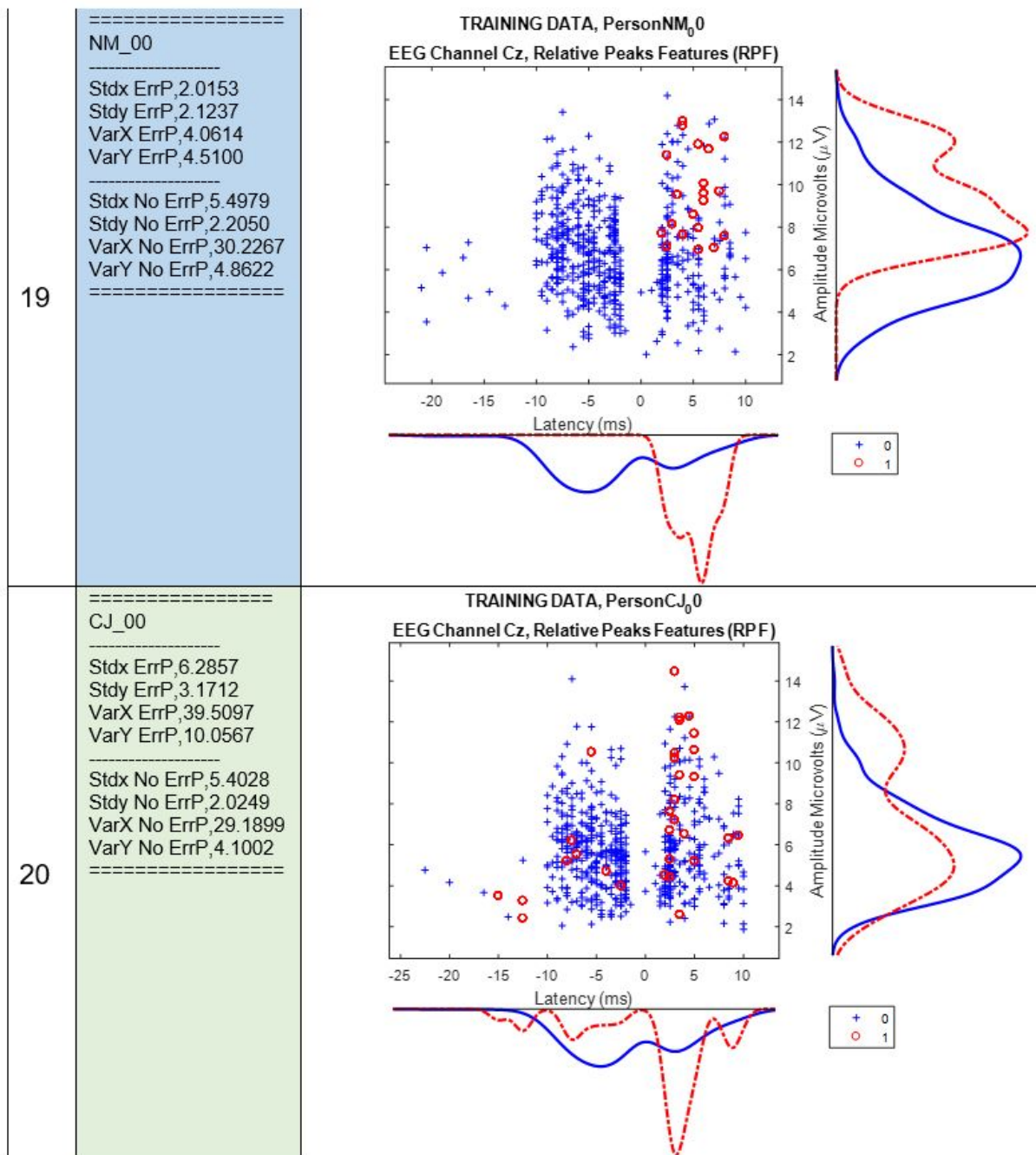


Figure B.22: Dataset young participants relative peak features ID 19,20

B.2.11 Young participants 21 and 22, showing their relative peak features(RPF)

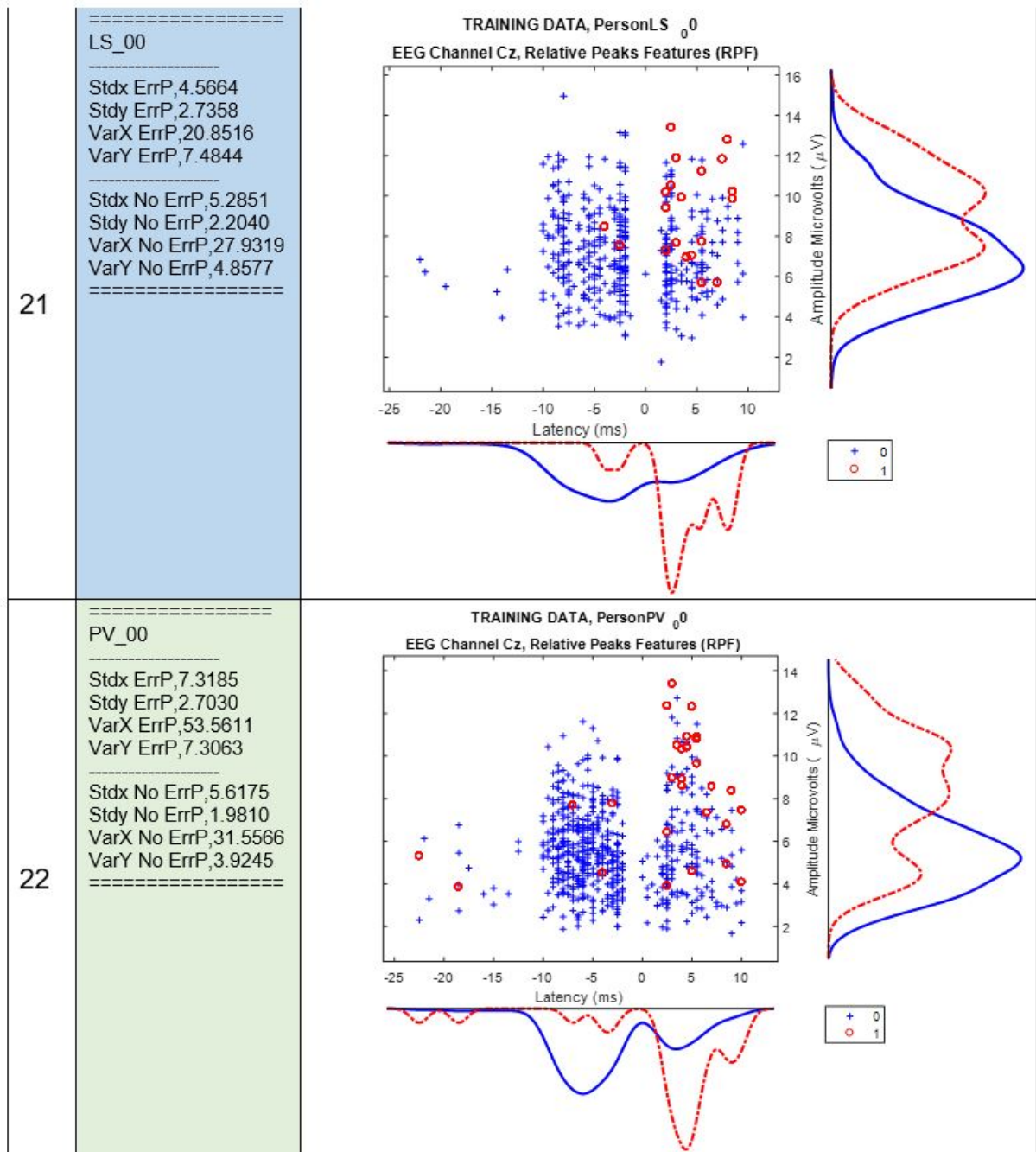


Figure B.23: Dataset young participants relative peak features ID 21,22

B.2.12 Young participants 23 and 24, showing their relative peak features(RPF)

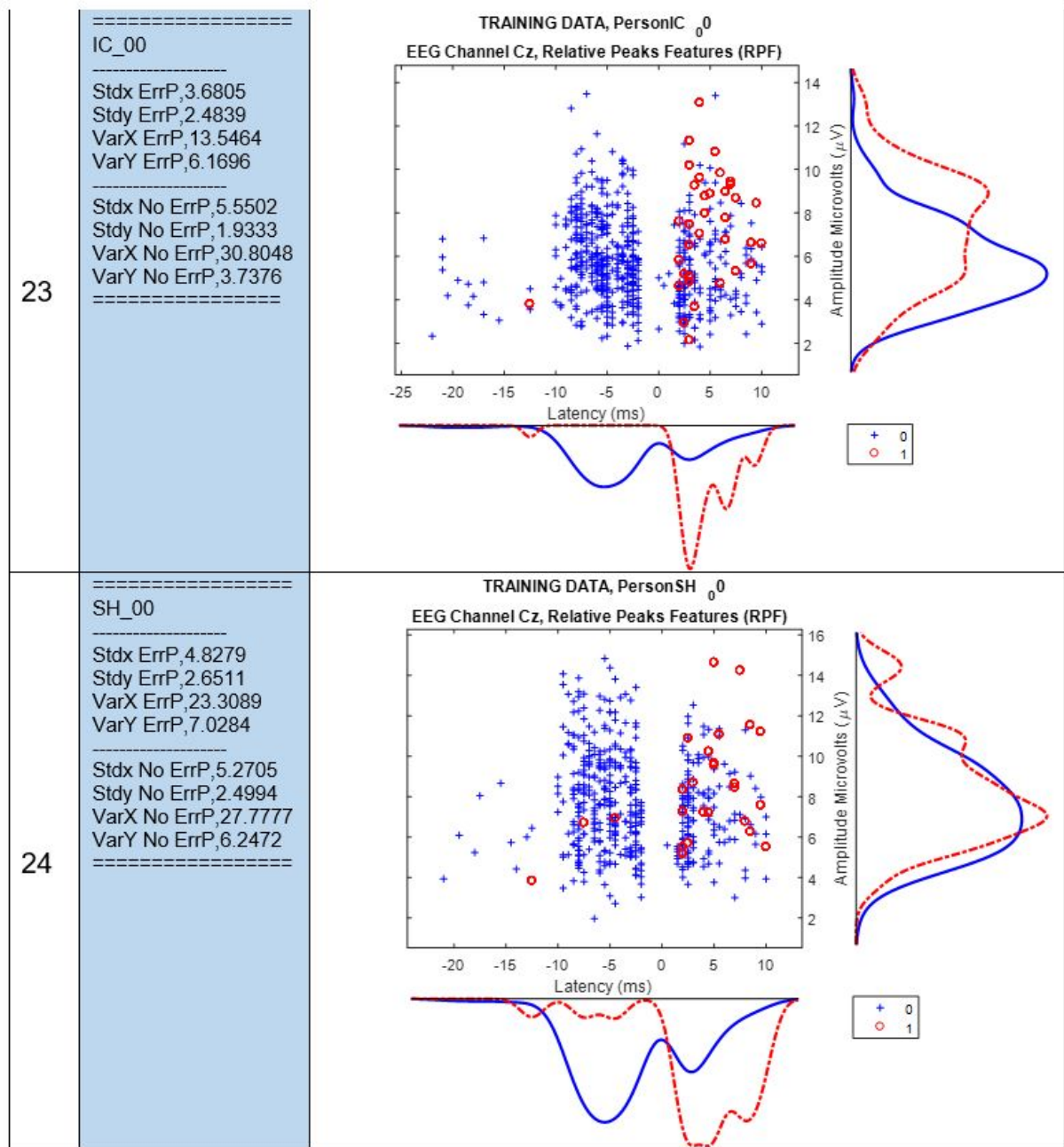


Figure B.24: Dataset young participants relative peak features ID 23,24

B.2.13 Young participants 25 and 26, showing their relative peak features(RPF)

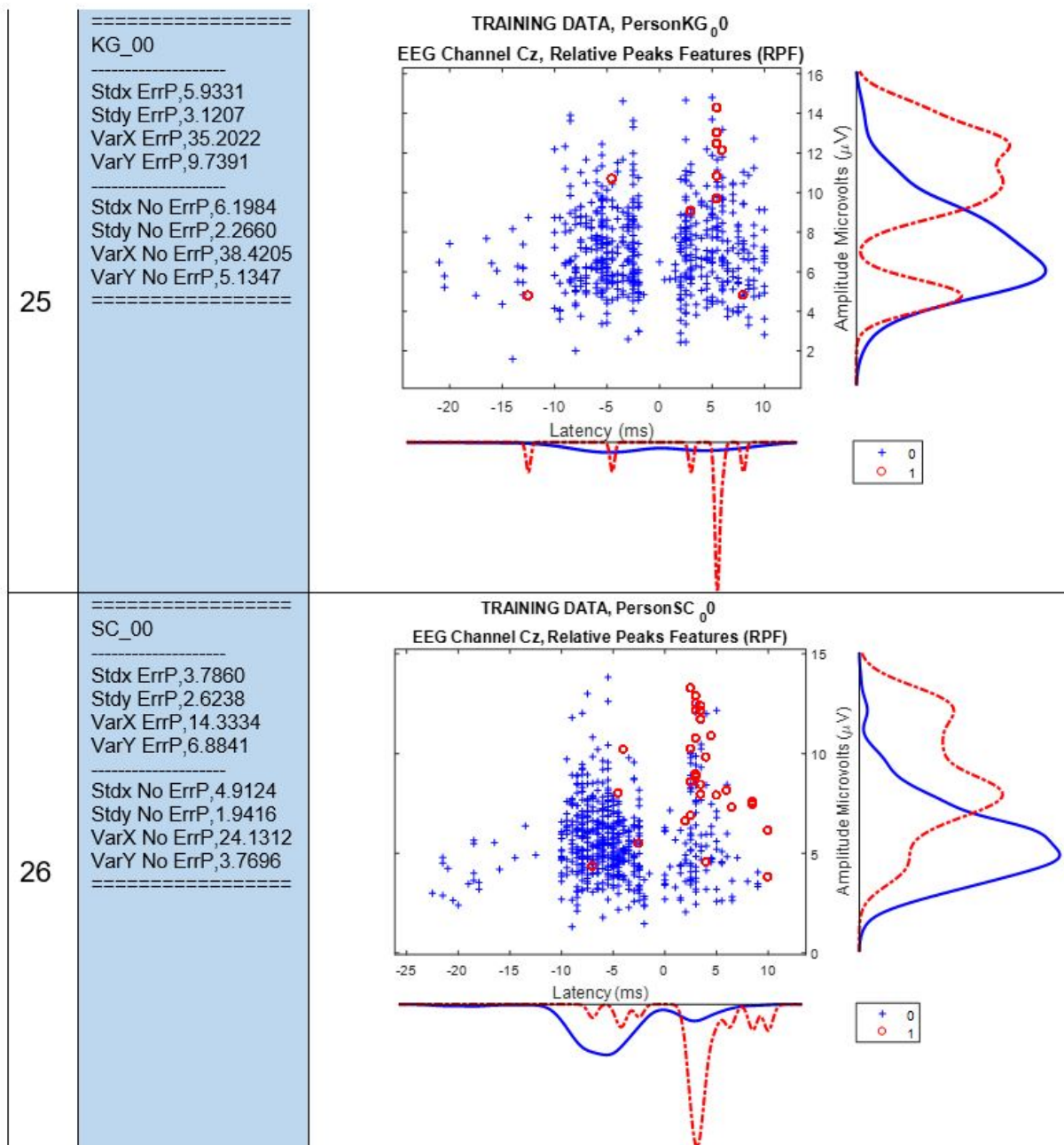


Figure B.25: Dataset young participants relative peak features ID 25,26

B.2.14 Young participant 27, showing their relative peak features(RPF)

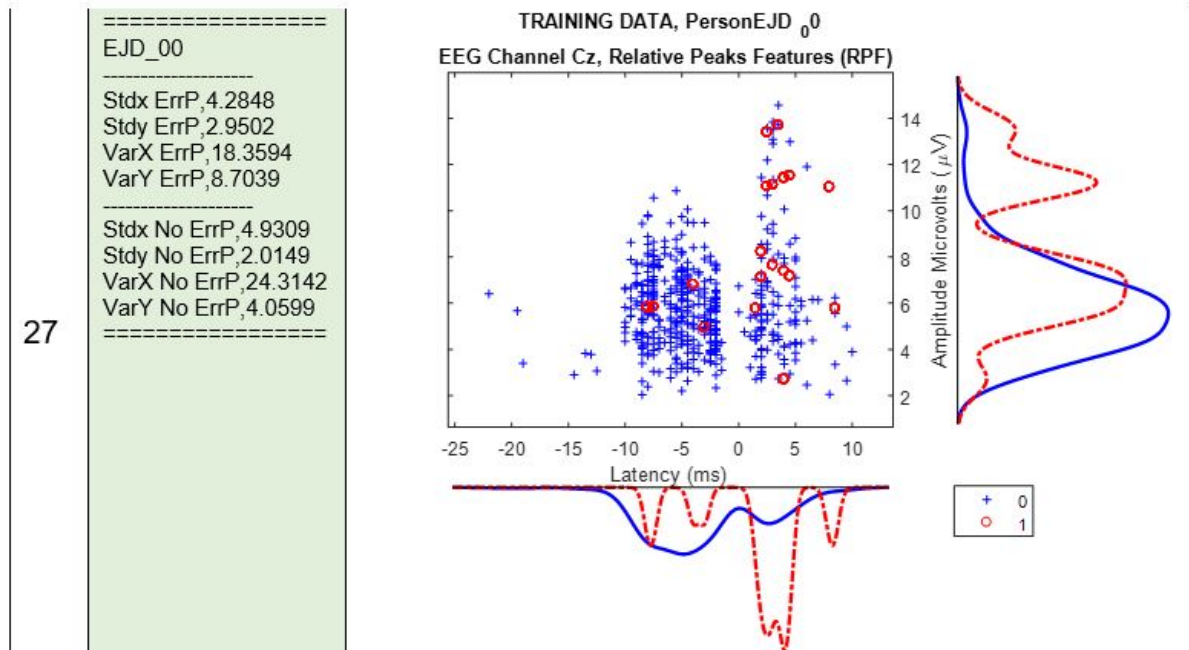


Figure B.26: Dataset young participants relative peak features ID 27

Appendix C

Standard deviation and variance

Standard deviation and variances

Table 1 Elderly participants standard deviation and variance : C.1

Table 2 Young participants Standard deviation and variance : C.2

C.1 Standard deviation and variances

C.1.1 Elderly participants Standard deviation and variance

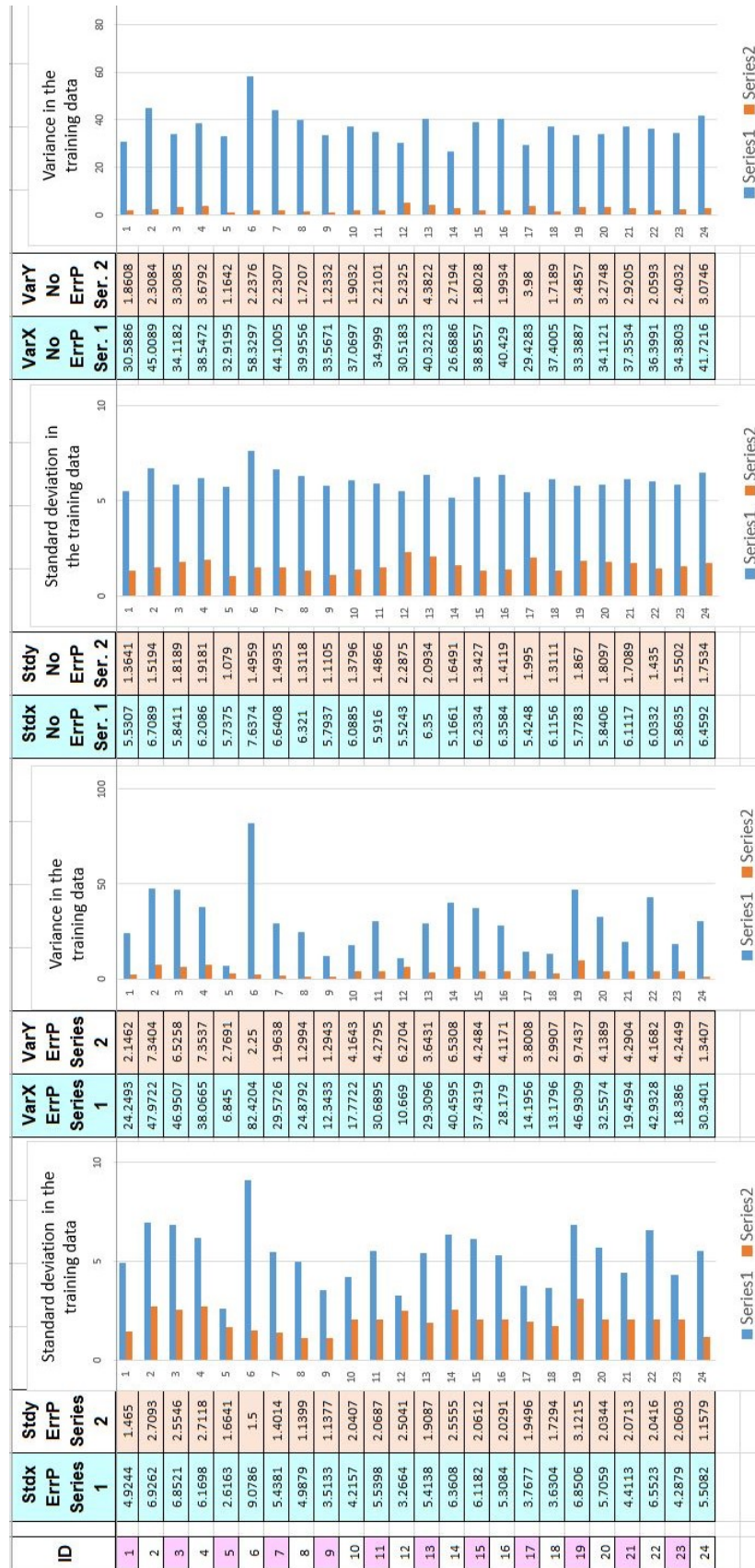


Figure C.1: Elderly participants Standard deviation and variance

C.1.2 Young participants Standard deviation and variance

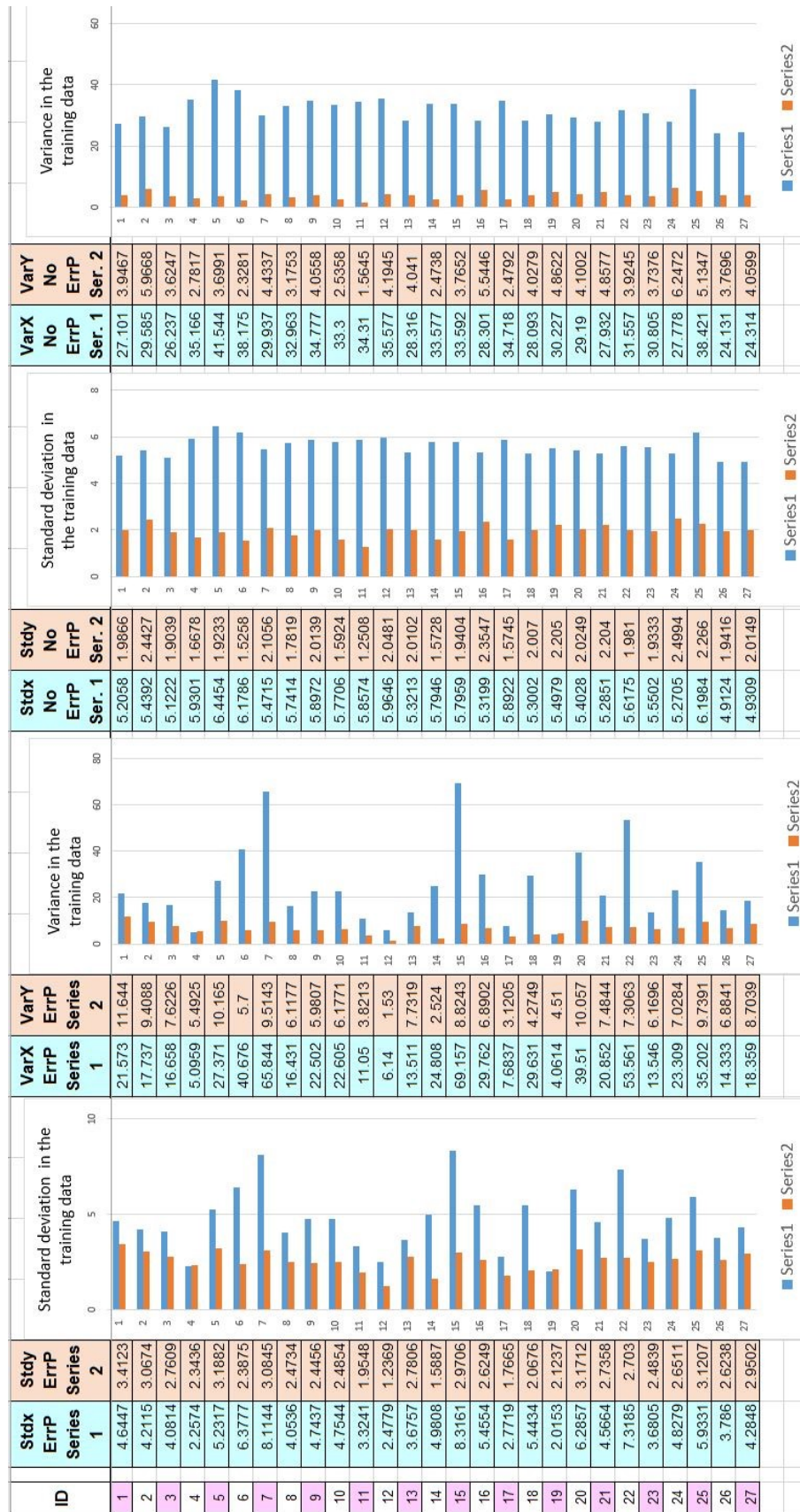


Figure C.2: Young participants Standard deviation and variance

Appendix D

Results Tables

Table 1 D.1 : Elderly participants classifier using temporal features results tables

Table 2 D.2 : Elderly participants classifier using relative features results tables

Young participants and temporal features

Table 3 D.3 : Young participants classifier using temporal features results tables P2-1

Table 4 D.4 : Young participants classifier using temporal features results tables P2-2

Young participants and relative peak features(RPF)

Table 5 D.5 : Young participants classifier using relative peak features results tables P2-1

Table 6 D.6 : Young participants classifier using relative peak features results tables P2-2

D.1.2 Elderly participants classifier using relative features results tables

Test 1 Person	Elderly Person Session ID	Error Accuracy [%]	Correct Accuracy [%]	Elderly Person Session ID	Error Accuracy [%]	Correct Accuracy [%]	Elderly Person Session ID	Error Accuracy [%]	Correct Accuracy [%]	Elderly Person Session ID	Error Accuracy [%]	Correct Accuracy [%]	Elderly Person Session ID	Error Accuracy [%]	Correct Accuracy [%]
1	MS_00	66.67	65.28	MS_00	61.54	63.89	MS_00	66.67	60.24	MS_00	66.67	60.24	MS_00	66.67	67.36
2	BL_00	51.35	86.92	BL_00	62.16	84.26	BL_00	51.35	86.92	BL_00	51.35	86.92	BL_00	51.35	85.81
3	CC2_00	69.23	78.15	CC2_00	76.92	70.28	CC2_00	76.92	77.97	CC2_00	76.92	77.97	CC2_00	69.23	77.97
4	CMD_00	50	74.77	CMD_00	50	57.85	CMD_00	50	65.54	CMD_00	50	65.54	CMD_00	50	76.62
5	ED_00	87.5	73.03	ED_00	62.5	79.56	ED_00	87.5	66.5	ED_00	87.5	66.5	ED_00	87.5	72.86
6	EOB_00	66.67	52.16	EOB_00	69.44	62.41	EOB_00	66.67	49.64	EOB_00	66.67	49.64	EOB_00	66.67	51.62
7	JD_00	77.78	63.73	JD_00	55.56	54.05	JD_00	77.78	55.63	JD_00	77.78	55.63	JD_00	66.67	66.37
8	JM_00	75	67.62	JM_00	100	43.43	JM_00	75	67.43	JM_00	75	67.43	JM_00	75	68.76
9	JS_00	86.96	67.13	JS_00	95.65	69.41	JS_00	86.96	64.34	JS_00	91.3	64.34	JS_00	86.96	67.31
10	KF_00	68.75	82.66	KF_00	81.25	79.51	KF_00	68.75	79.51	KF_00	75	79.68	KF_00	75	82.14
11	BC_00	100	80.37	BC_00	100	56.82	BC_00	100	76.64	BC_00	100	76.64	BC_00	100	80.56
12	KB_00	88.57	76.92	KB_00	85.71	76.57	KB_00	88.57	75.35	KB_00	94.29	75.35	KB_00	88.57	78.15
13	FM_00	88.24	71.69	FM_00	82.35	74.86	FM_00	88.24	71.88	FM_00	88.24	71.88	FM_00	88.24	74.49
14	DD_00	86.21	75.59	DD_00	86.21	73.4	DD_00	86.21	74.41	DD_00	86.21	74.41	DD_00	86.21	76.6
15	FMC_00	64.71	75.05	FMC_00	64.71	69.4	FMC_00	64.71	79.92	FMC_00	58.82	79.92	FMC_00	64.71	75.24
16	BJC_00	78.57	87.16	BJC_00	89.29	80.99	BJC_00	78.57	89.38	BJC_00	75	89.38	BJC_00	78.57	85.79
17	EID_00	100	73.63	EID_00	100	71.52	EID_00	100	72.15	EID_00	100	72.15	EID_00	100	73.84
18	SM_00	76.19	93.5	SM_00	76.19	91.52	SM_00	76.19	93.14	SM_00	80.95	93.14	SM_00	80.95	92.78
19	UL_00	77.78	70.02	UL_00	61.11	90.99	UL_00	77.78	80.5	UL_00	77.78	80.5	UL_00	77.78	70.44
20	MM_00	58.06	81.91	MM_00	70.97	77.53	MM_00	58.06	74.55	MM_00	77.42	74.55	MM_00	54.84	82.11
21	DMCH_00	70.59	78.62	DMCH_00	76.47	74.21	DMCH_00	70.59	81.13	DMCH_00	70.59	81.13	DMCH_00	70.59	80.19
22	GD_00	100	74.87	GD_00	100	65.77	GD_00	100	70.15	GD_00	100	70.15	GD_00	100	74.7
23	MM2_00	67.86	84.78	MM2_00	60.71	87.54	MM2_00	67.86	86.16	MM2_00	67.86	86.16	MM2_00	64.29	84.95
24	BL2_00	100	54.69	BL2_00	80	38.19	BL2_00	100	45.14	BL2_00	80	45.14	BL2_00	100	53.82
		AVERAGE	77.36%	74.59%	AVERAGE	77.03%	70.58%	AVERAGE	78.83%	72.68%	AVERAGE	77.41%	75.02%		
		Linear Discriminant Analysis (LDA)		Logistic Regression		KNN Coarse		SVM Linear		Subspace Discriminant					
		MATLAB classification learner		MATLAB classification learner		MATLAB classification learner		MATLAB classification learner		MATLAB classification learner		MATLAB classification learner			
		Relative Peak Features (RPF)		Relative Peak Features (RPF)		Relative Peak Features (RPF)		Relative Peak Features (RPF)		Relative Peak Features (RPF)		Relative Peak Features (RPF)			
		Date time : May 18, 2018		Date time : May 18, 2018		Date time : May 17, 2018		Date time : May 17, 2018		Date time : May 17, 2018		Date time : May 21, 2018			
		3:16:20.193		2:21:14.465 PM		4:34:21.922 PM		7:03:10.907 PM		7:03:10.907 PM		6:56:55.409 PM			

