

DEVELOPING ENHANCED CLASSIFICATION METHODS FOR ECG AND EEG SIGNALS

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

College of Engineering and Science

Victoria University

Australia

by

Roozbeh Zarei

March 2017

© 2017 Roozbeh Zarei

ALL RIGHTS RESERVED

ABSTRACT

A huge amount of biomedical data such as Electrocardiography (ECG) and Electroencephalography (EEG) signals are recorded daily from human body to assess and monitor human performance and physiological condition. The analysis of these signals is important for research as well as for medical diagnosis and treatment. Although ECG and EEG signals provide useful information about the heart and brain, the classification of these signals has not been well developed. Even now these signals are often examined manually by physicians. Hence, there is a need for developing automatic classification techniques that evaluate and assess these signals. This thesis presents enhanced methods for the classification of ECG and EEG signals in three areas: the detection of premature ventricular contraction (PVC), the identification of epileptic seizure, and the recognition of motor imagery (MI) tasks in Brain-Computer Interface (BCI).

The classification of ECG signals is essential to diagnosis of critical heart conditions. In this thesis, ECG signals are classified to detect PVC beats which are the most common arrhythmias in humans. A new replacing strategy is developed to study and check the effects of PVC and normal heartbeats on the variation of principal directions. It is observed that the PVC beats have a larger effect on principal directions compared to the normal ones. Based on this property, an online PVC detection method is proposed to identify the PVC beats in the real-time. The experimental results indicate that the variation of principal directions caused by PVC beats can be used to identify PVCs accurately with an accuracy of more than 98%. The results also show the capability of the proposed method to effectively and efficiently detect PVC beats in online manner.

Epilepsy is one of the most common serious neurological disorder in humans that can be identified by analysing the concealed patterns of EEG signals. In this thesis, a

novel feature extraction technique based on the Douglas-Peucker algorithm (DP) and the principal component analysis (PCA) is presented for the identification of epileptic seizure in multi-category EEG signals. The DP is used to extract the most representative samples from a large number of EEG data, and then the PCA is applied to reduce the dimensionality of the extracted samples. Four machine learning techniques: random forest classifier (RF), k -nearest neighbour algorithm (k -NN), support vector machine (SVM), and decision tree classifier (DT) are employed to evaluate the effectiveness of the proposed method. The results indicate that the DP is effective for extracting representative samples from EEG data and the RF classifier is one of the best techniques for detection of the epileptic EEG signal. The proposed feature extraction method with the RF classifier achieves the overall classification accuracy of 99.85%.

BCI are devices that use mental activity, generated by a user such as a patient with motor disabilities, to control a computer or an external device via EEG signals. The identification of different categories of MI tasks based EEG signals is a main, yet very challenging, step in developing BCI systems. In this thesis, a robust feature extraction method based on the PCA and cross-covariance method (CCOV) is introduced for the extraction of discriminatory information from the MI-based EEG signals. The proposed method extracts a temporal pattern of MI tasks by discovering the inter-channel information in EEG signals using PCA. Then, it identifies the relationship between the EEG signals and the temporal pattern via the CCOV method to provide more discriminative information about those signals. To find the optimal feature set, the correlation-based variable selection method is applied to the extracted features. Three machine learning techniques: multilayer perceptron neural networks (MLP), least square support vector machine (LS-SVM), and logistic regression (LR) are employed on the obtained features to verify the robustness of the proposed method. The results show the proposed method is able to get representative information from mental states EEG signals in BCI appli-

cations. Furthermore, Both MLP and LS-SVM classifiers perform well (above 99% overall classification accuracy) in the MI-based EEG signals classification.

All three proposed methods have been tested and evaluated on publicly available databases and compared to their counterparts found in the literature. The results demonstrated in this thesis indicate the performance advancement of the proposed methods.

DECLARATION

I, Roozbeh Zarei, declare that the PhD thesis entitled '*Developing Enhanced Methods for Detecting Abnormalities in ECG and EEG Signals*' is no more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work.

Signature:



Date: 31/03/2017

*This thesis is dedicated to my parents
for their love, endless support
and encouragement.*

ACKNOWLEDGEMENTS

I would like to thank my principal supervisor Prof. Yanchun Zhang for his knowledge, guidance, patience and support. I would also like to thank my associate supervisors Prof. Jing He and Dr. Guangyan Huang for all their valuable feedback, advice, encouragement and support.

The research presented in this thesis was conducted in the College of Engineering and Science at Victoria University (2013-2017). I was supported by the Australian Postgraduate Research Scholarship (VUPRS) offered by the College of Engineering and Science. This financial support is gratefully acknowledged.

I have been warmed and encouraged by the enthusiastic support of my family and friends. I thank my parents Fariba and Iraj who often had more faith in me than I had in myself. I am grateful for the encouragement and support of my friends Leila Rahimi, Soheil Gohari, Peng Zhang and Lee Sun.

Last, but not least, I thank Dr. Siuly who provided me feedback and support during my PhD study.

LIST OF PUBLICATIONS

Based on this research work, the candidate has written the following papers, which have been published or resubmitted for publication in international journals.

1. **Roozbeh Zarei**, Jing He, Guangyan Huang, and Yanchun Zhang, “Effective and efficient detection of premature ventricular contractions based on variation of principal directions.”, *Digital Signal Processing* (50): 93-102, 2016.
2. **Roozbeh Zarei**, Jing He, Siuly Siuly, and Yanchun Zhang, “A PCA Aided Cross-Covariance Scheme for Discriminative Feature Extraction From EEG Signals”, *Computer Methods and Programs in Biomedicine*, (146): 47-57, 2017.
3. **Roozbeh Zarei**, Jing He, Siuly Siuly, and Yanchun Zhang, “Exploring Douglas-Peucker Algorithm in the Detection of Epileptic Seizure from Multi-category EEG Signals”, Submitted to *Engineering Applications of Artificial Intelligence*, 2017.

LIST OF ADDITIONAL PUBLICATIONS

Additional related publications by the author of this thesis are listed below.

1. Siuly Siuly, **Roozbeh Zarei**, Hua Wang, and Yanchun Zhang. “A New Data Mining Scheme for Analysis of Big Brain Signal Data.” *Databases Theory and Applications. ADC 2017*, pp. 151-164, 2017.
2. Jing He, **Roozbeh Zarei**, Jie Cao, Kersi Taraporewalla, Michale Steyn, Andre Van Zundert, Guangyan Huang, Haolan Zhang, and Chi-Hung Chi. “Building the Computational Virtual Reality Environment for Anaesthetists’ Training and Practice.” *IEEE International Conference on Services Computing (SCC)*, pp. 242-248, 2015.
3. Hao Lan Zhang, **Roozbeh Zarei**, Chaoyi Pang, and Xiaohui Hu. “Discovering New Analytical Methods for Large Volume Medical and Online Data Process-

ing.” *In International Conference on Health Information Science*, pp. 220-228. Springer International Publishing, 2014. (Best Paper Award)

4. Zhijun Xie, Guangyan Huang, **Roozbeh Zarei**, Jing He, Yanchun Zhang, and Hongwu Ye. “Wireless sensor networks for heritage object deformation detection and tracking algorithm.” *Sensors* 14, no. 11: 20562-20588, 2014.

TABLE OF CONTENTS

Declaration	iii
Dedication	iv
Acknowledgements	v
List of Publications	vi
Table of Contents	viii
List of Tables	x
List of Figures	xi
1 Introduction	1
1.1 Introduction	1
1.2 Motivations and Problems	2
1.3 Objectives and Contributions	5
1.4 Thesis Structure	7
2 ECG and EEG Signals-Background Knowledge and Classification	9
2.1 Background Knowledge Related to ECG Signals	9
2.1.1 Human Heart: Anatomy and Function	9
2.1.2 The Conduction System of the Heart	12
2.1.3 Electrocardiogram	15
2.1.4 Premature Ventricular Contraction	21
2.2 Background Knowledge Related to EEG Signals	23
2.2.1 Human Brain: Structure and Function	23
2.2.2 Neurophysiology of Human Brain	26
2.2.3 Electroencephalography	30
2.2.4 Epilepsy and Epileptic Seizures	33
2.2.5 Brain-Computer Interfaces	36
2.3 Review of the ECG and EEG Classification	39
2.3.1 PVC Beat Classification Methods	40
2.3.2 Epileptic Classification Methods	44
2.3.3 MI Signal Classification Methods in the BCI	47
3 Effective and Efficient Detection of Premature Ventricular Contractions Based on Variation of Principal Directions	51
3.1 Introduction	51
3.2 ECG Data	54
3.3 Method for Online PVC Heartbeat Detection	55
3.3.1 Heartbeat Detection	57
3.3.2 Segmentation and Normalisation	57
3.3.3 PVC Detection via Variation of Dominant Principal Directions	58
3.4 Classification Performance Measures	67
3.5 Results	69
3.5.1 Performance of the PVC Classifier	69

3.5.2	Comparison with Published Studies	70
3.6	Conclusions	75
4	Exploring Douglas-Peucker Algorithm in the Detection of Epileptic Seizure from Multi-category EEG Signals	76
4.1	Introduction	76
4.2	Previous Work	81
4.3	Proposed Approach	83
4.3.1	Data Segmentation	84
4.3.2	Douglas-Peucker Algorithm	85
4.3.3	Dimension Reduction by PCA	88
4.3.4	DP_PCA Feature Set	90
4.3.5	Classification by the RF, k -NN, SVM, and DT	90
4.4	Data and Implementation	94
4.4.1	The Epileptic EEG Data	94
4.4.2	Implementation	95
4.5	Results and Discussions	101
4.5.1	Classification Results for Each Classifier	102
4.5.2	Comparison	108
4.6	Conclusion	109
5	A PCA Aided Cross-Covariance Scheme for Classification of MI Tasks	111
5.1	Introduction	112
5.2	Materials and Method	114
5.2.1	EEG Datasets	114
5.2.2	Proposed Method	116
5.3	Performance Evaluation	127
5.4	Experiments and Results	128
5.4.1	Parameter Selection	128
5.4.2	The MI Classification Results for Each Classifier Testing Different Features Fs_1 and Fs_2	129
5.4.3	Performance Comparisons Between the Proposed Methods and the Existing Techniques	134
5.5	Conclusions	137
6	Conclusions and Future Work	139
6.1	Summary	139
6.2	Future Work	142
	Bibliography	144

LIST OF TABLES

2.1	Significance of EEG in different frequency bands.	34
3.1	Heartbeat classes associated with the extracted beats for the full database, dataset 1 (Ds1) and dataset 2 (Ds2) from the MIT-BIH arrhythmia database.	55
3.2	The resulting angle of each heartbeat in the sample ECG.	63
3.3	The effect of different values of k on resulting angles of different heartbeats classes for nine records in the training dataset DS1.	66
3.4	Performance measures used in this study for distinguishing PVC from non-PVC.	68
3.5	Overall performance for nine records in the training set (DS1) using different values of r_1 and r_2 in the classification function.	69
3.6	Classification performance on each recording of DS1 using the AAMI recommended performance measures.	71
3.7	Confusion matrix on DS1 data set.	72
3.8	Classification performance on each recording of DS2 using the AAMI recommended performance measures.	73
3.9	Confusion matrix on DS2 data set.	74
3.10	Comparison between this study and published studies.	74
4.1	Summary of the epileptic EEG data.	95
4.2	The obtained value of ϵ for each <i>Segm</i> in each of the five classes	97
4.3	The representative samples chosen by DP for each <i>Segm</i> in each of the five classes.	99
4.4	Classification results on the epileptic EEG data.	103
4.5	Obtained false positive rate (<i>FPR</i>) for each of the proposed classifiers. .	106
4.6	Comparison with the existing methods on epileptic EEG database. . . .	110
5.1	Set of features extracted from each cross-covariance sequence.	123
5.2	The selected features by BestFirst+CfsSubsetEval for each subject in datasets (IVa and IVb).	124
5.3	The obtained values for the LS-SVM parameters (γ and σ^2) for each subject in dataset IVa.	130
5.4	The obtained values for the LS-SVM parameters (γ and σ^2) for the subject in dataset IVb.	130
5.5	The classification results of each classifier based on the eleven-feature set (Fs1) in dataset IVa.	131
5.6	The classification results of each classifier based on the two-features set (Fs2) in dataset IVa.	131
5.7	The classification results of each classifier based on the eleven-feature set (Fs1) and the two-feature set (Fs2) in dataset IVb.	133
5.8	Performance comparisons with the existing methods for dataset IVa . .	137

LIST OF FIGURES

2.1	Basic heart anatomy schema (Adapted from [35]).	10
2.2	The AV plane and the four valves view from above with the atria and major vessels removed (Adapted from [36]).	11
2.3	The conduction system of the heart (Adapted from [35]).	13
2.4	Ion exchanges at the surface of the cell membrane that generate the action potential (Adapted from [39]).	14
2.5	The cardiac action potential for each of the specialised cells found in the heart (Adapted from [35]).	15
2.6	A typical ECG waveform (Adapted from [35]).	16
2.7	Standard 12-lead ECG placement (Adapted from [43]).	17
2.8	Standard limbs lead positions of Einthoven (Adapted from [43]).	19
2.9	Augmented lead positions and connections (Adapted from [43]). R stands for resistors which are usually set to 5 K.	19
2.10	Precordial lead positions (Adapted from [43]). R stands for resistors which are normally set to 5 K.	20
2.11	Frontal and horizontal planes (Adapted from [46]).	21
2.12	Typical ECG beat:	22
2.13	Schematic midsagittal view of the human brain (Adapted from [57]).	24
2.14	The four major lobes of the cerebrum (Adapted from [35]).	26
2.15	CNS nerve cells: (Adapted from [59])	27
2.16	Action potentials: (Adapted from [59])	28
2.17	Electrical dipole measured by EEG electrode (Adapted from [60]).	30
2.18	The International 10-20 electrode placement system (Adapted from [61]).	32
2.19	Example of EEG rhythms in different frequency bands (Adapted from [64]).	33
2.20	General BCI architecture (Adapted from [71]).	37
3.1	Schematic representation of the proposed PVC detection.	56
3.2	A segment of ECG recording 119, with the R-wave peaks of non-PVC beats marked by 'N' and PVC beats marked by 'V'.	56
3.3	First segment extracted from the sample ECG signal shown in Figure 3.2.	58
3.4	The Construction of normal data matrix \mathbf{M} and data matrix \mathbf{M}'_2	61
3.5	The effects of replacing a non-PVC or a PVC beat on the principal directions.	62
4.1	Block diagram of the proposed method for the classification of epileptic EEG signals.	84
4.2	An example of determining <i>Segms</i> from an EEG signals of a class.	85
4.3	An example of DP sample point extraction.	86
4.4	The structure of random forest classifier.	91
4.5	Example of five different sets of EEG signals taken from different subjects.	96

4.6	The typical results of DP for the healthy subject:	98
4.7	The typical results of DP for the epileptic patient:	99
4.8	The cumulative eigenvalues for all 100 eigenvectors.	100
4.9	Individual classification performances of each of the ten folds in each class for the proposed classifiers: (a) RF, (b) k -NN, (c) SVM, (d) J48 . .	104
4.10	The overall classification accuracy (OCA) in each of the ten folds. . . .	105
4.11	The AUC for the proposed classifiers.	107
4.12	Kappa statistics values for the proposed classifiers	108
5.1	Block diagram of the proposed method for the MI tasks classification in BCIs.	116
5.2	The dominant temporal patterns extracted from:	119
5.3	Typical cross-covariance sequences obtained for the right-hand and the right-foot MI signal for the subject aa in dataset IVa.	121
5.4	Typical cross-covariance sequences obtained for the right-foot and the left-hand MI signal for dataset IVb.	122
5.5	The total execution time of the proposed method based on:	135
5.6	ROC area for the MLP, LS-SVM and LR classifiers with feature set $Fs2$ in dataset IVa and IVb.	136

CHAPTER 1

INTRODUCTION

1.1 Introduction

Due to the rapid development of technology and increased usage of portable monitoring devices, a significant amount of biomedical data is recorded daily to monitor and observe physiological condition of human body. These biomedical signals quantify the physiological activities of different organs such as heart, brain, muscles, cornea, etc. They are usually obtained by placing one or more electrodes on the organ of interest. Electrocardiogram (ECG) and Electroencephalogram (EEG) are the most common biomedical signals recorded from heart and brain, respectively.

ECG signal is an electrical signal generated by cardiac muscles. ECG signals reflect the electrical activity of the heart over time and provide useful information about the state of the heart. EEG signal is a recording of the electrical activities of the brain from electrodes on the scalp. EEG signals provide information about how the brain functions act over time. A detailed discussion about ECG and EEG is provided in Chapter 2. As EEG and ECG are fully safe and non-invasive procedures with no possible risk, they have become valuable tools for research purposes and disease diagnosis. ECG is commonly used for diagnosis of heart disease and abnormalities such as arrhythmia [1], detecting obstructive sleep apnea [2], and biometric identification [3]. EEG has been used in different fields, including brain disease diagnosis [4], fatigue monitoring [5], and brain computer interface [6].

The analysis and classification of ECG and EEG signals are vital both for research and for medical diagnosis and treatment. However, it is a challenging task to effectively

and efficiently analyse these biomedical signals due to the complexity of these signals. The evaluation of these signals is generally conducted by trained medical experts who visually scan the ECG and EEG records. The visual scanning of ECG and EEG signals is very time-consuming and costly; it may be inaccurate, very complex, subject to judgement and human error as these signals usually contain a huge amount of data. Therefore, there is an increasing need for developing automated classification methods not only to alleviate the expert's burden of analysing long-term ECG and EEG signals but also to ensure a proper diagnosis and evaluation of diseases of the heart and brain.

This thesis focuses on the classification of ECG and EEG signals in three areas, namely, premature ventricular contraction (PVC) classification, epileptic seizure detection, and classification of motor imagery tasks for BCI system.

1.2 Motivations and Problems

PVC is the most common arrhythmias in humans which can be linked to mortality associated with myocardial infarction [7]. It is considered as a precursor of more severe cardiac arrhythmias such as ventricular fibrillation [8], and atrial fibrillation [9]. Therefore, their immediate detection and classification are essential to prevent the possible beginning of life-threatening cardiac conditions. PVCs detection and classification can be performed by monitoring and analysing the ECG signals as they generate a different ECG morphology than a normal heart beat. A description about PVC is provided in Chapter 2.

The accurate classification of ECG signals can provide sufficient information to detect PVC arrhythmias and aid physicians to find best treatment for patients. In recent years, various methods for the automatic detection of ECG beats have been developed

for this purpose. From the literature, it is observed that there are still some limitations in the PVC detection and classification methods. As ECG signals commonly exhibit inter- and intra-patient variability in morphology and timing, one challenge in the current ECG classifications is how to achieve high classification accuracy over a large number of patients. A classifier can perform well and achieve high classification performance on relatively small data sets but it could easily fail when faced with a big dataset containing various ECG waveforms from different patients. Developing a simple classifier which is capable of classifying PVC arrhythmia on real-time is another challenge in ECG arrhythmia classification. Most of the existing techniques employ many complicated mathematical tools such as wavelet transform and artificial neural network to extract the relevant information from ECG signals for arrhythmia classification. Thus, they are not suitable for real time implementation where the lower computation complexity is expected. The efficiency of these methods is mainly accompanied by high complexity and long computational time. To overcome these problems, in this thesis, a method was proposed for online PVC classification which is evaluated on the benchmark MIT-BIH arrhythmia database.

Two of the most common applications of EEG are in epilepsy and BCI. Epilepsy is one of the most common neurological disorders of the human brain that affects approximately 65 million people in the world [10]. It is characterised by spontaneously recurrent seizures [11]. A description about epilepsy and epileptic seizures can be found in Chapter 2. In most cases, seizures occur unexpectedly, without a sign of warning to alert and prepare the person for an onset of a seizure. Such abrupt and uncontrollable nature of the disease can cause physical injury due to loss of motor control, loss of consciousness, or delayed reactivity during seizures. Epileptic seizures are caused by rapid changes in the electrical functioning of the brain. This electrical epileptic activity can generate abnormalities in EEG signals. Detection of these abnormalities is a vital

component in diagnosis and treatment of epilepsy.

BCI is a communication system that provides a direct communication channel for transmitting messages from the human brain to computers by analysing the brain's mental activities [12]. BCI systems make the use of EEG signals to translate a subject's thought or intention into a control signal that allows a subject, such as a disabled person, to communicate with a device, such as a wheelchair. BCI systems can only detect and classify specific activity patterns in continuous brain signals associated with certain events or tasks. In recent years, there has been increased attention to the classification of motor imagery (MI) tasks for BCI applications. MI task is a mental process by which a user simulates or imagine a given action such as movements of hands and feet, without performing it. A detailed discussion of BCI systems is provided in Chapter 2.

The design of an effective EEG classification is a challenging task due to complex characteristics of EEG signals such as non-stationary, aperiodic, and poor signal-to-noise ratio. The main challenge of any detection method is the extraction of representative characteristics or features from EEG signals as it significantly affects the performance and the computational time of the classification method. In the past two decades, various algorithms have been developed for the automatic detection of EEG signals to identify different categories of EEG data. From the literature, it is observed that the performances of most existing methods are limited regarding success rate and effectiveness. Some methods are not suitable for practical applications as they are complex and have high computational time to obtain the relevant information from EEG signals. Although most of the EEG recordings are multi-categories in a real clinical application, most of the current methods are applied for binary EEG classification problems (Normal signal vs. epileptic signal), and only a few methods focus on multi-class EEG classification. By considering the aforementioned issues, this thesis focuses on developing robust EEG

classification methods for epileptic seizure detection, and motor imagery classification in BCI.

1.3 Objectives and Contributions

This research aims to develop new methods and techniques for the classification of ECG and EEG signals. The main objectives of this thesis are:

- To develop a low complexity algorithm for online premature ventricular contractions detection from ECG signals.
- To propose a method for accurate epileptic seizure detection in multi-category EEG signal.
- To introduce a method for classification of motor imagery based EEG signals in BCI system.
- To empirically evaluate the classification performance of the proposed methods on different benchmark ECG and EEG databases with some state-of-the-art methods presented in the literature.

In summary, the main contributions of this thesis are:

- *A Method for Detection of Premature Ventricular Contractions*

An online PVC detection method based on the variation of principal directions is proposed to effectively and efficiently detect PVC beats in an online manner. The method uses a replacing strategy to check the effects of each heartbeat on the variation of principal directions. In order to reduce the computational complexity

of the proposed method, an efficient computation for estimating principal directions is also proposed. The proposed method is evaluated on two non-overlapping data sets from the MIT-BIH arrhythmia database to assess its performance over large data. The experimental results showed that the proposed PVC detection method can achieve high classification accuracy and outperforms some existing methods [13, 14, 15, 16]. This approach can help clinicians to diagnose PVC arrhythmia by processing ECG signal in an online manner.

- *A Method for Identification of Epileptic Seizure*

A novel feature extraction technique based on Douglas-Peucker algorithm (DP) and principal component analysis (PCA) is developed for classification of epileptic seizure in multi-category EEG signals. In this research, the effectiveness of the proposed method is evaluated by applying four machine learning techniques: random forest classifier (RF), k -nearest neighbour algorithm (k -NN), support vector machine (SVM), and decision tree classifier (DT) on a benchmark epileptic EEG dataset. The results indicate that the proposed method can effectively extract the representative features set from EEG signals and the random forest classifier achieves the best performance on the proposed features set compared to the other classifiers tested. The proposed method also outperforms some recently reported methods [17, 18, 4, 19, 20, 21]. This method can help neurologists and researchers to classify the epileptic EEG signals.

- *A Method for Classification of Motor Imagery Tasks*

To improve classification performances and to reduce the time complexity, a robust feature extraction method based on the PCA and cross-covariance method (CCOV) is introduced for the extraction of discriminatory features from the motor imagery based EEG signals. A correlation-based variable selection method is applied to find an optimal feature set from the extracted features. To ver-

ify the effectiveness of the proposed method, three machine learning techniques: multilayer perceptron neural networks (MLP), least square support vector machine (LS-SVM), and logistic regression (LR) are employed for the classification of the extracted MI features. The proposed method is tested on two benchmark datasets, IVa and IVb of BCI Competition III. The results show all three classifiers achieve high classification performance for the proposed feature set and the proposed method outperforms the most recently reported methods [22, 23, 24, 25, 6, 26, 27, 28, 29, 30] in terms of average accuracy. The proposed method can effectively extract characteristic information from MI-based EEG signals in BCI applications and also reduce the computational complexity of classifiers by reducing the number of extracted features. It also has a potential for developing an MI-based BCI analyses that can contribute in diagnosis and rehabilitation practices.

1.4 Thesis Structure

The rest of this thesis is organised as follows:

Chapter 2 provides background knowledge about ECG and EEG signals and reviews some existing methods for ECG and EEG classification. It first introduces the general concepts and background knowledge about the human heart, the fundamentals of ECG, PVC and the effect of PVC on the ECG signals. Then a background knowledge about the human brain is provided, together with the fundamentals of EEG, epilepsy and BCIs. Finally, this chapter reviews some methods for the classification of ECG and EEG signals.

Chapter 3 presents an online PVC detection method based on the variation of prin-

cipal directions. In this approach, the effects of non-PVC and PVC heartbeats on the variation of principal directions are studied and used for the classification of PVC beats. It also introduces an effective computation for estimating principal directions to reduce the computational complexity of the proposed method.

Chapter 4 introduces a novel feature extraction technique based on the DP and the PCA for detection of epileptic seizure from multi-category EEG signal. It also investigates the effectiveness of DP algorithm in the classification of epileptic seizure from EEG data and identifies a suitable classifier for the proposed feature method by evaluating four classification algorithms.

Chapter 5 presents a robust feature extraction method combining the PCA and the CCOV for the extraction of discriminatory features from the mental states based on EEG signals in BCI applications. It also investigates which features are the best for characterising the distribution of mental states signals and evaluates three classification algorithms to find out an efficient classifier for the MI signal classification.

Chapter 6 summarises and concludes the current works in this thesis and points out some directions for the future works.

CHAPTER 2

ECG AND EEG SIGNALS-BACKGROUND KNOWLEDGE AND CLASSIFICATION

This chapter provides background knowledge about ECG and EEG signals and reviews some existing methods for ECG and EEG classification. Section 2.1 provides the general concepts and background knowledge about ECG and PVC. A background knowledge about EEG, epilepsy and BCIs is given in Section 2.2. Section 2.3 reviews some methods for the classification of ECG and EEG signals in three areas: the detection of PVC, the identification of epileptic seizure, and the recognition of MI tasks in BCI.

2.1 Background Knowledge Related to ECG Signals

This section provides the background knowledge related to ECG signals. As ECG is generated from the heart, Section 2.1.1 gives a brief overview on the anatomy of the heart and its functions. Section 2.1.2 focuses on the electrical system of the heart and the sources behind its electrical activities. A review on ECG and its waves is discussed in Section 2.1.3. Section 2.1.4 briefly describes PVC and its effect on the ECG.

2.1.1 Human Heart: Anatomy and Function

The heart is a muscular organ that pumps blood continuously throughout the body via the blood vessels of the circulatory system, providing nutrients and oxygen to all the tissues of the body and removing metabolic wastes and carbon dioxide [31] . The heart is located in the centre of the chest with its apex toward the left and is surrounded by the

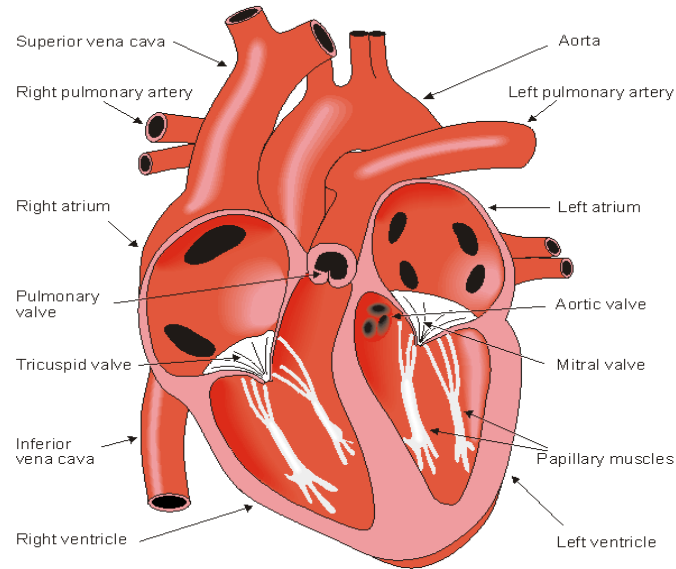


Figure 2.1: Basic heart anatomy schema (Adapted from [35]).

pericardium. The human heart is roughly the size of a clenched fist and weighs between about 280 to 340 grams in men and 230 to 280 grams in women [32].

The human heart consists of four chambers (as shown in Figure 2.1): upper left and right chambers (atria) and lower left and right chambers (ventricles) [33]. The right atrium and the right ventricle are together considered as right heart, and the left atrium and the left ventricle are considered as left heart [34]. The right heart pumps deoxygenated blood to lungs for oxygenation while the left heart pumps oxygenated blood to all body tissues for sustaining cellular respiration.

The heart has four valves namely, tricuspid valve, pulmonary valve, mitral valve and aortic valve, which are responsible for controlling the direction of blood flow through the heart [34]. The mitral valve connects the left atrium and the left ventricle while the tricuspid valve connects the right atrium and the right ventricle. The valve between the right ventricle and the pulmonary artery is the pulmonary valve, while The valve found between the left ventricle and the aorta is called the aortic valve. The four valves lie in

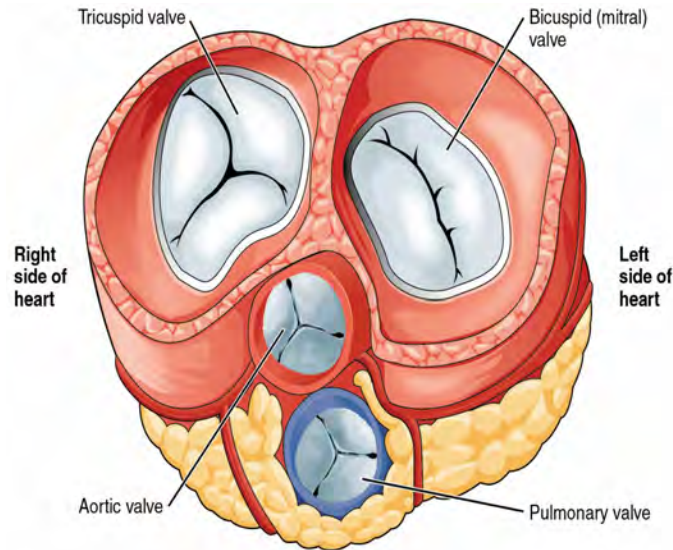


Figure 2.2: The AV plane and the four valves view from above with the atria and major vessels removed (Adapted from [36]).

a plane which is commonly known as the atrioventricular (AV) plane. Figure 2.2 shows an overview of the AV plane and all four valves. The four valves are surrounded by dense connective tissue which keeps the structure of the valves and electrically isolates the atria and ventricle.

The wall of the heart comprised of three layers: epicardium (outer layer), myocardium (middle layer), and endocardium (inner layer) [34]. The epicardium protects and lubricates the outside portion of the heart and is composed of connective tissues. The thickest part of the heart wall is the myocardium which is formed by cardiac muscle and is responsible for pumping the blood. The endocardium keeps blood from sticking and prevents the formation of harmful blood clots.

The human heart works regularly and continuously to distribute oxygen to the whole body. The blood is circulated through two pathways: the pulmonary circuit and the systemic circuit. In the pulmonary circuit, deoxygenated blood from the right ventricle flows through the pulmonary artery to the lungs. The deoxygenated blood releases car-

bon dioxide and receives oxygen in the lungs and returns as oxygenated blood to the left atrium through pulmonary vein. In the systemic circuit, oxygenated blood is forced out from the left ventricle to the systemic circulation via the aorta and then goes into smaller arteries and capillaries to provide oxygen to body tissues. From tissues, deoxygenated blood returns to the heart via the venae cavae and enters the right atrium. This blood circulation is achieved by a series of contractions and relaxation of the heart muscle in a cyclic pattern which can be described by the Cardiac Cycle.

2.1.2 The Conduction System of the Heart

The conducting system of the heart consists of a group of specialised cardiac muscle cells known as autorhythmic cells and conducting fibres. It is responsible for generating electrical impulses and conducting them through the heart. Figure 2.3 illustrates the conduction system of the heart. The autorhythmic cells serve as a pacemaker to start the normal cardiac cycle and provide the conduction system to coordinate the contractions of cardiac chambers. These cells periodically generate electrical impulses without nervous input. However, the rate of electrical impulses produced by these cells is controlled by the autonomous nervous system depending on needs of the body.

The conducting system of the heart propagates electrical waves over the myocardium (cardiac muscle) and depolarises cardiac cells known as myocytes to contract and then relax, resulting in the cardiac function [37]. As shown in Figure 2.3, the conducting system of the heart is made up of following elements: Sino-atrial (SA) node, Atrio-ventricular (AV) node, a bundle of His, left and right Bundle branches, and Purkinje fibres. The SA node is the natural pacemaker of the heart and is located at the junction of the superior vena cava and the right atrium. It initiates the cardiac cycle by releasing

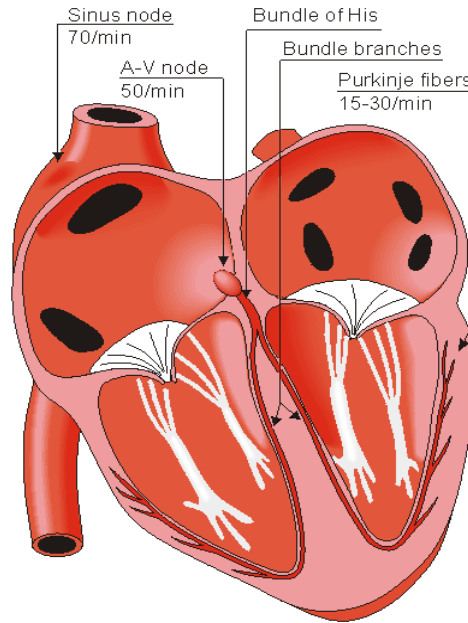


Figure 2.3: The conduction system of the heart (Adapted from [35]).

an electrical impulse that propagates as an electrical wave through both atria towards the AV nodes. The electrical wave depolarises the atrial myocytes and creates a wave of contraction on both atria which pumps the blood to the ventricles. The electrical impulse from SA node eventually reaches the AV node which is located at the junction of atria and ventricles and is delayed briefly (120 milliseconds) in order to allow the atria to contract and fully pump all the blood into the ventricles. Both the mitral valve and the tricuspid valve close once the atria pump all blood into the ventricles. At this phase, the atria start to refill, and the electrical impulse travels through the AV node and Bundle of His into the Bundle branches and Purkinje fibres. Finally, the electrical impulse is spread throughout the ventricular wall cells causing them to contract. Then the right ventricle pumps blood to the lungs while the left ventricle pumps blood into the aorta.

The electrical wave passes in all directions from one cell to another by modifying the ion concentration inside and outside of the cell membrane [37] as shown in Figure 2.4. When a cardiac cell receives a stimulus from the conduction system, sodium

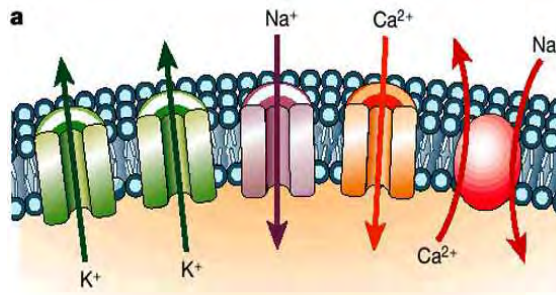


Figure 2.4: Ion exchanges at the surface of the cell membrane that generate the action potential (Adapted from [39]).

ions (Na^+) rush into the cell through a set of specialised ion channels, causing the inside of the cell to become more positive. Calcium ions (Ca^{2+}) also flood inside the cell through a different set of specialised channels. This release of calcium maintains the depolarised state of the cell membrane and initiates the contraction of the myocyte [38]. To balance the increased membrane voltage, potassium ions (K^+) inside the cell flow to the extracellular space. Once the cell is depolarised, it cannot respond to an electrical impulse anymore for a short period which is called refractory period. Following depolarisation, each cell must repolarise by re-equilibrating the ion concentrations to come back to its relax state. This sequence of change in electrical membrane potential (transient polarisation/depolarisation) is called action potential (AP), and the duration of an action potential is called Action Potential Duration (APD). Figure 2.5 illustrates the cardiac action potential for each of the specialised cells found in the heart. It can be seen from Figure 2.5, there are differences in the waveforms of action potentials recorded in different regions of the heart. There are also differences in the duration of the action potential in these areas. These potential differences can be detected and recorded by placing electrodes on the skin. The recorded signal is called the ECG. In the following section, the background information of ECG is provided in detail.

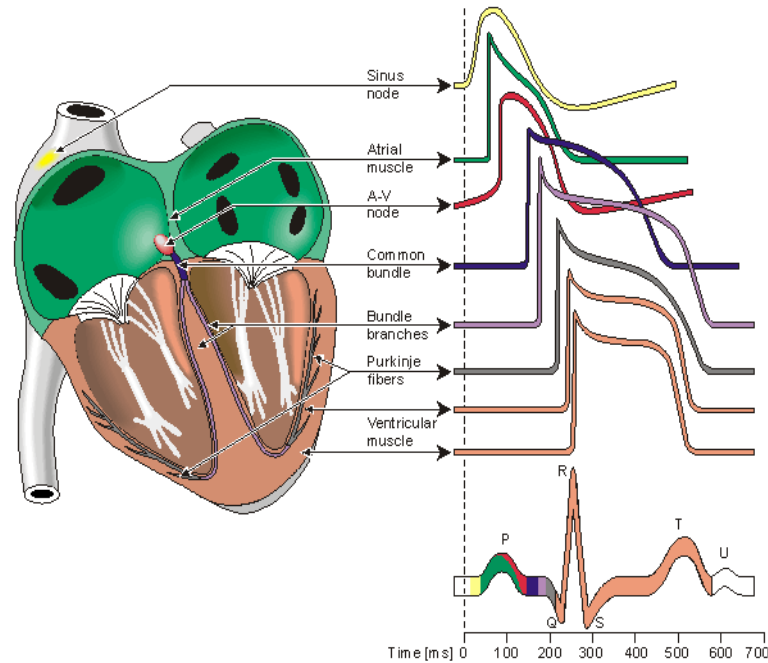


Figure 2.5: The cardiac action potential for each of the specialised cells found in the heart (Adapted from [35]).

2.1.3 Electrocardiogram

ECG is the non-invasive method used to record the electrical activity of the heart over time by attaching electrodes to the surface of the patient's skin. It is a diagnostic tool which provides information about the state of the heart. It is widely used by physicians to study heart function and to diagnose heart diseases [40]. ECG device records the electrical cardiac signals as characteristic waves which reflect different periods of the heart cycle. Figure 2.6 illustrates a typical ECG beat signal and its various waves.

2.1.3.1 ECG Waves and Interval

The normal ECG beat, known as Normal Sinus Rhythm (NSR), starts with a P-wave followed by a QRS complex and a T wave. The P wave is produced by the spread of

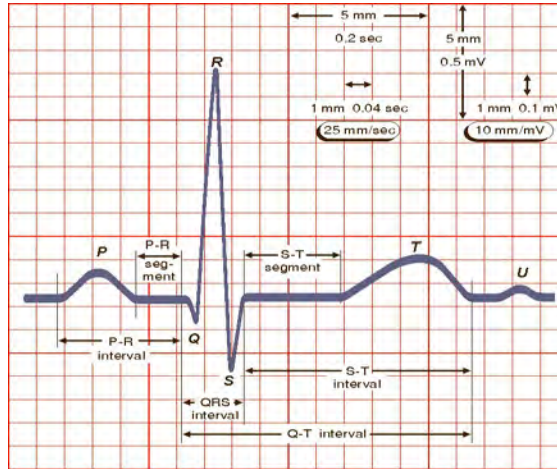


Figure 2.6: A typical ECG waveform (Adapted from [35]).

electric currents during the atrial depolarisation [40]. The QRS consists of the Q, R, and S waves and represents the electrical wave spreads over the ventricles during the ventricular depolarisation. The T wave represents the repolarisation of ventricles. The repolarisation of the atria occurs during ventricular depolarisation and is buried within the QRS complex, so it is not visible on ECG signal [41]. A U wave may be present after the T wave which represents the late repolarisation of Purkinje fibres in ventricles. The PR interval is measured from the beginning of the P wave to the beginning of the QRS complex. The PR interval represents the time required for the electrical impulse to travel from SA node through the AV node and to enter the ventricles. The QT interval is measured from the beginning of the QRS complex to the end of the T wave. The QT interval reflects the time for both ventricular depolarisation and repolarisation.