Figure D.2: Elderly participants classifier comparison results tables using relative features

D.1.3 Young participants classifier using temporal features results tables P2-1

ID	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]
1	AH_00	92	69.54	AH_00	88	87.65	AH_00	88	87.14	AH_00	96	87.82	AH_00	72	96.28
2	AI_00	86.11	76.04	AI_00	69.44	85.24	AI_00	72.22	84.9	AI_00	77.78	85.07	AI_00	66.67	92.71
3	AM_00	85.11	48.95	AM_00	72.34	67.07	AM_00	72.34	64.29	AM_00	72.34	65.51	AM_00	46.81	87.8
4	AU_00	78.57	87.06	AU_00	78.57	88.4	AU_00	71.43	84.2	AU_00	71.43	88.91	AU_00	57.14	97.31
5	BW_00	86.36	76.48	BW_00	77.27	81.39	BW_00	84.09	81.56	BW_00	79.55	79.7	BW_00	63.64	93.4
6	CD_00	73.08	69.26	CD_00	53.85	74.56	CD_00	57.69	68.9	CD_00	53.85	72.61	CD_00	42.31	93.99
7	COC2_00	75	89.38	COC2_00	75	93.25	COC2_00	70	93.93	COC2_00	60	95.78	COC2_00	60	94.77
8	DS_00	71.05	33.5	DS_00	52.63	63.3	DS_00	63.16	53.54	DS_00	55.26	64.14	DS_00	23.68	88.05
9	GG_00	70.59	70.97	GG_00	82.35	74.36	GG_00	80.39	74.36	GG_00	82.35	77.08	GG_00	58.82	80.98
10	HJ_00	73.33	40.94	HJ_00	66.67	69.8	HJ_00	53.33	71.81	HJ_00	66.67	64.43	HJ_00	20	87.25
11	JG_00	83.87	75.76	JG_00	64.52	78.31	JG_00	67.74	78.81	JG_00	54.84	79.15	JG_00	45.16	92.88
12	LF_00	82.35	84.07	LF_00	70.59	91.02	LF_00	76.47	90.68	LF_00	58.82	93.22	LF_00	35.29	97.63
13	MA_00	75	87.21	MA_00	50	90.91	MA_00	75	87.88	MA_00	50	92.59	MA_00	33.33	96.97
14	MB_00	78.95	68.91	MB_00	52.63	82.18	MB_00	57.89	83.19	MB_00	63.16	84.87	MB_00	47.37	92.1
15	NIK_00	73.68	82.26	NIK_00	86.84	84.12	NIK_00	84.21	82.94	NIK_00	78.95	85.47	NIK_00	52.63	96.11
16	KM_00	79.31	65.2	KM_00	68.97	74.95	KM_00	65.52	78.01	KM_00	72.41	74.38	KM_00	34.48	92.54
17	NK_00	73.91	76.01	NK_00	73.91	78.02	NK_00	73.91	83.05	NK_00	82.61	76.01	NK_00	56.52	93.46
18	NL_00	75.86	52.69	NL_00	62.07	75.59	NL_00	62.07	76.43	NL_00	58.62	75.08	NL_00	51.72	92.76
19	NM_00	82.14	79.97	NM_00	78.57	83	NM_00	75	84.01	NM_00	75	83.84	NM_00	71.43	95.45
20	CJ_00	87.88	45.8	CJ_00	72.73	63.81	CJ_00	75.76	60.55	CJ_00	69.7	66.72	CJ_00	48.48	87.65
21	LS_00	43.9	72.49	LS_00	39.02	81.14	LS_00	39.02	81.49	LS_00	34.15	81.31	LS_00	43.9	92.91
22	PV_00	90.91	71.55	PV_00	72.73	76.83	PV_00	77.27	77	PV_00	72.73	79.56	PV_00	68.18	89.95
23	IC_00	80.56	80.94	IC_00	75	86.34	IC_00	77.78	87.52	IC_00	75	88.03	IC_00	69.44	93.42
24	SH_00	64.44	50.7	SH_00	53.33	66.96	SH_00	53.33	67.48	SH_00	48.89	68.18	SH_00	33.33	87.94
25	KG_00	78.57	84.26	KG_00	64.29	84.95	KG_00	71.43	81.83	KG_00	71.43	83.56	KG_00	21.43	96.37
26	SC_00	86.27	84.09	SC_00	88.24	88.83	SC_00	90.2	88.16	SC_00	86.27	88.32	SC_00	72.55	95.09
27	EJD_00	70	81.48	EJD_00	75	87.21	EJD_00	75	86.87	EJD_00	80	87.88	EJD_00	62.5	96.63
Average		77.73	70.57	Average	69.06	79.97	Average	70.75	79.28	Average	68.44	80.34	Average	50.33	92.68
		<div style="border: 1px solid black; padding: 5px;"> Subspace discriminant MATLAB® Ver.(2018a)+App Classification learner Temporal Features Date time : June 03, 2018 2:32:08.358 PM </div>													
		<div style="border: 1px solid black; padding: 5px;"> Linear Discriminant LDA MATLAB® Ver.(2018a)+App Classification learner Temporal Features Date time : June 03, 2018 2:31:05.577 PM </div>													
		<div style="border: 1px solid black; padding: 5px;"> SVM Linear MATLAB® Ver.(2018a)+App Classification learner Temporal Features Date time : June 03, 2018 2:33:49.909 PM </div>													
		<div style="border: 1px solid black; padding: 5px;"> Logistic Regression MATLAB® Ver.(2018a)+App Classification learner Temporal Features Date time : June 03, 2018 2:31:09.403 PM </div>													
		<div style="border: 1px solid black; padding: 5px;"> SVM Coarse Gaussian MATLAB® Ver.(2018a)+App Classification learner Temporal Features Date time : June 03, 2018 2:33:25.777 PM </div>													

Figure D.3: Young participants classifier comparison results tables using temporal features

D.1.4 Young participants classifier using temporal features results tables P2-2

Young Person Session	Error Accuracy [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]
AH_00	84	94.25	AH_00	52	92.22
AI_00	77.78	90.1	AI_00	72.22	86.11
AM_00	46.81	89.9	AM_00	46.81	90.77
AU_00	67.86	94.29	AU_00	64.29	85.21
BW_00	75	87.31	BW_00	63.64	86.13
CD_00	53.85	84.63	CD_00	50	81.98
COC2_00	70	93.76	COC2_00	35	94.94
DS_00	26.32	82.49	DS_00	63.16	48.99
GG_00	56.86	86.25	GG_00	66.67	69.1
HJ_00	20	87.25	HJ_00	66.67	51.68
JG_00	80.65	71.02	JG_00	61.29	63.9
LF_00	70.59	91.69	LF_00	64.71	74.75
MA_00	75	91.92	MA_00	83.33	86.87
MB_00	68.42	91.26	MB_00	47.37	82.02
NIK_00	63.16	95.44	NIK_00	68.42	78.72
KM_00	65.52	82.22	KM_00	41.38	75.53
NK_00	56.52	92.28	NK_00	52.17	91.28
NL_00	65.52	88.22	NL_00	79.31	55.56
NM_00	78.57	95.45	NM_00	71.43	88.55
CJ_00	57.58	76.5	CJ_00	78.79	41.17
LS_00	56.1	89.62	LS_00	51.22	71.11
PV_00	77.27	91.82	PV_00	68.18	80.58
IC_00	72.22	89.04	IC_00	58.33	91.57
SH_00	46.67	83.22	SH_00	53.33	61.54
KG_00	92.86	72.15	KG_00	35.71	84.95
SC_00	84.31	94.92	SC_00	86.27	91.37
EJD_00	65	93.43	EJD_00	60	91.75
Average	64.98	88.16	Average	60.80	77.72

KNN Coarse MATLAB® Ver.(2018a)+App Classification learner Temporal Features Date time : June 03, 2018 2:30:21.203 PM	Decision Tree Simple MATLAB® Ver.(2018a)+App Classification learner Temporal Features Date time : June 03, 2018 7:35:23.275 PM
--	--

Figure D.4: Young participants classifier comparison results tables using temporal features

D.1.5 Young participants classifier using relative peak features results tables P2-1

ID	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]
1	AH_00	95.45	76.34	AH_00	95.45	74.44	AH_00	95.45	70.81	AH_00	95.45	74.61	AH_00	95.45	79.62
2	AI_00	75	79.91	AI_00	75	79.22	AI_00	75	82.88	AI_00	75	80.14	AI_00	70.83	82.65
3	AM_00	87.8	74.75	AM_00	82.93	76.77	AM_00	82.93	76.77	AM_00	78.05	80.4	AM_00	82.93	78.59
4	AU_00	92.86	84.6	AU_00	92.86	84.26	AU_00	96.43	79.53	AU_00	92.86	86.13	AU_00	92.86	85.45
5	BW_00	69.05	78.9	BW_00	61.9	81.86	BW_00	69.05	79.11	BW_00	61.9	80.8	BW_00	69.05	77.22
6	CD_00	76.92	85.19	CD_00	73.08	84.04	CD_00	76.92	85.38	CD_00	73.08	83.27	CD_00	61.54	87.69
7	COC2_00	89.47	85.89	COC2_00	89.47	83.82	COC2_00	89.47	82.27	COC2_00	89.47	83.99	COC2_00	94.74	88.81
8	DS_00	75	70.86	DS_00	75	70.86	DS_00	75	71.77	DS_00	75	72.68	DS_00	71.88	73.95
9	GG_00	68.09	75.62	GG_00	63.83	77.35	GG_00	65.96	77.54	GG_00	65.96	77.93	GG_00	65.96	78.5
10	HJ_00	75	71.52	HJ_00	75	72.5	HJ_00	75	63.67	HJ_00	75	72.18	HJ_00	75	69.72
11	JG_00	70.97	80.1	JG_00	67.74	80.61	JG_00	67.74	80.61	JG_00	67.74	81.29	JG_00	67.74	81.12
12	LF_00	70	86.75	LF_00	70	86.57	LF_00	60	88.38	LF_00	60	89.84	LF_00	80	89.47
13	MA_00	100	73.33	MA_00	100	72.94	MA_00	100	65.29	MA_00	100	74.12	MA_00	100	64.71
14	MB_00	89.47	77.15	MB_00	89.47	73.91	MB_00	89.47	74.55	MB_00	89.47	75.04	MB_00	89.47	77.47
15	NIK_00	68.75	87.77	NIK_00	65.63	83.09	NIK_00	62.5	90.65	NIK_00	65.63	80.04	NIK_00	62.5	94.06
16	KM_00	90	66.97	KM_00	90	66.97	KM_00	95	67.88	KM_00	90	67.65	KM_00	90	70.62
17	NK_00	86.96	75.92	NK_00	86.96	75.92	NK_00	78.26	79.26	NK_00	78.26	78.21	NK_00	73.91	81.9
18	NL_00	96.55	65.36	NL_00	96.55	65.02	NL_00	96.55	67.24	NL_00	96.55	65.36	NL_00	86.21	73.89
19	NM_00	95.45	83.97	NM_00	95.45	83.97	NM_00	95.45	85.97	NM_00	95.45	87.25	NM_00	95.45	84.7
20	CJ_00	85.19	66.44	CJ_00	85.19	66.61	CJ_00	88.89	67.3	CJ_00	85.19	66.26	CJ_00	74.07	78.89
21	LS_00	80	73.92	LS_00	86.67	75.12	LS_00	80	71.53	LS_00	86.67	75.84	LS_00	80	72.01
22	PV_00	81.82	85.45	PV_00	81.82	85.28	PV_00	77.27	86.97	PV_00	81.82	84.77	PV_00	77.27	85.96
23	IC_00	83.87	78.73	IC_00	83.87	77.7	IC_00	83.87	75.3	IC_00	80.65	79.25	IC_00	80.65	79.59
24	SH_00	80	71.23	SH_00	82.86	68.81	SH_00	82.86	68.81	SH_00	82.86	69.01	SH_00	80	77.26
25	KG_00	100	91.92	KG_00	100	91.56	KG_00	100	94.79	KG_00	100	91.2	KG_00	100	94.08
26	SC_00	79.17	88.68	SC_00	77.08	89.54	SC_00	70.83	90.91	SC_00	68.75	90.91	SC_00	77.08	90.91
27	EJD_00	88.46	80.67	EJD_00	88.46	79.83	EJD_00	88.46	76.89	EJD_00	88.46	79.62	EJD_00	88.46	80.46
Average		83.38	78.44	Average	82.68	78.10	Average	82.16	77.85	Average	81.45	78.81	Average	80.85	80.71

Subspace discriminant MATLAB® Ver.(2018a)+App Classification learner Relative Peak Features Date time : June 03, 2018 2:32:08.358 PM	Linear Discriminant LDA MATLAB® Ver.(2018a)+App Classification learner Relative Peak Features Date time June 03, 2018 2:28:55.219 PM	SVM Linear MATLAB® Ver.(2018a)+App Classification learner Relative Peak Features Date time : June 03, 2018 8:07:29.710 PM	Logistic Regression MATLAB® Ver.(2018a)+App Classification learner Relative Peak Features Date time : June 03, 2018 7:37:09.297 PM	SVM Coarse Gaussian MATLAB® Ver.(2018a)+App Classification learner Relative Peak Features Date time : June 03, 2018 7:39:27.022 PM
--	--	---	--	--

Figure D.5: Young participants classifier comparison results tables using relative peak features

D.1.6 Young participants classifier using relative peak features results tables P2-2

Young Person Session	Error Accuracy y [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy y [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy y [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy y [%]	Correct Accuracy [%]	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]	Test 1 Person	Young Person Session	Error Accuracy [%]	Correct Accuracy [%]
AH_00	95.45	79.62	AH_00	90.91	92.57	AH_00	90.91	87.74	AH_00	90.91	89.81	1	AH_00	100.00	94.13	100.00	89.81	
AI_00	70.83	82.65	AI_00	75	79.91	AI_00	70.83	84.25	AI_00	75	83.56	2	AI_00	91.67	78.44	91.67	83.56	
AM_00	82.93	78.59	AM_00	85.37	75.96	AM_00	75.61	84.24	AM_00	82.93	82.02	3	AM_00	79.49	83.08	82.93	82.02	
AU_00	92.86	85.45	AU_00	89.29	78.17	AU_00	96.43	84.26	AU_00	92.86	91.88	4	AU_00	89.29	95.06	92.86	91.88	
BW_00	69.05	77.22	BW_00	45.24	89.66	BW_00	47.62	85.86	BW_00	30.95	95.36	5	BW_00	76.92	81.53	30.95	95.36	
CD_00	61.54	87.69	CD_00	65.38	88.08	CD_00	61.54	87.31	CD_00	61.54	89.62	6	CD_00	60.00	82.88	61.54	89.62	
COC2_00	94.74	88.81	COC2_00	94.74	88.12	COC2_00	78.95	91.74	COC2_00	68.42	93.46	7	COC2_00	78.95	89.75	68.42	93.46	
DS_00	71.88	73.95	DS_00	78.13	68.12	DS_00	68.75	78.51	DS_00	81.25	70.86	8	DS_00	75.00	75.90	81.25	70.86	
GG_00	65.96	78.5	GG_00	63.83	78.12	GG_00	65.96	78.89	GG_00	70.21	76.78	9	GG_00	84.78	81.12	70.21	76.78	
HJ_00	75	69.72	HJ_00	58.33	79.87	HJ_00	75	71.69	HJ_00	58.33	86.25	10	HJ_00	66.67	74.42	58.33	86.25	
JG_00	67.74	81.12	JG_00	70.97	80.1	JG_00	67.74	86.9	JG_00	64.52	91.67	11	JG_00	70.00	86.27	64.52	91.67	
LF_00	80	89.47	LF_00	70	87.48	LF_00	60	91.83	LF_00	40	94.74	12	LF_00	87.5	92.37	40	94.74	
MA_00	100	64.71	MA_00	100	69.8	MA_00	70	71.45	MA_00	100	71.96	13	MA_00	90	91.72	100	71.96	
MB_00	89.47	77.47	MB_00	84.21	76.18	MB_00	78.95	75.85	MB_00	84.21	78.44	14	MB_00	83.33	80.52	84.21	78.44	
NIK_00	62.5	94.06	NIK_00	78.13	88.13	NIK_00	81.25	89.21	NIK_00	65.63	90.83	15	NIK_00	77.42	90.38	65.63	90.83	
KM_00	90	70.62	KM_00	95	66.74	KM_00	90	67.88	KM_00	85	67.65	16	KM_00	84.21	81.53	85	67.65	
NK_00	73.91	81.9	NK_00	69.67	84.53	NK_00	91.3	73.46	NK_00	65.22	72.76	17	NK_00	73.91	82.56	65.22	72.76	
NL_00	86.21	73.89	NL_00	96.55	62.8	NL_00	89.66	68.94	NL_00	79.31	58.53	18	NL_00	89.66	75.14	79.31	58.53	
NM_00	95.45	84.7	NM_00	95.45	82.51	NM_00	86.36	84.15	NM_00	90.91	87.8	19	NM_00	95.24	86.00	90.91	87.8	
CJ_00	74.07	78.89	CJ_00	92.59	60.9	CJ_00	66.67	83.22	CJ_00	70.37	71.45	20	CJ_00	80.77	72.56	70.37	71.45	
LS_00	80	72.01	LS_00	86.67	72.49	LS_00	80	75.6	LS_00	80	76.08	21	LS_00	80	81.93	80	76.08	
PV_00	77.27	85.96	PV_00	72.73	88.66	PV_00	81.82	87.65	PV_00	77.27	77.83	22	PV_00	71.43	84.39	77.27	77.83	
IC_00	80.65	79.59	IC_00	74.19	79.42	IC_00	80.65	75.81	IC_00	77.42	77.53	23	IC_00	83.33	84.89	77.42	77.53	
SH_00	80	77.26	SH_00	82.86	68.81	SH_00	80	71.43	SH_00	74.29	72.84	24	SH_00	85.71	73.95	74.29	72.84	
KG_00	100	94.08	KG_00	100	76.12	KG_00	81.82	88.15	KG_00	100	94.08	25	KG_00	90	90.68	100	94.08	
SC_00	77.08	90.91	SC_00	66.67	92.97	SC_00	79.17	88.85	SC_00	58.33	93.83	26	SC_00	80	86.93	58.33	93.83	
EJD_00	88.46	80.46	EJD_00	80.77	85.08	EJD_00	88.46	85.29	EJD_00	80.77	86.55	27	EJD_00	84.62	86.67	80.77	86.55	
Average	80.85	80.71	Average	80.85	78.83	Average	80.33	80.78	Average	75.37	81.86	Average	74.28	82.38	AVERAGE	81.85	83.88	

SVM Coarse Gaussian
MATLAB® Ver.(2018a)+App
Classification learner
Relative Peak Features
Date time : June 03, 2018
7:39:27.022 PM

KNN Coarse
MATLAB® Ver.(2018a)+App
Classification learner
Relative Peak Features
Date time : June 03, 2018
7:36:43.203 PM

Quadratic Discriminant QDA
MATLAB® Ver.(2018a)+App
Classification learner
Relative Peak Features
Date time : June 04, 2018
5:51:11.471 PM

SVM Medium Gaussian
MATLAB® Ver.(2018a)+App
Classification learner
Relative Peak Features
Date time : June 03, 2018
8:10:32.722 PM

Decision Tree Simple
MATLAB® Ver.(2018a)+App
Classification learner
Relative Peak Features
Date time : June 03, 2018
7:35:23.275 PM

Adaptive Decision Surface ADS
MATLAB® Ver.(2018a)+
Relative Peak Features (RPF)
Data 29/01/2018
File
test1_participant_RESULTS_COLLE
CTION_v2_29_01_2018_adaptive_v2

Figure D.6: Young participants classifier comparison results tables using relative peak features

Appendix E

Results of 23 motor imagery classifier

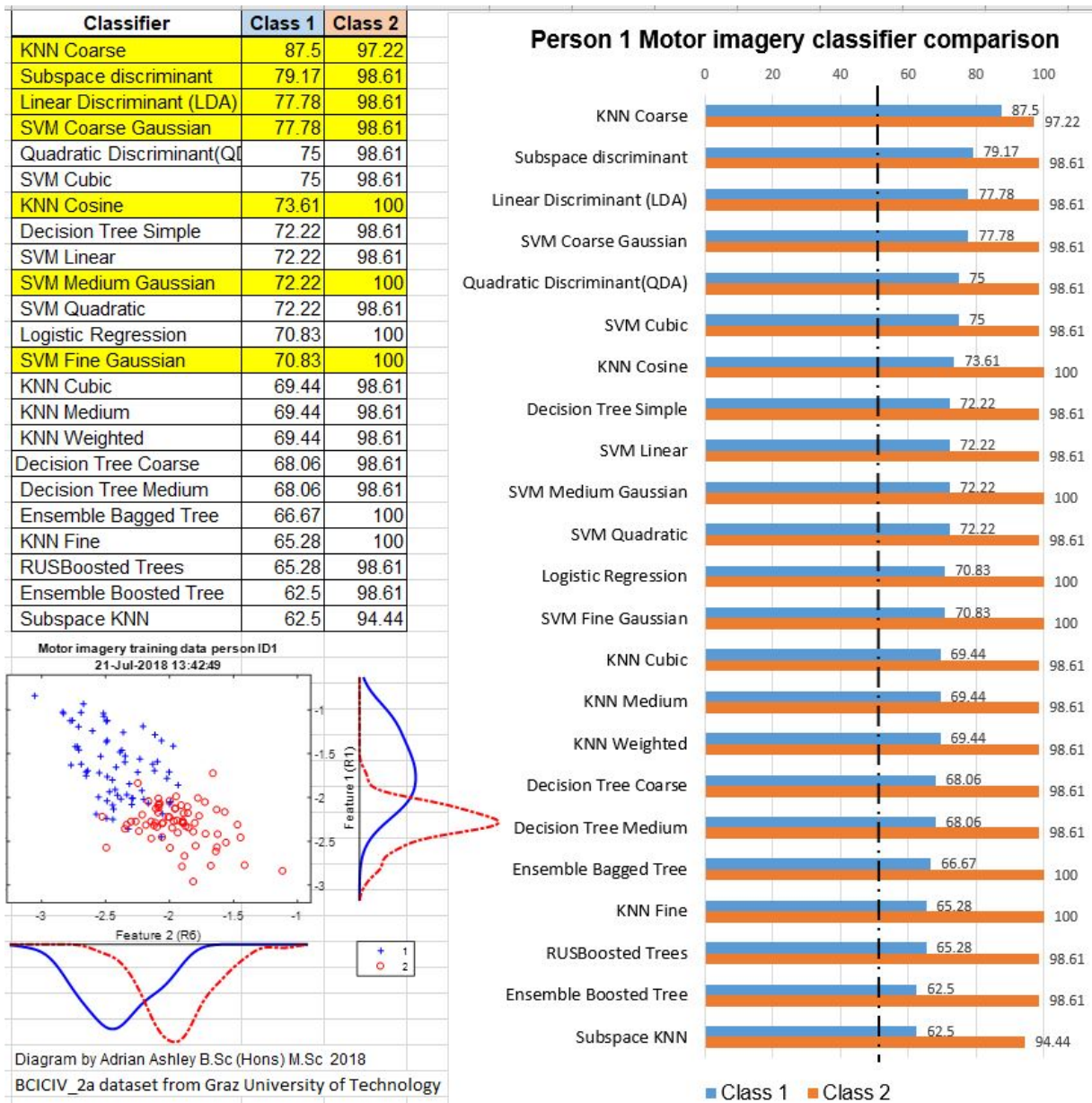


Figure E.1: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison for person 1

Classifier	Class 1	Class 2
SVM Linear	97.22	0
SVM Medium Gaussian	80.56	36.11
SVM Quadratic	75	40.28
SVM Fine Gaussian	73.61	47.22
Decision Tree Simple	72.22	50
Ensemble Boosted Tree	70.83	47.22
RUSBoosted Trees	70.83	51.39
SVM Coarse Gaussian	69.44	48.61
SVM Cubic	69.44	44.44
Ensemble Bagged Tree	63.89	47.22
KNN Cubic	63.89	61.11
KNN Weighted	63.89	44.44
Linear Discriminant (LDA)	63.89	50
Logistic Regression	63.89	50
Decision Tree Coarse	62.5	50
KNN Cosine	62.5	54.17
KNN Fine	62.5	50
KNN Medium	61.11	62.5
Subspace discriminant	61.11	48.61
KNN Coarse	59.72	50
Quadratic Discriminant(QDA)	59.72	52.78
Decision Tree Medium	58.33	56.94
Subspace KNN	31.94	69.44

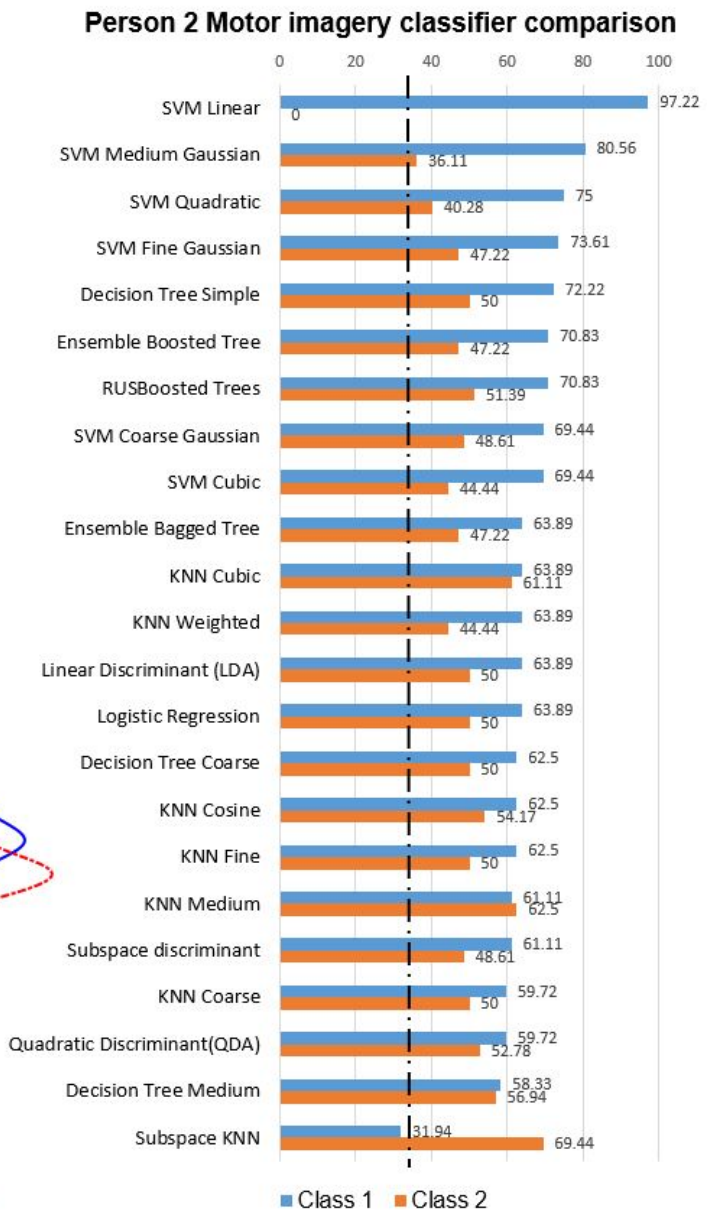
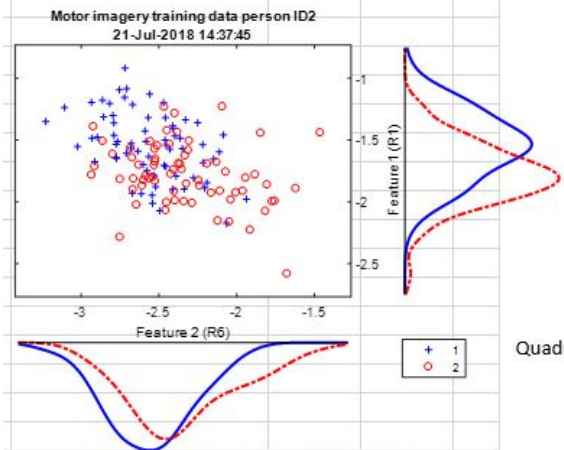


Figure E.2: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison for person 2