Analysing the shape and lengths of the ECG waves (morphologic features of ECG signal) provides crucial information about the cardiac electrophysiology. Any malfunction in heart changes the morphological pattern of ECG signal and results in a change in the normal size and shape of the ECG waves. The medical experts can diagnose heart diseases and malfunctions of heart by checking the change in clinical signatures

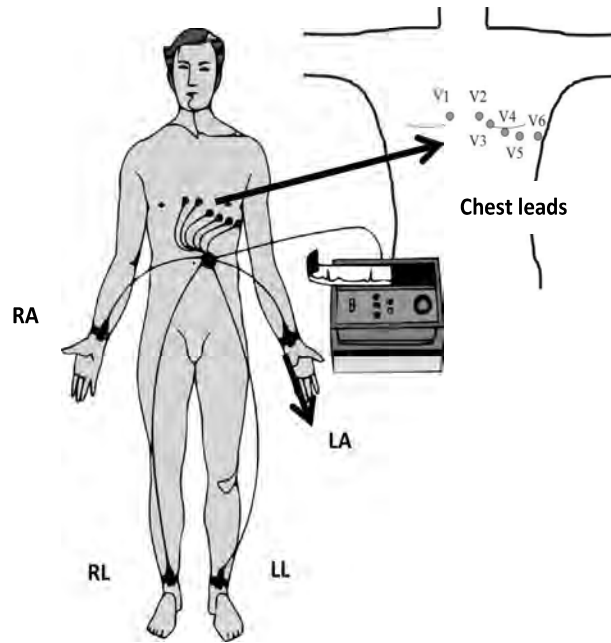


Figure 2.7: Standard 12-lead ECG placement (Adapted from [43]).

of heart and comparing them to their nominal range. It worths to mention that other parameters such as age, sex and gene are also taken into account for the actual clinical diagnosis [42].

2.1.3.2 ECG Signal Acquisition

The electrical activity of the heart can be measured from the body surface using surface electrodes. The standard 12-lead ECG is the most common and accepted method for representation of the heart's electrical activity [35]. The 12-lead ECG system uses ten electrodes to provides 12 different views of the heart. Four electrodes are placed on the limbs (right arm (RA), left arm (LA), right leg (RL), and left leg (LL)), and six electrodes are placed on the patient's chest. Figure 2.7 depicts the location of these electrodes.

The 12-lead ECG consist of twelve leads: three standard limb leads (Leads I, II, and III), three augmented limb leads (Leads aVR, aVL, and aVF) and six precordial leads (leads V1, V2, V3, V4, V5, and V6). These leads are categorised in bipolar and unipolar. The bipolar leads are measured between two electrodes (positive and negative electrodes) and unipolar leads are measured between one electrode (positive electrode) and a reference point. The standard limb leads are bipolar while the augmented limb leads and the precordial leads are unipolar.

Einthoven [44] introduced the lead positions and connections for the standard (bipolar) limb leads. The electrodes are placed on the right arm, left arm and left leg of a person as shown in Figure 2.8. The standard limb leads are defined as:

$$\begin{aligned}
 I &= \Phi_{LA} - \Phi_{RA} \\
 II &= \Phi_{LL} - \Phi_{RA} \\
 III &= \Phi_{LL} - \Phi_{LA}
 \end{aligned} \tag{2.1}$$

where the VLA is the potential of the left arm, the VRA is the potential of the right arm, and the VLL is the potential of the left leg.

Goldberger developed the augmented (unipolar) limb lead electrode positions [45]. The position of electrodes in the augmented limb leads is similar to the electrode positions in the standard limb leads. The electrode positions and connections for the three augmented limb leads are given in Figure 2.9. The augmented limb leads (aVL, aVR and aVF) are calculated as the potential difference between one of the electrodes and the average of remaining two electrodes:

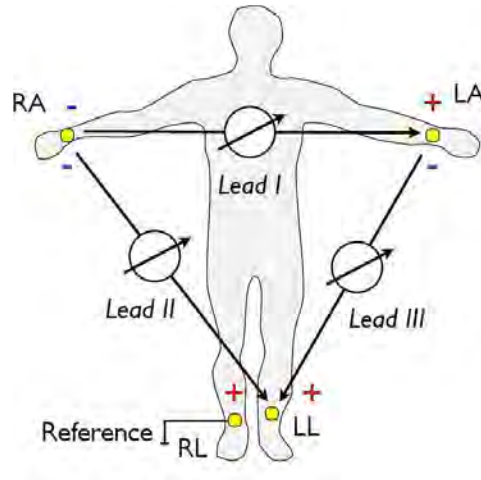


Figure 2.8: Standard limbs lead positions of Einthoven (Adapted from [43]).

$$\begin{aligned}
 aVR &= \Phi_{RA} - \frac{\Phi_{LA} + \Phi_{LL}}{2} \\
 aVL &= \Phi_{LA} - \frac{\Phi_{RA} + \Phi_{LL}}{2} \\
 aVF &= \Phi_{LL} - \frac{\Phi_{RA} + \Phi_{LA}}{2}
 \end{aligned} \tag{2.2}$$

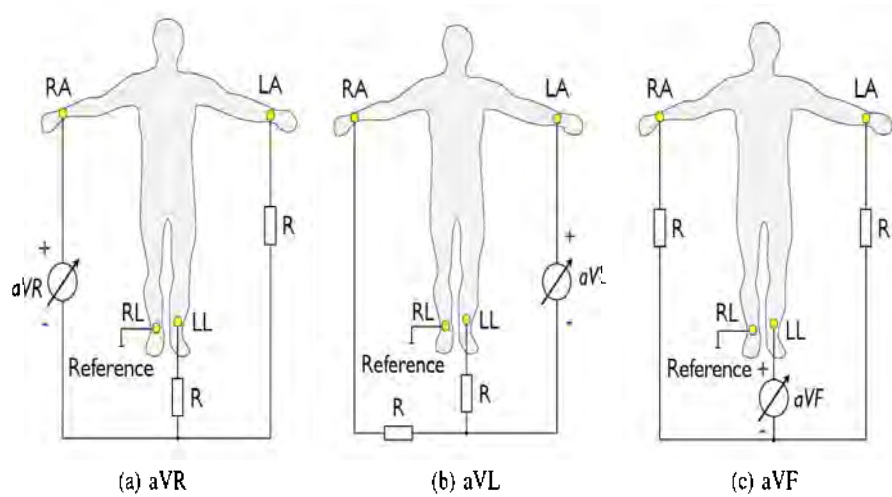


Figure 2.9: Augmented lead positions and connections (Adapted from [43]). R stands for resistors which are usually set to 5 K.

The six precordial leads (leads V1, V2, V3, V4, V5, and V6) are placed in sequence

across the chest as shown in Figure 2.10. These leads are defined as the potential difference between one of the electrodes and a common reference point which is commonly called Wilson Central Terminal (WCT). The WCT is obtained by a combination of the three standard limb leads and can be calculated as follows:

$$\Phi_{WCT} = \frac{\Phi_{RA} + \Phi_{LA} + \Phi_{LL}}{3} \quad (2.3)$$

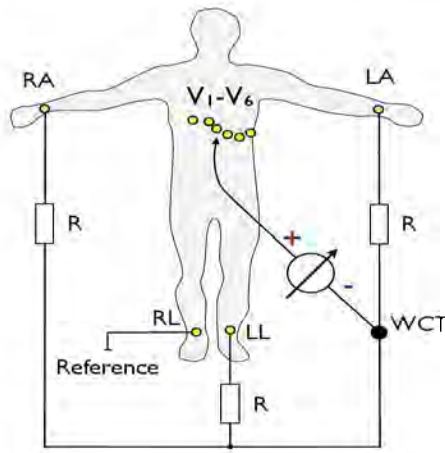


Figure 2.10: Precordial lead positions (Adapted from [43]). R stands for resistors which are normally set to 5 K.

The combination of the six limb leads (the three standard limb leads and the three augmented leads) and the six precordial leads provides a three-dimensional view of the heart as shown in Figure 2.11. The six limb leads provide information about the electrical activity of the heart in the frontal plane, while the six precordial leads provide information about the electrical activity of the heart in the horizontal plane. Additional information on ECG lead systems can be found at [46].

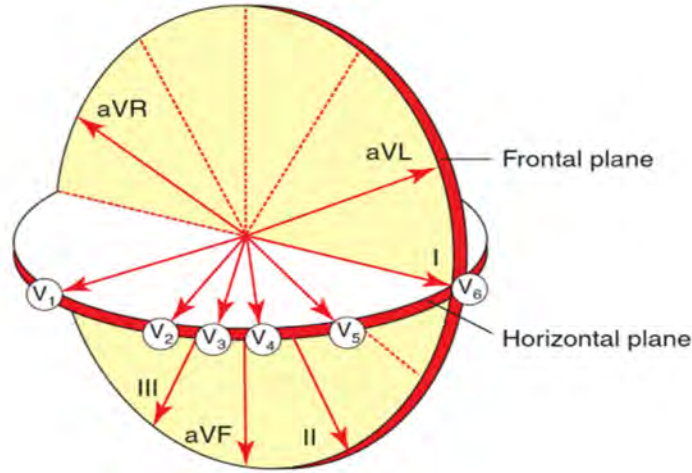
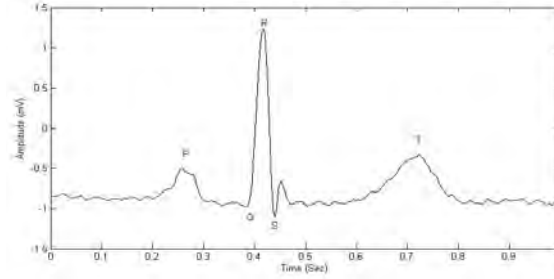


Figure 2.11: Frontal and horizontal planes (Adapted from [46]).

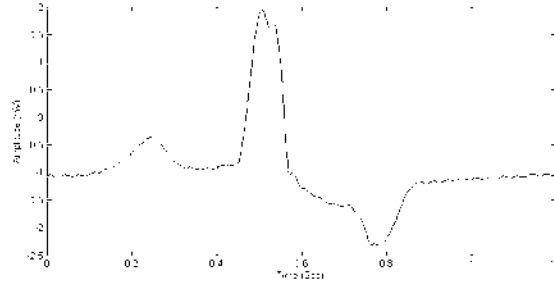
2.1.4 Premature Ventricular Contraction

Cardiovascular diseases (CVD) can be considered as one of the leading causes of death in many developed countries. According to recent reports, CVD accounted for 54.5% and 30% of all deaths in the United States [47] and in Australia [48], respectively. Cardiac arrhythmias are one of the most common CVD. They are abnormal cardiac rhythms caused by a disturbance in the regularity, in the frequency, or in the cardiac electrical impulse. Among the various heart abnormalities, PVC is the most common arrhythmias in humans with an estimated occurrence of 1-4% in the general population on standard 12-lead ECG and between 40-75% on routine 24- to 48-hour Holter monitoring [49].

The PVC is an irregular heartbeat which initiated by the secondary pacemakers (the ectopic foci) in the ventricular area of the heart. These foci cause the ventricles to contract prematurely. The PVCs may be present in a healthy person as well as a patient with structural heart diseases such as cardiomyopathy and myocarditis disease. Although PVCs was considered as benign in the absence of structural heart disease for many years [49, 50], recent studies have suggested that PVCs could be considered as



(a) Normal beat.



(b) PVC beat

Figure 2.12: Typical ECG beat:

a precursor of more severe cardiac arrhythmias such as ventricular fibrillation [8], and atrial fibrillation [9, 51]. Several studies have shown that PVCs occurred during exercise and recovery phases were associated with heart failure and sudden death in even adults without structural heart disease [52, 53]. Multiple frequent PVCs usually cause poor blood circulation which may lead to a temporary loss of consciousness or a dizziness [54, 55]. Although PVCs are usually detected incidentally, they may manifest some symptoms such as palpitations, fatigue, shortness of breath and chest pain.

As PVCs can be linked to various serious cardiac arrhythmias [7], their immediate detection and classification are essential to prevent the possible beginning of life-threatening cardiac conditions. PVC detection and classification can be performed by monitoring and analysing the ECG signals as they generate a different ECG morphology than a normal heart beat. Figure 2.12 shows a typical normal and PVC beat. There are morphological differences between PVC beats and normal beats in both the ventricular

depolarisation phase (QRS complex) and repolarisation phase (mainly the T-wave). The amplitude of the QRS complex for the PVC beats is either much larger or lower than that of a normal beat. Also, there is distortion in the T-wave for the PVC beats. For assisting clinicians to diagnose PVC arrhythmias, this dissertation aims to develop a method that can identify PVCs by processing ECG signal in real-time.

2.2 Background Knowledge Related to EEG Signals

This section provides the background knowledge related to brain, neural system and EEG signals. Section 2.2.1 presents a brief overview on the anatomy of the brain and its function. The sources behind brain's electrical activities are discussed in Section 2.2.2. Section 2.2.3 presents a description of EEG and its nature. Epilepsy and its effect on EEG are described in Section 2.2.4. Section 2.2.5 provides the information about BCI structure and its applications.

2.2.1 Human Brain: Structure and Function

The human brain is the central part of the nervous system that governs the functions of various organs in the body. The brain is separated into four major parts: *cerebellum*, *brain stem*, *diencephalon*, and *cerebrum* [56] as shown in Figure 2.13. A brief explanation of these parts is given below:

The *cerebellum* is the part of the brain which is located at the posterior bottom part of the brain. It has two hemispheres and contains hundreds of millions of neurons. The *cerebellum* is the centre of motor control within the brain that receives sensory information from different organs through sensory systems of the spinal cord and integrates this

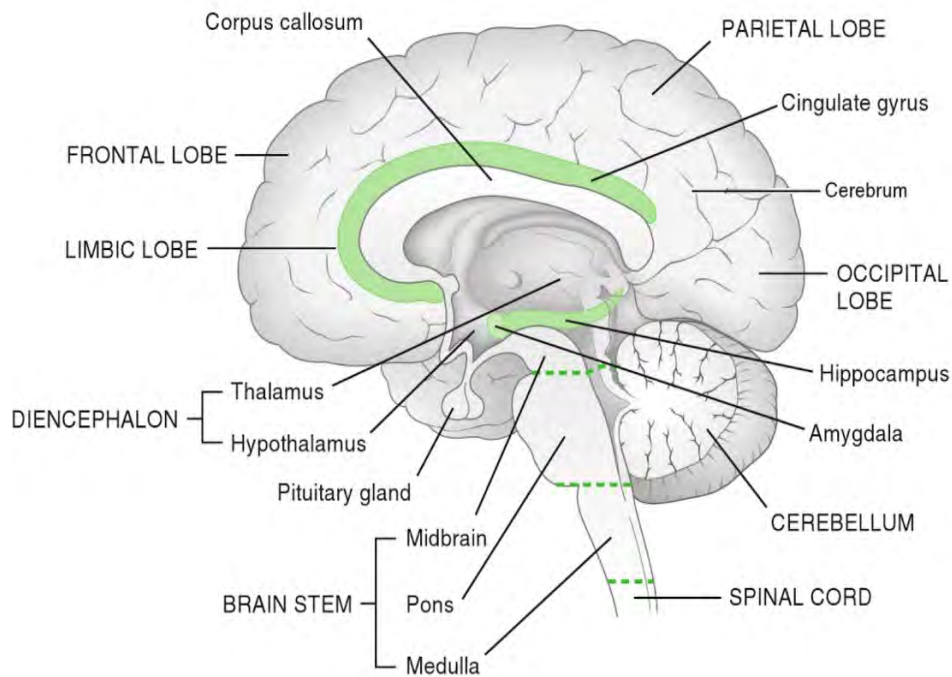


Figure 2.13: Schematic midsagittal view of the human brain (Adapted from [57]).

information to coordinates the body's movements. It is also responsible for posture, voluntary muscle movements, and balance. Damaging this brain region produces disorders in the coordination of body movement.

The *brain stem* is the lowermost part of the brain that connects the cerebrum to the spinal cord. It is responsible for regulating and controlling vital functions of the body such as breathing, heartbeat, awareness, consciousness and blood pressure. The *brain stem* is composed of three parts: medulla oblongata, pons, and midbrain. The medulla oblongata contains various vital centres for regulating heartbeat and respiration and also has reflex centres for coughing, sneezing, swallowing, and vomiting. The Pons works with the medulla oblongata to control breathing rate. The midbrain has reflex centres for visual and auditory responses.

The *diencephalon* is located deep in the brain between the cerebrum and midbrain. It consists of two main parts: hypothalamus and thalamus. The hypothalamus has several functions. It maintains homeostasis by regulating body temperature, hunger, blood pressure, and water balance. It also regulates the pituitary gland and controls the appetite. On the other hand, the thalamus acts as a relay station for sensory impulses travelling to the cerebral cortex. It is involved in many functions such as the regulation of alertness, sleep, wakefulness, and consciousness.

The *cerebrum* is the largest and foremost portion of the human brain which is associated with various brain functions such as emotions, thoughts, motor functions and movements. The outer layer of the *cerebrum* is made up of grey matter and known as the cerebral cortex. The *cerebrum* is divided into two hemispheres: right and left hemispheres. Each hemisphere controls the activities of the side of the body opposite that hemisphere. Hemispheres are connected to each other by a bridge of white matter called the corpus callosum. Each hemisphere can be further subdivided into four lobes (as shown in Figure 2.14): frontal, parietal, occipital and temporal [56]. Each lobe is associated with a variety of bodily functions. The frontal lobe is responsible for personality, problem-solving, planning, emotions, parts of speech and movement, and reasoning. The parietal lobe is associated with sensation (e.g. pain, touch), orientation, perception of stimuli, and recognition. The occipital lobe is involved in visual processing. The temporal lobe is responsible for recognition and perception of auditory stimuli, speech and memory.

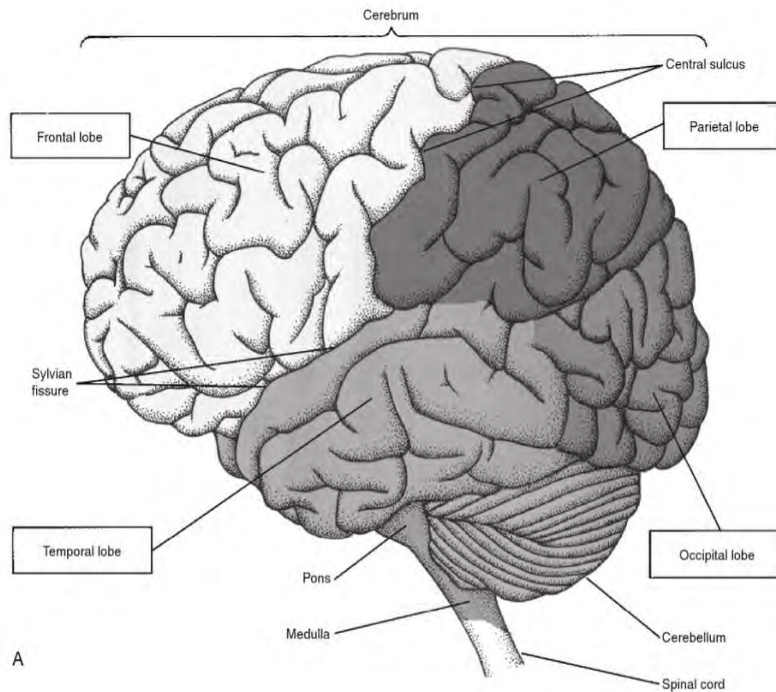
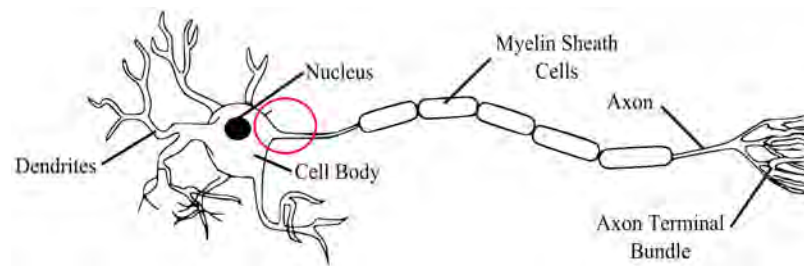


Figure 2.14: The four major lobes of the cerebrum (Adapted from [35]).

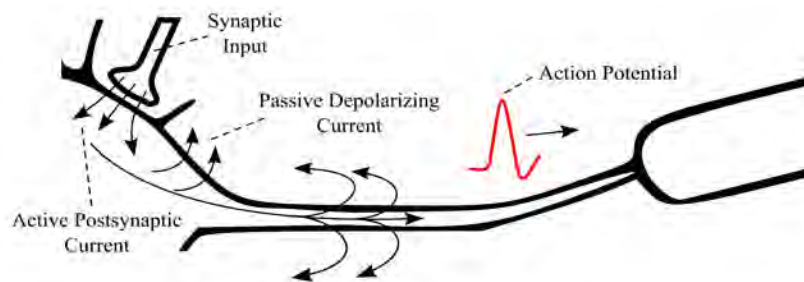
2.2.2 Neurophysiology of Human Brain

The human body utilises its nervous system to interact with the physical world. The human nervous system is divided into two major parts: the central nervous system (CNS) and the peripheral nervous system (PNS). The CNS consists of all the nerves inside the brain and spinal cord. It is accountable for processing the information received from all parts of the body. The PNS contains all the nerves outside of CNS. It acts a communication relay connecting the CNS to the limbs and organs.

The cells of the CNS can be divided into nerve cells (neurons) and neuroglia cells [58]. Each nerve cell has three basic parts: axon, dendrites and cell body (soma) as shown in Figure 2.15(a). The nerve cells carry information in the form of electrical impulses (action potential (AP)) throughout the body. The dendrite is a short section of



(a) anatomy

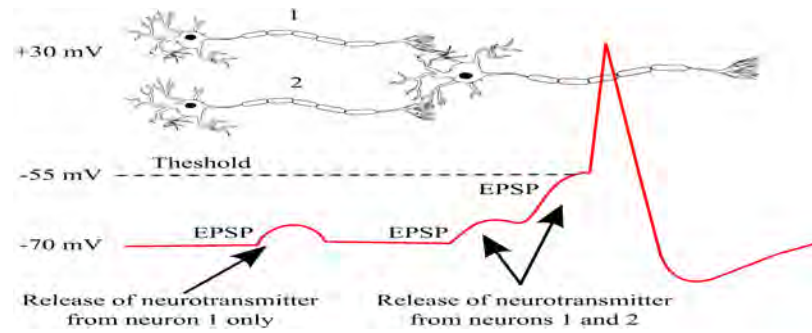


(b) electrical impulse generation

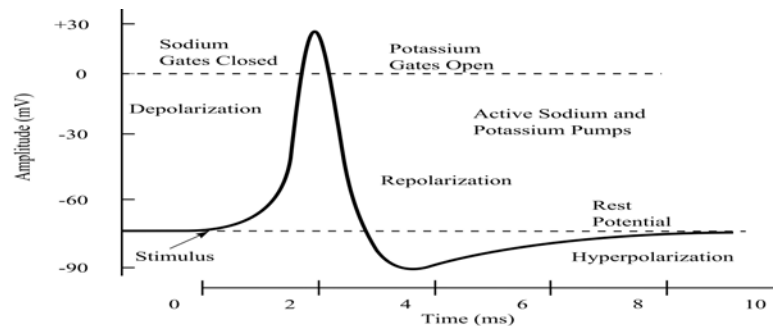
Figure 2.15: CNS nerve cells: (Adapted from [59])

the neuron which receives the electrical impulse from other nerve cells. The cell body is the heart of the cell which processes and integrates the impulse. The axon is a slender portion of the neuron that conducts the impulse to other cells.

Various electrical impulses may occur within a nerve cell. When an impulse arrives at a synapse from a presynaptic neuron (an activated neuron), a chemical substance called a neurotransmitter is released into a synaptic space. The neurotransmitter triggers ionic shifts that either hyperpolarise or depolarise the membrane of postsynaptic neuron (the resting neuron). The first type is associated with depression of the postsynaptic neuron and its potential change is known as an inhibitory postsynaptic potential (IPSP). The depolarising type produces a potential change, known as an excitatory postsynaptic potential (EPSP). The EPSP increases the neurons membrane potential while the IPSP decreases it. The AP is a temporary change in the membrane potential which is caused by an exchange of ions across the neuron membrane. The AP is initiated when the EPSP



(a) threshold limit



(b) ion exchange during generation

Figure 2.16: Action potentials: (Adapted from [59])

surpasses the threshold conduction level for the postsynaptic neuron. The propagation of the AP along the axon involves two current flows: the active and passive currents as shown in Figure 2.15(b). The active current occurs through ion channels across the membrane while the passive current occurs inside the axon. Multiple presynaptic action potentials (from one presynaptic neuron or multiple presynaptic neurons) are almost always required for the EPSP to reach the threshold and to initiate an AP as seen in Figure 2.16(a). The membrane potential depolarises (becomes more positive) during the exchange of ions and generates a spike. The membrane repolarises after the spike and becomes more negative than the original membrane potential, known as the resting membrane potential, and then it returns to the normal.

As shown in Figure 2.16(b), the AP spike is generated when a neuron is stimulated. The main ions involved in generating the AP spike are sodium (Na^+) and potassium

(K⁺) ions. When the neuron receives the stimulus, the Na⁺ channels open and allow Na⁺ ions to move to the inside of the neuron cell membrane. Therefore, the membrane potential increases from the membrane resting potential (70 mV) up to the threshold potential (55 mV). When the threshold potential is reached, more Na⁺ channels (known as voltage-gated channels) open and allow more Na⁺ to move to the inside of the neuron membrane. This causes the membrane potential to increase to +30 mV. This process is known as depolarisation. Before reaching the peak, the Na⁺ channels become inactive and the K⁺ channels open to allow K⁺ ions to move to the outside of the neuron membrane. The membrane starts to repolarise and its potential decreases back to the rest potential. The repolarisation passes the resting membrane potential which is known as hyperpolarisation. The hyperpolarisation prevents the neuron to receive another stimulus which may cause another AP in the opposite direction. Therefore, the signal travels in one direction. The membrane potential then returns to its resting membrane potential. The entire process takes around 8ms.

In the brain, neurons produce APs which contribute to the generation of neural activity. The human brain consists of about 100 billion neurons. The neural activity is produced by the summation of tens of thousands of neurons activities. A current flow known as extracellular current is generated during the excitation of pyramidal neurons. This current produces electrical dipoles between the soma and apical dendrites. EEG is a method to measure and record the electrical dipole generated in the cerebral cortex and extended to the scalp surface [60] (see Figure 2.17). The EEG electrodes record the signals that contain both the neural activity of the source located below the electrode and a summation of neural activity conducted in different parts of the brain. The background information of EEG is provided in detail in the following section.

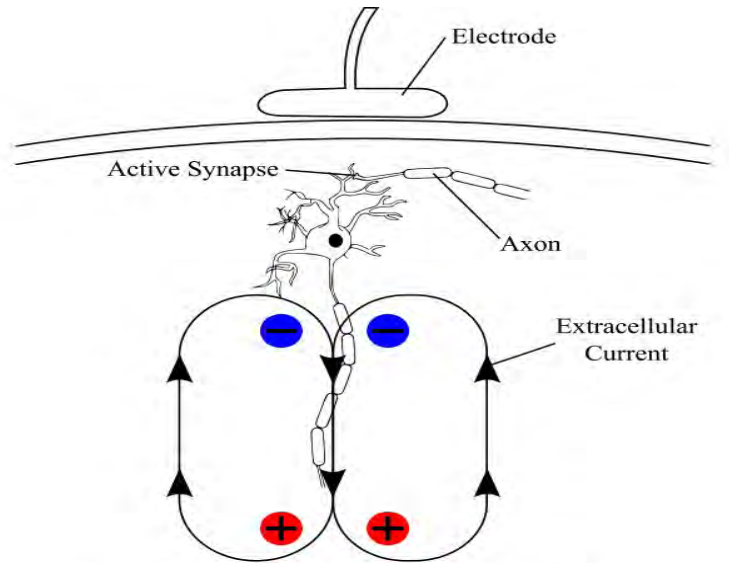


Figure 2.17: Electrical dipole measured by EEG electrode (Adapted from [60]).

2.2.3 Electroencephalography

EEG is a non-invasive method to measure the electrical activity of the human brain. It is a safe procedure that provides information about how the brain functions over time. Physicians and scientists utilise the EEG to study brain functions and also to diagnose neurological disorder such as epilepsy, encephalitis, head injuries, brain tumours, sleep disorders, memory problems, dementia, and stroke. EEG is also used for studying the behavioural disturbances, learning problems, language delay and developmental delay. During an EEG test, a number of small flat metal discs called electrodes are used to capture the signals by placing them on different parts of the scalp. Then the signals captured by electrodes are transmitted to an EEG recording machine via amplifiers. Finally, the signals are sent to a computer to record the results. One pair of electrodes is referred to as a channel. Each channel produces a single EEG signal. Depending on the different usage of EEG, a different number of electrodes from 1 to 512 electrodes can be used to record the EEG signals in parallel (multichannel EEG recordings). The EEG

signal has a low amplitude that typically ranges from about 1 to 100 μ V in a normal adult. The EEG can vary depending on the location of the recording due to non-uniformity of brain architecture.

The international 10-20 electrode system is the conventional method for the placement of electrodes on the scalp. Figure 2.18 illustrates the electrodeposition of 21 electrodes on the brain according to the International 10-20 system. In this system, the positions of electrodes are determined based on two lines: the line from left side to right side of the skull and the line from nasion to inion points where the nasion is the point between the forehead and the nose and the inion is the lowest point of the skull at the back of the head. The 10-20 represents that the actual distances between electrodes are either 10% or 20% of the total distance of the two predefined lines. The letters F, T, C, P and O are used to refer to electrodes position on the Frontal, Temporal, Central, Parietal and Occipital, respectively. An electrode placed on the midline is labelled with z letter. Even and odd numbers refer to electrode positions on the right hemisphere and the left hemisphere, respectively. Although the international 10-20 electrode system is used to define the location of only 21 electrodes, a system with higher resolution setting can be obtained by adding more intermediate electrodes to the 10-20 electrode system. The new system is called Modified Combinatorial Nomenclature (MCN).

Since EEG measures the voltage difference between two electrodes, there are several ways for measuring the voltage difference and viewing EEG electrode which are called montages. A brief explanation of common montages is given below [62]:

Bipolar montage: the data in each channel represents the voltage difference between two adjacent electrodes. The entire montage contains a series of bipolar electrode pairs. For instance, the voltage difference between the Fp1 electrode and the F3 electrode is measured and recorded as the channel “Fp1-F3”.

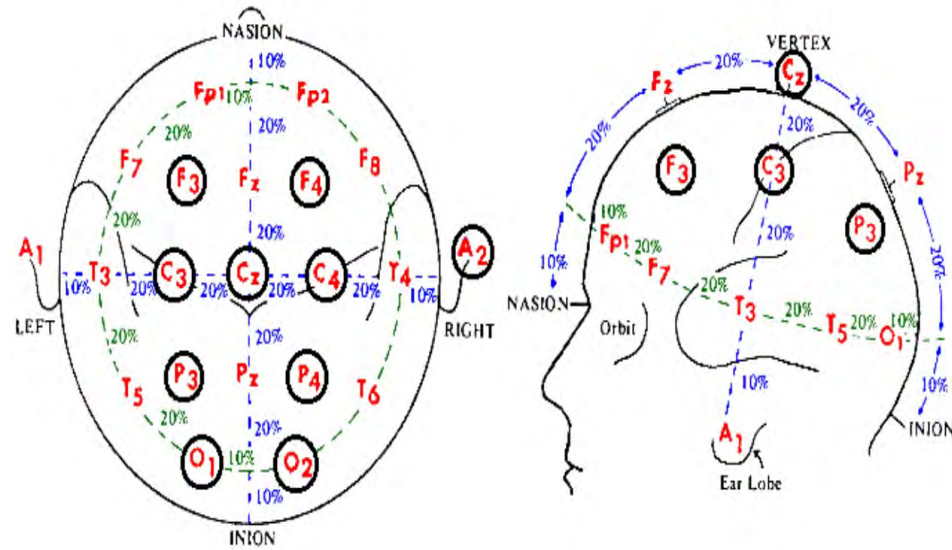


Figure 2.18: The International 10-20 electrode placement system (Adapted from [61]).

Referential montage: the data in each channel represents the voltage difference between a certain electrode and a designated reference electrode. As there is no standard position for this reference electrode, it can be an electrode placed on the chin, earlobes, or nose. The electrodes placed on the midline are often considered as a reference electrode because they do not amplify the signal in one hemisphere versus the other.

Average reference montage: the data in each channel represents the voltage difference between an active electrode and the average reference which is calculated using the average of the outputs of all the electrodes.

The rhythmic activity of EEG signals is divided into six bands in terms of frequency. These frequency bands are important for understanding brain activities and for assessing abnormalities in EEGs. The rhythmic activity in different frequency range can be associated with various mental activities or conscious states. They also have a different distribution over the scalp. The six bands are called alpha (α), theta (θ), beta (β), delta (δ), gamma (γ), and mu (μ) [63, 59]. Figure 2.19 illustrates examples of these

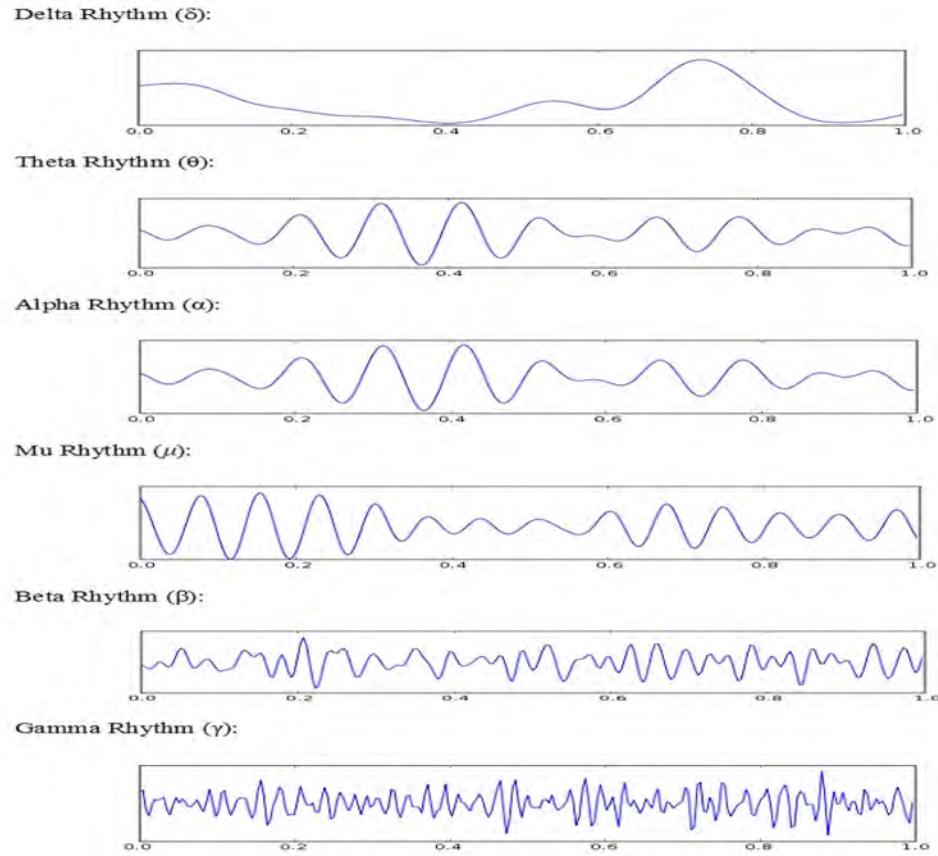


Figure 2.19: Example of EEG rhythms in different frequency bands (Adapted from [64]).

EEG rhythms, and Table 2.1 shows the frequency range of each band and some related mental functions these bands correspond to.

2.2.4 Epilepsy and Epileptic Seizures

Epilepsy is one of the most common serious neurological disorders in humans that affects people of all ages. About 1% of people worldwide (70 million) have epilepsy [65], and nearly 80% of cases occur in developing countries. More than 250,000 people in Australia are diagnosed with epilepsy [66], and approximately 3% of Australians will

Table 2.1: Significance of EEG in different frequency bands.

Band	Frequency (Hz)	Location	Mental State
Delta (δ)	0.5-4	Everywhere	Deep sleep, severe brain disorder and lack of attention.
Theta (θ)	4-8	Temporal and parietal	Emotional stress, creative inspiration and deep meditation.
Alpha (α)	8-13	Occipital lobe	Eyes closed and relaxation;
Beta (β)	13-30	Frontal regions	active thinking, active attention, focus, solving problems
Gamma (γ)	30-100	Somatosensory cortex	various cognitive and motor functions
Mu (μ)	8-13	Frontal (motor cortex)	Suppression during motor activity

experience epilepsy at some point in their lives. Epilepsy is characterised by spontaneously recurrent seizures [11]. Seizures are caused by rapid changes in the electrical functioning of the brain, which result in sudden behaviour changes such as jerky movements, temporarily loss of memory, losing consciousness and muscle spasm. These occur in the outside rim of the brain (Called the cortex).

Epilepsy may develop due to many different factors such as an imbalance of nerve signalling chemicals called neurotransmitters, an abnormality in brain wiring, or some combination of these factors. Neurons normally release electrochemical impulses that

influence on other neurons, glands, and muscles to generate human thoughts, feelings and actions. The normal neuronal activity pattern becomes disturbed in epilepsy. This causes strange emotions sensations and behaviours that sometimes lead to muscle spasms, convulsions, and loss of consciousness [67]. Epilepsy may be caused by a sudden surge of electrical activity from neurones that are inherently unstable due to a genetic defect (as in the different types of inherited epilepsy), or from neurones that are unstable by metabolic abnormalities such as very high or low blood glucose and alcohol. Moreover, the abnormal discharge can come from a localised area of the brain, such as in patients with epilepsy caused by head injury or brain tumour. During a seizure, neurons may fire as many as 500 times a second, much faster than the normal rate of about 80 times a second. This occasionally occurs in some people while may occur up to hundreds of times a day for others.

EEG recordings provide valuable information for understanding epilepsy. Epileptic activities cause abnormalities in EEG signals. The detection of these abnormalities in EEG (during seizures) is a vital component in the diagnosis and treatment of epilepsy. Two types of abnormal activity can be detected in an EEG signal: ictal and interictal. The ictal represents the period of seizure while the interictal represents the intermediate period between two seizures. A sudden change of frequency in the EEG signal often considered as the beginning of a seizure. It lies in the alpha wave frequency band with an increase in amplitude (but a decrease in frequency) during the seizure period. The epileptiform transient waveforms (Spikes and sharp waves) are frequently observed in EEG in between seizures. For assisting the diagnosis and treatment of epilepsy, this dissertation aims to develop a method for classification of both epileptic EEG signals during seizure activity and during the seizure-free time.

2.2.5 Brain-Computer Interfaces

A Brain-Computer Interface (BCI) is a system that allows to control an external device by brain activity only, without the need for any physical effort or muscle control [12]. It provides an alternative communication channel between the human brain and a computer [68]. Therefore, it allows a user to send messages and commands (encoded into the user's brain activity) through a direct channel to a computer.

The BCI uses only brain activities to command a machine. An electrode cap is placed on a user's head to measure EEG signals (users' brain activities). Then, the user imagines different specific tasks such as movements of various body parts (hand, legs, etc.) and composing of words to command the machine. These tasks affect the EEG signals and produce different EEG patterns. Then computer classifies these patterns into different classes to control a machine (such as a wheelchair) or a computer application (e.g. cursor movement). There are two types of BCIs: invasive and non-invasive. The invasive BCIs record signals from electrodes implanted over the brain cortex (requiring surgery) while non-invasive ones record signals from electrodes placed on the scalp (outside the head) [12]. In most cases, the non-invasive technique is preferable than the invasive technique since it is harmless and easy to use.

2.2.5.1 BCI System Architecture

BCI system consists of six basic components shown in Figure 2.20. They include signal acquisition, signal pre-processing, feature extraction, classification, translation into a command and feedback [69, 70].

Measuring brain activity is one of the main steps in any BCI system. The brain activities reflect user's intentions which can be measured as electrical signals. Based

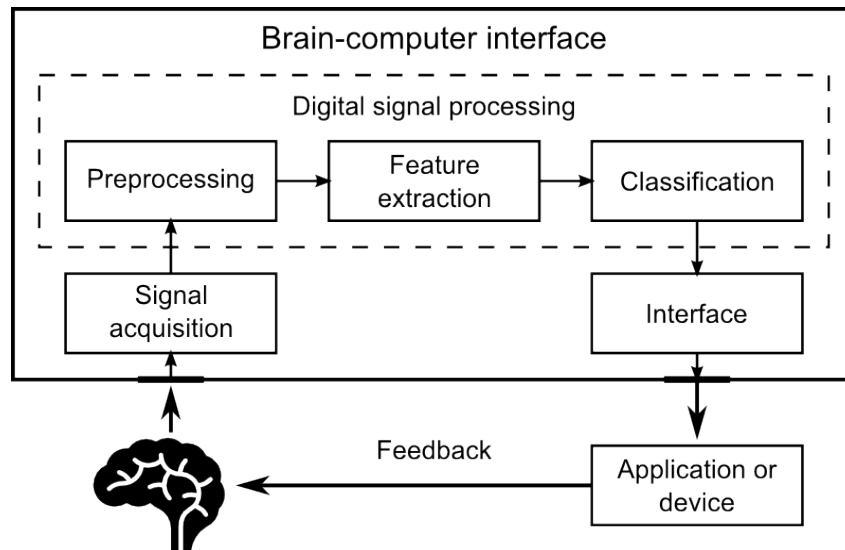


Figure 2.20: General BCI architecture (Adapted from [71]).

on different types of electrodes, various methods have been used for signal acquisition. Then these signals are digitised and prepared for next step. In this thesis, EEG are utilised for measuring brain electrical activities. The recorded signals are sent to the pre-processing step for signal enhancement and noise reduction. In this step, the recorded signals are often filtered and cleaned in order to improve the relevant information embedded in the signals [72]. Feature extraction step generates the discriminative characteristics (called features) from the brain patterns. These features represent the properties of the respective brain patterns. The classification step, also known as feature translation, tries to detect or classify certain brain states by assigning a class (associated with the type of mental states) to a set of features extracted from the signals. Once the mental state (class) is determined, BCI translates this mental state into a command (associated with this mental state) to control a given application such as a computer. Finally, some feedback is provided to the users regarding the identified mental state in order to help them to control their brain activities and increase their performances.