Classifier	Class 1	Class 2
KNN Coarse	100	87.5
KNN Cubic	100	97.22
KNN Medium	100	97.22
Quadratic Discriminant(QDA)	100	95.83
Subspace discriminant	100	90.28
SVM Coarse Gaussian	100	90.28
SVM Cubic	100	97.22
SVM Linear	100	95.83
SVM Medium Gaussian	100	95.83
SVM Quadratic	100	95.83
KNN Cosine	98.61	94.44
SVM Fine Gaussian	98.61	91.67
Decision Tree Coarse	97.22	97.22
Decision Tree Medium	97.22	97.22
Decision Tree Simple	97.22	97.22
Ensemble Bagged Tree	97.22	94.44
KNN Fine	97.22	90.28
KNN Weighted	97.22	95.83
Linear Discriminant (LDA)	97.22	95.83
Logistic Regression	97.22	95.83
Subspace KNN	79.17	94.44
Ensemble Boosted Tree	0	100
RUSBoosted Trees	0	100

Motor Imagery training data person ID3
 Diagram by A Ashley, Reference C-SP software by Dr Mahnaz Arvanah
 19-Jul-2018 16:10:12

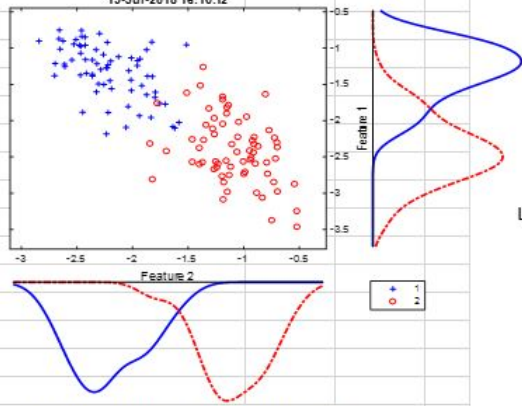


Diagram by Adrian Ashley B.Sc (Hons) M.Sc 2018
 BCICIV_2a dataset from Graz University of Technology

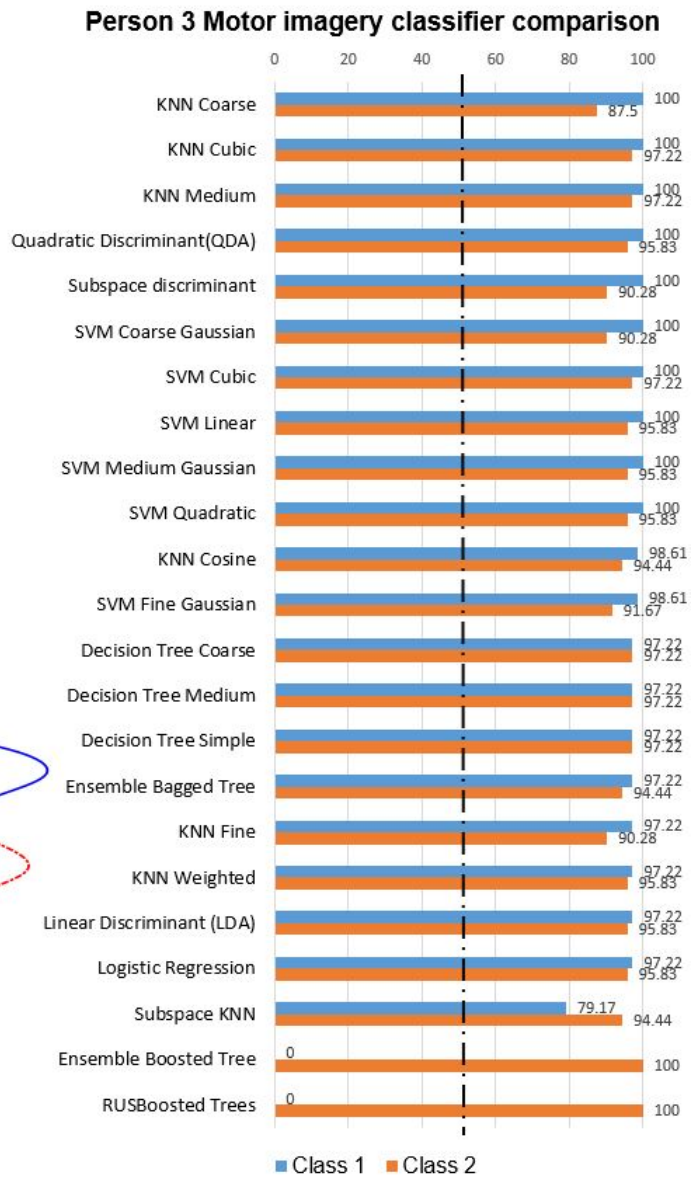


Figure E.3: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison for person 3

Classifier	Class 1	Class 2
SVM Cubic	86.11	50
KNN Coarse	84.72	59.72
Subspace discriminant	83.33	62.5
SVM Coarse Gaussian	81.94	66.67
SVM Fine Gaussian	81.94	52.78
Ensemble Boosted Tree	79.17	55.56
Linear Discriminant (LDA)	79.17	66.67
Logistic Regression	79.17	68.06
Quadratic Discriminant(QDA)	79.17	62.5
SVM Linear	79.17	68.06
SVM Medium Gaussian	79.17	66.67
Decision Tree Coarse	77.78	61.11
Decision Tree Medium	77.78	61.11
KNN Weighted	77.78	62.5
SVM Quadratic	77.78	68.06
KNN Fine	75	61.11
RUSBoosted Trees	73.61	61.11
Ensemble Bagged Tree	72.22	68.06
KNN Cosine	72.22	73.61
KNN Medium	70.83	77.78
Decision Tree Simple	69.44	77.78
KNN Cubic	69.44	79.17
Subspace KNN	51.39	79.17

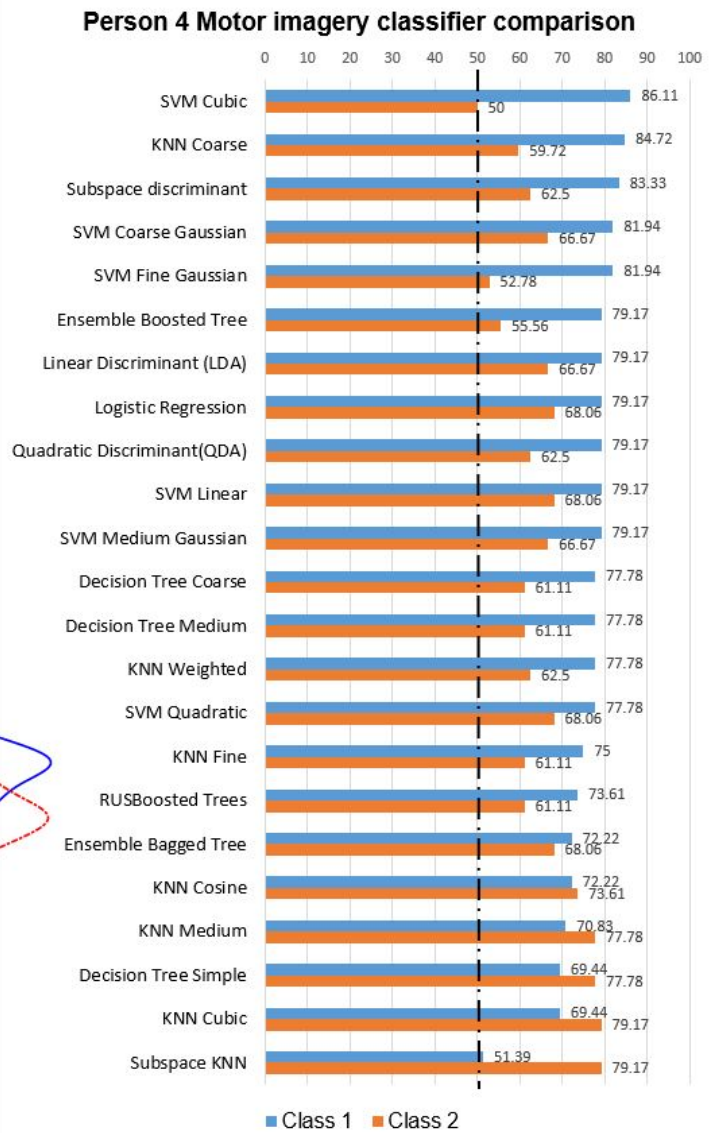
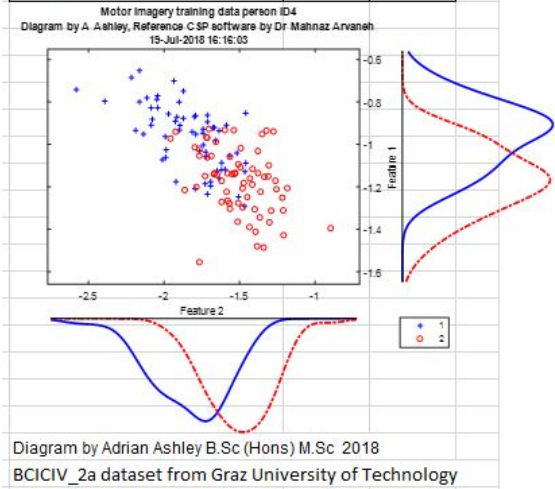


Figure E.4: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison for person 4

Classifier	Class 1	Class 2
Ensemble Bagged Tree	86.11	52.78
Ensemble Boosted Tree	86.11	45.83
SVM Quadratic	86.11	62.5
KNN Weighted	84.72	50
Logistic Regression	84.72	62.5
Quadratic Discriminant(QDA)	84.72	59.72
Subspace discriminant	84.72	62.5
SVM Coarse Gaussian	84.72	63.89
SVM Fine Gaussian	84.72	47.22
SVM Linear	84.72	62.5
SVM Medium Gaussian	84.72	61.11
RUSBoosted Trees	83.33	52.78
SVM Cubic	83.33	63.89
KNN Coarse	81.94	63.89
KNN Cubic	81.94	68.06
KNN Fine	81.94	41.67
KNN Medium	81.94	68.06
Linear Discriminant (LDA)	81.94	62.5
Decision Tree Simple	77.78	50
Subspace KNN	77.78	50
Decision Tree Coarse	76.39	50
Decision Tree Medium	76.39	50
KNN Cosine	72.22	63.89

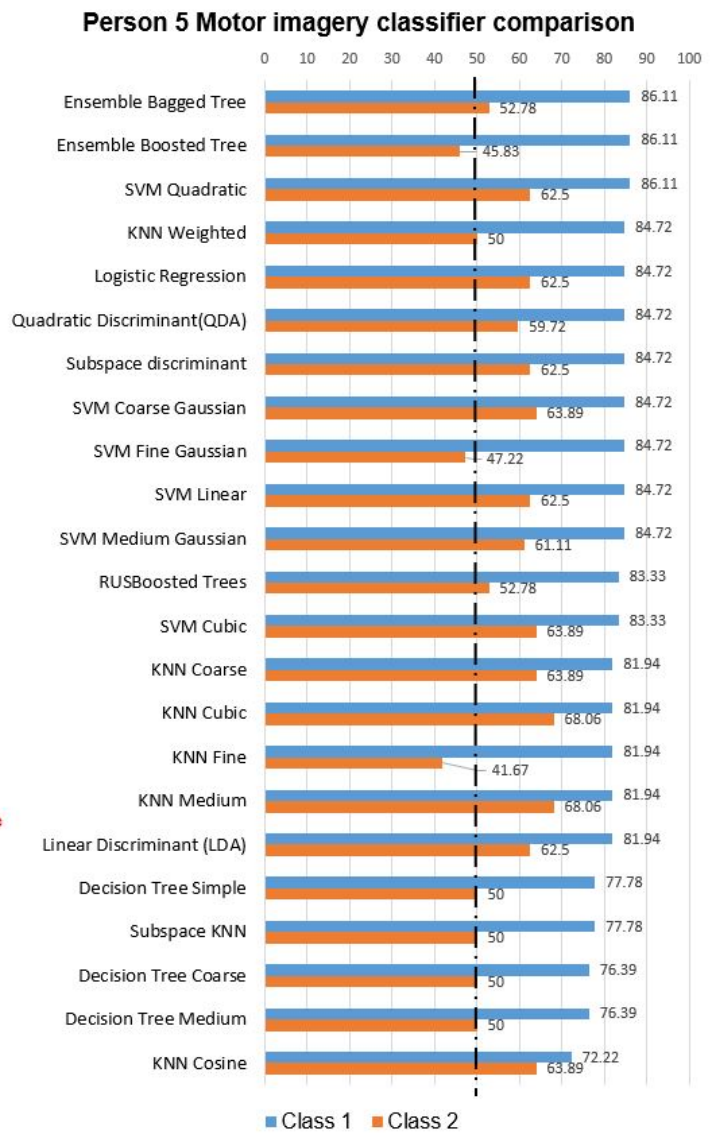
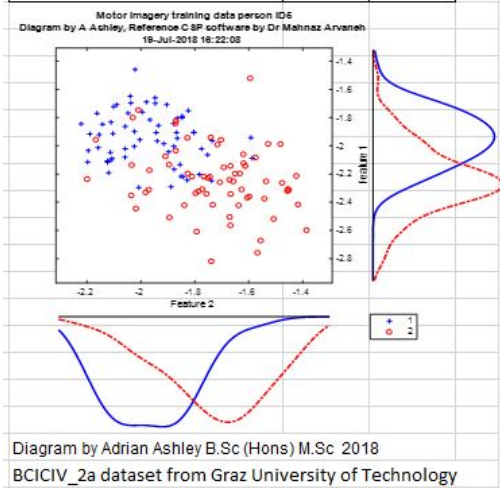


Figure E.5: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison for person 5

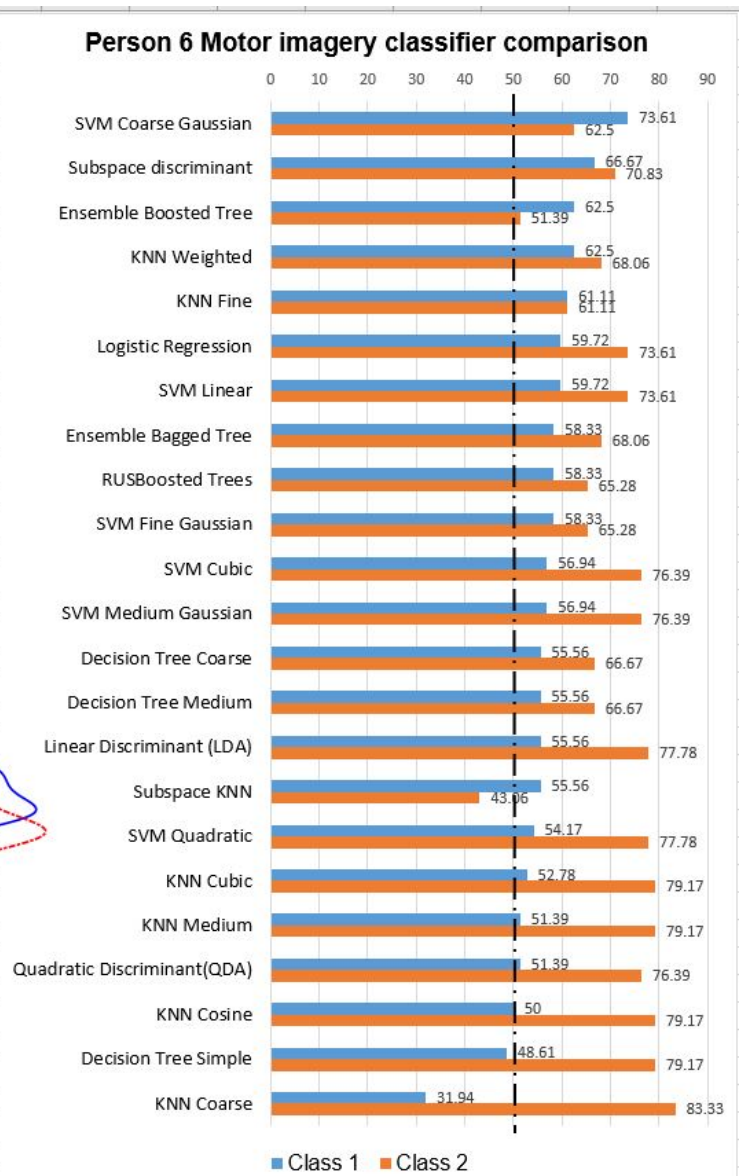
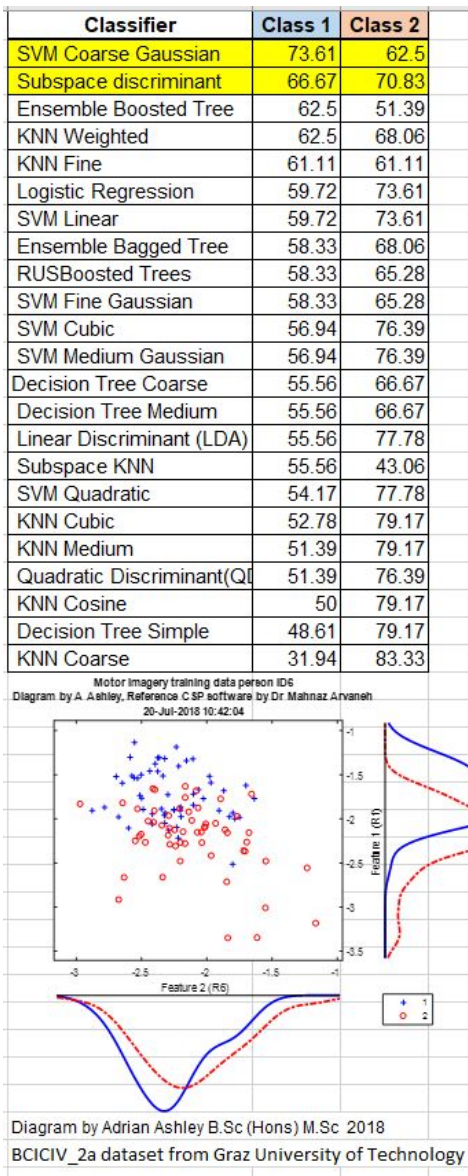


Figure E.6: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison for person 6

Classifier	Class 1	Class 2
RUSBoosted Trees	77.78	76.39
KNN Cubic	76.39	88.89
KNN Medium	76.39	90.28
Logistic Regression	75	86.11
Quadratic Discriminant(QDA)	75	87.5
Subspace discriminant	75	90.28
Ensemble Boosted Tree	73.61	86.11
Linear Discriminant (LDA)	73.61	87.5
SVM Coarse Gaussian	73.61	86.11
SVM Linear	73.61	86.11
SVM Medium Gaussian	73.61	86.11
Decision Tree Coarse	72.22	87.5
Decision Tree Medium	72.22	87.5
KNN Coarse	72.22	90.28
KNN Cosine	72.22	91.67
KNN Weighted	72.22	87.5
Ensemble Bagged Tree	70.83	88.89
Decision Tree Simple	69.44	84.72
SVM Cubic	69.44	84.72
SVM Quadratic	69.44	90.28
SVM Fine Gaussian	66.67	87.5
KNN Fine	65.28	81.94
Subspace KNN	63.89	80.56

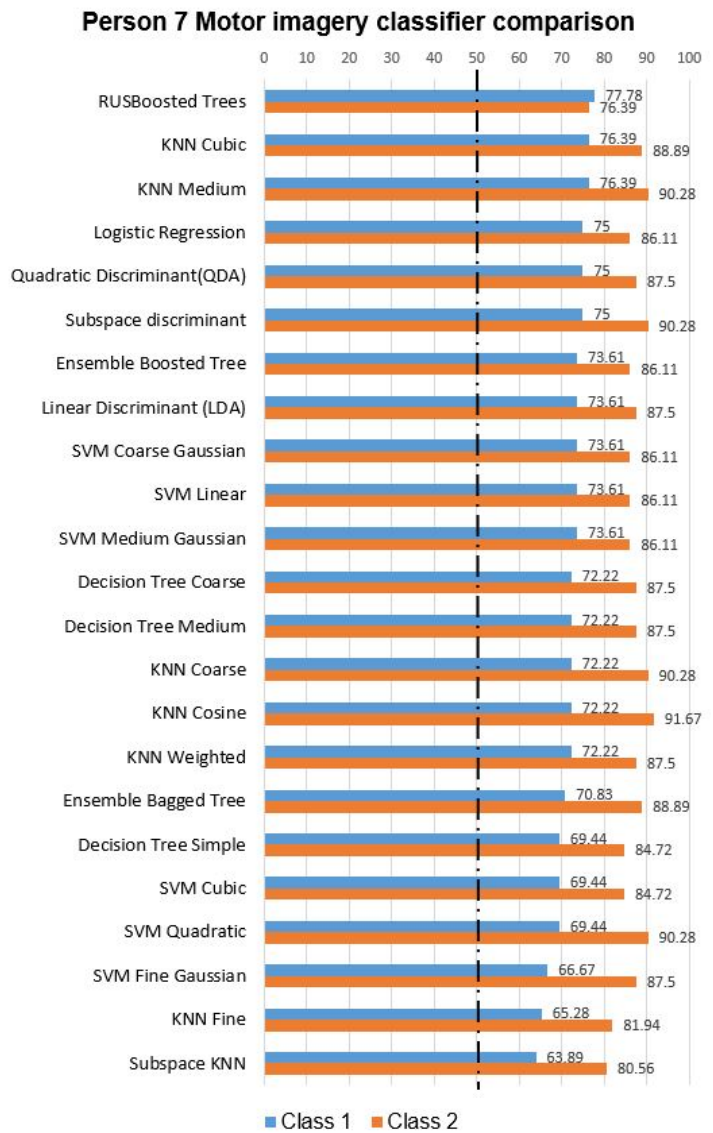
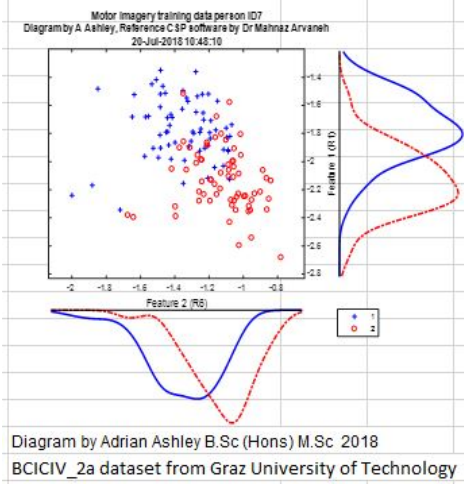


Figure E.7: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison for person 7

Classifier	Class 1	Class 2
Subspace KNN	95.83	86.11
KNN Cosine	93.06	100
KNN Cubic	93.06	100
KNN Fine	93.06	100
KNN Medium	93.06	100
KNN Weighted	93.06	100
Quadratic Discriminant(QDA)	93.06	100
SVM Coarse Gaussian	93.06	100
SVM Cubic	93.06	100
SVM Fine Gaussian	93.06	98.61
SVM Linear	93.06	100
SVM Medium Gaussian	93.06	100
SVM Quadratic	93.06	100
Subspace discriminant	90.28	100
Linear Discriminant (LDA)	88.89	100
KNN Coarse	87.5	100
Logistic Regression	84.72	100
Decision Tree Coarse	80.56	100
Decision Tree Medium	80.56	100
Decision Tree Simple	80.56	100
Ensemble Bagged Tree	80.56	100
Ensemble Boosted Tree	0	100
RUSBoosted Trees	0	100

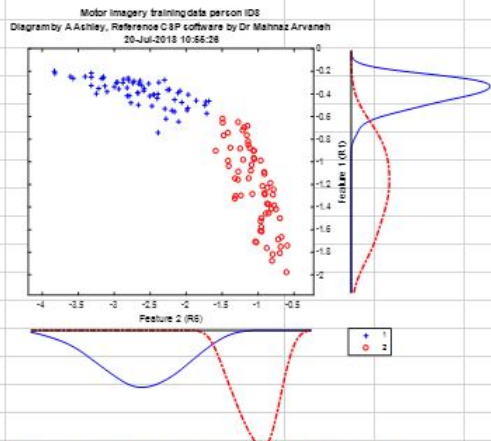


Diagram by Adrian Ashley B.Sc (Hons) M.Sc 2018
BCICIV_2a dataset from Graz University of Technology

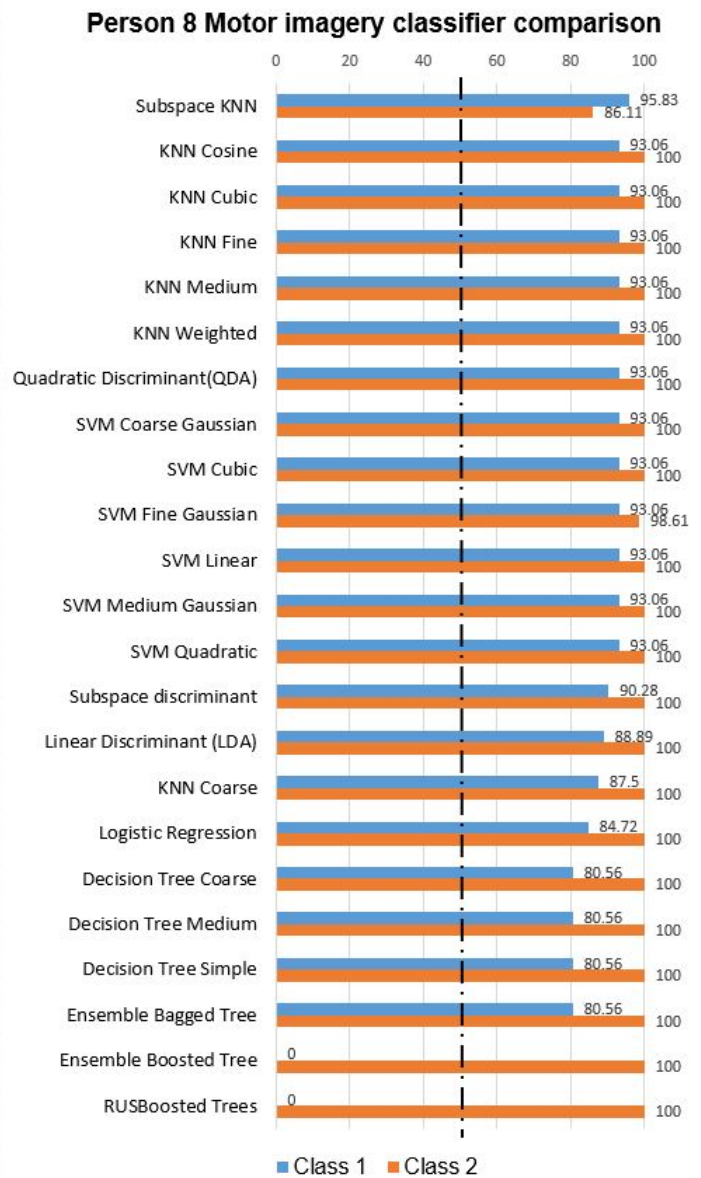


Figure E.8: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison for person 8

Classifier	Class 1	Class 2
Ensemble Boosted Tree	100	0
RUSBoosted Trees	100	0
Linear Discriminant (LDA)	98.61	72.22
Quadratic Discriminant(QDA)	98.61	81.94
Subspace discriminant	98.61	75
SVM Linear	98.61	75
Ensemble Bagged Tree	97.22	81.94
KNN Coarse	97.22	75
KNN Cosine	97.22	80.56
KNN Cubic	97.22	83.33
KNN Fine	97.22	69.44
KNN Medium	97.22	84.72
KNN Weighted	97.22	79.17
Logistic Regression	97.22	83.33
SVM Coarse Gaussian	97.22	77.78
SVM Cubic	97.22	84.72
SVM Fine Gaussian	97.22	81.94
SVM Medium Gaussian	97.22	77.78
SVM Quadratic	97.22	80.56
Decision Tree Coarse	95.83	88.89
Decision Tree Medium	95.83	88.89
Decision Tree Simple	95.83	88.89
Subspace KNN	88.89	87.5

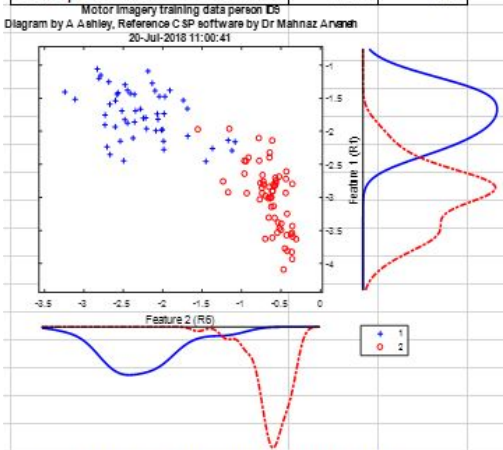


Diagram by Adrian Ashley B.Sc (Hons) M.Sc 2018
BCICIV_2a dataset from Graz University of Technology

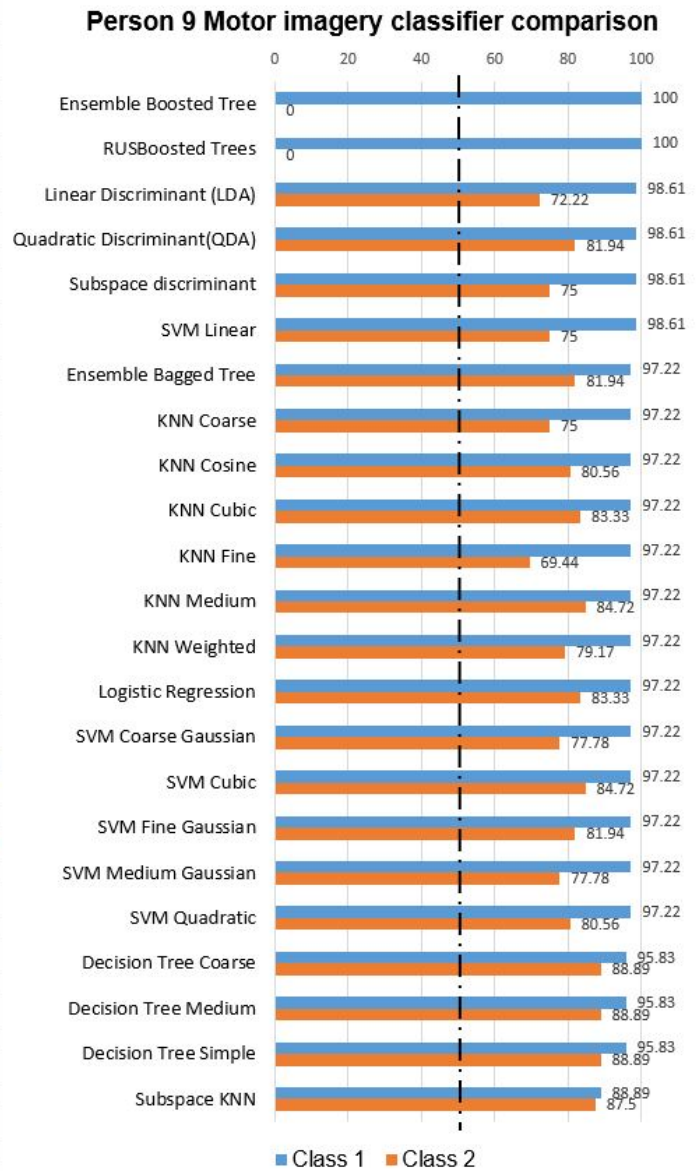


Figure E.9: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison for person 9