BCI systems can only detect and classify specific activity patterns in continuous

brain signals associated with certain events or tasks. The way a BCI user has to produce these patterns is determined by the mental strategy. The mental strategy is the foundation of any brain-computer communication. The mental strategy determines the necessary actions that have to be performed by the user to generate specific activity patterns that the BCI can interpret. The most common mental strategies are selective (focused) attention and MI [70]. BCIs based on selective attention require external stimuli. These stimuli are provided by the BCI system. On the other hand, BCIs based on motor imagery do not need any external stimuli. MI is a mental process by which a user simulates or imagine a given action without actually performing it. The BCI based on motor imagery is considered in this research.

2.2.5.2 BCI Applications

BCIs have contributed and involved in various fields of research such as medical, games and entertainment, educational and self-regulation and smart environment. A variety of BCI applications have been developed ranging from assistive technologies for patients with severe motor disabilities to entertainment devices. The main application for BCI is in assistive technology which allows disabled people to communicate without moving a muscle. For example, patients suffer from Amyotrophic Lateral Sclerosis (ALS), who are paralysed or locked-in their paralysed body, can use spellers (a BCI application) to communicate with other people [73]. BCIs can also provide the communication and control for patients with less severe motor disabilities. BCI can help stroke patients to learn how to prevent unwanted patterns by visualising their current brain state. BCIs could also help patients with spinal cord injuries to control an orthosis or wheelchair [74], or even a telepresence device [75]. Apart from assistive technology, BCI has been developed in many other fields [76]. BCI can be utilised as a measure-

ment device to evaluate users' mental states such as attention [77] or workload mental fatigue [78], in order to prevent human failure in critical environments and to improve workplace conditions. BCIs have been studied in self-regulation such as emotional regulation [79] and controlling depression feeling [80]. BCI can also be used for gaming and entertainment [81, 82]. For instance, in the video game named BrainArena, players can play a football match by means of two BCIs (imagining left or right-hand movements).

Most of the BCI applications run on conventional computers that host a BCI system and its application. Communication or control based BCI technologies require specific brain activity patterns which can be generated by a subject and distinguished by a computer system. The subject needs to concentrate on a mental task (e.g. the imagination of a hand movement) to produce these specific patterns that are classified with the desired control. The user performs different mental tasks which generate different wave pattern in EEG signals. These wave patterns can be identified and translated into a control code-book for the user. In this thesis, a method is proposed to classify different mental tasks (MI tasks) for developing a communication-based BCI systems.

2.3 Review of the ECG and EEG Classification

The classification of ECG and EEG signals plays a significant role in biomedical research and medical diagnosis. The ECG signal classification is essential for the diagnosis and treatment of heart diseases while the classification of EEG signals is vital in diagnosis of brain diseases and in developing BCIs. Accurate ECG and EEG classifications can provide clinicians substantial information to help them in the diagnosis and treatment of patients. The classification of these signals can reveal information about the underlying causes of heart and brain diseases. As both ECG and EEG recordings con-

tain a large amount of data, one of the main challenges is how to represent these signals for classification. It is important first to extract discriminatory information from raw ECG and EEG signals and then use this information to classify ECG and EEG signals.

The classification of ECG and EEG signals is divided into mainly two parts: feature extraction and classification. The feature extraction part aims to extract the most important features from the signals. The classification part seeks to use the extracted features to determine the class of signals. In the feature extraction, the features are extracted to characterise the signals. These features are a smaller number of values that describe the properties of the signals. Different features could be extracted from the signals based on various classification purposes. In classification part, the extracted signal features are fed into a classifier for decision making i.e. identifying which class each signal belongs to.

In past two decades, several techniques have been developed for ECG and EEG classifications in different fields. Since the focus of this thesis is on developing methods for the classification of PVC beats, the epileptic EEG signals, and the MI-based EEG signals in BCI systems, an overview of classification methods in these three areas is provided. Section 2.3.1 reports the method used in the PVC beat classification. Section 2.3.2 reviews techniques for feature extraction and the classification in the epileptic EEG signal classification. The review of methods for the MI tasks classification is presented in Section 2.3.3.

2.3.1 PVC Beat Classification Methods

In the past decades, a significant amount of research work has been conducted for automatic classification of PVC beats based on different feature extraction and classifica-

tion methods. Feature extraction methods include QRS complex morphology [83, 84], interbeat R-R intervals [83], filter bank [85], Hermite functions [86], wavelet transform [13, 14, 87], higher order cumulant features [88], Karhunen-Loeve expansion of ECG morphology [89], and correlation dimension and largest Lyapunov exponent [90, 91]. Several classifiers have been used for heartbeats classification including decision trees [85], nearest neighbors [92], Linear Discriminant [83], self-organizing maps [86, 93], neural networks [94, 95, 14], fuzzy neural network [13], support vector machines [96], and Optimum-path forest [97]. Some of these methods have been developed for the classification between PVC and normal beats while others are developed to identify several arrhythmias into different classes. As our proposed method for PVC beat classification is evaluated in the MIT-BIH Arrhythmia database [98] (The description of the database is provided in Section 3.2), only some of the recent methods that evaluated on this database are reviewed.

A PVC classification method based on filter bank features and decision trees presented by Wieben et al. [85] achieved a sensitivity of 85.3% and a positive predictivity of 85.2%. In [94] author developed a classifier based on principal component analysis and neural networks. The method achieved a sensitivity of 98.1% and a positive predictivity of 94.7% for 14 records of the MIT-BIH database. Lagerholm et al. [86] built a classifier by combining Hermite functions and self-organizing maps neural networks for clustering of QRS complexes into 25 groups. They achieved a very low degree of misclassification (about 1.5% beats misclassified). A real time QRS complex classification method based on QRS complex morphology and Mahalanobis distance was developed in [84]. Using only 44 records of the MIT-BIH database, they achieved a sensitivity of 90.74% and a positive predictivity of 96.55%. In [95] authors introduced a classifier based on neural network classifier for classification of normal QRS complexes and PVC. Feeding the classifier by 26 parameters extracted from two ECG leads, they achieved

98.5% of correct detection of premature ventricular beats and 99.7% of normal beats for entire MIT-BIH database. Shyu et al. [13] presented a PVC detection method based on wavelet transform and fuzzy neural network. The algorithm achieved an accuracy of 99.79% for six records of the MIT-BIH database. An arrhythmia beat classification method based on the RR-interval features was proposed by Tsipouras et al. [99]. They achieved 98% of accuracy. In [14] authors presented a neural network-based PVC classifier based on wavelet transform and timing interval features. The algorithm achieved an accuracy of 96.82% and 95.16% for 22 and 40 records of the MIT-BIH database, respectively. A method based on linear discriminants classifier and ECG morphology and heartbeat interval features was developed by De Chazal et al. [83]. They reported a specificity of 98.8% and a sensitivity of 77.7%. A technique based on a nonlinear trimmed moving average filter proposed in [100] achieved 99.7% of specificity and 97.8% of sensitivity for 34 records of the MIT-BIH database. In [101] authors proposed a PVC classification method based on QRS template matching using different complex features such as cross-correlation, frequency, and morphological of the QRS complex. The classifier was evaluated on 48 records of the MIT-BIH database, and the results were 98.4% of sensitivity and 98.86% of specificity. Lim [15] developed a PVC classifier based on the neural network with weighted fuzzy membership functions. Using the six records of the MIT-BIH database, the algorithm achieved a specificity of 99.93% and a sensitivity of 99.21%. Talbi and Charef [93] proposed a PVC detection method based on the QRS power spectrum and self-organizing maps. They achieved a specificity of 95.18% and a sensitivity of 82.20% for 46 records. A Real-time Cardiac Arrhythmia Classification based on layered hidden Markov model proposed in [102] obtained a sensitivity of 97.75% and a positive predictivity of 96.63% for 16 records of the MIT-BIH database. Daamouche et al. [103] developed an arrhythmia beat classification method based on Wavelet transform and particle swarm optimisation technique. Using

only 20 records of the MIT-BIH database, they achieved a sensitivity of 91.75% and a positive predictivity of 96.14%. A low-complexity data-adaptive approach for PVC identification based on template matching was developed by Li et al. [16]. The classifier was tested on 22 records of the MIT-BIH database, and the results were 93.1% of sensitivity and 81.4% of positive predictivity. In [104] authors developed a PVC detection method based on Lyapunov exponents and LVQ neural network. The algorithm achieved an accuracy of 98.90%, sensitivity of 90.26%, positive predictability of 92.31% for 20 records of the MIT-BIH database. Talbi and Ravier [105] developed a PVC classification method based on the fractional linear prediction. Using only 43 records of the MIT-BIH database, they evaluated several classifiers and achieved the maximum accuracy of 96% for Neural network trained by LevenbergMarquardt rule classifier with a Specificity of 98% and a sensitivity of 74%.

From the literature, it is observed that there are still some limitations in the PVC detection and classification methods. Most of the existing techniques are not suitable for online PVC detection as they employ many complicated mathematical tools such as wavelet transform and artificial neural network. The efficiency of these methods is mainly accompanied by high complexity and long computational time. Moreover, some of the existing high detection results are based on relatively small data sets or used overlapped training and testing data sets. The efficiency of these methods over a large number of records (patients) is a difficult problem to deal with. To overcome these problems, this study aims to introduce an online PVC detection method to detect PVC beats accurately in an online manner.

2.3.2 Epileptic Classification Methods

The key challenge of any automated epileptic seizure detection method is the extraction of the distinguishing features from EEG signals as it significantly affects the performance of the classifier. Representative characteristics or features extracted from EEG data can describe the key properties or morphologies of the signals for perfect detection of epileptic seizure [106]. In past two decades, a variety of methods has been developed for feature extraction in epileptic EEG data. These feature extraction techniques can be divided into four groups [107]: *time domain*, *frequency domain*, *timefrequency domain* and *nonlinear methods*.

The simplest features are extracted in the *time domain*. Three popular time domain methods are PCA, independent component analyses (ICA) and linear discriminant analysis (LDA). PCA is an unsupervised method that transforms high-dimensional data into a low-dimensional feature subspace. In epileptic EEG signals classification, PCA can be used for feature enhancement [108] and feature reduction [109, 110]. Xie and Krishnan [111] developed an efficient feature extraction based on dynamic PCA with nonoverlapping moving window to separate EEGs from interictal and ictal for seizure detection. ICA is a feature extraction method that transforms a multivariate random signal into a signal with independent components [109]. It also can be utilised to remove artefacts and to separate brain signals sources from EEG signals [112]. LDA is another method for dimension reduction. It finds a linear combination of features that can serve as a classifier to separate two or more classes [107]. Subasi and Gursoy [109] employed all three methods: PCA, ICA, and LDA and compared the effectiveness of them in epileptic EEG classification.

The *Frequency domain* methods are based on spectral analysis methods. The spectral analysis methods estimate power spectral density (PSD) which is the power dis-

tributions of a signal across the frequency. These methods can be divided into two categories [107]: the non-parametric methods which calculate spectra directly from signals by using Fourier Transform (FT) method, and the parametric methods which derive spectral quantities from a statistical model of the signals. The periodogram and the correlogram are two common non-parametric methods. These methods can provide high resolution for adequately long data lengths. However, they suffer from spectral leakage effects and poor frequency resolution due to the use of windows and the finite length of data [17]. These limitations are overcome by the parametric method. Three parametric models are: the moving average (MA) model which is suitable for modelling spectra with broad peaks and sharp nulls, the autoregressive (AR) model which is suitable for modelling spectra with narrow peaks, and the autoregressive moving average (ARMA) model which is suitable for modelling spectra with both sharp peaks and deep nulls [17]. Polat and Gunes [113] used spectral parameters based on the Fourier transformation of EEG signals to detect epileptic seizures in EEG signals. Übeyli [114] classified normal and ictal EEG time series using the AR method for feature extraction. Srinivasan et al. [115] proposed a two-class epilepsy detection methods based on EEG features extracted in both *time domain* and *frequency domain*.

Since the EEG signals are non-stationary [116], many *timefrequency* based methods have been developed for epileptic seizure classification from EEG signals. Unlike *time-* and *frequency-domain* methods, the *time-frequency* based methods do not impose the quasi-stationarity assumption on the data [117]. These methods include wavelet transform [118, 117, 109, 113, 119], time-frequency distribution [120, 121], multi-wavelet transform [122], and empirical mode decomposition (EMD) [123, 124]. As the wavelet transform provides a representation of the signal in both the time and frequency domains, they can accurately detect and localise transient features in the signals like epileptic spikes [125]. Discrete Wavelet Transform (DWT), Wavelet Packet Decompo-

sition (WPD), and Continuous Wavelet Transform (CWT) are the three types of wavelet transforms [107]. Among these methods, the DWT stands out in terms of efficiency and algorithmic elegance [126]. Lee et al. [127] used features based on the Euclidian distance calculated from phase space representation (PSR) of wavelet coefficient to detect normal and epileptic seizure EEG signals. Pachori and Patidar [128] proposed a method for classification of ictal and seizure-free EEG signals using the 95% confidence ellipse area measured from the second-order difference plot (SODP) of intrinsic mode functions (IMFs) as a feature. Samiee et al. [129] proposed a feature extraction technique based on a rational discrete short-time Fourier transform for classification of normal and epileptic segments.

Although *frequency domain* methods can detect rhythmic oscillations in a signal, they have some limitation in capturing phase locking and nonlinear coupling between harmonics in the spectrum [130]. To better represent and reflect the characteristics of the EEG signals, various nonlinear parameters have been proposed as features using different *nonlinear methods* for classification of epileptic seizure EEG signals. The nonlinear parameters namely, Higher Order Spectra (HOS) [131, 132], Correlation Dimension (CD) [133, 133, 134], Lyapunov Exponent (LE) [19, 135], Hurst Exponent (H) [130], Fractal Dimension (FD) [136], Approximate Entropy (ApEn) [137, 138, 139], Sample Entropy (SampEn) [140, 138, 141], and Recurrence Quantification Analysis (RQA) [142, 143, 143] have been used in classification of epileptic seizure EEG signals. The HOS is the spectral representation of higher order statistic i.e. moments and cumulants of third and higher order [116]. Acharya et al. [132] used features extracted from HOS for classification of the normal, interictal, and ictal EEG signals. The measure of CD has been used to quantify the complex neural activity of the human brain during interictal and ictal activities [133]. The LE has been used for discrimination of epileptic seizure EEG signals as it provides significant details about changes in EEG

activity [19]. The FD parameter has been used effectively to discriminate the chaotic nature of EEG signals in interictal and ictal activities [136]. Acharya et al. [130] used the CD, H, and ApEn parameters for the automated detection of epilepsy. The value of ApEn has been found to show a strong connection with the synchronous discharge of large groups of neurons during seizure [139]. Acharya [142] used the RQA features for classification of the normal, interictal, and ictal EEG signals.

Once features are extracted from EEG signals, a classifier is employed to differentiate between normal and epileptic EEG. In past two decades, many classification methods have been used for seizure detection such as artificial neural networks (ANNs) [144, 141], multilayer perceptron neural network (MLPNN) [144], recurrent neural networks (RNNs) [135], Fuzzy Sugeno Classifier (FSC) [138, 145], SVM [146, 132, 147, 109], LS-SVM [148, 149], Gaussian mixture model (GMM) [130, 110], DT [123, 150], k -NN [151, 152], Mixture expert model [119], RF [20].

2.3.3 MI Signal Classification Methods in the BCI

In BCI application, a variety of methods have been developed and employed to extract features from the MI based EEG signals. The feature extraction methods include Hjorth parameters [153], Band Powers [154], autoregressive (AR) and its adaptive version (AAR) [155, 156, 157, 158, 159], Fast Fourier transform [113, 160], power spectral density (PSD) [161], common spatial patterns (CSP) [162, 163, 164], wavelet coefficients [165, 166], phase locking value (PLV) and spectral coherence [167, 168, 169, 170]. Obermeier [153] used Hjorth parameters including activity, mobility, and complexity from a single trial EEG signal for the classification of a right and left imagery problem. Spectral powers of alpha (α) and beta (β) frequency bands were

used in [154] as features to discriminate right and left imagery movements. The Band Powers were also used in [171, 172] for the classification of mental tasks. The feature extraction method based on AR model and Bayesian logistic regression model has been developed in [155] to classify motor imagery tasks. Tavakolian et al. [173] used AAR model parameters with five different machine learning techniques for classification of mental tasks. Many researchers have used the PSD as features in BCI systems [174, 175, 176, 30]. Lachaux et al. [168] proposed the PLV which is the commonly used feature extracted from phase information. A number of researchers have attempted to use PLV in BCI systems [167, 177, 160]. The combination of phase information and conventional features such as BP and AR-based features have been used in [177] for the classification of motor imagery tasks. Krusienski et al. [160] used the phase information with other spectral features extraction by the Fast Fourier Transform or Magnitude-Squared Coherence to control cursor movement. Authors in [178, 6, 179] used the CSP to discriminate the spatial patterns for two classes of motor imagery in EEG-based BCIs.

Various classification methods have been studied and used in a BCI system design. They can be divided into five categories: *linear classifiers*, *neural networks*, *non-linear Bayesian classifiers*, *nearest neighbour classifiers* and *combinations of classifiers* [180, 181]. *Linear classifiers* are the most popular algorithms for BCI applications which use linear functions to discriminate classes. Linear Discriminant Analysis (LDA) also known as Fisher's Linear Discriminant Analysis (FLDA) and SVM are two widely used linear classifiers. Both methods construct a hyperplane or set of hyperplanes to separate the data representing different classes [182]. These classifiers have been used with success in many BCI system [183, 184, 185]. LDA is a simple classifier with the low computational requirement, which makes it suitable for online BCI system. However, it can provide poor results on complex nonlinear EEG data due to its linearity [186]. SVM

is robust to dimensionality due to regularisation, which means a large training set is not required for good results, even with very high dimensional feature vectors [187]. This advantage is gained at the expense of a low speed of execution [181].

Neural Networks (NN) are non-linear classifiers that have been broadly used in BCI research [188, 181]. The most widely used NN for BCI is the MLP which has been applied to mental task classification [189, 171, 190] and motor imagery [191]. As the MLP is universal approximators, it is sensitive to overtraining, especially with such noisy and nonstationary data as EEGs [192]. Therefore, careful architecture selection and regularisation is required [193]. Besides MLP, different types of ANN architecture have been used in the design of BCI systems such as Gaussian classifier [189], Probabilistic Neural Networks [194, 195], Probability estimating Guarded Neural Classifiers [196], RBF Neural Network [197], Fuzzy ARTMAP Neural Networks [198], Finite Impulse Response Neural Networks (FIRNN) [199], and Bayesian Logistic Regression Neural Network [155]. The performance of MLP Neural Network with Driven Pattern Replication (MLP-DPR), Probabilistic Neural Network Delta band (PNN-DB), Hierarchical Model (HM), and Modular Multi-net system (MMN) have been studied in [200] to select the best classifier to control a mobile robot. The HM with statistical implementation achieved the best result with an accuracy of 91%.

Bayes quadratic and Hidden Markov Model (HMM) are two types of *Non-linear Bayesian classifiers* used in BCI system [180, 181]. Both classifiers generate nonlinear decision boundaries. As these classifiers are generative, they enable to perform more efficient solutions of uncertain samples than discriminative classifiers. However, they are not very popular as *linear classifiers* or *Neural Networks* in BCI applications [180, 181]. They have been applied to motor imagery and mental task classification [201, 191, 202]. There are two types of nearest neighbour classifiers used in BCI

systems; k Nearest Neighbour (kNN) and Mahalanobis Distance [180, 181]. KNN assign an unseen point to the dominant class among its k nearest neighbours inside the training set. These nearest neighbours are obtained using a metric distance as Euclidean distance [203]. KNN algorithms are not widespread in BCI applications due to being very sensitive to the curse-of-dimensionality [180], which made them fail in several BCI experiments [204]. However, they prove to be efficient In BCI system with low-dimensional feature vectors [205]. Mahalanobis Distance classifier is a simple yet robust classifier with good performances [204, 206, 207]. However, it has been hardly used in the BCI system [180, 181].

The *combination of classifiers* is a new trend which uses several classifiers and aggregates them in different ways to reduce the variance and the classification error and achieve better results compared to the single classifier. Boosting, voting and stacking are the classifier combination strategies used in BCI applications [180, 181]. Boosting consists of using some classifiers in cascade, each classifier focusing on the errors committed by the previous ones [182]. Boosting with MLP [208] and Ordinary Least Square [209] have been attempted. In voting, several classifiers are used with each of them assigning an input feature vector to a class. The final class will be the majority of them [193]. Voting had been experimented with LVQ NN [210], MLP [211] or SVM [212]. Stacking consists of using several classifiers (called level-0 classifiers), each of them classifying the input feature vector. The output of each classifier is given as input to a meta-classifier which makes the final decision [213]. Stacking has been used in BCI research using HMM as level-0 classifiers, and an SVM as meta-classifier [214].

CHAPTER 3

EFFECTIVE AND EFFICIENT DETECTION OF PREMATURE VENTRICULAR CONTRACTIONS BASED ON VARIATION OF PRINCIPAL DIRECTIONS

The classification of ECG data stream is essential to diagnosis of critical heart conditions. It is vital to accurately detect abnormality in the ECG in order to prevent possible beginning of life-threatening cardiac symptoms. This chapter focuses on identifying PVC which is one of the most common heart rhythm abnormalities. A “Replacing” strategy is used to check the effects of each individual heartbeat on the variation of principal directions. Based on this idea, an online PVC detection method is proposed to classify the new arriving PVC beats in the real-time and online manner. The proposed approach is tested on the MIT-BIH arrhythmia database (MIT-BIH-AR). These results show that the proposed method is effective in terms of computation time and is capable of accurately classify PVCs in an online manner.

3.1 Introduction

The ECG is the non-invasive method used to detect heart disease. It reflects the electrical activity of the heart over time and provides information about the state of the heart. The ECG is commonly used to detect cardiac rhythm abnormalities also known as arrhythmias. Arrhythmias can be divided into two groups. The first group is life-threatening and require immediate treatment with a defibrillator such as ventricular fibrillation and tachycardia. The second group is not imminently life-threatening but may require treatment to prevent further problems such as premature ventricular contractions (PVCs).

Recently portable ambulatory ECG monitoring has attracted increasing attention due

to its ability to meet the growing clinical needs for personal healthcare and remote monitoring, etc. [215]. Its aim is to help clinicians to diagnose critical cardiac conditions by processing ECG signal in real-time. PVCs are premature heartbeats originating from the ventricles of the heart. They are one of the most common heart rhythm abnormalities which can be linked to mortality associated with myocardial infarction [7]. Therefore, their immediate detection and treatment is essential to prevent possible beginning of life-threatening cardiac conditions.

Many methods for automatic classification of various arrhythmias have been developed based on different features extraction and classification methods. Features methods include ECG morphology [216, 1], heartbeat interval features [89, 14], frequency-based features [88, 14], higher order cumulant features [217], Karhunen-Loeve expansion of ECG morphology [89], and hermite polynomials [86]. Several classifiers have been used for heartbeats classification including linear discriminants [89, 218], artificial neural networks [13, 14, 15], and self-organizing networks [86]. Although these methods accurately identify PVCs, they have long computational time. Therefore, due to high complexity of their algorithms, they are not suitable for portable ambulatory ECG devices where the computation resources to process ECG in real-time is limited. In addition, some of existing high detection results are based on small data sets or used overlapped training and testing data sets. For example, wavelet feature extraction in tandem with fuzzy neural network classification was used in [13] to classify PVC beat with 97.04% accuracy. However the method was tested on only seven files of the MIT/BIH arrhythmia database [219].

To address the aforementioned problems, this study focuses on developing an online PVC detection algorithm with low-complexity. It is observed that replacing a non-PVC beat with a PVC beat will cause a larger effect on principal directions than replacing a

non-PVC beat with another non-PVC beat. From this observation, the dominant principal direction is extracted from training data set which consists of K normal beats to characterise the normal profile for the dataset. Then a “Replacing” procedure is applied to check the effects of each individual new arriving heartbeat on the variation of principal directions. The beat is considered as PVC beat if this score is greater than a certain threshold. Since the proposed method needs to calculate the principal directions n times for an ECG signal with n heartbeats to classify the beats, the power method [220] is applied to alleviate this heavy loading and reduce the computational complexity. The proposed method is evaluated on two non-overlapping data sets (DS1 and DS2) from the MIT-BIH arrhythmia database (MIT-BIH-AR) [219] in order to assess its performance over a large data set.

The main contributions of this chapter are summarised as follows: (1) a low-complexity PVC detection method based on the variation of dominant principal direction is developed; (2) the proposed method uses a small training data set to train the classifier; (3) it achieves high accuracy detection rates over large data set and demonstrates the capability to detect abnormal PVCs in online manner.

The remainder of this chapter is organised as follows: The ECG database is discussed in Section 3.2. Section 3.3 presents the proposed PVC classification method. Section 3.4 describes evaluation criteria which is used to evaluate the performance of proposed method. Section 3.5 presents the experimental results, including comparisons with previously published work. Finally, Section 3.6 concludes this chapter.

3.2 ECG Data

This study uses a publicly available benchmark dataset, MIT-BIH arrhythmia database [219] to test the effectiveness of the proposed method. The database consists of 48 two-lead (denoted lead A and B) recordings of approximately 30 minutes and sampled at 360 Hz. Note that only recordings of lead A (modified-lead II (MLII)) were used in this study. In accordance with the AAMI recommended practice, four recordings (# 102, # 104, # 107 and # 217) containing paced beats were discarded since these beats do not retain sufficient signal quality for reliable processing [221]. The remaining 44 recordings were divided into two datasets (DS1 and DS2) with each dataset containing ECG data from 22 recordings. For comparison purposes, the same dataset division scheme as [218, 83, 16] is used for the DS1 and DS2.

The data consist of 5 different heartbeat classes. Class N contains beats originating in the sinus node (normal and bundle branch block beat types), class S contains supraventricular ectopic beats, class V contains premature ventricular contraction beats (The ventricular escape beats were also classified as the PVC beats in this study [218, 83, 221, 16]), class F contains beats that result from fusing normal and ventricular ectopic beats, and class Q contains unknown beats. Table 3.1 shows the breakdown of each dataset by heartbeat classes. The first dataset (DS1) was used to evaluate the performance of different candidate classifiers. The second dataset (DS2) was used for a final performance evaluation of the proposed method.

Table 3.1: Heartbeat classes associated with the extracted beats for the full database, dataset 1 (Ds1) and dataset 2 (Ds2) from the MIT-BIH arrhythmia database.

Dataset	Recordings	N	S	V	F	Q	Total
Full Database	44 recordings	90127	2745	7008	802	17	100699
DS1	101,106,108,109,112	45880	909	3788	414	10	51001
	114,115,116,118,119						
	122,124,201,203,205,207						
	208,209,215,220,223,230						
DS2	100,103,105,111,113	44247	1836	3220	388	7	49698
	117,121,123,200,202						
	210,212,213,214,219,221						
	222,228,231,232,233,234						

Note: N= Normal Beat; S= Supraventricular ectopic beats; V= Premature ventricular contraction beats; F= Fusion of ventricular and normal beat; Q=unknown beats. This chapter focuses on identifying V class (premature ventricular contractions (PVCs) abnormalities).

3.3 Method for Online PVC Heartbeat Detection

Figure 3.1 depicts the stages of the proposed system for online detection of PVC beats. It consists of two stages: a processing stage and an abnormality detection stage. The ECG lead MLII is applied at the input to the processing stage. A sample ECG signal from recording # 119 in the MIT-BIH arrhythmia database is shown in Figure 3.2. The processing stage consists of heartbeat detection, segmentation and normalisation modules. The heartbeat detection module attempts to locate all heartbeats. The segmentation module extracts single beats from the entire ECG. The normalisation module is concerned with normalizing heartbeats that are processed by the classifier stage. The abnormality detection stage contains a classifier unit that classifies heartbeats into two

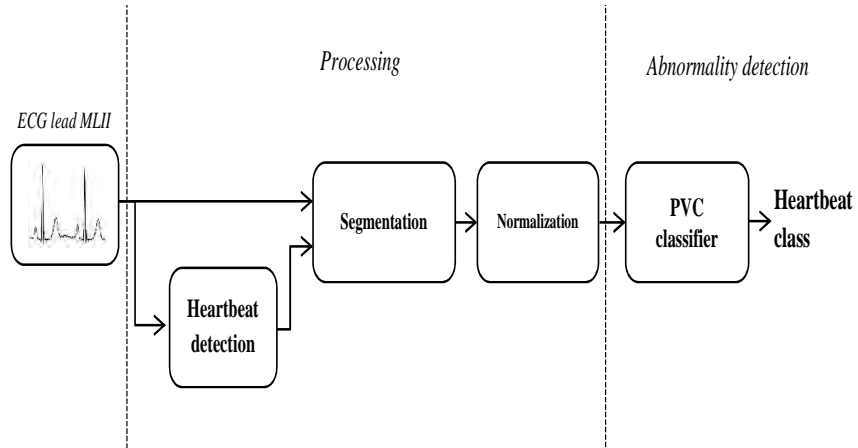


Figure 3.1: Schematic representation of the proposed PVC detection.

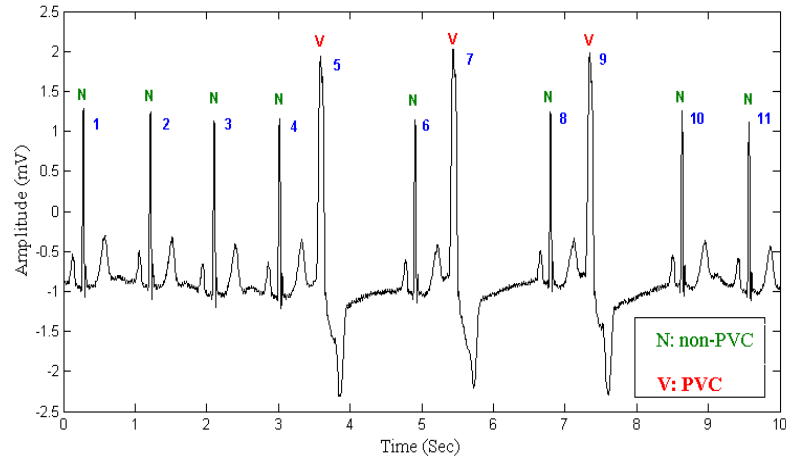


Figure 3.2: A segment of ECG recording 119, with the R-wave peaks of non-PVC beats marked by 'N' and PVC beats marked by 'V'.

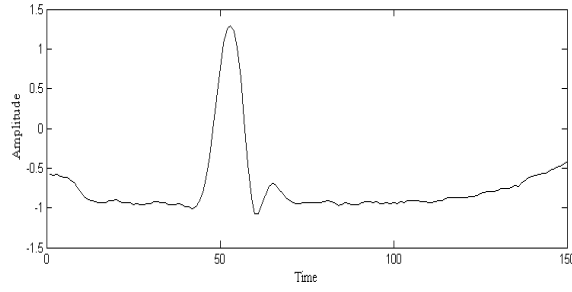
classes: Non-PVC and PVC heartbeat. The classifier utilises variation of dominant principal directions to determine PVC heartbeats. It contains some parameters which are set during the training phase to optimise the classification performance. The modules forming these stages are discussed in more details below.

3.3.1 Heartbeat Detection

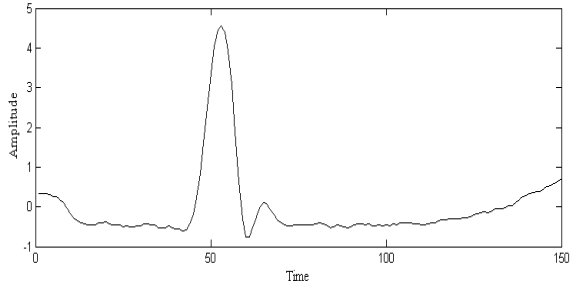
The “Pan and Tompkins” QRS detection algorithm [222] was used to detect the location of QRS complexes. This method is chosen due to its computational simplicity and ease in implementation. The Pan-Tompkins algorithm identifies the R-point based on digital analyses of the slope, amplitude, and width of the ECG data. It consists of the following steps: band pass filtering of the ECG (to reduce interference present in the ECG signal), differentiating the signal (to provide the QRS complex slope information), squaring of the data samples (to emphasise peaks), moving average filtering (to smooth close-by peaks) and detecting R point. The algorithm automatically adjusts the thresholds and parameters periodically to adapt to changes in QRS morphology and heart rate.

3.3.2 Segmentation and Normalisation

After detection of R-points, the ECG signal was segmented such that each segment contain a portion of signal before and after each R-point. In this study, each segment consists of 50 samples (138.88 ms) before R-point, 99 samples (275 ms) after R-point. Each of these 150 samples (416.67 ms) segment are considered as a single beat segment. Figure 3.3.2 depicts first segment extracted from the sample ECG signal shown in Figure 3.2. Then each segment is normalised to have a mean of zero and a standard deviation of one in order to eliminate the offset effect and to decrease the effect of false classification due to personal or instrumental differences. Figure 3.3(b) shows the effect of normalisation on a segment obtained from the patient record # 119.



(a) Before Normalisation.



(b) After Normalisation.

Figure 3.3: First segment extracted from the sample ECG signal shown in Figure 3.2.

3.3.3 PVC Detection via Variation of Dominant Principal Directions

In this section, first a brief review of the PCA algorithm is presented. Then based on the replacing strategy, the effects of PVC beats on the derived dominant principal direction (a direction that captures the largest variability in the data) and the proposed PVC detection method are presented. Finally, an effective computation for estimating principal directions in replacing strategy is proposed.

3.3.3.1 Principal Component Analysis

PCA is a linear dimensionality reduction technique, which determines the principal directions of the data distribution. It involves constructing the data covariance matrix and calculating its dominant eigenvectors to obtain the principal directions. These directions contain the most important aspects of the data. Let $\mathbf{A} = [x_1^T; x_2^T; \dots; x_n^T] \in R^{n \times p}$ where each row x_i represents a data instance in a p dimensional space, and n is the number of the instances. PCA can be formulated as the following optimisation problem:

$$\max_{\mathbf{U} \in R^{p \times k}, \|\mathbf{U}\|=I} \sum_{i=1}^n \mathbf{U}^T (x_i - \mu)(x_i - \mu)^T \mathbf{U} \quad (3.1)$$

where \mathbf{U} is a matrix consisting of k dominant eigenvectors. This problem can be solved by deriving an eigenvalue decomposition problem of the covariance data matrix.

$$\hat{\mathbf{C}}_{\mathbf{x}} \mathbf{U} = \mathbf{U} \Lambda \quad (3.2)$$

where

$$\hat{\mathbf{C}}_{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T \quad (3.3)$$

is the covariance matrix, μ is the global mean, and the resulting \mathbf{U} is the eigenvector set of $\hat{\mathbf{C}}_{\mathbf{x}}$. In order to reduce the dimension of the data, only few principal directions are selected to represent the data. These principal components explain the majority of the variance.

It is found that the dominant principal direction will be affected by the presence of PVC beats in the data due to correlation change. This property is used to accurately identify PVC beats.

3.3.3.2 The Use of Principal Component Analysis for PVC Detection

In this section, the variation of principal directions are studied when replacing one heart beat with a new heart beat into the data matrix, utilizing this property to determine PVC beats.

Figure 3.4 and Figure 3.5 are used to explain the above observation. Every ECG signal can be represented as a row vector $E : E = (x_0, x_1, \dots, x_n)$ where each x_i represent a data instance in time i .

After detecting the R-point using QRS detection algorithm [222], ECG signal is segmented and each segment is normalised. Let $S = [s_1, s_2, \dots, s_{n-1}, s_n]$ where each segment $s_i = (s_{i,1}, \dots, s_{i,150})$ consist of 150 data samples extracted from ECG signal (50 points before and 100 points after each R-point). For the ECG recoding #119 in Figure 3.2, $S = [s_1, s_2, \dots, s_{10}, s_{11}]$ since there are 11 heartbeats in that ECG sample. Then a normal data matrix \mathbf{M} is constructed from k normal segmented heartbeats. the normal data matrix $\mathbf{M} = [s_1, s_2, s_3]^T$ is constructed from three normal segmented heartbeats (beat number 1, 2, and 3 from the sample ECG shown in Figure 3.2) and the dominate principal direction u is extracted from it by using Equation 3.2.

The blue triangles in Figure 3.3.3.2 represent normal data samples \mathbf{M} (see Figure 3.4) constructing from three normal segmented heartbeats (beat number 1, 2, and 3 from the sample ECG shown in Figure 3.2), and the black arrow is the dominant principal direction. Figure 3.5(b) and 3.5(c) show data samples matrix $\mathbf{M}'_1 = [s_1, s_2, s_4]^T$ and $\mathbf{M}'_2 = [s_1, s_2, s_5]^T$ constructing from replacing segmented heartbeat number 3 with heartbeat number 4 and 5 in normal data samples matrix and resulting dominant principal directions u'_1 and u'_2 (red arrows). The Construction of normal data matrices \mathbf{M} and \mathbf{M}'_2 are shown in Figure 3.4.

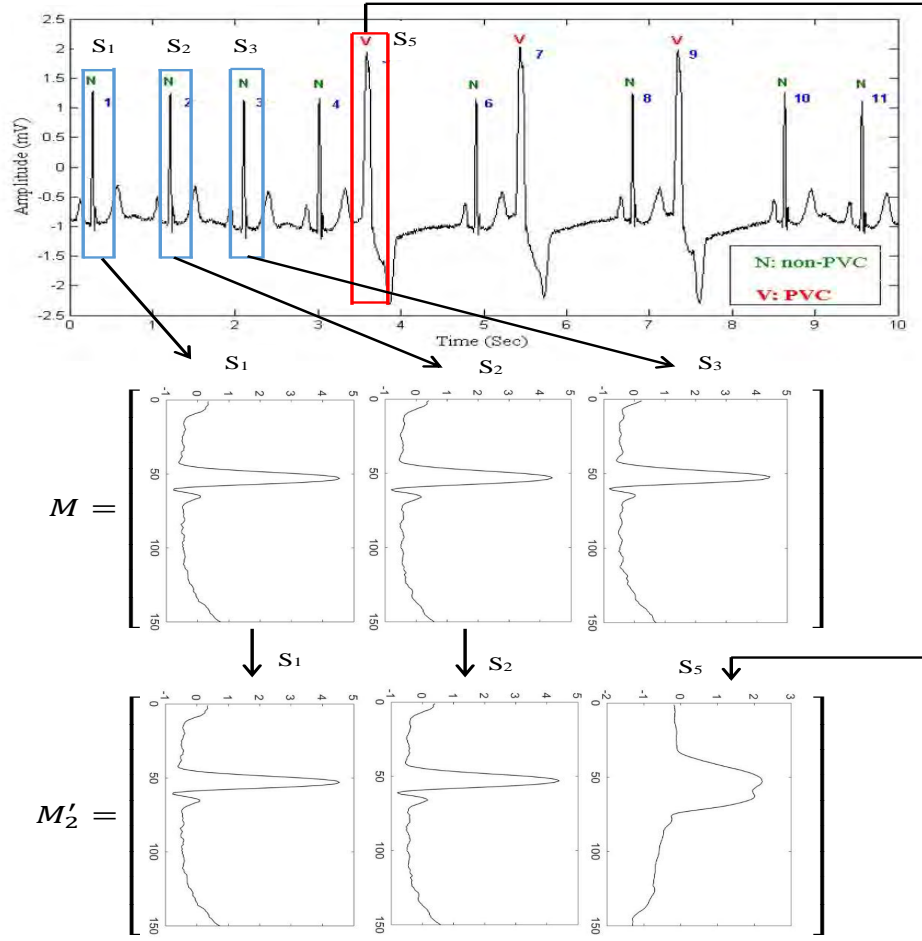


Figure 3.4: The Construction of normal data matrix M and data matrix M'_2

Figure 3.5(b) shows that the dominant principal direction is not affected by replacing non-PVC beat (beat number 3) with non-PVC beat (beat number 4) and only produces a small angle between the old first principal direction (black arrow) and the new resulting one (red arrow). However, the dominant principal direction is deviated by replacing non-PVC (beat number 3) beat with PVC beat (beat number 5). As shown in Figure 3.5(c), the dominant principal direction is changed and forms a larger angle between the old dominant principal direction (black arrow) and the new resulting one (red arrow). Via this observation, this property is used to identify PVC beats using replacing strategy.