	Person 1		Person 2		Person 3		Person 4		Person 5		Person 6		Person 7		Person 8		Person 9		Average		
	Classifier	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2
1	Decision Tree Coarse	68.06	98.61	62.5	50	97.22	97.22	77.78	61.11	76.39	50	55.56	66.67	72.22	87.5	80.56	100	95.83	88.89	76.24	77.78
2	Decision Tree Medium	68.06	98.61	58.33	56.94	97.22	97.22	77.78	61.11	76.39	50	55.56	66.67	72.22	87.5	80.56	100	95.83	88.89	75.77	78.55
3	Decision Tree Simple	72.22	98.61	72.22	50	97.22	97.22	69.44	77.78	77.78	50	48.61	79.17	69.44	84.72	80.56	100	95.83	88.89	75.92	80.71
4	Ensemble Bagged Tree	66.67	100	61.11	45.83	97.22	94.44	72.22	68.06	86.11	52.78	58.33	68.06	70.83	88.89	80.56	100	97.22	81.94	76.70	77.78
5	Ensemble Boosted Tree	62.5	98.61	70.83	47.22	0	100	79.17	55.56	86.11	45.83	62.5	51.39	73.61	86.11	0	100	100	0	59.41	64.97
6	KNN Coarse	87.5	97.22	59.72	50	100	87.5	84.72	59.72	81.94	63.89	31.94	83.33	72.22	90.28	87.5	100	97.22	75	78.08	78.55
7	KNN Cosine	73.61	100	62.5	54.17	98.61	94.44	72.22	73.61	72.22	63.89	50	79.17	72.22	91.67	93.06	100	97.22	80.56	76.85	81.95
8	KNN Cubic	69.44	98.61	63.89	61.11	100	97.22	69.44	79.17	81.94	68.06	52.78	79.17	76.39	88.89	93.06	100	97.22	83.33	78.24	83.95
9	KNN Fine	65.28	100	61.11	62.5	97.22	90.28	75	61.11	81.94	41.67	61.11	61.11	65.28	81.94	93.06	100	97.22	69.44	77.62	72.84
10	KNN Medium	69.44	98.61	61.11	62.5	100	97.22	70.83	77.78	81.94	68.06	51.39	79.17	76.39	90.28	93.06	100	97.22	84.72	77.93	84.26
11	KNN Weighted	69.44	98.61	63.89	44.44	97.22	95.83	77.78	62.5	84.72	50	62.5	68.06	72.22	87.5	93.06	100	97.22	79.17	79.78	76.23
12	Linear Discriminant (LDA)	77.78	98.61	63.89	50	97.22	95.83	79.17	66.67	81.94	62.5	55.56	77.78	73.61	87.5	88.89	100	98.61	72.22	79.63	79.01
13	Logistic Regression	70.83	100	68.06	50	97.22	95.83	79.17	68.06	84.72	62.5	59.72	73.61	75	86.11	84.72	100	97.22	83.33	79.17	79.94
14	Quadratic Discriminant(QDA)	75	98.61	68.06	52.78	100	95.83	79.17	62.5	84.72	59.72	51.39	76.39	75	87.5	93.06	100	98.61	81.94	79.63	79.47
15	RUSBoosted Trees	65.28	98.61	68.06	52.78	0	100	73.61	61.11	83.33	52.78	58.33	65.28	77.78	76.39	0	100	100	0	58.49	67.44
16	Subspace discriminant	79.17	98.61	63.89	48.61	100	90.28	83.33	62.5	84.72	62.5	66.67	70.83	75	90.28	90.28	100	98.61	75	82.41	77.62
17	Subspace KNN	62.5	94.44	63.89	48.61	79.17	94.44	51.39	79.17	77.78	50	55.56	43.06	63.89	80.56	95.83	86.11	88.89	87.5	67.44	76.08
18	SVM Coarse Gaussian	77.78	98.61	69.44	48.61	100	90.28	81.94	66.67	84.72	63.89	73.61	62.5	73.61	86.11	93.06	100	97.22	77.78	83.49	77.16
19	SVM Cubic	75	98.61	69.44	44.44	100	97.22	86.11	50	83.33	63.89	56.94	76.39	69.44	84.72	93.06	100	97.22	84.72	81.17	77.78
20	SVM Fine Gaussian	70.83	100	73.61	47.22	98.61	91.67	81.94	52.78	84.72	47.22	58.33	65.28	66.67	87.5	93.06	98.61	97.22	81.94	80.55	74.69
21	SVM Linear	72.22	98.61	66.67	52.78	100	95.83	79.17	68.06	84.72	62.5	59.72	73.61	73.61	86.11	93.06	100	98.61	75	80.86	79.17
22	SVM Medium Gaussian	72.22	100	80.56	36.11	100	95.83	79.17	66.67	84.72	61.11	56.94	76.39	73.61	86.11	93.06	100	97.22	77.78	81.94	77.78
23	SVM Quadratic	72.22	98.61	75	40.28	100	95.83	77.78	68.06	86.11	62.5	54.17	77.78	69.44	90.28	93.06	100	97.22	80.56	80.56	79.32

Diagram by
Adrian Ashley B.Sc (Hons) M.Sc
22/07/2018
BCICIV_2a dataset from Graz
University of Technology

Figure E.10: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison table

	Classifier	Class 1	Class 2
1	SVM Coarse Gaussian	83.49	77.16
2	Subspace discriminant	82.41	77.62
3	SVM Medium Gaussian	81.94	77.78
4	SVM Cubic	81.17	77.78
5	SVM Linear	80.86	79.17
6	SVM Quadratic	80.56	79.32
7	SVM Fine Gaussian	80.55	74.69
8	KNN Weighted	79.78	76.23
9	Linear Discriminant (LDA)	79.63	79.01
10	Quadratic Discriminant(QDA)	79.63	79.47
11	Logistic Regression	79.17	79.94
12	KNN Cubic	78.24	83.95
13	KNN Coarse	78.08	78.55
14	KNN Medium	77.93	84.26
15	KNN Fine	77.62	72.84
16	KNN Cosine	76.85	81.95
17	Ensemble Bagged Tree	76.70	77.78
18	Decision Tree Coarse	76.24	77.78
19	Decision Tree Simple	75.92	80.71
20	Decision Tree Medium	75.77	78.55
21	Subspace KNN	67.44	76.08
22	Ensemble Boosted Tree	59.41	64.97
23	RUSBoosted Trees	58.49	67.44

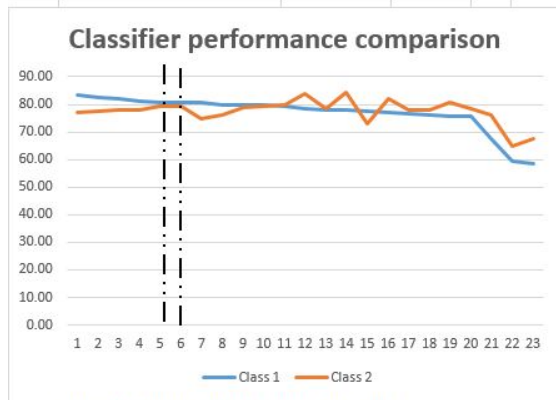


Diagram by Adrian Ashley B.Sc (Hons) M.Sc 2018
 BCICIV_2a dataset from Graz University of Technology

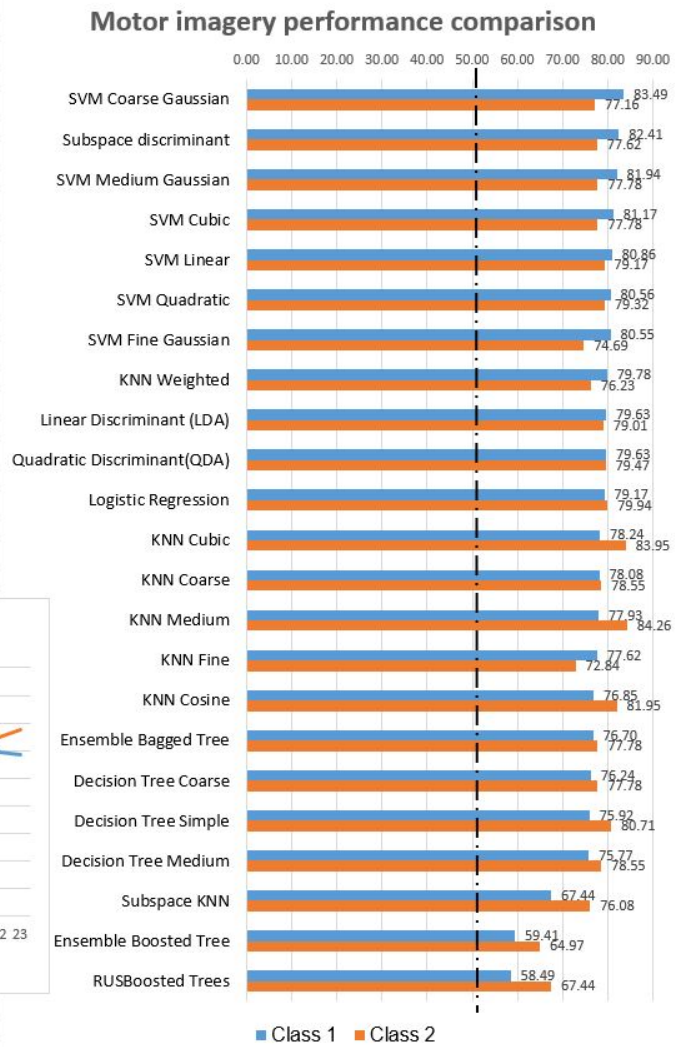


Figure E.11: BCI competition 2008 Graz dataset (A) Motor imagery classifier performance comparison

Appendix F

Results of classifier comparisons using a novel feature extraction method

F.0.1 Results of the statistical analysis using a subspace discriminant classifier using young participants

A 2(features) X 2(Actions: error action vs correct) repeated ANOVA was performed on the classification results of young participants. The statistical results showed a statistically significant difference between the different methods of feature extraction, temporal features compared to relative peak features. In this statistical analysis we used ANOVA with repeated measures using software called Statistical Package for the Social Sciences (SPSS) from the IBM® Corporation, a **subspace discriminant classifier** and the dataset from **young participants**. (Please see column **subspace discriminant classifier** in:

Table 3 D.4 : Young participants classifier using temporal features results tables P2-2

Table 5 D.6 : Young participants classifier using relative peak features results tables P2-2

F.0.2 Tests of Within-Subjects Effects

The tests of Within-Subjects Effects indicate if there was an overall significant difference between the means of the different feature extraction methods.

Table F.1: Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Subspace_Discriminant	Sphericity Assumed	1233.092	1	1233.092	20.482	<0.0001
	Greenhouse-Geisser	1233.092	1.000	1233.092	20.482	<0.0001

From above table F.1 we are able to discover the **F** value for the “**feature extraction method**” factor, and its significance level

We observe the values in the “**Greenhouse-Geisser**” row because the Mauchly’s Test of Sphericity to evaluated the assumption of sphericity. The test of Sphericity indicated the assumption of sphericity has been violated. The violation of sphericity occurs because of the unequal variance in the different permutation of the groups shown in [124].

Table F.2: Mauchly’s Test of Sphericity

Within Subjects Effect	Mauchly’s W	Approx. Chi-Square	df	Sig.
Subspace_Discriminant	1.000	.000	0	.

We report that when using an ANOVA with repeated measures with a **Greenhouse-Geisser** correction, the mean scores for a **subspace discriminant classifier** using the different feature extraction methods were statistically significantly different.

$$(F(1.0, 20.482) = 19.482, \rho < 0.0005$$

(Please see Table Tests of Within-Subjects Effects F.1 and Table, Tests of Within-Subjects Contrasts F.6)

Table F.3: Tests of Within-Subjects Contrasts

Source	Subspace_Discriminant	Type III Sum of Squares	df	Mean Square	F	Sig.
Subspace Discriminant	Temporal Features vs Relative Peak Features	1233.092	1	1233.092	20.482	<0.0001

We further report that when using an ANOVA with repeated measures “Tests of Within-Subject Contrasts” using a **subspace discriminant classifier** using “Temporal features vs Relative Peak features” were statistically significantly different.

$$(F(1.0, 20.482) = 19.482, \rho < 0.0005$$

Profile Plots

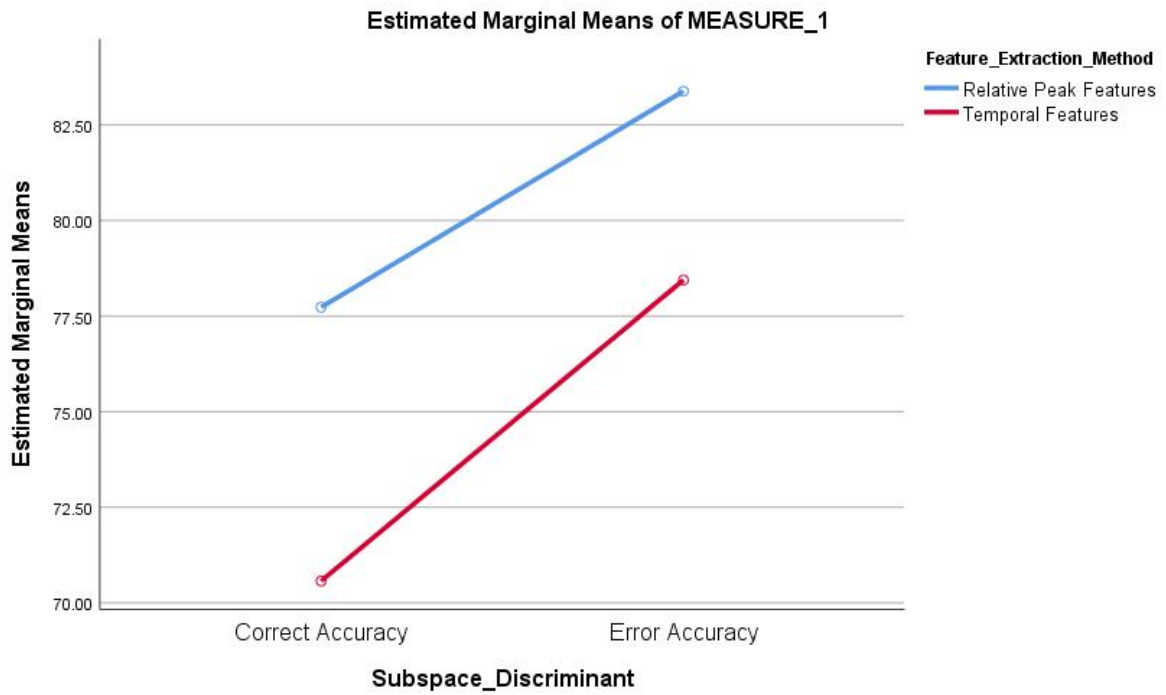


Figure F.1: Statistical analysis using a subspace discriminant classifier

F.0.3 Results of the statistical analysis using a subspace discriminant classifier using elderly participants

A 2(features) X 2(Actions: error action vs correct) repeated ANOVA was performed on the classification results of young participants. The statistical results showed a statistically significant difference between the different methods of feature extraction, temporal features compared to relative peak features. The data was processed with a **subspace discriminant classifier** and the dataset from **elderly participants**. (Please see column **subspace discriminant classifier** in: Table 1 D.1 : Elderly participants classifier using temporal features results tables Table 2 D.2 : Elderly participants classifier using relative features results tables

F.0.4 Tests of Within-Subjects Effects

The tests of Within-Subjects Effects indicate if there was an overall significant difference between the means of the different feature extraction methods.

Table F.4: Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Subspace_Discriminant	Sphericity Assumed	2217.412	1	2217.412	23.144	<0.0001
	Greenhouse-Geisser	2217.412	1.000	2217.412	23.144	<0.0001

From above table F.4 we are able to discover the **F** value for the “**feature extraction method**” factor, and its significance level

We observe the values in the “**Greenhouse-Geisser**” row because the Mauchly’s Test of Sphericity to evaluated the assumption of sphericity. The test of Sphericity indicated the assumption of sphericity has been violated. The violation of sphericity occurs because of the unequal variance in the different permutation of the groups shown in [124].

Table F.5: Mauchly’s Test of Sphericity

Within Subjects Effect	Mauchly’s W	Approx. Chi-Square	df	Sig.
Subspace_Discriminant	1.000	.000	0	.

We report that when using an ANOVA with repeated measures with a **Greenhouse-Geisser** correction, the mean scores for a **subspace discriminant classifier** using the different feature extraction methods were statistically significantly different.

$$(F(1.0, 23.144) = 22.144, \rho < 0.0001$$

(Please see Table Tests of Within-Subjects Effects F.1 and Table, Tests of Within-Subjects Contrasts F.6)

Table F.6: Tests of Within-Subjects Contrasts

Source	Subspace_Discriminant	Type III Sum of Squares	df	Mean Square	F	Sig.
Subspace Discriminant	Temporal Features vs Relative Peak Features	2217.412	1	2217.412	23.144	<0.0001

We further report that when using an ANOVA with repeated measures “Tests of Within-Subject Contrasts” using a **subspace discriminant classifier** using “Temporal features vs Relative Peak features” were statistically significantly different.

$$(F(1.0, 23.144) = 22.144, \rho < 0.0001$$

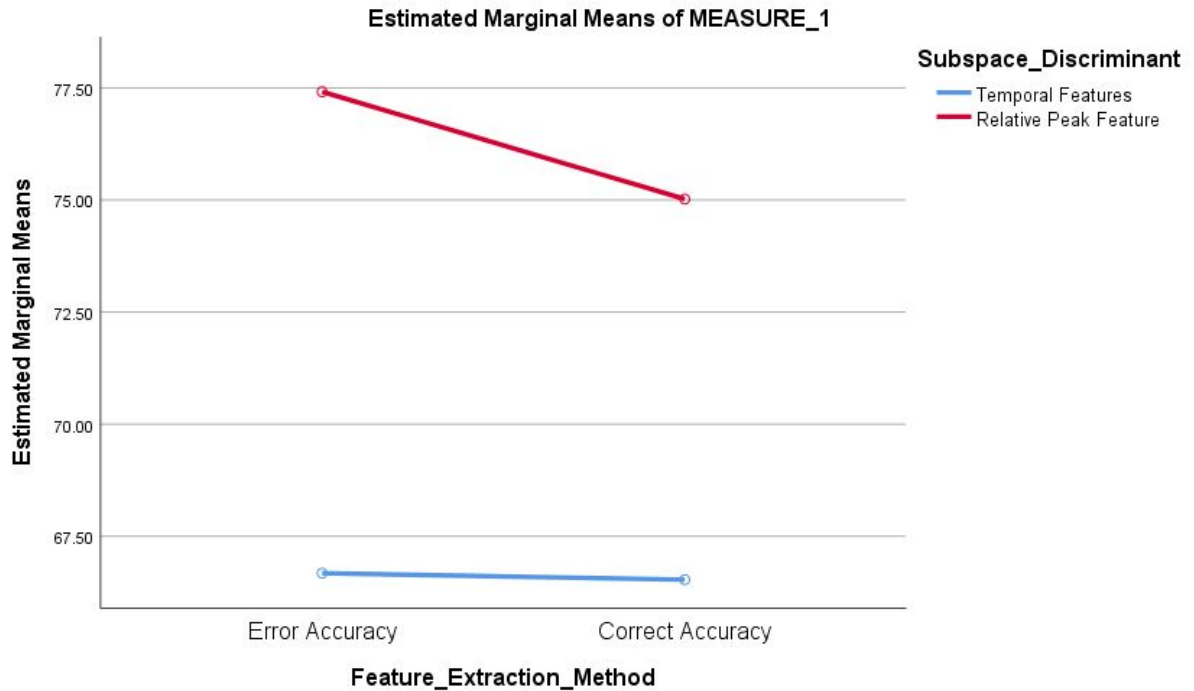


Figure F.2: Statistical analysis using a subspace discriminant classifier using elderly participants

F.0.5 Results of the statistical analysis using LDA classifier using elderly participants

A 2(features) X 2(Actions: error action vs correct) repeated ANOVA was performed on the classification results of young participants. The statistical results showed a statistically significant difference between the different methods of feature extraction, temporal features compared to relative peak features. The data was processed with a **Linear Discriminant Analysis(LDA) classifier** and the dataset from **elderly participants**.

(Please see column **Linear Discriminant Analysis(LDA)** in:

Table 1 D.1 : Elderly participants classifier using temporal features results tables

Table 2 D.2 : Elderly participants classifier using relative features results tables

F.0.6 Tests of Within-Subjects Effects and tests Within-Subjects Contrasts

The tests of Within-Subjects Effects indicate if there was an overall significant difference between the means of the different feature extraction methods.

Table F.7: Tests of Within-Subjects Effects and tests Within-Subjects Contrasts

Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Linear Discriminant Analysis (LDA)	Greenhouse-Geisser	3528.618	1	3528.618	25.061	<0.0001

Tests of Within-Subjects Contrasts

Source	Linear Discriminant Analysis (LDA)	Type III Sum of Squares	df	Mean Square	F	Sig
Linear Discriminant Analysis (LDA)	Temporal Features vs Relative Peak Features	3528.618	1	3528.618	25.061	<0.0001

From above table F.7 we are able to discover the **F** value for the “**feature extraction method**” factor, and its significance value.

We report that when using an ANOVA with repeated measures with a **Greenhouse-Geisser** correction, the mean scores for an **LDA classifier** using the different feature extraction methods were statistically significantly different.

$$(F(1.0, 25.061) = 24.061, \rho < 0.0001$$

We further report in table F.7 that when using an ANOVA with repeated measures “Tests of Within-Subject Contrasts” using a **Linear Discriminant Analysis(LDA)** using “Temporal features vs Relative Peak features” were statistically significantly different.

$$(F(1.0, 25.061) = 24.061, \rho < 0.0001$$

Please see the estimated marginal means plot for the different feature extraction methods shown in Fig F.3

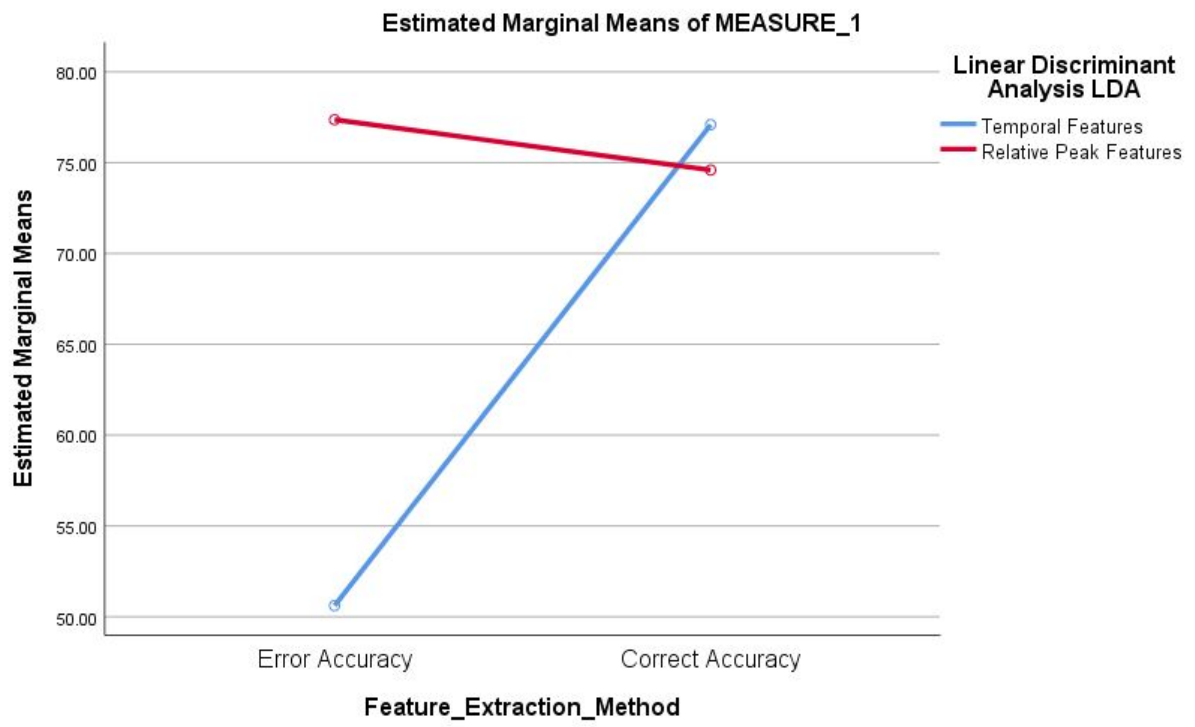


Figure F.3: Statistical analysis using an LDA classifier using elderly participants

F.0.7 Results of the statistical analysis using LDA classifier using young participants

A 2(features) X 2(Actions: error action vs correct) repeated ANOVA was performed on the classification results of young participants. The statistical results showed a statistically significant difference between the different methods of feature extraction, temporal features compared to relative peak features. The data was processed with a **Linear Discriminant Analysis(LDA) classifier** and the dataset from **elderly participants**.

(Please see column **Linear Discriminant Analysis(LDA)** in:

Young participants and temporal features

Table 3 D.3 : Young participants classifier using temporal features results tables P2-1

Young participants and relative peak features(RPF)

Table 5 D.5 : Young participants classifier using relative peak features results tables P2-1

F.0.8 Tests of Within-Subjects Effects and tests Within-Subjects Contrasts

The tests of Within-Subjects Effects indicate if there was an overall significant difference between the means of the different feature extraction methods.

Table F.8: Tests of Within-Subjects Effects and tests Within-Subjects Contrasts

Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Linear Discriminant Analysis (LDA)	Greenhouse-Geisser	930.982	1	930.982	12.220	0.002

Tests of Within-Subjects Contrasts

Source	Linear Discriminant Analysis (LDA)	Type III Sum of Squares	df	Mean Square	F	Sig.
Linear Discriminant Analysis (LDA)	Temporal Features vs Relative Peak Features	930.982	1	930.982	12.220	0.002

From above table F.8 we are able to discover the F value for the “**feature extraction method**” factor, and its significance value.

We report that when using an ANOVA with repeated measures with a **Greenhouse-Geisser** correction, the mean scores for an **LDA classifier** using the different feature extraction methods were statistically significantly different.

$$(F(1.0, 12.220) = 11.220, \rho = 0.002$$

We further report in table F.8 that when using an ANOVA with repeated measures “Tests of Within-Subject Contrasts” using a **Linear Discriminant Analysis(LDA)** using “Temporal features vs Relative Peak features” were statistically significantly different.

$$(F(1.0, 12.220) = 11.220, \rho = 0.002$$

Please see the estimated marginal means plot for the different feature extraction methods shown in Fig F.4

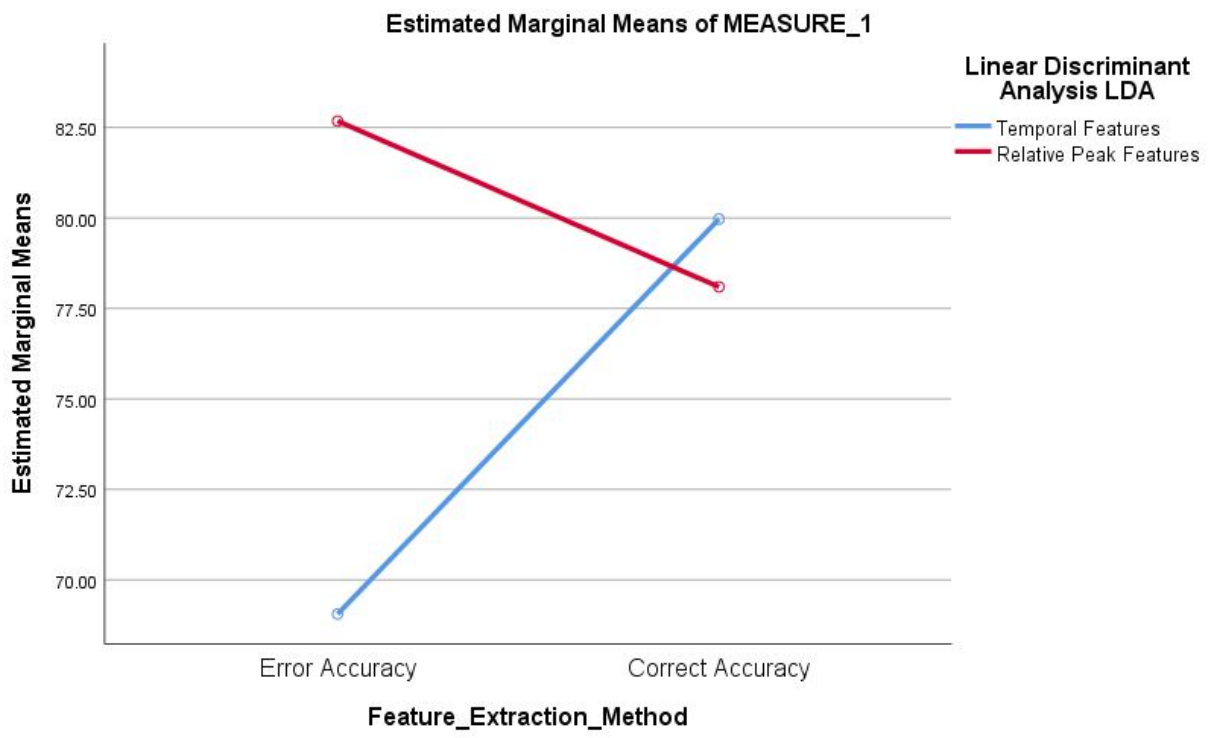


Figure F.4: Statistical analysis using an LDA classifier using different feature extraction methods

F.0.9 Results of the statistical analysis using SVM Linear classifier using young participants

A 2(features) X 2(Actions: error action vs correct) repeated ANOVA was performed on the classification results of young participants. The statistical results showed a statistically significant difference between the different methods of feature extraction, temporal features compared to relative peak features. The data was processed with a **SVM Linear classifier** and the dataset from **elderly participants**.