The idea of combining PCA and the replacing strategy for PVCs detection is pre-

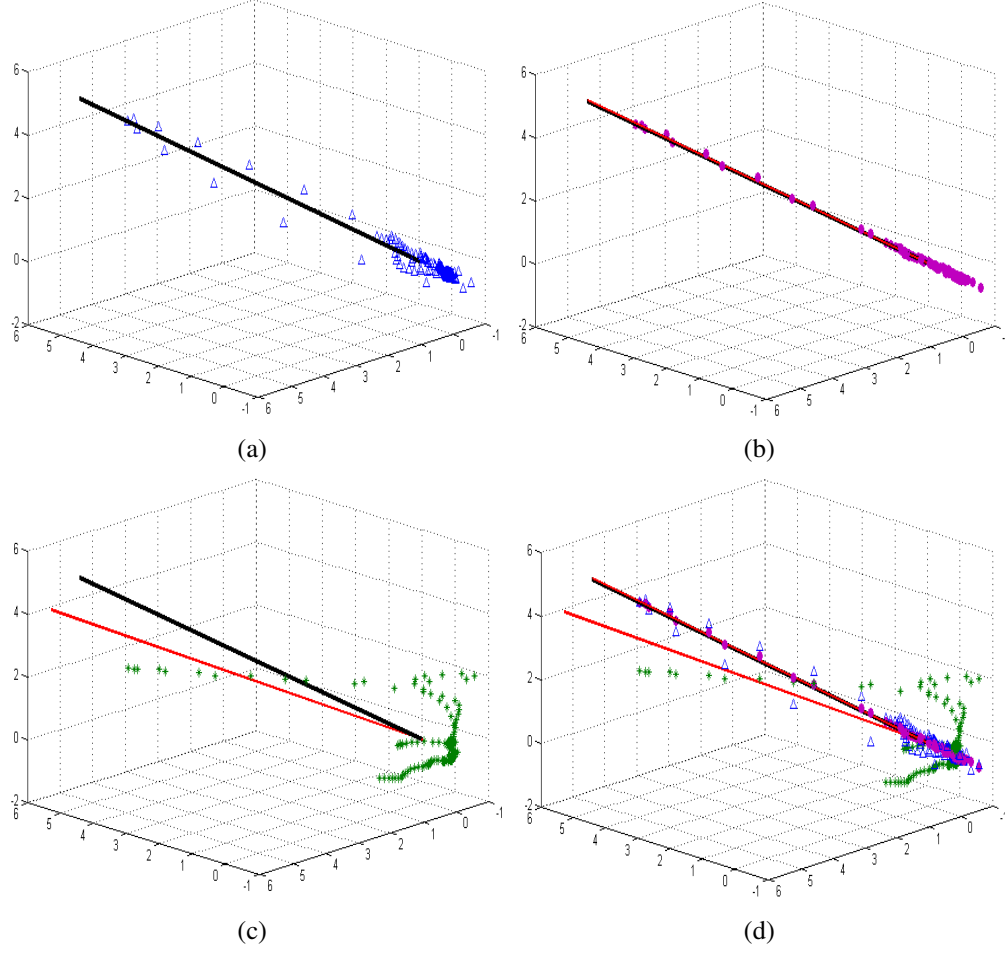


Figure 3.5: The effects of replacing a non-PVC or a PVC beat on the principal directions.

sented at this stage. Given an ECG signal with n segmented heartbeats, k normal beats are selected from the first 2 min of each recording. It is worth to mention that the beginning of some records in both DS1 and DS2 contain a lot of V, S, F, Q beats and noise. Therefore the first 2 min of each recording is used to make sure sufficient k normal beats can be obtained. Then a matrix $\mathbf{M} = [s_1 \dots s_k]^T \in R^{150 \times k}$ is built from the k normal beats. The dominant principal direction u is extracted from matrix \mathbf{M} . When receiving a new segmented heartbeat s_x (the rest of the beats in each record regardless of the k normal beats), the leading principal direction u' is computed from matrix \mathbf{M}' which is constructed by replacing s_x with one of the normal beats in matrix \mathbf{M} . To

Table 3.2: The resulting angle of each heartbeat in the sample ECG.

Beat number (i)	4	5	6	7	8	9	10	11
Angle θ_i (degree)	0.26	7.23	0.13	7.04	0.09	7.29	0.08	0.24

classify the PVC beats in an ECG signal, this procedure is repeated $n-k$ times with the replacing strategy (one for each remaining segmented beat).

Once these leading principal directions \mathbf{u}' are obtained, the angle between \mathbf{u} and \mathbf{u}' is calculated to measure the variation of the principal directions, i.e.

$$f(\mathbf{u}, \mathbf{u}') = \arccos\left(\frac{\mathbf{u} \cdot \mathbf{u}'}{|\mathbf{u}| \times |\mathbf{u}'|}\right), f \in [0, \pi] \quad (3.4)$$

For example the resulting angle of each heartbeat in the sample ECG shown in Figure 3.2, is presented in Table 3.2. The matrix \mathbf{M} is constructed from the first three normal beats number 1, 2 and 3 ($k = 3$) in the sample ECG. As it can be seen, PVC beats number 5, 7 and 9 have larger angle as compared to non-PVC beats.

Based on the variation of the principal directions, two adaptive thresholds are proposed: threshold T_1 for updating the dominant principal direction \mathbf{u} and threshold T_2 for classifying the PVC beats.

$$\begin{aligned} T_1 &= \text{mean}(\theta_u) \times r_1 \\ T_2 &= \text{mean}(\theta_{non-PVC}) \times r_2 \end{aligned} \quad (3.5)$$

where r_1 and r_2 are unknown parameters which are set during training process, both θ_u and $\theta_{non-PVC}$ are two sets containing angles with initial value of θ_0 (the angle between the normal dominant principal direction \mathbf{u} and a dominant principal direction resulting

of replacing a normal beat u'). The value of θ_0 is calculated for each record during training process.

Both T_1 and T_2 are updated by the newly received non-PVC or PVC beats. Let θ_x be an angle for a new segmented heartbeat s_x , the s_x will be classified as PVC beat if its θ_x is above the threshold T_2 . However, if the value of θ_x is less than T_2 , the segment s_x will be classified as non-PVC beat and the θ_x value will be added into $\theta_{non-PVC}$ set. Therefore, the threshold T_2 will be updated.

In order to adapt the algorithm to changes in QRS morphology and heart rate, the dominant principal direction u is replaced by the u_x (the leading principal direction after replacing one normal beat in matrix M with s_x) when the θ_x is below the threshold T_1 . When the value of θ_x is less than T_1 , The θ_x value will be added to the θ_u set and T_1 will be updated

3.3.3.3 The Choice of Parameters (k , r_1 and r_2)

Different normal data matrices M with different values of k were constructed by selecting k normal beats from each record in the training dataset DS1 and then the resulting angle of replacing the rest of the beats in each record is calculated. It is worth to note that only those records in the training dataset DS1 are used that contain more than one hundred PVC beats in order to reduce the number of experiments. Table 3.3 shows the effect of different values of k on resulting angles of different heartbeats classes for nine records in the training dataset DS1. The results of each record is reported in terms of mean \pm standard deviation of the resulting angles for different heartbeats classes.

As it is shown in the Table 3.3, the resulting angles for PVC beats (V) are larger than the resulting angles for normal beats (N) in each record for all different values of k .

The results also show that the resulting angles of PVC beats (V) decrease by increasing the number of normal beats k in matrix M which mean that the differences between the resulting angles for normal beats and PVC beats are also decreased by increasing the number of normal beats k in matrix M . It can be seen from Table 3.3, the maximum differences between the resulting angles for normal beats and PVC beats are achieved by choosing $k=10$. For this reason, and also in order to use a small training data set, the value of k is chosen to be 10 in this study.

The selection of parameters r_1 and r_2 play an important role in developing the above PVC classifier, which was the main purpose of the training process. The PVC classification performances for nine records (# 106, # 116, # 119, # 201, # 203, # 207, # 208, # 215, and # 223) in the training data set (DS1) were obtained for different r_1 and r_2 . The r_1 is set from 1 to 5 and obtained the performance of the classifier for different values of r_2 (6, 7, 8, 9 and 10 were assigned to r_2 in this study). They were then compared with the cardiologist's annotations, with the overall accuracy (see Table 3.4) calculated. Table 3.5 shows the classification performance for nine records in DS1 with different values of r_1 and r_2 . It can be seen from Table 3.5, the best classification performance was obtained by setting $r_1 = 3$ and $r_2 = 9$. Thus, the r_1 is selected to be 3 and r_2 to be 9 to evaluate the performance of the classifier in the testing data set (DS2).

3.3.3.4 Power Method for Estimating Principal Directions

PCA can be determined by solving an eigenvalue decomposition problem. In the proposed PVC classifier with replacing strategy, there is a need to calculate the principal directions n times for an ECG signal with n heartbeats. In order to alleviate this heavy loading, the power method [220] is applied to extract the first principal direction. This method is a simple iterative algorithm which does not compute matrix decomposition.

Table 3.3: The effect of different values of k on resulting angles of different heart-beats classes for nine records in the training dataset DS1.

Rec.	classes	$\theta(\text{mean} \pm \text{standard deviation})$						
		k=9	k=10	k=11	k=12	k=13	k=14	k=15
106	N	0.43±0.33	0.20±0.19	0.19±0.18	0.19±0.17	0.38±0.26	0.41±0.27	0.39±0.26
	V	8.37±5.92	7.74±5.56	7.44±5.29	7.17±5.06	7.12±4.86	6.96±4.66	6.71±4.48
116	N	0.47±0.72	0.22±0.60	0.52±0.67	0.47±0.62	0.48±0.60	0.50±0.60	0.47±0.55
	V	9.90±5.01	9.38±4.67	9.21±4.40	8.81±4.15	8.47±3.96	8.27±3.78	7.81±3.66
	S	2.76	1.01	2.71	2.51	2.48	2.49	2.75
119	N	0.61±0.04	0.59±0.04	0.59±0.03	0.58±0.04	0.55±0.04	0.52±0.04	0.49±0.03
	V	8.09±0.39	7.93±0.37	7.43±0.35	7.33±0.34	7.14±0.33	6.93±0.31	6.75±0.30
201	N	0.88±0.21	0.90±0.21	0.80±0.19	0.79±0.19	0.77±0.18	0.76±0.18	0.74±0.18
	V	11.48±2.64	10.44±3.85	10.70±2.22	10.34±2.23	9.86±2.24	9.5±2.41	9±2.76
	S	4.99±3.34	5.99±3.43	4.56±3.04	4.59±2.93	4.34±2.80	4.44±2.72	4.67±2.84
	F	29.18±0.81	19.17±12.66	26.07±0.72	24.92±0.70	23.74±0.67	22.85±0.63	15.35±10.02
203	N	1.90±3.62	1.67±3.47	1.59±3.36	1.98±3.67	2.14±3.74	2.07±3.58	2.04±3.44
	V	31.27±5.65	29.99±4.59	28.56±4.33	27.09±4.82	26.15±4.77	24.98±4.64	24.14±4.41
	S	36.28±0.01	34.28±0.01	32.63±0.01	31.38±0.08	30.44±0.06	29.15±0.07	28.17±0.08
	F	9.29	7.87	8.16	14.47	14.79	14.06	14.30
	Q	29.68±13.52	28.15±13.73	20.80±13.73	18.89±8.41	19.62±7.32	18.82±7.00	18.94±6.32
207	N	0.14±0.42	0.23±0.74	0.20±0.66	0.22±0.70	0.20±0.66	0.20±0.63	0.19±0.60
	V	15.30±6.02	16.44±5.40	15.38±5.16	15.01±4.83	14.32±4.61	13.86±4.41	13.17±4.26
	S	5.71±10.13	1.26±3.99	2.17±5.65	1.18±3.61	1.33±3.96	1.22±3.31	1.59±4.45
208	N	0.51±0.38	0.38±0.54	0.54±0.68	0.36±0.52	0.51±0.55	0.76±0.87	0.67±0.22
	V	10.12±3.76	11.88±1.78	10.69±2.05	10.49±2.06	8.14±3.08	8.44±2.85	7.84±2.81
	S	13.85±0.001	1.97±0.01	2.05±0.002	1.84±0.0014	11.38±0.00027	3.48±0.0064	10.72±0.0018
	F	14.68±1.18	7.45±5.23	8.21±5.41	8.28±5.20	12.16±0.84	8.71±3.06	11.27±0.82
	Q	24.80±1.19	23.37±1.07	22.07±1.02	21.09±0.92	20.09±0.87	17.51±0.59	18.35±0.91
215	N	0.22±0.21	0.21±0.20	0.21±0.20	0.21±0.19	0.21±0.21	0.48±0.14	0.20±0.20
	V	10.43±2.89	10.10±2.69	9.83±2.55	9.39±2.45	9.11±2.35	8.99±2.30	8.64±2.18
	S	1.86±0.07	1.88±0.09	1.67±0.15	1.69±0.25	1.76±0.07	4.20±0.13	1.64±0.06
	F	21.55	20.47	19.08	18.60	17.82	17.22	16.17
223	N	0.21±0.12	0.09±0.09	0.09±0.08	0.17±0.09	0.19±0.10	0.19±0.10	0.07±0.07
	V	10.24±12.09	8.90±10.77	8.43±10.15	8.79±10.23	8.58±9.85	8.29±9.46	7.27±8.56
	S	2.23±5.94	4.38±10.61	4.62±10.59	2.36±6.15	1.98±4.86	1.93±4.67	3.99±8.96
	F	4.90±9.40	13.14±16.93	12.47±16.10	4.26±8.03	4.25±7.66	4.12±7.36	10.63±13.64

Note: N= Normal Beat; S= Supraventricular ectopic beats; V= Premature ventricular contraction beats; F= Fusion of ventricular and normal beat; Q=unknown beats. This chapter focuses on identifying V class (premature ventricular contractions (PVCs) abnormalities).

Given a matrix \mathbf{M} , this method starts with a vector \mathbf{u}_0 , which may be an approximation to the dominant eigenvector or a nonzero random vector, then iteratively computes the \mathbf{u}_{k+1} as follows:

$$\mathbf{u}_{k+1} = \frac{\mathbf{M}\mathbf{u}_k}{\|\mathbf{M}\mathbf{u}_k\|} \quad (3.6)$$

The sequences $\{\mathbf{u}_k\}$ converge to an eigenvector associated with the dominant eigenvalue of \mathbf{M} . The power method only requires matrix multiplications to calculate the dominant principal direction. Therefore the use of the power method can reduce the computational complexity of the proposed PVC classifier. In this study, the tolerance is set to 0.001 and the number of iterations to 10 in order to stop the power method process.

In order to check the performance lost by using the power method, the root mean square error (RMSE) is calculated between estimated dominant principal direction extracted by power method and dominant principal direction extracted by PCA. The RMSE is very small 2.8051×10^{-9} which indicates that the performance lost by using the power method is negligible.

3.4 Classification Performance Measures

The performance of the proposed classifier was quantified by accuracy (Acc), sensitivity (Se), positive predictivity (+P) and false positive rate (FPR). Their formulas are given below:

Table 3.4: Performance measures used in this study for distinguishing PVC from non-PVC.

		Algorithm label	
		n	v
Reference label	N	N_n	N_v
	S	S_n	S_v
	V	V_n	V_v
	F	F_n	F_v
	Q	Q_n	Q_v

Note: N= Normal Beat; S= Supraventricular ectopic beats; V= Premature ventricular contraction beats; F= Fusion of ventricular and normal beat; Q=unknown beats; n=non-PVC beats (recognised); v= PVC beats (recognised).

$$TN = N_n + S_n + F_n + Q_n$$

$$FN = V_n$$

$$TP = V_v$$

$$FP = N_v + S_v$$

$$Se = TP / (TP + FN) \quad (3.7)$$

$$+P = TP / (TP + FP)$$

$$FPR = FP / (TN + FP)$$

$$Acc = (TP + TN) / (TP + TN + FP + FN)$$

$$t_t = t_{QRS} + t_{PVC}$$

Table 3.4 shows how the full classification matrix is used to calculate them. It is worth noting that the AAMI-recommended [223, 221] calculations shown in Table 3.4 do not reward or penalise a classifier for the classification of ventricular fusion (F) or

Table 3.5: Overall performance for nine records in the training set (DS1) using different values of r_1 and r_2 in the classification function.

Acc(%)					
r1 \ r2	1	2	3	4	5
6	96.91	95.24	97.19	95.5	84.96
7	97.32	96.67	97.62	96.39	86.54
8	96.21	95.81	97.84	96.96	88.1
9	95.63	95.41	97.9	97.83	88.68
10	91.61	91.54	93.32	93.38	90.41

unknown beats (Q) as PVC beats (V). In addition, execution time is measured to be an indicator of the computational cost. The execution time of the QRS detection algorithm (t_{QRS}), the execution time of the PVC classifier (t_{pvc}) and the total execution time of the proposed method (t_t) are measured separately. The proposed method was tested on Intel R CoreTM i5 CPU 3.2 GHZ computer with 8.00GB RAM using 32-bit OS and 32-bit MATLAB R signal processing tool.

3.5 Results

3.5.1 Performance of the PVC Classifier

After the training process, the performance of the PVC classifier was evaluated in the DS2 (testing data set). Table 3.6 and Table 3.8 show the PVC identification performance for each recording from DS1 and DS2. The confusion matrix of the classifications performed on the DS1 and DS2 are shown in Table 3.7 and 3.9. The overall accuracy

(Acc) was 98.77%, with sensitivity (Se) of 96.12%, positive predictivity (+P) of 86.48% and false positive rate (FPR) of 1.04% for the 22 files of the testing set (DS2). The PVC classifier performed slightly better in the training data set (DS1) than that in the testing set (DS2) in terms of A, +P and FPR. The proposed method shows the ability to accurately classify PVC beats with low complexity. The average execution time of the proposed method (t_t) for a 30 min ECG recording was 3.83 s (t_{QRS} (3.36 s) + t_{pvc} (0.47 s)). The average processing time of the proposed method to classify a heart beat is 1.67 (ms/beat), while the average time interval for a heart beat is 0.79 (s/beat). Therefore, the proposed method has the ability to detect a heart beat before the next upcoming heart beat arrives.

3.5.2 Comparison with Published Studies

The performance results of the proposed method were compared with 4 published studies where the entire or part of the MIT-BIH-AR database were used [13, 14, 15, 16]. The ECG recordings in common with [13, 14, 15, 16] were used, respectively. The performance results of the proposed method were compared directly with the corresponding reported results. Table 3.10 shows the comparison of the proposed method with 6 published studies. The proposed PVC classifier outperformed other methods where the high-complexity approaches were used.

Shyu et al. [13] proposed a PVC classification method based on wavelet transform (WT) and fuzzy neural network (FNN). They used a test-set of 7 recordings to determine the performance of their classifier (see Table 3.10). The performance of the classifier on these recordings was: accuracy 97.04% and sensitivity 99.02%. Using the 7 recordings in common with this study, the proposed classifier achieved better accuracy (99.35%)

Table 3.6: Classification performance on each recording of DS1 using the AAMI recommended performance measures.

Rec.	Number of beats					Measures						
	N	S	V	F	Q	Acc	Se	+P	FPR	t_{QRS}	t_{pvc}	t_t
101	1858	1	0	0	2	99.89	–	–	0.11	3.1912	0.3773	3.5686
106	1507	0	520	0	0	99.75	99.42	99.61	0.13	3.3046	0.4272	3.7318
108	1738	4	17	2	0	94.95	11.76	2.63	4.24	3.9321	0.3719	4.3041
109	2491	0	38	2	0	99.56	94.74	80	0.36	2.4894	0.5007	2.9901
112	2537	2	0	0	0	99.84	–	–	0.16	3.0206	0.4969	3.5176
114	1820	10	43	4	2	99.31	100	76.78	0.71	3.2952	0.4070	3.7023
115	1952	0	0	0	0	99.95	–	–	0.05	3.3668	0.4237	3.7906
116	2301	1	109	0	0	98.67	100	77.3	1.39	3.4511	0.5105	3.9617
118	2166	96	16	0	0	99.34	100	51.61	0.66	3.2979	0.4810	3.779
119	1542	0	444	0	0	100	100	100	0	3.1119	0.4133	3.5252
122	2475	0	0	0	0	99.84	–	–	0.16	2.9834	0.5249	3.5083
124	1564	2	47	5	0	99.88	100	95.92	0.13	2.9525	0.3345	3.2870
201	1635	127	198	2	0	98.52	86.87	98.29	0.17	3.2711	0.4020	3.6732
203	2528	2	444	1	4	95.33	95.05	78.29	4.62	3.3985	0.5916	3.9902
205	2570	3	71	11	0	99.96	100	98.61	0.04	3.3190	0.5173	3.8364
207	1543	106	210	0	0	94.35	59.05	86.71	1.15	3.2831	0.3945	3.6776
208	1585	2	992	372	2	99.68	99.8	99.3	0.38	3.3107	0.5578	3.8685
209	2621	383	1	0	0	100	100	100	0	3.6115	0.5845	4.1961
215	3195	3	164	1	0	99.7	99.39	94.77	0.28	4.3961	0.7106	5.1067
220	1952	94	0	0	0	100	–	–	0	3.1271	0.4107	3.5379
223	2045	73	473	14	0	94.85	75.9	94.72	0.94	3.3247	0.5338	3.8586
230	2255	0	1	0	0	100	100	100	0	3.7101	0.4764	4.1865
Total	45880	909	3788	414	10	98.81	92.85	91.33	0.71	73.1498	10.4491	83.5989

Table 3.7: Confusion matrix on DS1 data set.