(Please see column **SVM Linear classifier** in:

Young participants and temporal features

Table 3 D.3 : Young participants classifier using temporal features results tables P2-1

Young participants and relative peak features(RPF)

Table 5 D.5 : Young participants classifier using relative peak features results tables P2-1

F.0.10 Tests of Within-Subjects Effects and tests Within-Subjects Contrasts

The tests of Within-Subjects Effects indicate if there was an overall significant difference between the means of the different feature extraction methods.

Table F.9: Tests of Within-Subjects Effects and tests Within-Subjects Contrasts

Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
SVM Linear	Greenhouse-Geisser	673.201	1	673.201	8.488	0.007

Tests of Within-Subjects Contrasts

Source	Feature Extraction Method	Type III Sum of Squares	df	Mean Square	F	Sig.
SVM Linear	Temporal Features vs Relative Peak Features	1346.402	1	1346.402	8.488	0.007

From above table F.8 we are able to discover the **F** value for the “**feature extraction method**” factor, and its significance level

We report that when using an ANOVA with repeated measures with a **Greenhouse-Geisser** correction, the mean scores for an **SVM Linear classifier** using the different feature extraction methods were statistically significantly different.

$$(F(1.0, 8.488) = 7.488, \rho = 0.007$$

We further report in table F.9 that when using an ANOVA with repeated measures “Tests of Within-Subject Contrasts” using a **SVM Linear** using “Temporal features vs Relative Peak features” were statistically significantly different.

$$(F(1.0, 8.488) = 7.488, \rho = 0.007$$

Please see the estimated marginal means plot for the different feature extraction methods shown in Fig F.5

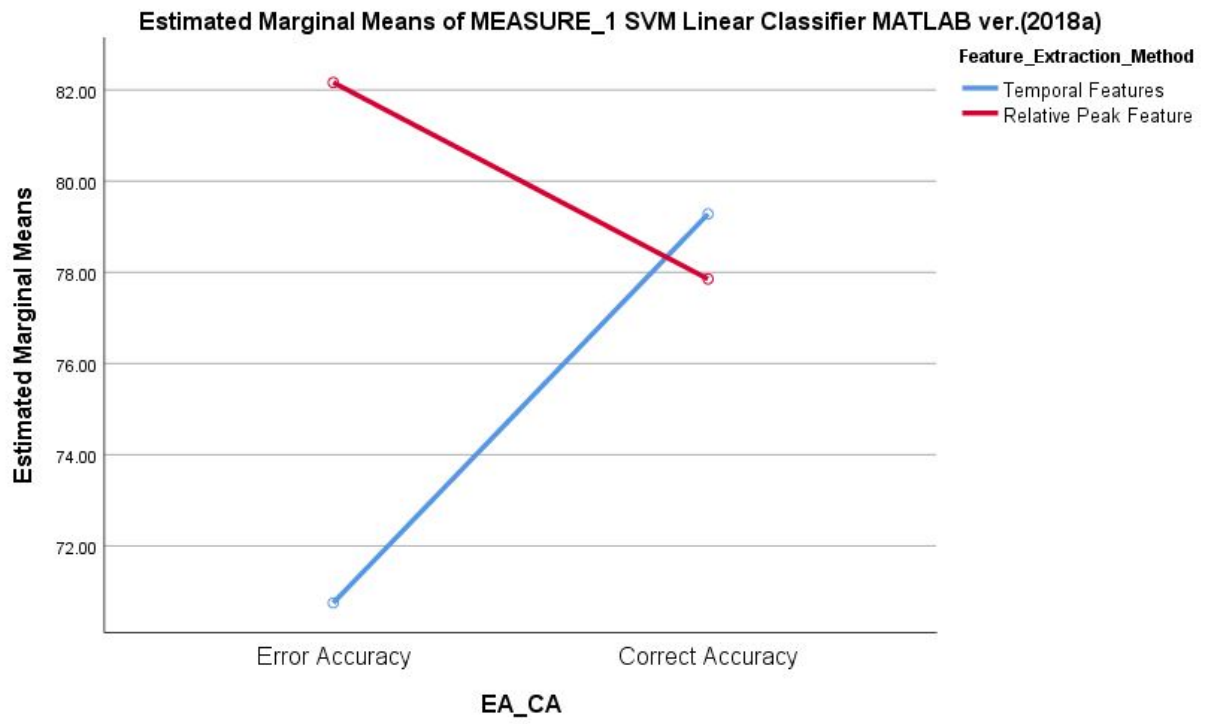


Figure F.5: Statistical analysis using an SVM Linear classifier using different feature extraction methods

F.0.11 Results of the statistical analysis using SVM Linear classifier using elderly participants

A 2(features) X 2(Actions: error action vs correct) repeated ANOVA was performed on the classification results of elderly participants. The statistical results showed a statistically significant difference between the different methods of feature extraction, temporal features compared to relative peak features. The data was processed with a **SVM Linear classifier** and the dataset from **elderly participants**.

(Please see column **SVM Linear classifier** in:

Table 1 D.1 : Elderly participants classifier using temporal features results tables

Table 2 D.2 : Elderly participants classifier using relative features results tables

F.0.12 Tests of Within-Subjects Effects and tests Within-Subjects Contrasts

The tests of Within-Subjects Effects indicate if there was an overall significant difference between the means of the different feature extraction methods.

Table F.10: Tests of Within-Subjects Effects and tests Within-Subjects Contrasts

Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
SVM Linear	Greenhouse-Geisser	2197.942	1	2197.942	21.398	< 0.0001

Tests of Within-Subjects Contrasts

Source	Feature Extraction Method	Type III Sum of Squares	df	Mean Square	F	Sig
SVM Linear	Temporal Features vs Relative Peak Features	4395.884	1	4395.884	21.398	< 0.0001

From above table F.10 we are able to discover the F value for the “**feature extraction method**” factor, and its significance level

We report that when using an ANOVA with repeated measures with a **Greenhouse-Geisser** correction, the mean scores for an **SVM Linear classifier** using the different feature extraction methods were statistically significantly different.

$$(F(1.0, 21.398) = 20.398, \rho = < 0.0001$$

We further report in table F.10 that when using an ANOVA with repeated measures “Tests of Within-Subject Contrasts” using a **SVM Linear** using “Temporal features vs Relative Peak features” were statistically significantly different.

$$(F(1.0, 21.398) = 20.398, \rho = < 0.0001$$

Please see the estimated marginal means plot for the different feature extraction methods shown in Fig F.6

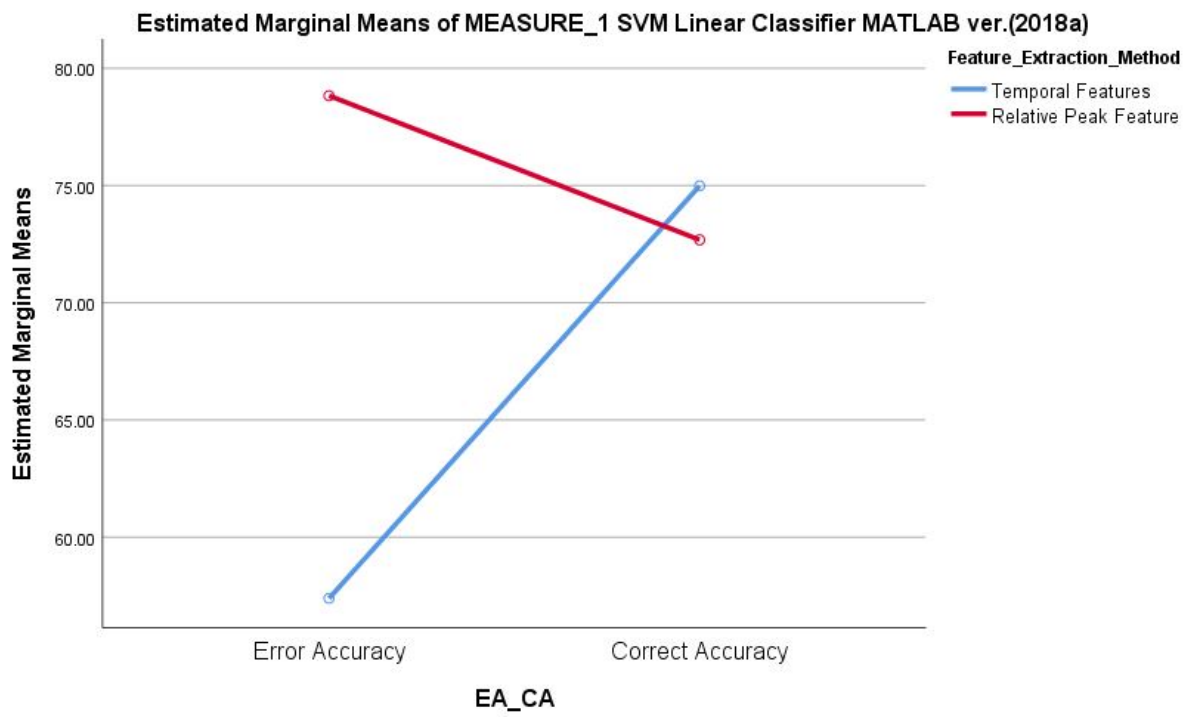


Figure F.6: Statistical analysis using an SVM Linear classifier using different feature extraction methods

F.0.13 Paired T-test Results of the statistical analysis using SVM Linear classifier with young and elderly participants

T-Test

Group Statistics						
	Young	Elderly	N	Mean	Std. Deviation	Std. Error Mean
EA TF	Young		27	70.7500	11.54762	2.22234
	Elderly		24	57.3917	21.67984	4.42538
CA TF	Young		27	79.2789	9.61276	1.84998
	Elderly		24	74.9854	10.99830	2.24502
EA RPF	Young		27	82.1615	11.82314	2.27536
	Elderly		24	78.8338	13.95883	2.84934
CA RPF	Young		27	77.8541	8.56374	1.64809
	Elderly		24	72.6829	11.98116	2.44564

Independent Samples Test						
		Levene's Test for Equality of Variances				
		F	Sig.	t	df	Sig. (2-tailed)
EA TF	Equal variances assumed	4.605	.037	2.790	49	.007
	Equal variances not assumed			2.698	34.142	.011
CA TF	Equal variances assumed	.574	.452	1.488	49	.143
	Equal variances not assumed			1.476	46.055	.147
EA RPF	Equal variances assumed	.186	.668	.922	49	.361
	Equal variances not assumed			.913	45.367	.366
CA RPF	Equal variances assumed	1.691	.200	1.788	49	.080
	Equal variances not assumed			1.753	41.130	.087

Figure F.7: Statistical analysis T-Test results using SPSS

F.0.13.1 Independent samples test summary

Error accuracy with an SVM Linear classifier using temporal features:

Equal variance not assumed is **statistically significant** $\rho = 0.011$

Correct accuracy with an SVM Linear classifier using temporal features:

Equal variance assumed is not statistically significant $\rho = 0.143$

Error accuracy with an SVM Linear classifier using **relative peak features**:

Equal variance not assumed is statistically not significant $\rho = 0.366$
 Correct accuracy with an SVM Linear classifier using **relative peak features**:
 Equal variance assumed is not statistically significant $\rho = 0.080$

F.0.13.2 T-Test Box Plot

Young Elderly

Case Processing Summary

		Valid		Cases Missing		Total	
		N	Percent	N	Percent	N	Percent
EA TF	Young	27	100.0%	0	0.0%	27	100.0%
	Elderly	24	100.0%	0	0.0%	24	100.0%

EA TF

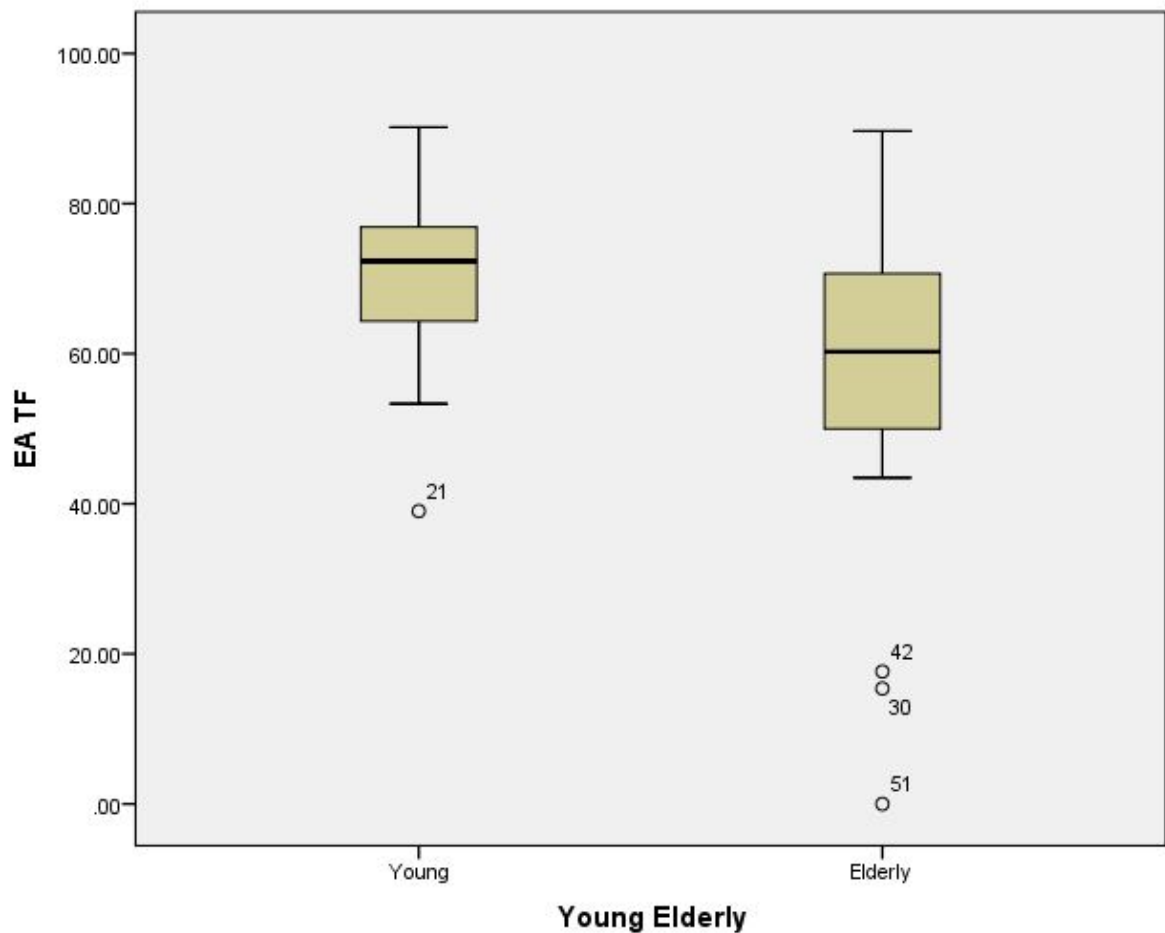


Figure F.8: Statistical analysis T-Test results using SPSS

F.0.13.3 T-Test Profile Plot

5. Young Elderly * Feature_Extraction_Method * EA_CA

Measure: MEASURE_1

Young Elderly	Feature_Extraction_Method	EA_CA	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Young	1	1	70.750	3.285	64.148	77.352
		2	79.279	1.980	75.301	83.257
	2	1	82.161	2.477	77.184	87.139
		2	77.854	1.984	73.867	81.841
Elderly	1	1	57.392	3.484	50.390	64.394
		2	74.985	2.100	70.766	79.205
	2	1	78.834	2.627	73.555	84.113
		2	72.683	2.104	68.454	76.912

Profile Plots

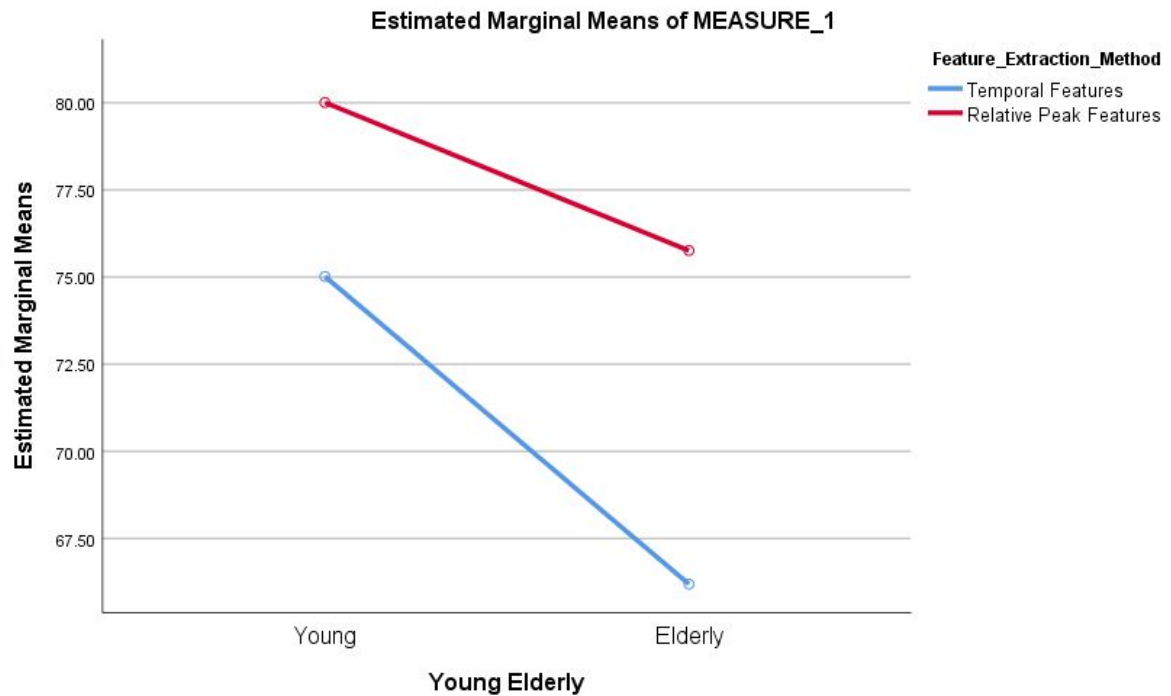


Figure F.9: SPSS data used for statistical analysis using an SVM Linear classifier using different feature extraction methods

F.0.14 LDA Classifiers Independent samples test summary

		Independent Samples Test				
		Levene's Test for Equality of Variances				
		F	Sig.	t	df	Sig. (2-tailed)
EA TF	Equal variances assumed	3.157	.082	3.623	49	.001
	Equal variances not assumed			3.506	34.486	.001
CA TF	Equal variances assumed	.000	.984	1.145	49	.258
	Equal variances not assumed			1.139	46.920	.260
EA RPF	Equal variances assumed	1.900	.174	1.460	49	.151
	Equal variances not assumed			1.436	42.387	.158
CA RPF	Equal variances assumed	.972	.329	1.478	49	.146
	Equal variances not assumed			1.450	41.389	.155

Figure F.10: Statistical analysis LDA T-Test results using SPSS

F.0.14.1 LDA T-Test results

Error accuracy with an LDA classifier using temporal features:

Equal variance not assumed is **statistically significant** $\rho = 0.001$

Correct accuracy with an LDA classifier using temporal features:

Equal variance assumed is not statistically significant $\rho = 0.258$

Error accuracy with an LDA classifier using **relative peak features**:

Equal variance not assumed is statistically not significant $\rho = 0.158$

Correct accuracy with an LDA classifier using **relative peak features**:

Equal variance assumed is not statistically significant $\rho = 0.146$

F.0.15 Logistic regression classifiers Independent samples test summary

		Independent Samples Test				
		Levene's Test for Equality of Variances		t	df	Sig. (2-tailed)
		F	Sig.			
EA TF	Equal variances assumed	2.137	.150	2.802	49	.007
	Equal variances not assumed			2.742	39.990	.009
CA TF	Equal variances assumed	.045	.833	1.555	49	.126
	Equal variances not assumed			1.555	48.268	.126
EA RPF	Equal variances assumed	.817	.370	1.335	49	.188
	Equal variances not assumed			1.318	43.989	.194
CA RPF	Equal variances assumed	.840	.364	1.677	49	.100
	Equal variances not assumed			1.651	43.010	.106

Figure F.11: Statistical analysis T-Test results using SPSS and a Logistic regression classifier

F.0.15.1 Logistic regression classifiers T-Test significance results

Error accuracy with a Logistic regression classifier using temporal features:

Equal variance not assumed is **statistically significant** $\rho = 0.009$

Correct accuracy with a Logistic regression classifier using temporal features:

Equal variance assumed is not statistically significant $\rho = 0.126$

Error accuracy with a Logistic regression classifier using **relative peak features**:

Equal variance not assumed is statistically not significant $\rho = 0.194$

Correct accuracy with a Logistic regression classifier using **relative peak features**:

Equal variance assumed is not statistically significant $\rho = 0.100$

F.0.16 KNN Coarse classifiers Independent samples test summary

		Independent Samples Test				
		Levene's Test for Equality of Variances		t	df	Sig. (2-tailed)
		F	Sig.			
EA TF	Equal variances assumed	.272	.604	1.709	49	.094
	Equal variances not assumed			1.698	46.582	.096
CA TF	Equal variances assumed	11.054	.002	2.462	49	.017
	Equal variances not assumed			2.369	31.582	.024
EA RPF	Equal variances assumed	1.522	.223	.995	49	.325
	Equal variances not assumed			.981	43.290	.332
CA RPF	Equal variances assumed	3.192	.080	2.620	49	.012
	Equal variances not assumed			2.552	37.759	.015

Figure F.12: Statistical analysis T-Test results using SPSS and a KNN Coarse classifier

F.0.16.1 KNN Coarse classifiers T-Test significance results

Error accuracy with a KNN coarse classifier using temporal features:
 Equal variance not assumed is not statistically significant $\rho = 0.0954$
 Correct accuracy with a KNN coarse classifier using temporal features:
 Equal variance assumed is statistically significant $\rho = 0.017$

Error accuracy with a KNN coarse classifier using **relative peak features**:
 Equal variance not assumed is statistically not significant $\rho = 0.332$
 Correct accuracy with a KNN coarse classifier using **relative peak features**:
 Equal variance assumed is statistically significant $\rho = 0.012$

F.0.17 Subspace discriminant classifier Independent samples test summary

		Independent Samples Test				
		Levene's Test for Equality of Variances				
		F	Sig.	t	df	Sig. (2-tailed)
EA TF	Equal variances assumed	4.998	.030	2.676	49	.010
	Equal variances not assumed			2.583	33.270	.014
CA TF	Equal variances assumed	.612	.438	.892	49	.377
	Equal variances not assumed			.887	46.678	.380
EA RPF	Equal variances assumed	4.148	.047	1.717	49	.092
	Equal variances not assumed			1.679	39.731	.101
CA RPF	Equal variances assumed	.491	.487	1.456	49	.152
	Equal variances not assumed			1.432	42.669	.159

Figure F.13: Statistical analysis T-Test results using SPSS and a Subspace discriminant classifier

F.0.17.1 Subspace discriminant classifiers T-Test significance results

Error accuracy with a Subspace discriminant classifier using temporal features:

Equal variance not assumed is **statistically significant** $\rho = 0.014$

Correct accuracy with a Subspace discriminant classifier using temporal features:

Equal variance assumed is not statistically significant $\rho = 0.377$

Error accuracy with a Subspace discriminant classifier using **relative peak features**:

Equal variance not assumed is not statistically significant $\rho = 0.101$

Correct accuracy with a Subspace discriminant classifier using **relative peak features**:

Equal variance assumed is not statistically significant $\rho = 0.152$

Appendix G

Results of the performance gains summary

G.0.1 Results of the performance gains summary

LDA Classifier	Error Accuracy		Correct Accuracy	
	Young	Elderly	Young	Elderly
Temporal Features	69.06%	50.61%	79.97%	77.09%
Relative Peak Features RPF	82.68%	77.36%	78.10%	74.59%
Average performance gains	13.62%	26.75%	-1.87%	-2.50%

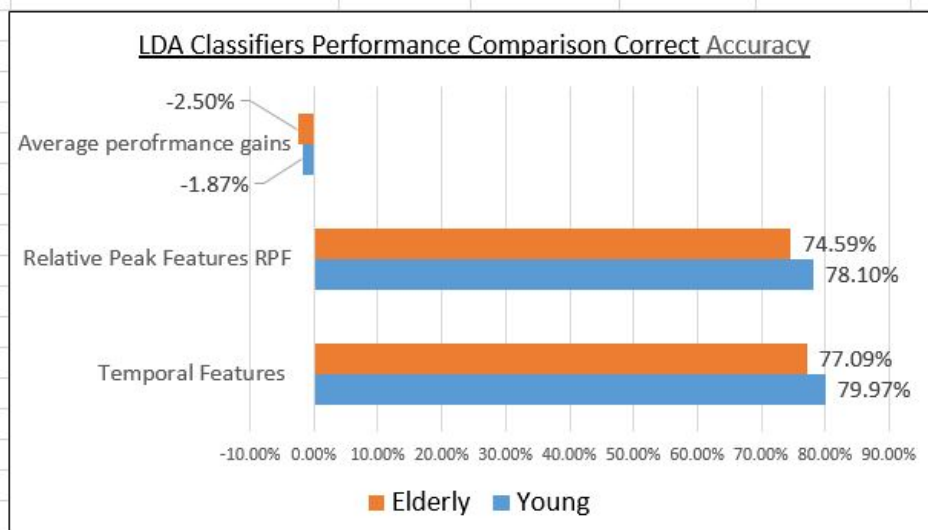
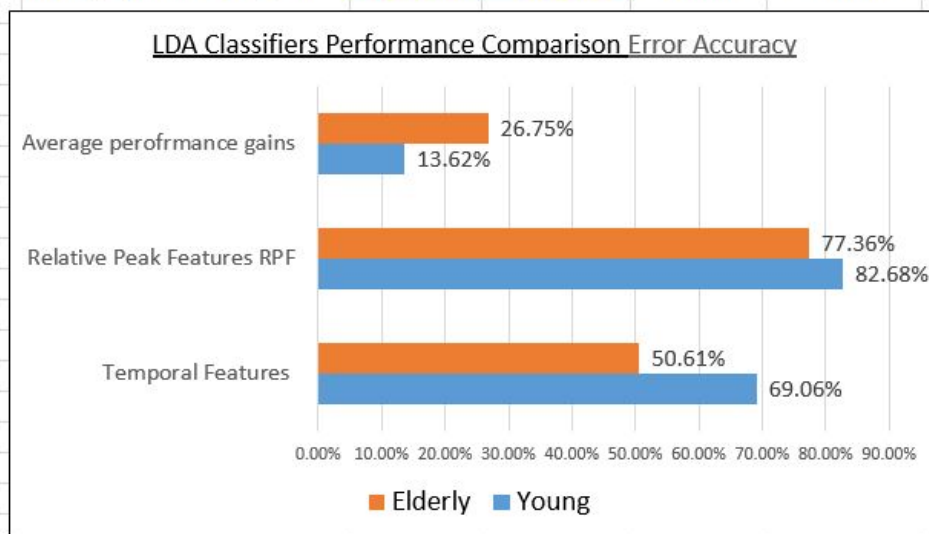


Figure G.1: Compare performance gains

Logistic Regression	Error Accuracy		Correct Accuracy	
	Young	Elderly	Young	Elderly
Temporal Features	68.44%	55.15%	80.34%	76.34%
Relative Peak Features RPF	81.45%	76.41%	78.81%	74.91%
Average performance gains	13.01%	21.26%	-1.53%	-1.43%

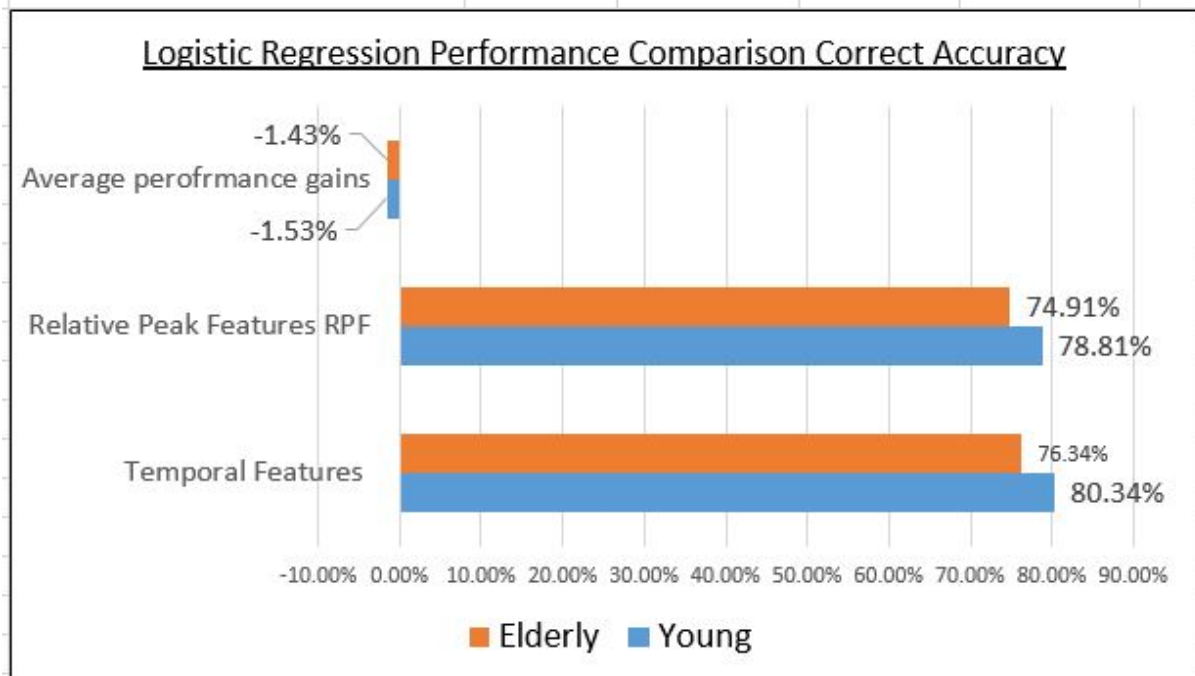
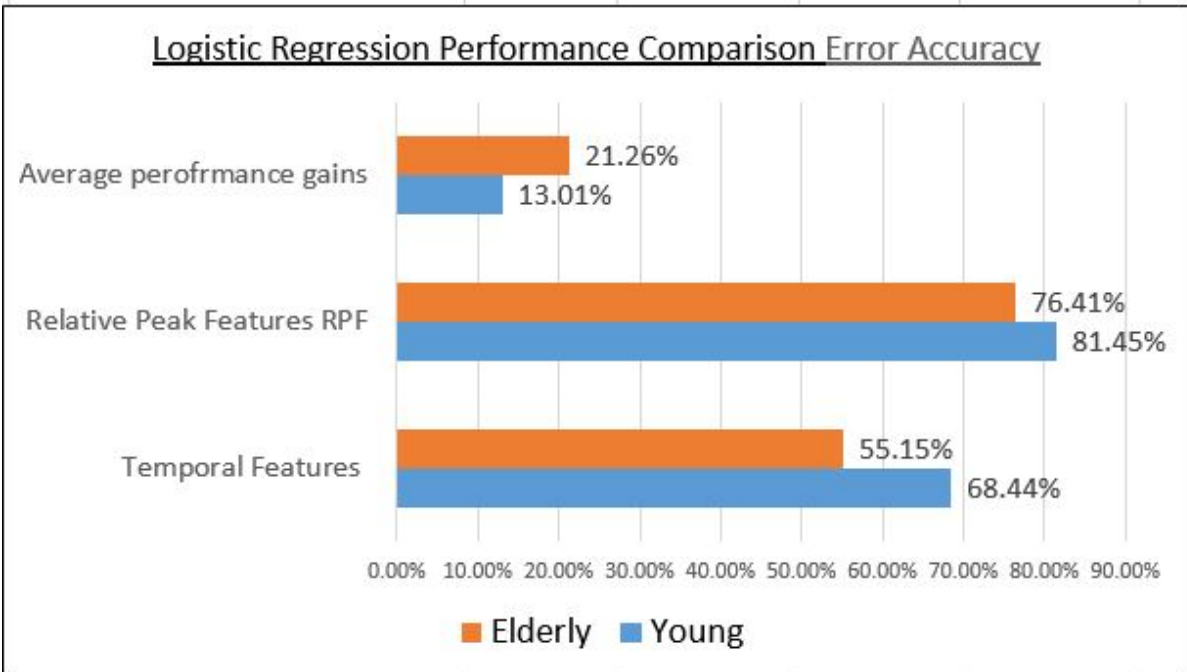


Figure G.2: Compare performance gains

KNN Coarse	Error Accuracy		Correct Accuracy	
	Young	Elderly	Young	Elderly
Temporal Features	64.98%	56.57%	88.16%	80.56%
Relative Peak Features RPF	80.85%	77.03%	78.83%	70.58%
Average performance gains	15.87%	20.46%	-9.33%	-9.98%

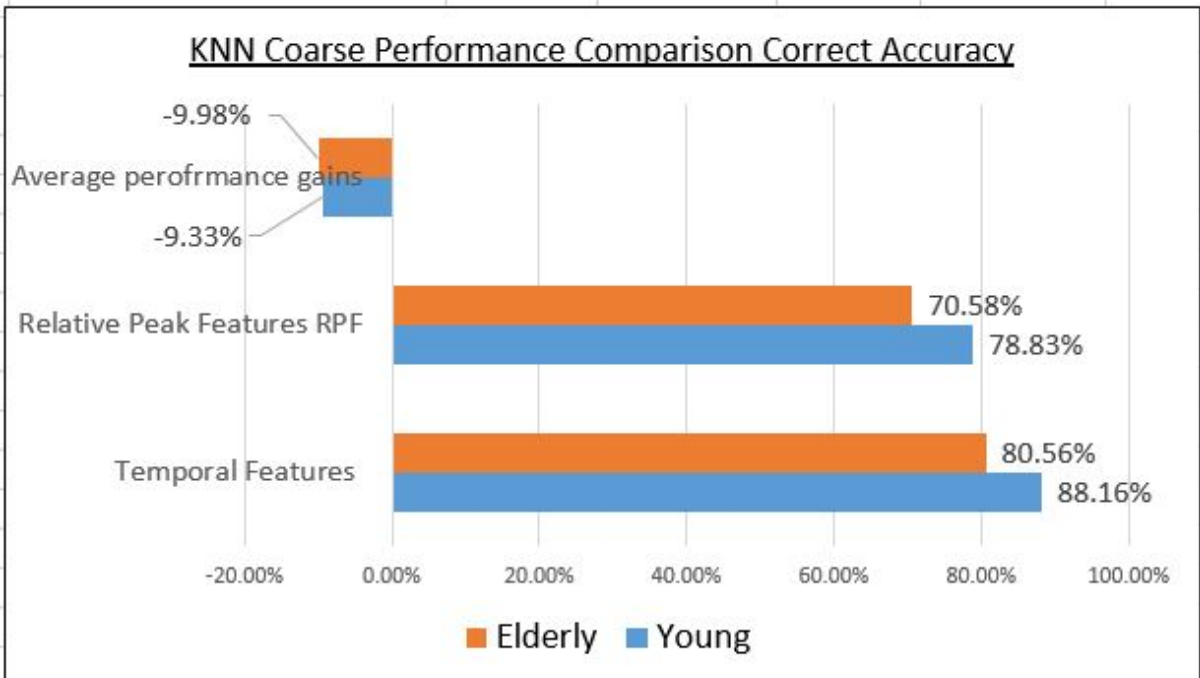
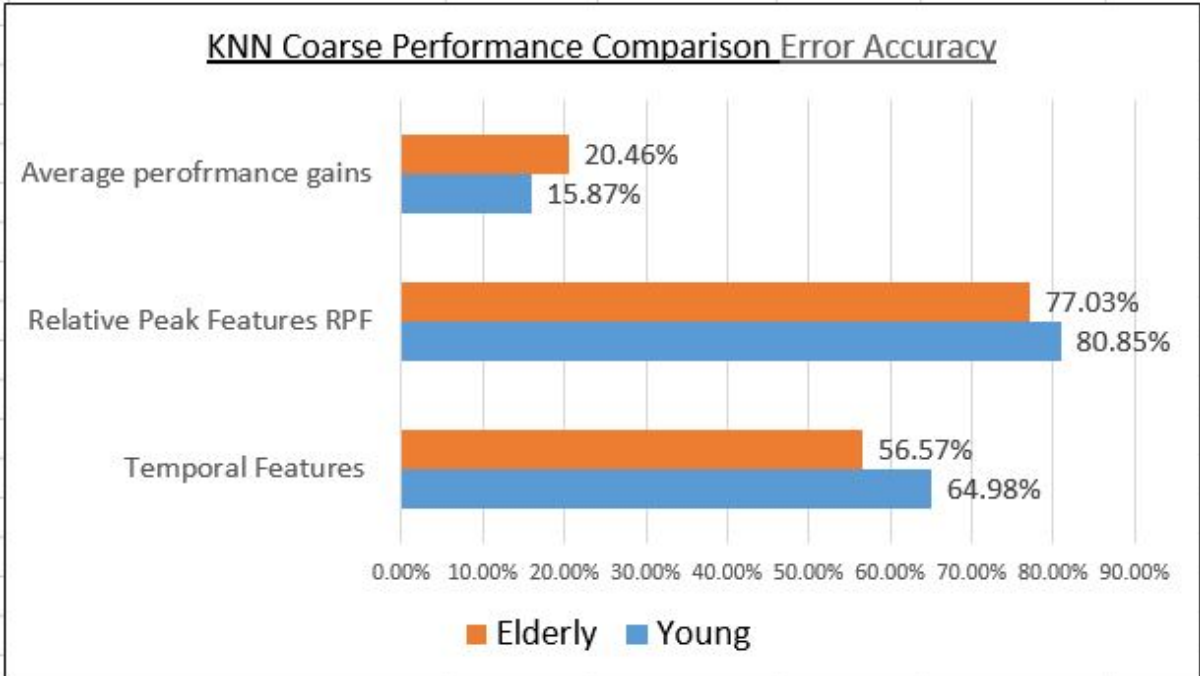


Figure G.3: Compare performance gains

SVM Linear	Error Accuracy		Correct Accuracy	
	Young	Elderly	Young	Elderly
Temporal Features	70.75%	57.39%	79.28%	74.98%
Relative Peak Features RPF	82.16%	78.83%	77.86%	72.68%
Average performance gains	11.41%	21.44%	-1.42%	-2.30%

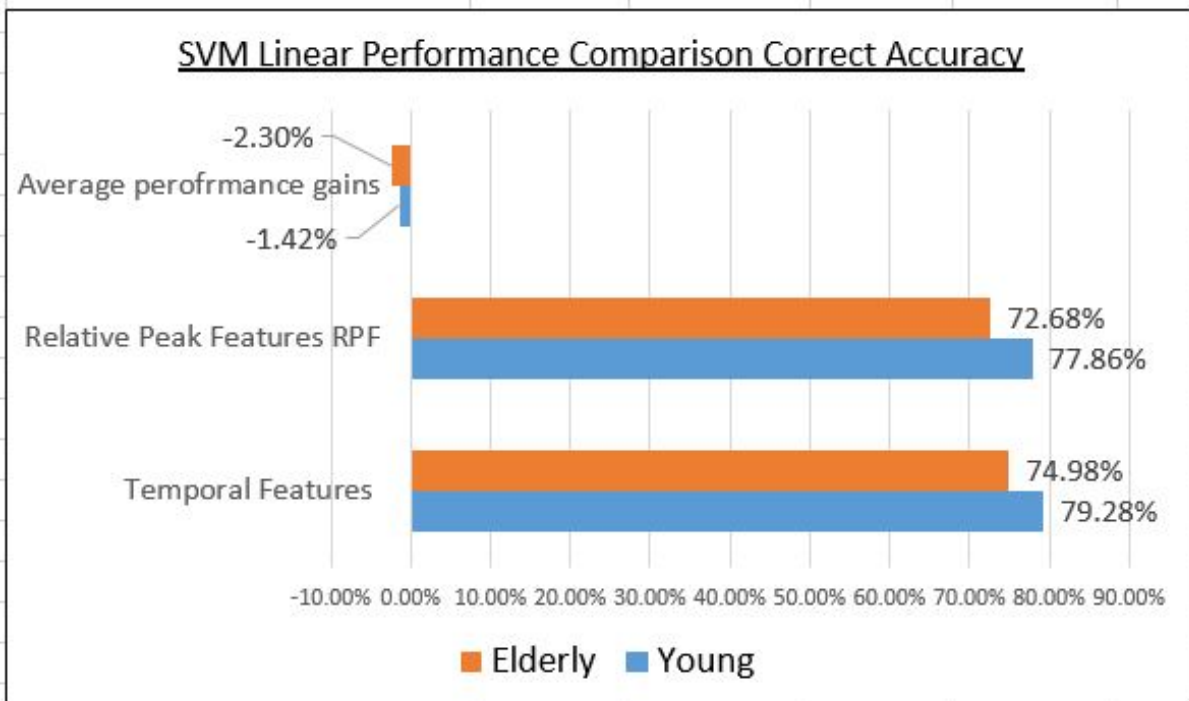
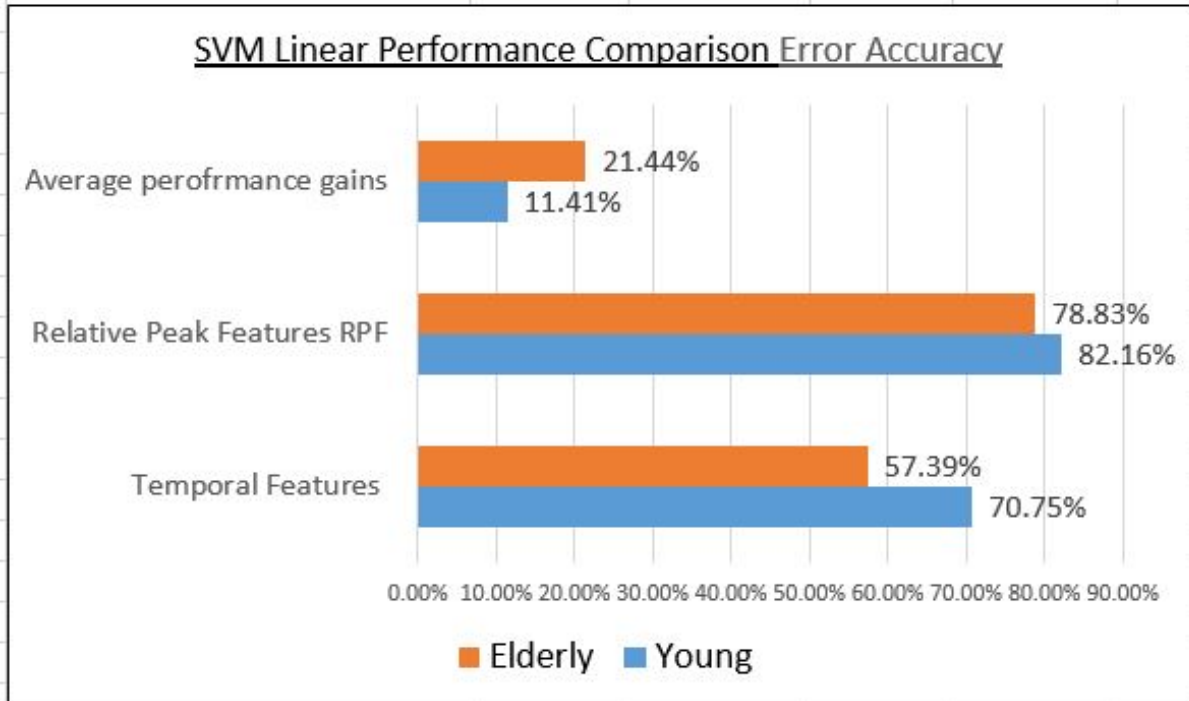


Figure G.4: Compare performance gains

Subspace discriminant	Error Accuracy		Correct Accuracy	
	Young	Elderly	Young	Elderly
Temporal Features	77.73%	67.21%	70.57%	66.63%
Relative Peak Features RPF	83.38%	77.41%	78.44%	75.02%
Average performance gains	5.65%	10.20%	7.87%	8.39%

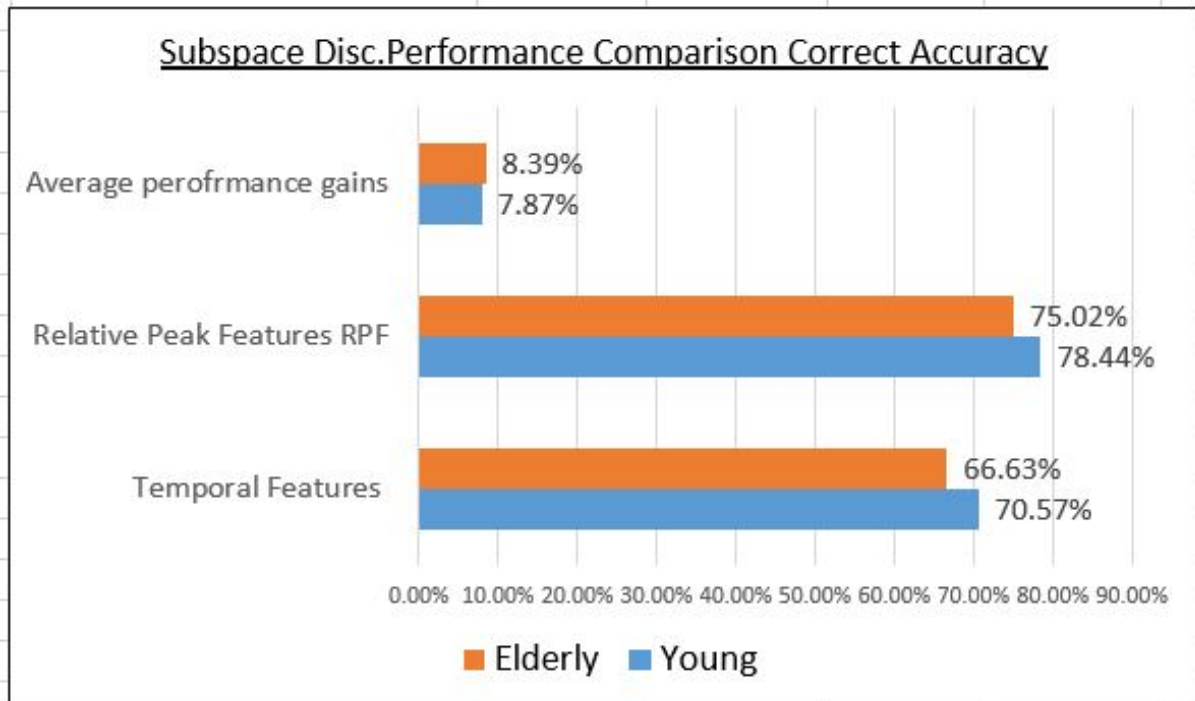
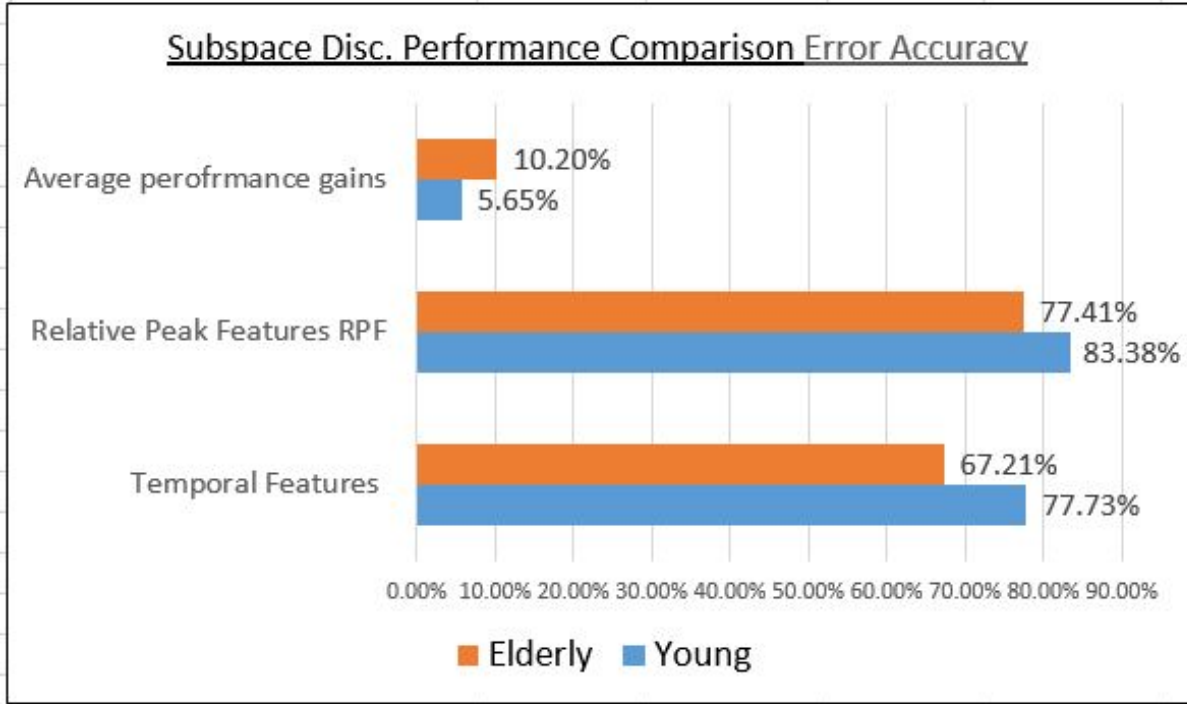


Figure G.5: Compare performance gains

G.0.2 Comparing the standard deviation and variance

The standard deviation and variance was calculated from the training data for each person from each group. The summary table are shown in Table 1 Elderly participants standard deviation and variance : C.1 and Table 2 Young participants Standard deviation and variance : C.2

The variance on feature 1 or the amplitude for the **VarY ErrP** or the y-axis of the relative peak plot showed a greater variance for the younger participants. In addition, the variance on feature 1 or the amplitude for the **VarY No ErrP** or the y-axis of the relative peak plot also showed a larger variation. Further standard deviation of feature 1 or the amplitude for the **Stdy No ErrP** or the y-axis of the relative peak plot showed a greater standard deviation for the younger participants.

This insight into the training data's standard deviation and variance may be used to configure a classifier to enhance performance. Please see the following diagrams G.6, G.8 and G.7

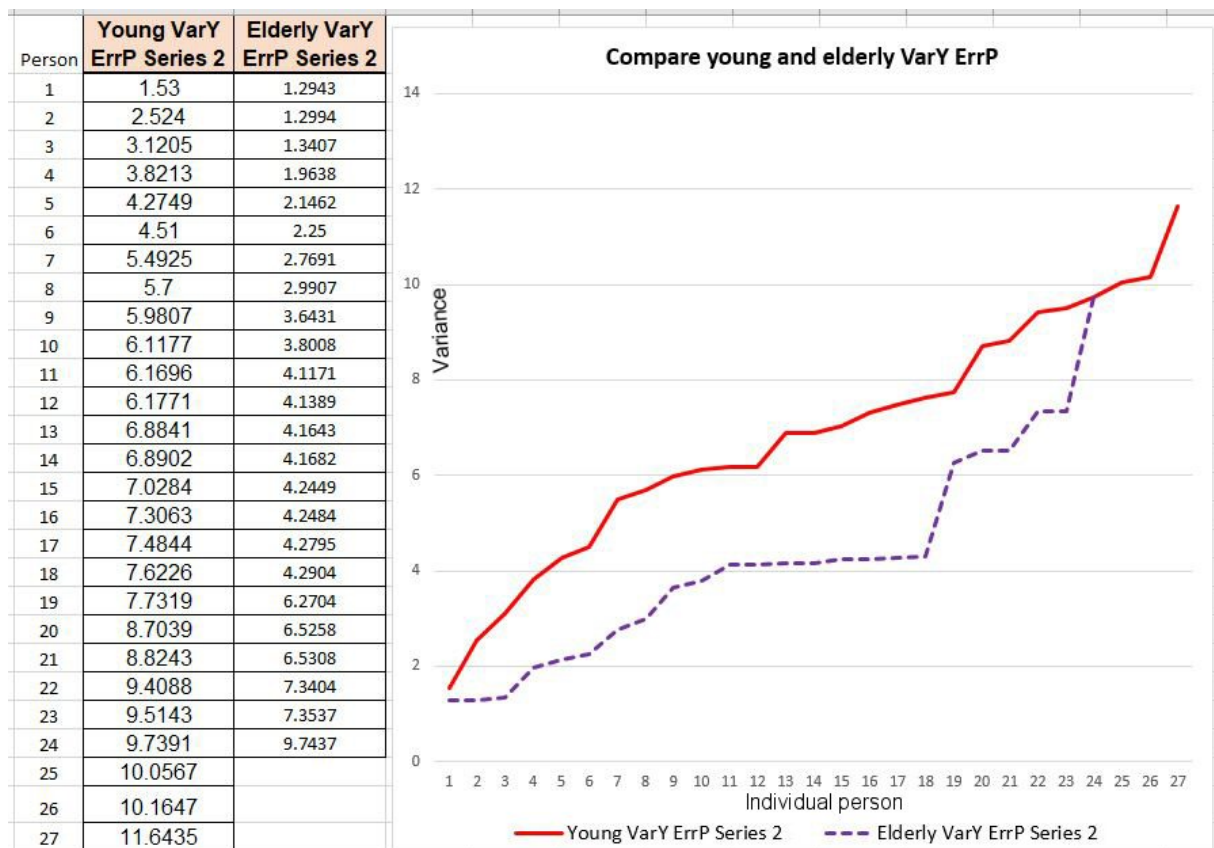


Figure G.6: Compare young elderly variance varY ErrP , shown in C.1 and C.2

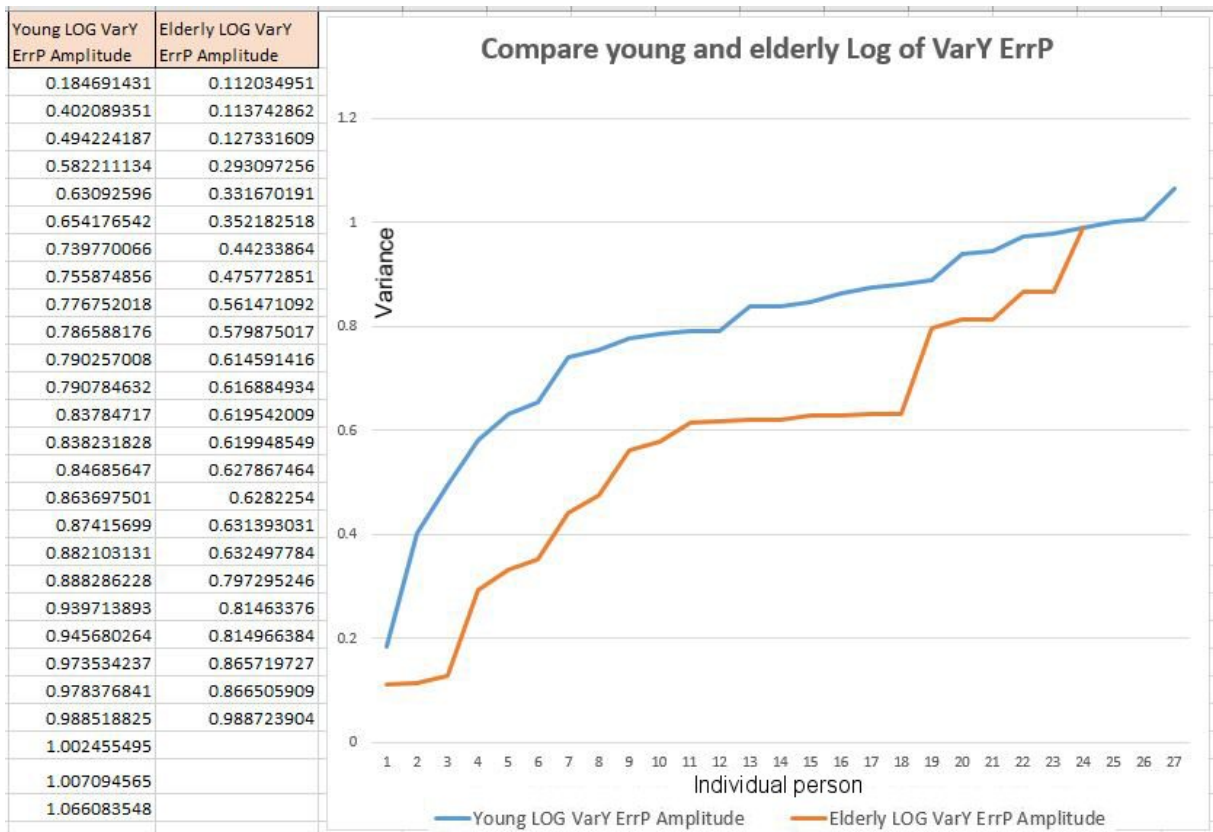


Figure G.7: Compare young elderly variance Log varY ErrP , shown in C.1 and C.2

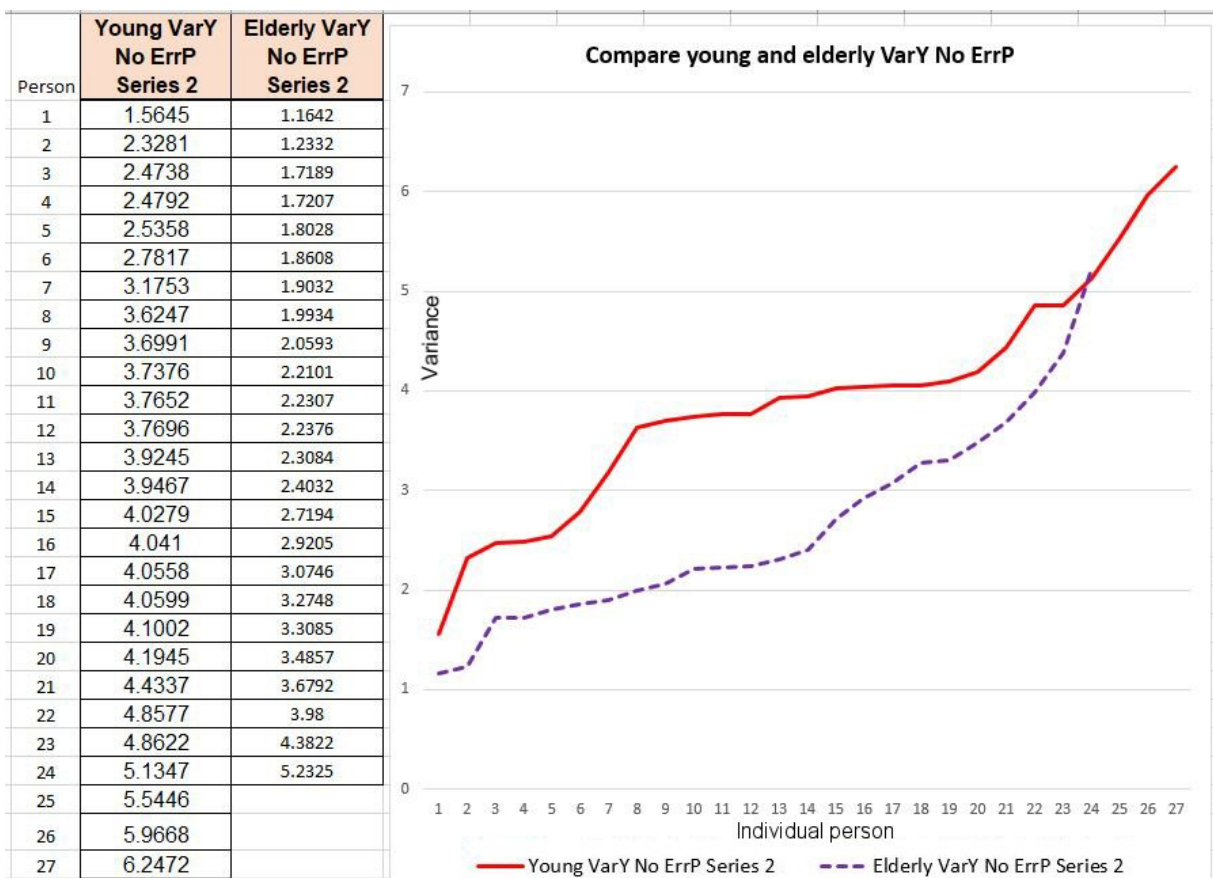


Figure G.8: Compare young elderly variance varY No ErrP , shown in C.1 and C.2

Appendix H

Mobility and empowerment

H.1 Re-establish a sense of empowerment

What sort of everyday daily tasks could be re-enabled for that person? What kind of assistance could make a positive supportive difference in that persons life? Moreover it would be enabling hope and independence.

H.1.1 Disabled persons daily life

Able bodied people would find the following, (see table H.1) daily life action easy and literally take any physical movement for granted in addition to securing the property at night in-terms or making the sure water taps are not left running and the gas is not on and doors, windows are locked and the fire alarms is fully operational with the batteries working.

In contrast a disabled person will encounter every physical task impractical to achieve. This could be solved by a very sophisticated 'smart' home full of automation however technology is not available to them when they are outside of their home going to work or shopping in town, what level of assistive support can be provided? How can a brain-computer interface know the needs of the disabled person? Moreover how can these needs be supported by an assistive robotic device.

Table H.1: Sample of daily life functions

open / close control	Basic functional control	Complex functional control
Doors,Windows	Heating system	Collecting the post
Fridge door,cupboards	TV / Radio channel selection	Making a phone call
kitchen bin		Cooking food
kitchen drawers		Cleaning up after eating
Bottle tops		Handling Food packaging
on / off control		Eating , Drinking
Room lights		Retrieving a utensils
Water taps		Personal hygiene

H.1.2 Understanding the problem and the difficulties

To understand the difficulties I will describe a scenario: A person is seated around a table with two cups of a beverage cup A has orange juice cup B has water. For a person with no movement difficulties they may take the following steps

- STEP 1:** Visually observe the choices in front of them:

△ Option A, cup of orange juice.

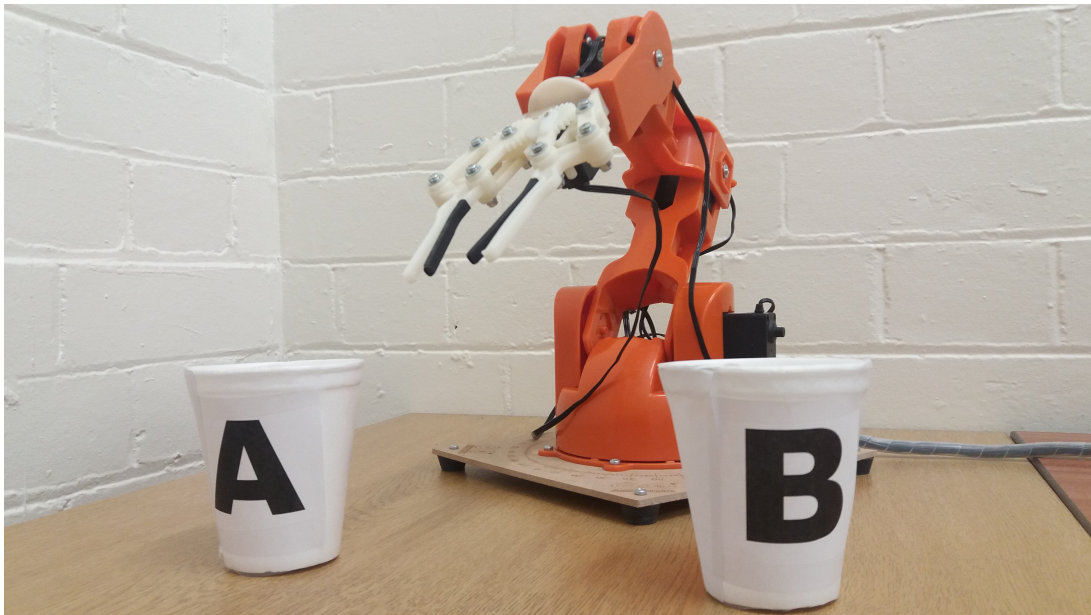


Figure H.1: LAB Assistive Robotic Arm (ARA)

△ Option B, cup of water.

● **STEP 2:** A decision is made, the person makes their selection

● **STEP 3:** The person commands their movements to their arm(s)

△ The person controls the movements of the cup to their mouth and take a drink

△ The person returns the cup to the table

All the above steps take place seamlessly and almost with no effort or thinking about it just a natural movement to take a drink from a cup.

However for a person with movement limitations such as no upper arm movement disability. How might those same steps be achieved?

● **STEP 1:** Visually observe the choices in front of them:

(Assumption that the person can see the options?)

△ Option A, cup of orange juice.

△ Option B, cup of water.

● **STEP 2:** A decision is made, the person makes their selection

(How do we know their choice? Who is going to be there to understand their choice?.)

(Moreover, how can an assistive robotic device be aware of their selection?)

● **STEP 3:** The person is unable to take a drink!

(How can an assistive robotic device support this step safely)

Appendix I

Statistical Analysis of Variance (ANOVA)

I.0.1 Statistical Analysis ANOVA Test 1, 13/01/2017

Acknowledgement Mathematics and Statistics Help (MASH) STEPS for ANOVA by Adrian Ashley Teaching instructor Miss Jo Rothwell MASH Staff 13/01/2017 Experiment Data , Dr Mahnaz Arvaneh 28/02/2017

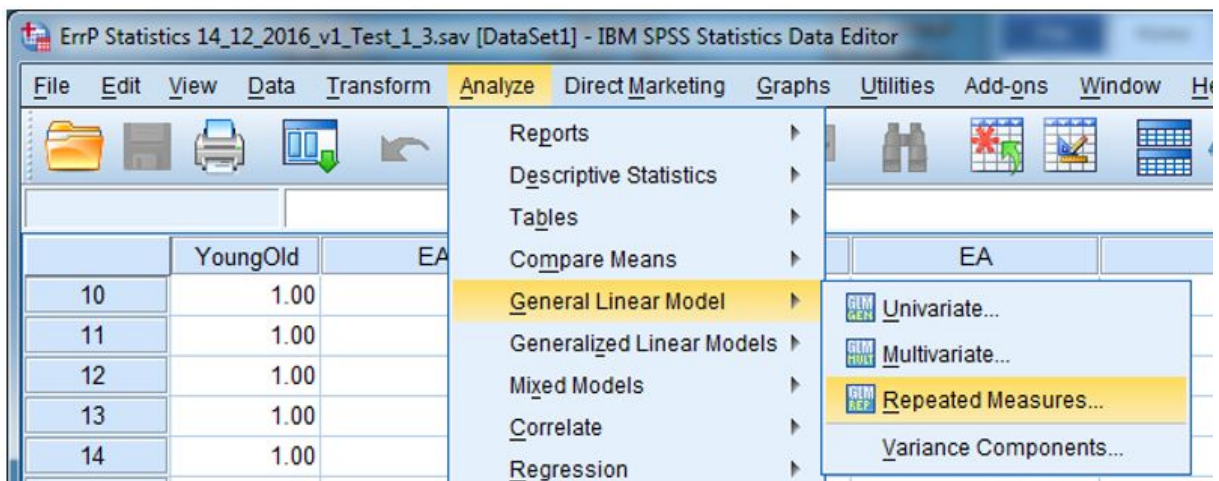


Figure I.1: SPSS menus Analyze - General Linear Model - Repeated Measures

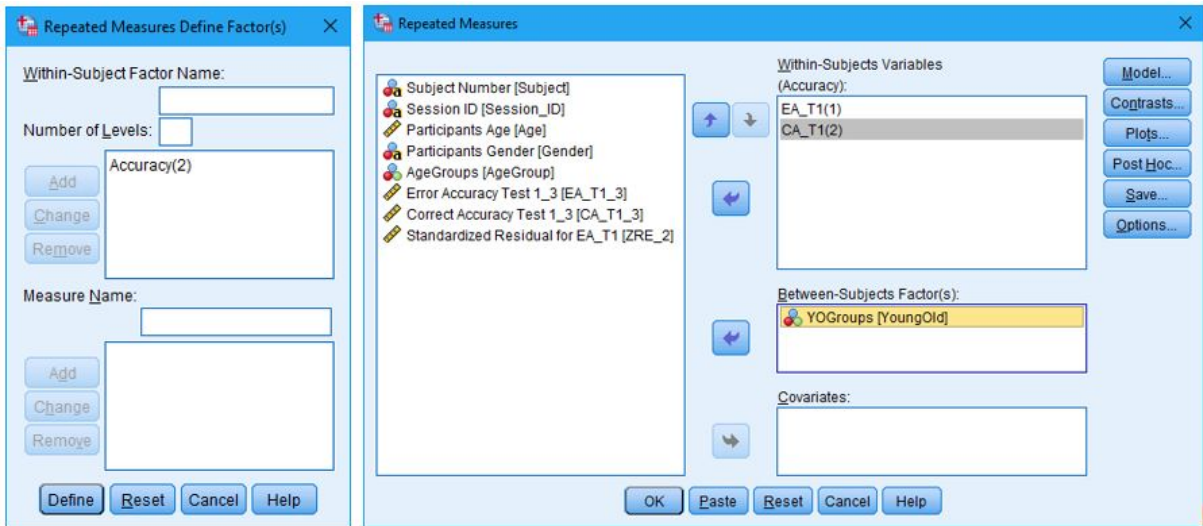


Figure I.2: SPSS Repeated Measures settings screen shot 1

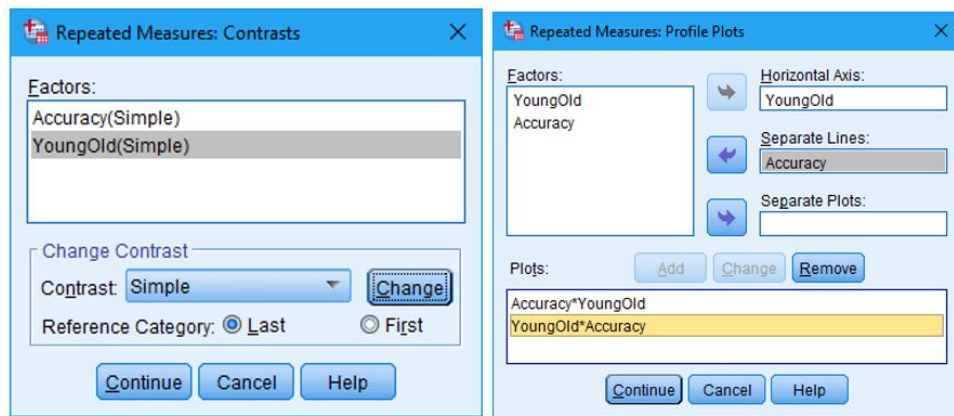


Figure I.3: SPSS Repeated Measures settings screen shot 2

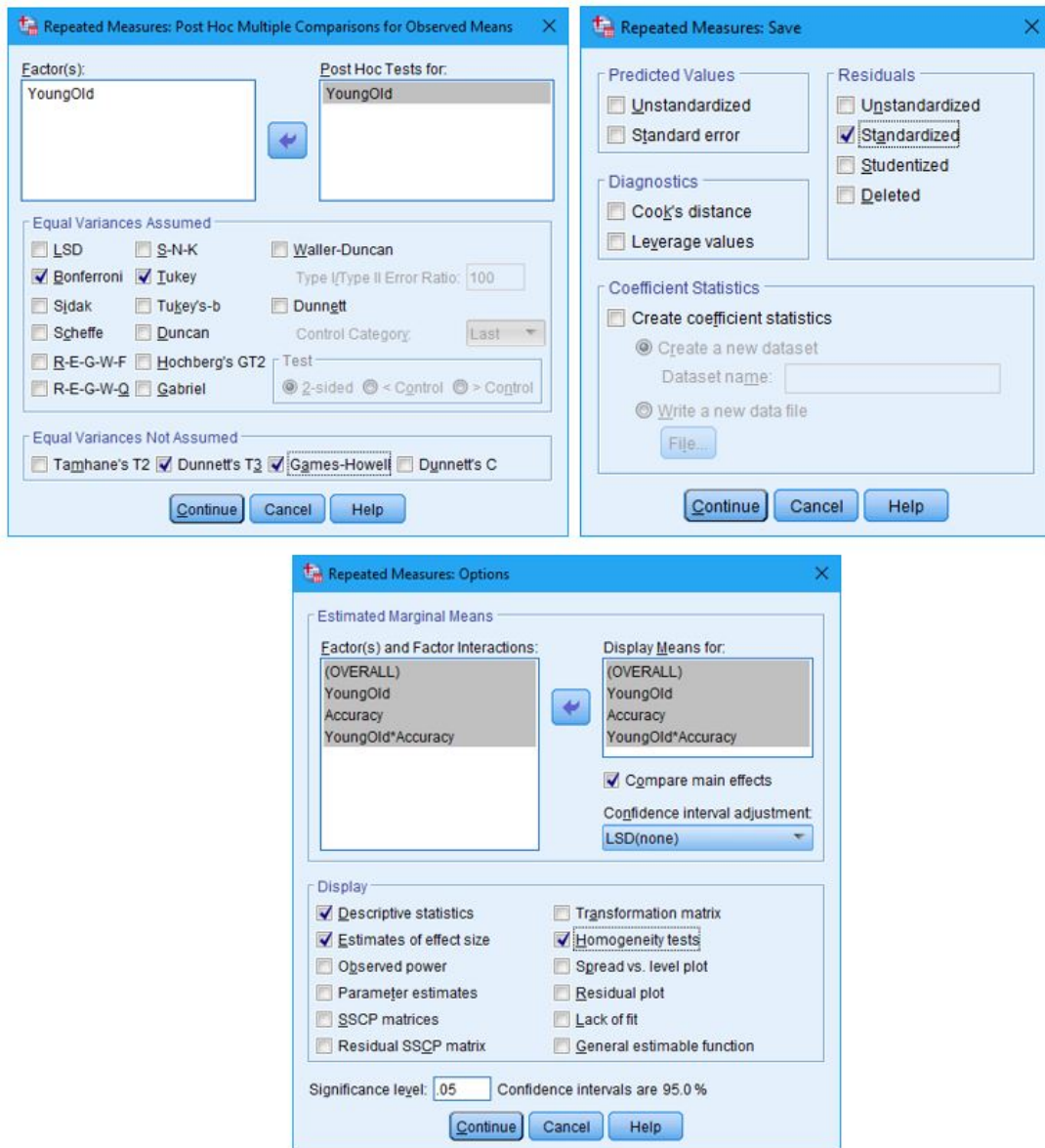


Figure I.4: SPSS Repeated Measures settings screen shot 3 and screen shot 4

I.0.2 SPSS Software Settings for statistical analysis of feature extraction methods

Part of my the data quality checks process will be to verify the SPSS Settings and the output with supervisors and MASH support . My project risk register will track the risks to the project.

The data used has been check

<>—< Risk register >—

Risk register example

Risk 1 , (suggest risks)

Data is generated from SPSS is faulty

Virus

Software faulty

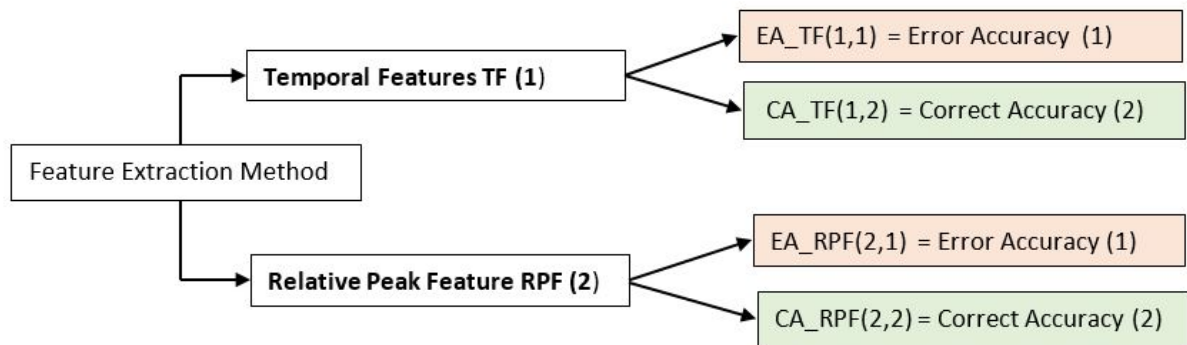
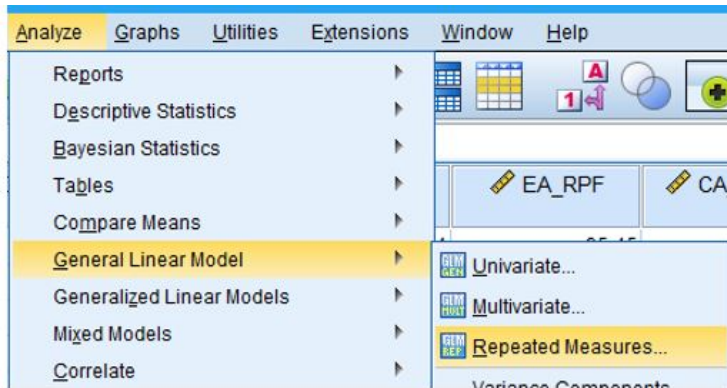
skill and experience of the user in statistics

Risk 1 Mitigation (Solutions to reduce the risk)

- 1.Received SPSS training and attended workshops 2016/2017
2. Check the setting and outputs with supervisions
- 3 Cross check the process with statistics staff at the MASH .

- 4 update windows
- 5 update to the last version of SPSS
- 6 update the computer's anti-virus software

<>—< Risk register >—



SPSS Settings, Feature Extraction Method

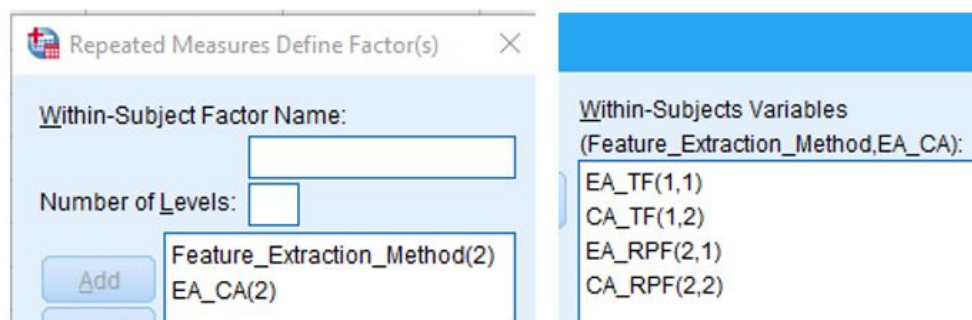


Figure I.5: Statistical analysis using SPSS for different feature extraction methods

Appendix J

LDA Methodology

J.1 LDA Methodology

Raschka [72] explained 5 steps of processing the data to achieve a linear

ERP Classification Blankertz, et al [73] presented a clear tutorial steps for ERP classification

(J.1)

k Trials

n Number of EEG Channels

T Samples time

Subset of EEG Channels

$$\mathbf{X} = \{x_1, \dots, x_n\} \quad (\text{J.2})$$

Compute Phase 1: Computer the mean vectors μ_i of each class $i = 1, 2$

$$\mu_i = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \quad (\text{J.3})$$

Compute Phase 2: Compute the Scatter Matrices [72]

Compute the within-class scatter matrix S_W

$$S_W = \sum_{i=1}^c S_i \quad (\text{J.4})$$

S_i is the scatter matrix for every class [72] where

$$S_i = \sum_{x \in D_i}^n (x - \mu_i) (x - \mu_i)^T \quad (\text{J.5})$$

and μ_i is the mean vector [72]

$$\mu_i = \frac{1}{n_i} \sum_{x \in D_i}^n \mathbf{X}_k \quad (\text{J.6})$$

Compute the between-class [72] scatter matrix S_B using the equation:

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu) (\mu_i - \mu)^T \quad (\text{J.7})$$

where

μ is the overall mean

μ_i is the sample means of each of the classes

N_i is the size of the classes

Compute Phase 3: Compute the linear discriminants by solving generalised eigenvalue problem for matrix below : [72]

$$S_W^{-1} S_B \quad (J.8)$$

Calculation objectives: The eigenvalues and eigenvectors can tell us about the linear transformations:

- Eigenvectors indicate the direction the effect.
- Eigenvalues indicate the magnitude of the effect.

The linear discriminant analysis LDA enables a dimensionality reduction by using the eigenvectors to indicate the new axis for the feature space. The eigenvalues enables a comparison of the eigenvectors.

J.1.0.1 Verification of the eigenvector-eigenvalues calculations

$$A v = \lambda v \quad (J.9)$$

where

$$A = S_W^{-1} S_B$$

v = Eigenvectors

λ = Eigenvalues

Compute Phase 4: Selecting the linear discriminants

The reason for obtaining the eigenvectors and eigenvalues was to find a new subspace to project the data on, in order to increase the class separation in addition to reducing the dimensionality of the feature space.

Select highest eigenvalue and associated eigenvector. The eigenvectors form the axes of the new feature space. Select the highest eigenvalue and associated eigenvector.

The set of eigenvectors and eigenvalues can be arranged into descending order assemble and W matrix comprising of $k \times d$ dimensions

Compute Phase 5: Transformation Use the following equation to transform the data to the new sub space :

$$Y = X \times W \quad (J.10)$$

where

X = $n \times d$ dimensional matrix representing n data samples

Y = transformed $n \times k$ dimensional samples in the new subspace [72]

Appendix K

Error Awareness Sustained Attention Task

K.1 Error Awareness Sustained Attention Task

Is secondary data from experiments from Ireland

K.1.1 Experiment Objectives

Investigate how the ErrP maybe different in amplitude according to the subjects age groups

K.1.1.1 Experiment Configuration

The objective of the data analysis was to investigate the awareness in a 'sustained attention' task. Statistical analysis shows that there is a significant difference in between the young and old participants.

Generally, the cause of the 'Error related potential' ErrP is that the subject is aware of their mistake during the sustained attention task. The configuration of the experiment is such that the subjects will evoke ErrP events during the experiment.

The experiment methodology for Test 1 and Test 1-3 had the following common configurations: EEG potentials using a 64 Channel Cap using the electrode location specified by the '10/20 international system. The equipment used was by 'Biosemi' with a sampling rate between 512 – 1024 Hz. Time intervals were 350 ms window starting 50 ms after the feedback and ending 400 ms after the feedback, The baseline interval was -200 to 0.

The criteria for artefact rejection was 30 microvolts. The referencing method used was the common average. A band pass of 1 to 10 Hz filtering was used as part of the pre-processing. Specific features are acquired from the EEG data and labelled with specific events and sent to the classifiers. During the session with each subject, generally 8 blocks of data were captured. Generally, blocks 1 to 4 were sent to the classifier for training and the other blocks of data from 5 to 8 were sent to the classifiers for testing.

The evaluation is checked in terms of 'Error Accuracy' EA and 'Correct Accuracy' CA. The Brain-Computer Interface BCI limitations are the potential for an incorrect classification of the user's intentions. The classifier method used was based on the Linear discriminant analysis (LDA) originating from the Fisher's linear discriminant. The differences between the tests were as follows;

K.1.1.2 Methodology

The subject is presented with a series of visual information presented in a certain sequence in order to capture the subjects awareness of an error

Table K.1: Experiment Configuration

Experiment Configuration	Test 1	Test 1_3
<i>Different configurations</i>		
Channels:	Fz, Cz	Fz,Cz,CPz
Resampling:	64 Hz	32 Hz
<i>Common configurations</i>		
Electrode location:	10/20 international system	
Sampling rate:	512 - 1024 Hz	
Time intervals:	350 ms window	
Starting:	50 ms after feedback	
Ending:	400 ms after feedback	
Baseline interval:	[-200 0]	
Referencing:	Common average	
Band pass filtering:	1 to 10 Hz	
Criteria for artefact rejection:	30 micro volts	

The experiment protocol was developed by as part of a research group [23] in Ireland based on a novel variant of Hester et al 'Error awareness task [110] The participants carried out a practice session in order to establish the suitability for taking part in the experiments:

An outline of the experiment is provided as follows: The visual stimuli has a grey background upon which a centrally positioned white fixation. This fixation cross remains in place while the colour dots change during the experiment sessions.

- (A) The system generates a random sequence of coloured dots from a given predefined set. [23]

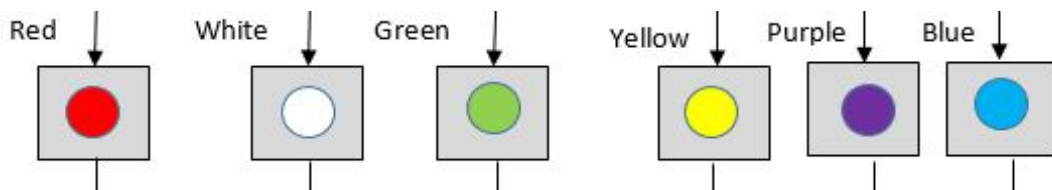


Figure K.1: See above for an example. Only one is displayed at any given time.

- (B) The Participant receives the visual stimuli to their visual cortex
- (C) The system waits for the participant's response
- (D) The participant should withhold from responding in the following two cases:
- **Colour condition.** If the colour of the dot is blue.
 - **Repeat condition.** If colour of the dot was identical to the previous dot
- (E) The participant will actuate their response from their motor cortex to click a button in the following way:
- A "go target" indicated by 'Event 3' in the EEG recordings requires the participant press the left mouse button with their right hand. This indicates an "Acknowledgement" indicated by Event 1 from the participant.
 - A "no go target" indicated by the following:
 - △ **Colour condition**, indicated by 'Event 4'
 - △ **Repeat condition**, indicated by 'Event 5'

- An **'Event 1'** after by **'Event 4'** or **'Event 5'** is an error made by the participant
- An **Error Correction Response** ECR indicated by **'Event 2'** in EEG recording is generated by participant pressing the right mouse button using their index finger of their right hand. This indicates that the participant is **"aware"** of the error the **'Event 2'** is after **'Event 1'**.

Appendix L

Error Selection Menu

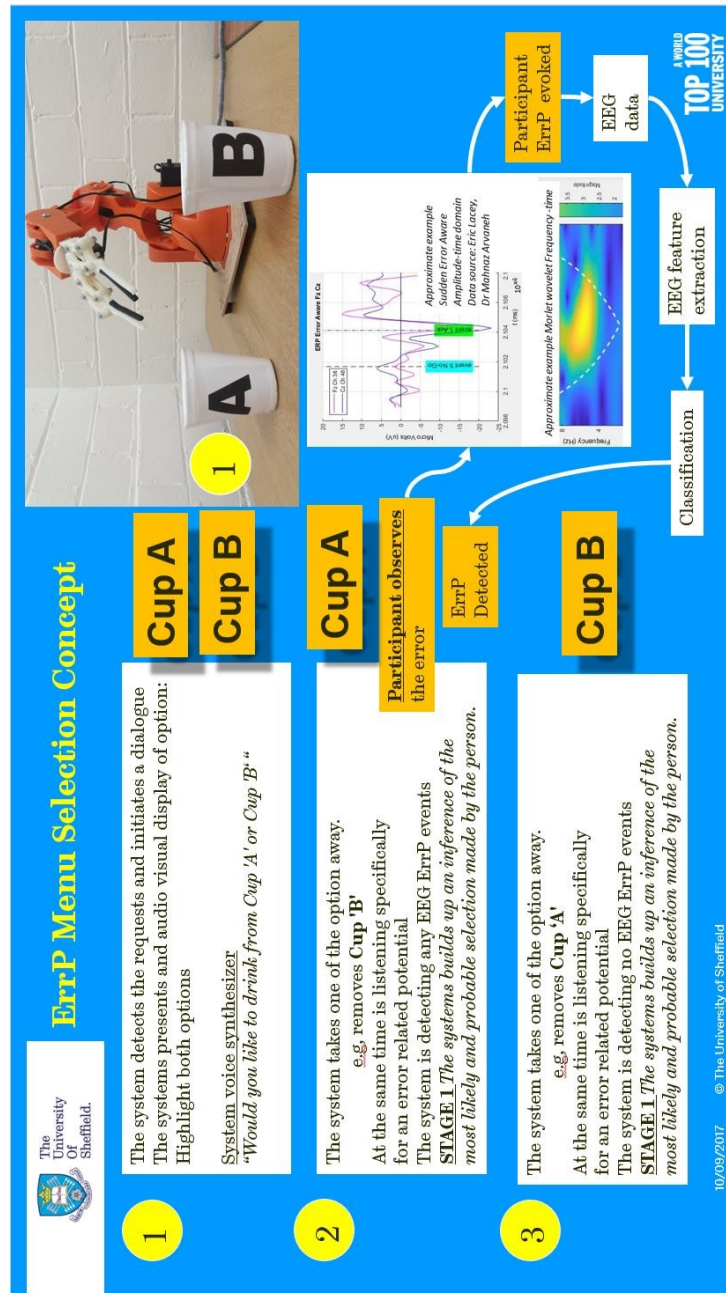


Figure L.1: ErrP menu selection process'

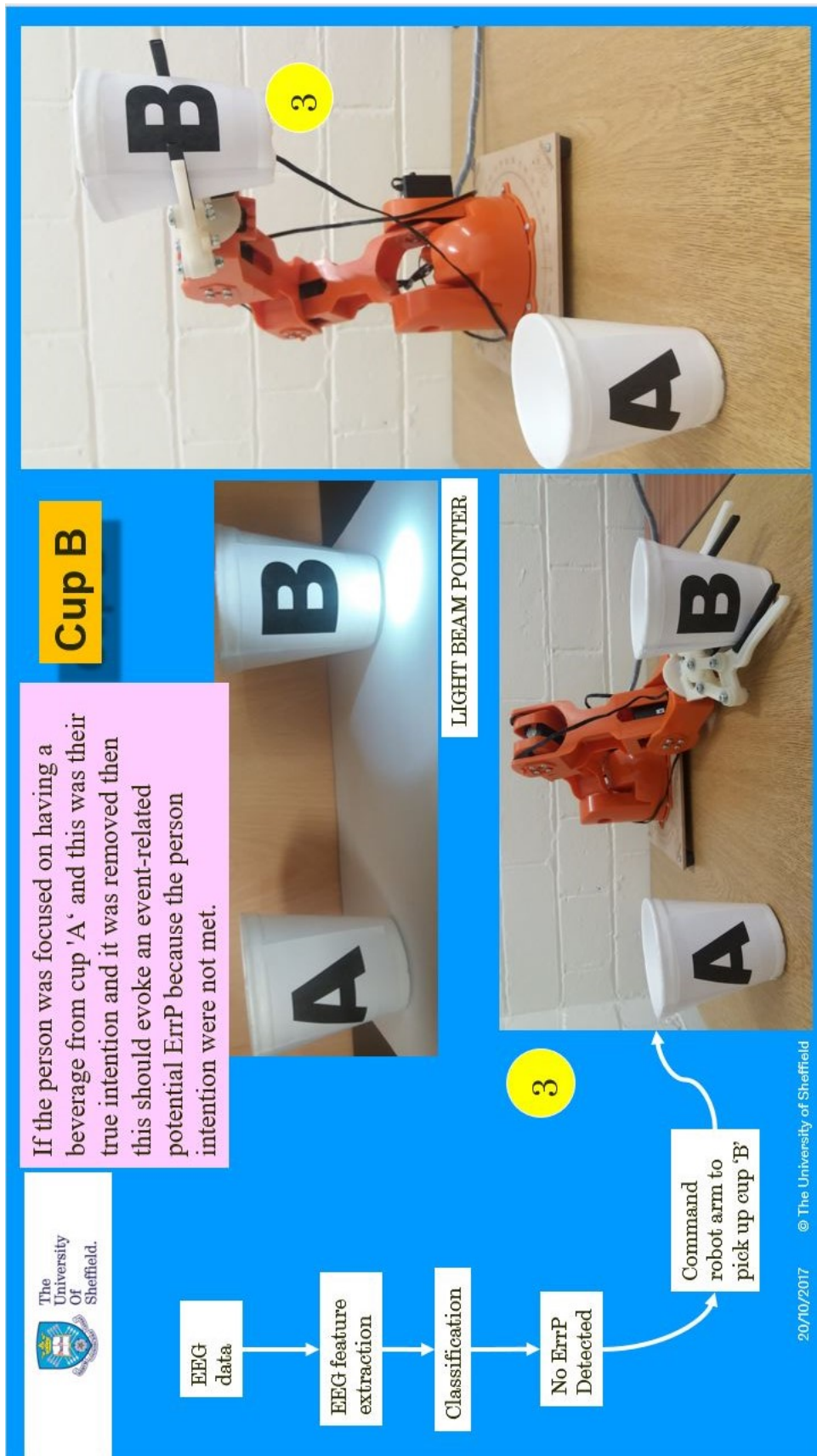


Figure L.2: ErrP menu selection process

Future possible assistive robotics



Fig. 1. SCAMP myoelectric hand in use by a child with below elbow deficiency.

**A Child's powered artificial limb (prosthesis)
Rehabilitation Medicine,
Mobility and Specialised Rehabilitation
Centre, Northern General Hospital Sheffield
Reference December 2016**

<https://publicdocuments.sth.nhs.uk/pil1630.pdf>

Using the prosthesis alone:

- Jigsaws with knobs
- Building bricks
- Holding food or sweets
(for example a packet of raisins, small yoghurt pots)

Using both hands:

- Use of knife and fork

Sheffield Teaching Hospitals
NHS Foundation Trust



Reference 1998: Northern General Hospital, Sheffield *Prosthetics and Orthotics International*, 1998. 22, 150-154 http://www.oandplibrary.org/poi/pdf/1998_02_150.pdf

18/09/2017 © The University of Sheffield

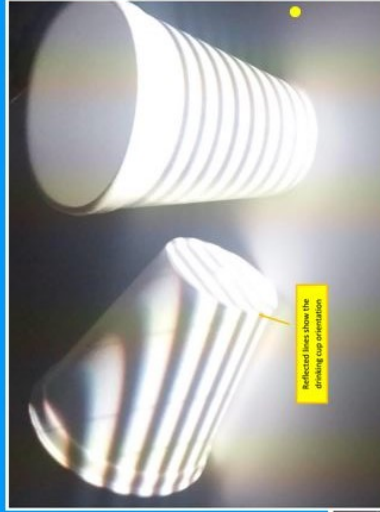
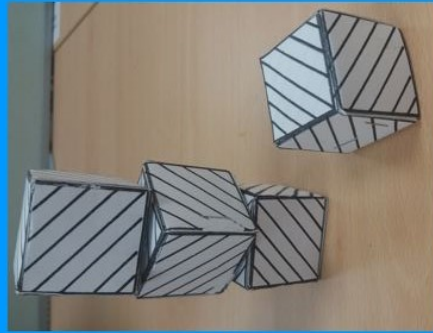
Figure L.3: Prosthetic upper limb



Future possible assistive robotics

Develop a brain-computer interface BCI such that the assistive robotic device works with all age groups during an interactive co-operative task to assist the person such as:

- **Building blocks**



- **Eating and drinking**

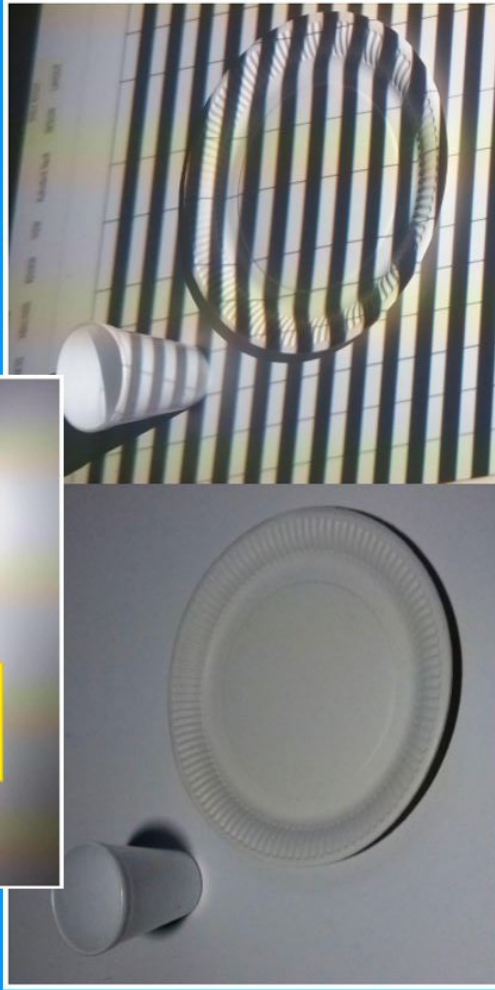


Figure L.4: Assistive Robotics Applications

References

- [1] University Of Sheffield MASH et al, "Mathematics and Statistics Help MASH," 2016.
- [2] B. N. Andrea Kubler, Boris Kotchoubey, Jochen Kaiser, Wolpaw J. R., "Brain-computer communication_Unlocking the locked in.pdf," vol. 127, no. 3, pp. 358–375, 2001.
- [3] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control.," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–91, 2002.
- [4] "BNCI Horizon 2020," 2017.
- [5] G. Pfurtscheller, B. Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, T. O. Zander, G. Mueller-Putz, C. Neuper, and N. Birbaumer, "The hybrid BCI," *Frontiers in Neuroscience*, vol. 4, no. April, p. 30, 2010.
- [6] P. W. Ferrez and J. del R Millan, "EEG-Based Brain-Computer Interaction: Improved Accuracy by Automatic Single-Trial Error Detection," *Proceedings of Neural Information Processing Systems Conference*, pp. 441–448, 2008.
- [7] L. Bi, X. A. Fan, and Y. Liu, "EEG-based brain-controlled mobile robots: A survey," *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 2, pp. 161–176, 2013.
- [8] a. Navarro-Cebrian, R. T. Knight, and a. S. Kayser, "Error-Monitoring and Post-Error Compensations: Dissociation between Perceptual Failures and Motor Errors with and without Awareness," *J Neurosci*, vol. 33, no. 30, pp. 12375–12383, 2013.
- [9] P. W. Ferrez, "You Are Wrong!—Automatic Detection of Interaction Errors from Brain Waves," *In Proceedings of the 19th international joint conference on Artificial intelligence*, 2005.
- [10] M. Ullsperger, C. Danielmeier, and G. Jocham, "Neurophysiology of Performance Monitoring and Adaptive Behavior," *Physiological Reviews*, vol. 94, no. 1, pp. 35–79, 2014.
- [11] J. Omedes, I. Iturrate, J. Minguez, and L. Montesano, "Analysis and asynchronous detection of gradually unfolding errors during monitoring tasks," *Journal of Neural Engineering*, vol. 12, no. 5, p. 056001, 2015.
- [12] J. R. Wolpaw, "Brain-computer interfaces as new brain output pathways," *Journal of Physiology*, vol. 579, no. 3, pp. 613–619, 2007.
- [13] R. Rupp, "Challenges in clinical applications of brain computer interfaces in individuals with spinal cord injury," *Frontiers in Neuroengineering*, vol. 7, no. September, p. 38, 2014.
- [14] M. Alimardani, S. Nishio, and H. Ishiguro, "Effect of biased feedback on motor imagery learning in BCI-teleoperation system," *Frontiers in Systems Neuroscience*, vol. 8, no. April, pp. 1–8, 2014.

- [15] P. W. Ferrez, J. Del, and R. Millán, "EEG-Based Brain-Computer Interaction: Improved Accuracy by Automatic Single-Trial Error Detection," in *Proceedings of Neural Information Processing Systems Conference*, pp. 441–448, Neural Information Processing Systems Conference, 2008.
- [16] M. Bryan, J. Green, M. Chung, L. Chang, R. Scherer, J. Smith, and R. Rao, "An adaptive brain-computer interface for humanoid robot control," *IEEE-RAS International Conference on Humanoid Robots*, pp. 199–204, 2011.
- [17] X. Perrin, R. Chavarriaga, F. Colas, R. Siegwart, and R. Millán, "Brain-coupled interaction for semi-autonomous navigation of an assistive robot," *Robotics and Autonomous Systems*, vol. 58, no. 12, pp. 1246–1255, 2010.
- [18] R. Chavarriaga, A. Sobolewski, and J. d. R. Millán, "Errare machinale est: The use of error-related potentials in brain-machine interfaces," 2014.
- [19] A. T. Sohaib, "An Empirical Study of Machine Learning Techniques for Classifying Emotional States from EEG Data," no. September, 2012.
- [20] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for EEG-based communication," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 3, pp. 386–394, 1997.
- [21] B. HB, "David Hubel and Torsten Wiesel Their contributions towards understanding," *Trends in neuroscience*, no. May, pp. 145–152, 1982.
- [22] M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein, "ERP components on reaction errors and their functional significance: A tutorial," 2000.
- [23] D. E. A. Lacey, "Error Awareness Dot Task (EADT)," tech. rep., Trinity College Dublin, School of Psychology, 2015.
- [24] I. Iturrate, L. Montesano, and J. M??niguez, "Robot reinforcement learning using EEG-based reward signals," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 4822–4829, 2010.
- [25] O. R. S. Blanco, H. Garcia, R. Quiñan Quiroga, L. Romanelli, "Stationarity of the EEG Series," *IEEE Engineering in Medicine and Biology*, 1995.
- [26] H. S. Anupama, N. K. Cauvery, and G. M. Lingaraju, "Brain Computer Interface and Its Types - A Study," *International Journal of Advances in Engineering & Technology*, vol. 3, no. 2, pp. 739–745, 2012.
- [27] B. Chambayil, R. Singla, and R. Jha, "Virtual keyboard BCI using Eye blinks in EEG," *2010 IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications, WiMob'2010*, pp. 466–470, 2010.
- [28] W. D. Penny, S. J. Roberts, E. A. Curran, and M. J. Stokes, "Approach," In *Proceedings of IEEE TRANSACTIONS ON REHABILITATION ENGINEERING*, vol. 8, no. 2, pp. 214–215, 2000.
- [29] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlögl, B. Obermaier, and M. Pregenzer, "Current trends in Graz Brain-Computer Interface (BCI) research," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 216–219, 2000.
- [30] I. Dokare, "Classification of EEG Signal for Imagined Left and Right Hand Movement for Brain Computer Interface Applications," vol. 2014, pp. 291–294, 2014.

- [31] R. Palaniappan, *Biological Signal Analysis*. bookboon.com Ltd., 1 ed., 2010.
- [32] Blausen, "Medical gallery of Blausen Medical 2014," *Wikiversity Journal of Medicine*, no. 2002-4436., 2014.
- [33] B. C. Oxley, "International 10-20 system."
- [34] N. K. Squires, K. C. Squires, and S. A. Hillyard, "TWO VARIETIES OF LONG - LATENCY POSITIVE WAVES EVOKED BY UNPREDICTABLE AUDITORY STIMULI IN MAN 1," pp. 387-401, 1975.
- [35] Y. Li, J. Long, T. Yu, Z. Yu, C. Wang, H. Zhang, and C. Guan, "An EEG-based BCI system for 2-D cursor control by combining Mu/Beta rhythm and P300 potential," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10 PART 1, pp. 2495-2505, 2010.
- [36] R. Dinteren, M. Arns, M. L. A. Jongsma, and R. P. C. Kessels, "P300 development across the lifespan: A systematic review and meta-analysis," *PLoS ONE*, vol. 9, no. 2, 2014.
- [37] B. Dal Seno, M. Matteucci, and L. Mainardi, "Online detection of P300 and error potentials in a BCI speller," *Computational Intelligence and Neuroscience*, vol. 2010, 2010.
- [38] I. Mischner, R. S. Schaefer, C. Gielen, and P. Desain, "Using multimodal frequency tagging for BCI," no. June 2014, p. 31, 2007.
- [39] C. Nangini, B. Ross, F. Tam, and S. J. Graham, "Magnetoencephalographic study of vibrotactile evoked transient and steady-state responses in human somatosensory cortex," *NeuroImage*, vol. 33, no. 1, pp. 252-262, 2006.
- [40] B. Ross, C. Borgmann, R. Draganova, L. E. Roberts, and C. Pantev, "A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones," *J Acoust Soc Am*, vol. 108, no. 2, pp. 679-691, 2000.
- [41] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, and D. Hu, "A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm," *Journal of Neural Engineering*, vol. 10, no. 2, 2013.
- [42] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b," *Frontiers in Neuroscience*, vol. 6, no. MAR, pp. 1-9, 2012.
- [43] L. C. Parra, C. D. Spence, A. D. Gerson, and P. Sajda, "Response error correction - A demonstration of improved human-machine performance using real-time EEG monitoring," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 173-177, 2003.
- [44] M. Wairagkar, "Motor Imagery based Brain Computer Interface (BCI) using Artificial Neural Network Classifiers,"
- [45] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 4, no. 2, pp. R1-R13, 2007.
- [46] M. Kaper, P. Meinicke, U. Grossekhoefer, T. Lingner, and H. Ritter, "BCI competition 2003 - Data set IIb: Support vector machines for the P300 speller paradigm," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1073-1076, 2004.
- [47] A. Haar, "Zur Theorie der orthogonalen Funktionensysteme," no. 2, 1909.

- [48] H. B. Kekre, S. D. Thepade, J. Jain, and N. Agrawal, "IRIS Recognition using Texture Features Extracted from Haarlet Pyramid," *International Journal of Computer Applications*, vol. 11, no. 12, pp. 1–5, 2010.
- [49] H. Ghaheri and A. Ahmadyfard, "Temporal windowing in CSP method for multi-class motor imagery classification," *ICEE 2012 - 20th Iranian Conference on Electrical Engineering*, pp. 1602–1607, 2012.
- [50] H. Parvin, H. Alizadeh, and B. Minaei-bidgoli, "MKNN : Modified K-Nearest Neighbor," *Proceedings of the World Congress on Engineering and Computer Science, WCECS*, pp. 22–25, 2008.
- [51] D. J. Mcfarland and J. R. Wolpaw, "Sensorimotor Rhythm-Based Brain – Computer Regression Improves Performance," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 3, pp. 372–379, 2005.
- [52] U. Orhan, M. Hekim, and M. Ozer, "EEG signals classification using the K-means clustering and a multilayer perceptron neural network model," *Expert Systems with Applications*, vol. 38, no. 10, pp. 13475–13481, 2011.
- [53] T. S. Körting, "How K-Means algorithm works," 2018.
- [54] P. Rani, C. Liu, N. Sarkar, and E. Vanman, "An empirical study of machine learning techniques for affect recognition in human-robot interaction," *Pattern Analysis and Applications*, vol. 9, no. 1, pp. 58–69, 2006.
- [55] T. S. Körting, "How kNN algorithm works."
- [56] Y. Liu, W. Wang, B. Lévy, F. Sun, D.-M. Yan, L. Lu, and C. Yang, "On centroidal voronoi tessellation—energy smoothness and fast computation," *ACM Transactions on Graphics*, vol. 28, no. 4, pp. 1–17, 2009.
- [57] I. Dokare and N. Kant, "Performance Analysis of SVM , k-NN and BPNN Classifiers for Motor Imagery," *International Journal of Engineering Trends and Technology*, vol. 10, no. 1, pp. 19–23, 2014.
- [58] D. J. Spiegelhalter and S. L. Lauritzen, "Techniques for Bayesian analysis in expert systems," *Annals of Mathematics and Artificial Intelligence*, vol. 2, no. 1-4, pp. 353–366, 1990.
- [59] D. Song, C. H. Ek, K. Huebner, and D. Kragic, "Multivariate discretization for Bayesian Network structure learning in robot grasping," *2011 IEEE International Conference on Robotics and Automation*, pp. 1944–1950, 2011.
- [60] K.-E. Ko, H.-C. Yang, and K.-B. Sim, "Emotion recognition using EEG signals with relative power values and Bayesian network," *International Journal of Control, Automation and Systems*, vol. 7, no. 5, pp. 865–870, 2009.
- [61] B. Moldovan, P. Moreno, D. Nitti, J. Santos-Victor, and L. de Raedt, "Relational affordances for multiple-object manipulation," *Autonomous Robots*, pp. 1–26, 2017.
- [62] A. Subasi and E. Erçelebi, "Classification of EEG signals using neural network and logistic regression," *Computer Methods and Programs in Biomedicine*, vol. 78, no. 2, pp. 87–99, 2005.
- [63] A. Hiraiwa, K. Shimohara, and Y. Tokunaga, "EEG Topography Recognition by Neural Networks," *IEEE Engineering in Medicine and Biology Magazine*, vol. 9, no. 3, pp. 39–42, 1990.

- [64] M. T. Hagan, *Neural network design*. Boston MA: PWS Publishing company, 1996.
- [65] R. Palaniappan, "Brain computer interface design using band powers extracted during mental tasks," *In Proceedings of 2nd International IEEE EMBS Conference on Neural Engineering*, vol. 2005, pp. 321–324, 2005.
- [66] C. Anderson and Z. Sijercic, "Classification of EEG signals from four subjects during five mental tasks," *Advances*, pp. 407–414, 1996.
- [67] E. Haselsteiner and G. Pfurtscheller, "Using time-dependent neural networks for EEG classification," *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 8, no. 4, pp. 457–463, 2000.
- [68] W. Penny and S. Roberts, "EEG-based communication via dynamic neural network models," *In Proceedings of IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339)*, vol. 5, pp. 3586–3590.
- [69] D. Maryanovsky, M. Mousavi, N. G. Moreno, and V. R. De Sa, "Csp-Nn: a Convolutional Neural Network Implementation of Common Spatial Patterns," *In Proceedings of the 7th Graz Brain-Computer Interface Conference 2017*, 2017.
- [70] M. Veres, M. Moussa, and G. W. Taylor, "Conditional Generative Models," vol. 2, no. 2, pp. 757–764, 2017.
- [71] C. Cortes, C. Cortes, V. Vapnik, and V. Vapnik, "Support Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [72] S. Raschka, "Linear Discriminant Analysis," 2014.
- [73] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K. R. Müller, "Single-trial analysis and classification of ERP components - A tutorial," *NeuroImage*, vol. 56, no. 2, pp. 814–825, 2011.
- [74] P. A. Herman, G. Prasad, and T. M. McGinnity, "Designing an Interval Type-2 Fuzzy Logic System for Handling Uncertainty Effects in Brain-Computer Interface Classification of Motor Imagery Induced EEG Patterns," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 1, pp. 29–42, 2017.
- [75] A. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, 2000.
- [76] J. Friedman, "On Bias, Variance, 0/1—Loss, and the Curse-of-Dimensionality," *Data Mining and Knowledge Discovery*, vol. 1, no. 1, pp. 55–77, 1997.
- [77] S. P. Layne, G. Mayer-Kress, and J. Holzfuss, "Problems Associated with Dimensional Analysis of Electroencephalogram Data BT - Dimensions and Entropies in Chaotic Systems: Quantification of Complex Behavior," pp. 246–256, Berlin, Heidelberg: Springer Berlin Heidelberg, 1986.
- [78] M. Falkenstein, J. Hohnsbein, J. Hoormann, and L. Blanke, "Effects of crossmodal divided attention on ERP components: Error processing in choice reaction tasks," *Electroencephalography and Clinical Neurophysiology*, vol. 78, no. 6, pp. 447–455, 1991.
- [79] W. J. Gehring, B. Goss, M. G. H. Coles, D. E. Meyer, and E. Donchin, "A Neural System For Error Detection and Compensation," *Psychological Science*, vol. 4, no. 6, pp. 385–391, 1993.

- [80] S. Debener, M. Ullsperger, M. Siegel, K. Fiehler, D. Y. von Cramon, and A. K. Engel, "Trial-by-Trial Coupling of Concurrent Electroencephalogram and Functional Magnetic Resonance Imaging Identifies the Dynamics of Performance Monitoring," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 25, no. 50, pp. 11730–11737, 2005.
- [81] M. J. Frank, B. S. Woroch, and T. Curran, "Error-related negativity predicts reinforcement learning and conflict biases," *Neuron*, vol. 47, no. 4, pp. 495–501, 2005.
- [82] J. R. Themanson, P. J. Rosen, M. B. Pontifex, C. H. Hillman, and E. McAuley, "Alterations in error-related brain activity and post-error behavior over time," *Brain and Cognition*, vol. 80, no. 2, pp. 257–265, 2012.
- [83] C. B. Holroyd and M. G. H. Coles, "The neural basis. of human error processing: Reinforcement learning, dopamine, and the error-related negativity," *Psychological Review-New York-*, vol. 109, no. 4, pp. 679–709, 2002.
- [84] S. F. Taylor, E. R. Stern, and W. J. Gehring, "Neural Systems for Error Monitoring," *The Neuroscientist*, vol. 13, no. 2, pp. 160–172, 2007.
- [85] D. M. Olvet and G. Hajcak, "Reliability of error-related brain activity," *Brain Research*, vol. 1284, pp. 89–99, 2009.
- [86] A. Riesel, A. Weinberg, T. Endrass, A. Meyer, and G. Hajcak, "The ERN is the ERN is the ERN? Convergent validity of error-related brain activity across different tasks," *Biological Psychology*, vol. 93, no. 3, pp. 377–385, 2013.
- [87] L. Boubchir, Y. Touati, B. Daachi, and A. A. Cherif, "EEG error potentials detection and classification using time-frequency features for robot reinforcement learning," 2015.
- [88] T. Zeyl, E. Yin, M. Keightley, and T. Chau, "Adding real-time Bayesian ranks to error-related potential scores improves error detection and auto-correction in a P300 speller," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 46–56, 2016.
- [89] A. Cruz, G. Pires, and U. J. Nunes, "Double ErrP Detection for Automatic Error Correction in an ERP-Based BCI Speller," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 4320, no. c, 2017.
- [90] M. Steinhauser, N. Yeung, T. A. Klein, and S. Shalgi, "Error awareness as evidence accumulation: effects of speed-accuracy trade-off on error signaling," 2012.
- [91] P. W. Ferrez and J. Del R. Millán, "Error-related EEG potentials generated during simulated brain-computer interaction," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 923–929, 2008.
- [92] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: A review of the first international meeting," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 164–173, 2000.
- [93] P. M. a. Rabbitt, "CHOICE-RESPONSE TASKS i," *Experimental Psychology*, vol. 71, no. 2, pp. 264–272, 1966.
- [94] X. Artusi, I. K. Niazi, M. F. Lucas, and D. Farina, "Performance of a simulated adaptive BCI based on experimental classification of movement-related and error potentials," 2011.

- [95] A. Llera, M. A. J. van Gerven, V. Gómez, O. Jensen, and H. J. Kappen, "On the use of interaction error potentials for adaptive brain computer interfaces," *Neural Networks*, vol. 24, no. 10, pp. 1120–1127, 2011.
- [96] R. Chavarriaga and J. D. R. Millán, "Learning from EEG error-related potentials in non-invasive brain-computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 381–388, 2010.
- [97] S. Wang, C. C.-J. Lin, C. Wu, and W. A. Chaovallitwongse, "Early Detection of Numerical Typing Errors Using Data Mining Techniques," *Systems, Man and ...*, vol. 41, no. 6, pp. 1199–1212, 2011.
- [98] T. O. Zander and C. Kothe, "Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general," *Journal of neural engineering*, vol. 8, no. 2, p. 025005, 2011.
- [99] T. O. Zander and S. Jatzev, "Context-aware brain-computer interfaces: exploring the information space of user, technical system and environment," *Journal of Neural Engineering*, vol. 9, no. 1, p. 016003, 2012.
- [100] G. Hajcak, J. S. Moser, N. Yeung, and R. F. Simons, "On the ERN and the significance of errors," *Psychophysiology*, vol. 42, no. 2, pp. 151–160, 2005.
- [101] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," vol. 139, pp. 23–37, 1995.
- [102] J. Brownlee, "A Tour of Machine Learning Algorithms," 2013.
- [103] L. Montesano, "Shared-control brain-computer interface for a two dimensional reaching task using EEG error-related potentials," pp. 6–10.
- [104] Z. Li, S. Zhao, C.-Y. Su, C. Yang, J. Duan, and X. Zhao, "Human Cooperative Wheelchair with Brain Machine Interaction Based on Shared Control Strategy," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 1–1, 2016.
- [105] C. M. Oddo, S. Raspopovic, F. Artoni, A. Mazzoni, G. Spigler, F. Petrini, F. Giambattistelli, F. Vecchio, F. Miraglia, L. Zollo, G. D. Pino, D. Camboni, M. C. Carrozza, E. Guglielmelli, P. M. Rossini, and U. Faraguna, "Intraneural stimulation elicits discrimination of textural features by artificial fingertip in intact and amputee humans," pp. 1–27, 2016.
- [106] G. Ghazaei, A. Alameer, P. Degenaar, G. Morgan, and K. Nazarpour, "Deep learning-based artificial vision for grasp classification in myoelectric hands," *Journal of Neural Engineering*, vol. 14, no. 3, 2017.
- [107] A. F. Salazar-Gomez, J. DelPreto, S. Gil, F. H. Guenther, and D. Rus, "Correcting robot mistakes in real time using EEG signals," *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6570–6577, 2017.
- [108] C. B. Peltz, G. Gratton, and M. Fabiani, "Age-related changes in electrophysiological and neuropsychological indices of working memory, attention control, and cognitive flexibility," *Frontiers in Psychology*, vol. 2, no. AUG, pp. 1–12, 2011.
- [109] World Medical Association, "World Medical Association Declaration of Helsinki," *Bulletin of the world health organization.*, vol. 79, no. 4, pp. 373–374, 2001.
- [110] R. Hester, J. J. Foxe, S. Molholm, M. Shpaner, and H. Garavan, "Neural mechanisms involved in error processing : A comparison of errors made with and without awareness," vol. 27, pp. 602–608, 2005.

- [111] N. Jatupaiboon, S. Pan-ngum, and P. Israsena, "Real-Time EEG-Based Happiness Detection System," *Scientific World Journal*, vol. 2013, 2013.
- [112] S. Chippa and S. Bengio, "HMM and IOHMM Modelling of EEG Rhythms for Asynchronous BCI systems," tech. rep., Dalle Molle Institute of Perceptual Artificial Intelligence, Valais Switzerland, 2003.
- [113] J. d. R. Millán and J. Mouriño, "Asynchronous BCI and local neural classifiers: An overview of the adaptive brain interface project," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 159–161, 2003.
- [114] F. Lotte, "Oscillatory EEG-based BCI design: signal processing and more."
- [115] M. Tangermann, K. R. Müller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, R. Leeb, C. Mehring, K. J. Miller, G. R. Müller-Putz, G. Nolte, G. Pfurtscheller, H. Preissl, G. Schalk, A. Schlögl, C. Vidaurre, S. Waldert, and B. Blankertz, "Review of the BCI competition IV," *Frontiers in Neuroscience*, no. JULY, 2012.
- [116] F. Lotte, M. Congedo, L. Anatole, F. Lotte, M. Congedo, and L. Anatole, "A review of classification algorithms for EEG-based brain – computer interfaces To cite this version : A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces," 2007.
- [117] F. Lotte and M. Congedo, "A review of classification algorithms for EEG-based brain – computer interfaces," *J Neural Eng*, vol. 4, pp. R1–R13, 2007.
- [118] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000.
- [119] D. J. McFarland, L. a. Miner, T. M. Vaughan, and J. R. Wolpaw, "Mu and beta rhythm topographies during motor imagery and actual movements.," *Brain topography*, vol. 12, no. 3, pp. 177–186, 2000.
- [120] A. Bashashati, M. Fatourehchi, R. K. Ward, and G. E. Birch, "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals," *Journal of Neural Engineering*, vol. 4, no. 2, 2007.
- [121] M. Zhang, Z. Zhang, Y. Chang, E.-s. Aziz, S. Esche, and C. Chassapis, "Recent Developments in Game-Based Virtual Reality Educational Laboratories Using the Microsoft Kinect," vol. 13, pp. 138–159, 2018.
- [122] K. H. Jian, SunRoss; Girshick,Shaoqing; Ren, "Object detection and classification in images," 2016.
- [123] A. Ashley, "Matrix Pattern Recognition Decision Making and Adaptive Learning Process," 2008.
- [124] Lund Research Ltd, "Laerd Statistics," 2018.