		Algorithm label		
		n	v	Total
Reference label	N	45576	304	45880
	S	879	30	909
	V	271	3517	3788
	F	278	136	414
	Q	3	7	10
	Total	47007	3994	51001

and sensitivity (100%).

Inan et al. [14] proposed a robust Neural-Network-Based PVC classifier. They combined wavelet transform and timing interval features as their feature set for classification. the performance of the proposed method was recalculated for 36 recordings (see Table 3.10) that are common to both studies in order to compare results. The results show that the proposed method is notably better than the Inan’s method.

A PVC classifier based on the neural network with weighted fuzzy membership functions (NEWFMs) was proposed in [15]. This method classifies PVC beats by the trained bounded sum of weighted fuzzy membership functions (BSWFMs) using wavelet transformed coefficients. Using the 6 recordings in common with this study, the proposed method achieved an accuracy of 99.74% which was slightly below the 99.8% achieved in [15]. However, the sensitivity of the proposed method was slightly better than the sensitivity of the approach in [15].

Li et al. [16] proposed a low-complexity data-adaptive approach for PVC identification based on template matching. Two beat-to-beat templates were proposed in this

Table 3.8: Classification performance on each recording of DS2 using the AAMI recommended performance measures.

Rec.	Number of beats					Measures						
	N	S	V	F	Q	Acc	Se	+P	FPR	t_{QRS}	t_{pvc}	t_t
100	2238	33	1	0	0	99.87	100	25	0.13	3.4436	0.4940	3.9376
103	2078	2	0	0	0	99.95	–	–	0.05	2.9467	0.4167	3.3634
105	2525	0	41	0	5	89.28	9.76	1.65	9.43	3.6198	0.5020	4.1218
111	2123	0	1	0	0	97.08	100	1.59	2.92	3.4092	0.4544	3.8636
113	1788	6	0	0	0	99.89	–	–	0.11	3.1659	0.3925	3.5584
117	1533	1	0	0	0	99.93	–	–	0.06	3.3579	0.3319	3.6898
121	1861	1	1	0	0	99.19	100	6.25	0.81	3.4061	0.3764	3.7825
123	1515	0	3	0	0	99.93	100	75	0.07	3.3571	0.3159	3.6731
200	1743	30	826	2	0	97.12	91.28	99.6	0.17	3.4980	0.5078	4.0058
202	2061	55	19	1	0	98.78	100	42.22	1.23	3.1592	0.4252	3.5844
210	2423	22	195	10	0	97.95	95.9	80.26	1.88	3.2601	0.5132	3.7733
212	2747	0	0	0	0	100	–	–	0	3.3646	0.5711	3.9357
213	2640	28	220	362	0	99.9	99.09	99.54	0.03	3.5973	0.6337	4.2310
214	2002	0	256	1	2	99.25	98.05	95.44	0.6	3.2300	0.4931	3.7232
219	2082	7	64	1	0	98.7	98.44	70	1.29	3.0130	0.4557	3.4688
221	2031	0	396	0	0	100	100	100	0	3.3542	0.4936	3.8478
222	2274	208	0	0	0	99.6	–	–	0.4	3.7511	0.5303	4.2814
228	1688	3	362	0	0	98.88	100	94.03	1.36	3.7520	0.4419	4.1940
231	1568	1	2	0	0	100	100	100	0	3.2480	0.3362	3.5843
232	398	1382	0	0	0	99.83	–	–	0.17	3.9879	0.3832	4.3712
233	2229	7	830	11	0	99.74	100	99.05	0.36	3.2988	0.6131	3.9120
234	2700	50	3	0	0	99.93	100	60	0.07	3.5279	0.5422	4.0701
Total	44247	1836	3220	388	7	98.77	96.12	86.48	1.04	74.7493	10.2250	84.9744

Table 3.9: Confusion matrix on DS2 data set.

		Algorithm label		
		n	v	Total
Reference label	N	43812	435	44247
	S	1787	49	1836
	V	125	3095	3220
	F	269	119	388
	Q	0	7	7
	Total	45993	3705	49698

Table 3.10: Comparison between this study and published studies.

Method	recordings	Measures			
		Acc (%)	Se (%)	+P (%)	FPR (%)
Shyu [13]	111,115,116,119,221,230,231	97.04	99.02	—	—
Proposed method		99.35	100	90.94	0.67
Inan [14]	100,101,103,105,106,109,112,113,114,115,116,118 119,121,122, 123,200,201,202,203,205,208,210,212	—	76.5	85.8	—
Proposed method		98.88	95.62	90.87	0.83
Lim [15]	115,116,119,221,230,231	99.8	99.2	—	—
Proposed method		99.74	100	96.65	0.28
Li, Peng [16]	All recordings of DS2	98.2	93.1	81.4	—
Proposed method		98.77	96.12	86.48	1.04

study based on the morphological differences of the PVC beats in the ventricular depolarisation phase (QRS complex) and repolarisation phase (mainly T-wave). The classifier was tested on the testing dataset (DS2). The performance of the classifier on these recordings was: accuracy 98.2%, sensitivity 93.1% and positive predictivity 81.4%. Using the same dataset (DS2) in common with this study, the proposed classifier achieved

better accuracy (98.77%) , sensitivity (96.12%) and positive predictivity (86.48%).

A limitation of the current study is that the proposed method is only able to classify PVC beats while there are a number of published works focusing on multi-class heartbeat classification [218, 83, 224, 103, 225, 226]. Although PVC detection is a simpler problem as compared to multi-class heartbeat classification, there is still need for further research to develop methods that can accurately identify PVC beats in online manner. Thus, the focus of this chapter is to develop a low-complexity PVC detection method with capability to detect PVC beats in online manner. As the proposed method focuses on PVC detection only, the experimental results of the proposed method was not compared with exciting studies focusing on multi-class heartbeat classification.

3.6 Conclusions

In this chapter, an online PVC detection method based on the variation of principal directions in “Replacing” strategy is proposed. This study targeted at one specific anomaly named PVC in ECG data stream. From the experimental results, it is demonstrated that the variation of principal directions caused by PVC beats can be used to accurately identify PVCs. An effective computation for estimating principal directions in “Replacing strategy is also proposed for reducing the computational complexity of the proposed method. 22 ECG recordings from the MIT-BIH arrhythmia database were used to train the PVC classifier and another 22 non-overlapping recordings from the same database were used to evaluate the performance of PVC classifier. The PVC detection accuracy was 98.77%, with the sensitivity and positive predictivity of 96.12% and 86.48%, respectively. The experimental results showed that the proposed method outperforms the existing methods in terms of accuracy sensitivity, positive predictivity and false positive rate. In addition, the proposed method is effective in term of computation time.

CHAPTER 4

EXPLORING DOUGLAS-PEUCKER ALGORITHM IN THE DETECTION OF EPILEPTIC SEIZURE FROM MULTI-CATEGORY EEG SIGNALS

Discovering the concealed patterns of EEG signals is a crucial part in efficient detection of epileptic seizure. The purpose of this chapter is to develop a novel feature extraction technique for determining discriminating features from EEG signals for the epileptic seizure classification. This chapter develops a new scheme based on the DP and the PCA for the extraction of representative and discriminatory information from epileptic EEG data. As the multi-channel EEG signals are highly correlated and are in large volumes, the DP algorithm is applied to extract the most representative samples from a large number of EEG data. The PCA is utilised to produce uncorrelated variables and to reduce the dimensionality of the DP samples for better recognition. To verify the robustness of the proposed method, four machine learning techniques: random forest classifier (RF), k -nearest neighbour algorithm (k -NN), support vector machine (SVM), and decision tree classifier (DT) are employed on the obtained features. Furthermore, the performance of the proposed method is accessed by comparing it with some recently reported algorithms. The major finding of this research is that the DP algorithm is very effective for getting representative samples from EEG data in the detection of epileptic seizure and the RF classifier is the best technique for the DP and PCA based features for efficient detection of the epileptic EEG signal.

4.1 Introduction

Epilepsy is one of the most common neurological disorders of the human brain that affects approximately 65 million people of the world [10]. It is characterised by unpro-

voked recurring seizures which are induced by abnormal and synchronous discharges of a group of neurons in the brain [227]. Transient and unexpected electrical disturbances of the brain are recognised as the possible causatives for epileptic seizures. In the majority of cases, seizures occur unexpectedly, without a sign of warning to alert and prepare the person for an onset of a seizure. Such abrupt and uncontrollable nature of the disease can cause physical injury due to loss of motor control, loss of consciousness, or delayed reactivity during seizures. Impairment of consciousness can be life-threatening, especially if they occur while the person is driving, swimming, climbing heights, or alone. EEG is most commonly used technique for diagnosis of epileptic seizure in the medical community [228]. EEG record electrical activity along the scalp, via the placement on the scalp of multiple electrodes; it measures voltage fluctuations resulting from ionic current flows within the brain [229, 230]. Epileptic activity can create clear abnormalities on a standard EEG and leaves its signature on it [149]. Epileptic seizure activities in the brain commonly manifest spikes or spike wave complexes in EEG signals which are usually analysed visually by expert or neurologists [231, 232]. However, the visual scanning of EEG signal is very time-consuming and costly; it may be inaccurate, very complex, subject to judgement and human error [151] as EEG signals contain a huge amount of data (in size and dimension). Therefore, there is an increasing need for developing automated epileptic seizure detection algorithms not only to alleviate the neurologist's burden of analysing long-term EEG signals but also to ensure a proper diagnosis and evaluation of neurological diseases.

In past two decades, several EEG signal processing techniques have been developed for automated epileptic seizure detection based on various feature extraction and classification techniques. The key challenge of any detection method is the extraction of the distinguishing features from EEG signals as it significantly affects the performance of the classifier. Representative characteristics or features extracted from EEG data can

describe the key properties or morphologies of the signals for perfect detection of epileptic seizure [106]. As feature extraction is the most important part of detection process which plays key role in the performance of a classifier, this study aims to develop a new efficient feature extraction technique for the classification of epileptic seizure from EEG signals.

Several feature extraction methods have been applied in epileptic seizure detection, such as correlation [233], linear prediction error energy [234], fast Fourier transform (FFT) [113], wavelet transform [117, 235, 236], empirical mode decomposition (EMD) [6, 237], Lyapunov exponent [135], Correlation dimension [118], approximate entropy (ApEn) [139, 238], clustering technique [239], Sampling technique [151, 240], Complex network [241, 230, 242], and Optimum allocation [149, 243]. These feature extraction techniques can be grouped into four categories [107] namely, time-domain [233, 234], frequency-domain [113], time-frequency domain [117, 235, 6, 237] and nonlinear methods [135, 118, 139]. Once features are extracted from EEG signals, a classifier is employed to differentiate between normal and epileptic EEG. Many classification methods have been proposed for seizure detection such as various types of artificial neural networks (ANNs) [141, 244, 144], SVM [233, 245, 246, 229], DT [247], k -NN [111], and RF [20].

Due to complex characteristics of EEG signals (e.g. non-stationary, aperiodic; poor signal-to-noise ratio,), sometimes it is very hard to achieve reasonable performance in the detection of epileptic seizure. Some of the existing feature extraction methods are not a good choice for obtaining characteristic features from non-stationary epileptic EEG data (e.g., Fourier transformation) [113, 248], and therefore most of their performances are limited regarding success rate and effectiveness [18, 4]. Moreover, the majority of the existing methods cannot appropriately handle large EEG data. Although

most of the EEG recordings are multi-categories in a real clinical application, most of the current methods are applied for binary EEG classification problems (Normal signal vs. ictal signal) [249, 138, 145, 144, 143, 250] and only a few methods focus on multi-class EEG classification [17, 18, 4, 19, 20, 21]. Considering these issues, this chapter proposes a new feature extraction technique based on DP and PCA for classification of multi-class EEG signals.

The DP [251] is the most well-known line simplification algorithm which is widely used in cartographic and computer graphic applications to reduce the complexity and storage requirements of curves by removing curve's no-characteristic points and extracting characteristic points [252, 253, 254, 255]. It is also applied in biomedical applications such as ECG signals compression [256, 257, 258]. The main theme of this algorithm is to shorten a line by detecting and preserving the most significant points of a line while neglecting less important points. Although the DP technique has a high capability to represent the original patterns of time series data and reduce the size of data, it has not been considered before for epileptic detection in the EEG signal analysis to the best of author's knowledge. Thus this study introduces the idea of using the DP methods for extracting representative sampling points from huge amount of raw EEG data for the first time.

The main aim of this research is to develop a novel feature extraction technique for detection of epileptic seizure from multi-category EEG signal for properly handling big size EEG data. Moreover, this chapter investigates the effectiveness of DP algorithm in the detection of epileptic seizure from EEG data and also discovers an effective classifier for the proposed features. In the proposed methodology, first the non-stationary epileptic EEG signals are partitioned into some non-overlapping segments (called *Segm*) to make them stationary (discussed in detail in Section 4.3.1). Then the DP algorithm

is effectively employed to extract representative sampling points from each *Segm* and also to reduce the size of each *Segm* by removing redundant points. At the next stage, the PCA is used to reduce the dimensionality of DP data and also to produce uncorrelated variables which are considered as features, denoted as DP_PCA feature set. In order to select an efficient classifier for DP_PCA feature set, this study employs four popular machine learning techniques namely, RF , k -NN, SVM, and DT classifier on the extracted features. To evaluate the consistency and performance of the proposed methods, ten-fold cross validation is applied to create training and testing set. The performance of each method is evaluated by sensitivity (Se), overall classification accuracy (OCA), false positive rate (FPR), kappa statistic, and receiver operating characteristic (ROC) curve area. In order to further evaluate the performances, the proposed method is compared with other six existing algorithms. The experiment results show that the RF classifier is the best classifier for DP_PCA feature set compared to other three classifiers. The results also indicate that the proposed method outperforms the existing methods [17, 18, 4, 19, 20, 21] regarding sensitivity, specificity, and overall classification accuracy (OCA).

The rest of the chapter is organised as follows: in Section 4.2, the prior studies in multi-class EEG signal classifications are described. Section 4.3 presents the methodology of the proposed method. Section 4.4 describes the experimental data and implementation. Section 4.5 discusses the experimental results and discussions. Finally, Section 4.6 draws the conclusion for this chapter.

4.2 Previous Work

In the last decade, various methods have been proposed for the classification of EEG signals [10, 227, 259, 260, 231, 232, 261, 111, 107, 233, 234, 113, 117, 235, 139]. However, only a few approaches have dealt with multi-class EEG classification problems [17, 18, 4, 19, 20, 21]. For comparative reasons, the most recent and relevant studies dealing with multi-class EEG classification problems on a benchmark epileptic EEG dataset [262, 263] are reviewed.

Most recently, Emigdio et al. [20] developed a method based on Holderian regularity and the Matching Pursuit (MP) algorithm for feature extraction in the epileptic EEG signal classification. The feature sets were constructed by combining features extracted from EEG signals through regularity analysis, the MP algorithm and simple time-domain statistical analysis. These feature sets were then fed to a Random Forests classifier for classification of epileptic states. The performance of the method was tested on the Bonn data set [262, 263] considering different classification problems (binary classification problems and multi-class classification problems). The results showed that the overall classification accuracy was 97.6% for the five-class classification problem.

Murugavel and Ramakrishnan [21] introduced an approach based on a hierarchical multi-class SVM (H-MSVM) with extreme learning machine (ELM) as the kernel for the classification of epileptic EEG signals. The wavelet transform was used to decomposed the EEG data into six sub-bands and then six features such as largest Lyapunov exponent, statistical values, and approximate entropy were extracted from each sub-band. The extracted features were employed as the input to the classifier. The artificial neural network (ANN) and multi-class SVM were also utilised to identify the five-category EEG signals. The experimental results showed that the H-MSVM classifier with ELM

kernel yielded a better performance regarding classification accuracy and computation complexity compared to the ANN and SVM classifiers. The H-MSVM achieved an overall classification accuracy of 94%.

Ubeyli [19] reported a method based on Lyapunov exponents and a probabilistic neural network (PNN) classifier for classification of EEG signals. The Lyapunov exponents were obtained from each EEG signal using Jacobi-based algorithms and considered as feature vectors. The statistic over the Lyapunov exponents was used to reduce the dimensionality of the extracted feature vectors. The selected features were fed to the PNN and multilayer perceptron neural network (MLPNN) classifiers. The classification results show that the PNN with Lyapunov exponents features achieved an overall classification accuracy of 98.05% while the MLPNN produced a 92.20% accuracy rate.

Ubeyli [17] presented automated diagnostic systems combined with spectral analysis techniques for classification of EEG signals. Eigenvector methods were used to calculate the wavelet coefficients and power spectral density (PSD) values which considered as features. The selected features then were fed to seven classification algorithms: SVM, PNN, mixture of experts (ME), modified mixture of experts (MME), recurrent neural networks (RNN), MLPNN, and combined neural networks (CNN). The experimental results showed that the SVM and MME classifiers achieved better performance compared to other five classifiers. The performances of the SVM, MME, PNN, ME, RNN, CNN and MLPNN were 99.20%, 98.68%, 95.30%, 95%, 94.85%, 93.48%, and 90.48%, respectively.

Ubeyli [18] developed a method based on multi-class SVMs with the error correcting output codes (ECOC) and eigenvector methods for the classification of EEG signals. The PSD values of the EEG signals were obtained using three different eigenvector methods such as the MUSIC, Pisarenko and minimum-norm. The statistics over the

set of the power levels of the PSDs were considered as features and fed to the multi-class SVMs. The MLPNN classifier was also applied to the same feature set. The total classification accuracy obtained by SVM with the ECOC and the MLPNN were 99.30% and 92.90%, respectively.

Guler and Ubeyli [4] proposed the multi-class SVM with the ECOC for the classification of multi-class EEG signals. They also tested the probabilistic neural network (PNN) and multilayer perceptron neural network (MLPNN) classifiers on the same epileptic EEG data. The wavelet coefficients and Lyapunov exponents were used to extract features from the EEG data. The extracted features were employed as the input of the three classifiers. The results showed that the multi-class SVM classifier achieved better performance than the other two classifiers. The total classification accuracies of the SVM, PNN, and MLPNN were 99.28%, 98.05%, and 93.63%, respectively.

4.3 Proposed Approach

The chapter introduces a novel method based on DP in the multi-class EEG signal classification. In this study, the DP approach is developed to select representative samples from the original EEG signals that reflect an entire database. Next, The PCA is used to reduce the dimension of the obtained DP sample set which is considered as a feature set. Finally, the extracted features are tested by four machine learning methods, including RF, k -NN, SVM, and DT classifiers. As shown in Figure 4.1, the entire process of proposed method is divided into five major parts: data segmentation, Douglas-Peucker algorithm, dimension reduction by PCA, DP_PCA feature set, and the classification part by the RF, k -NN, DT, and SVM. A detailed description of these five parts is provided in the following sections.

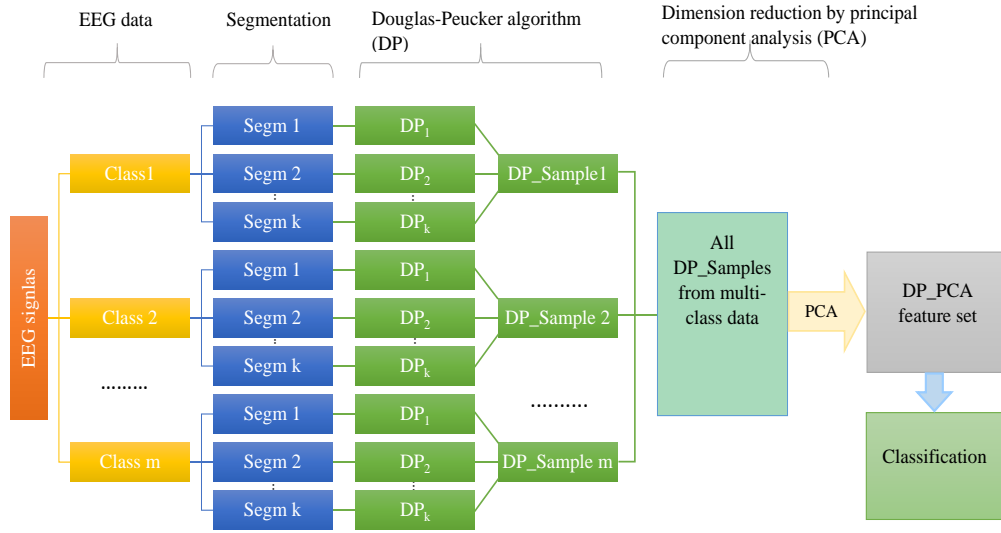


Figure 4.1: Block diagram of the proposed method for the classification of epileptic EEG signals.

4.3.1 Data Segmentation

Most of the EEG signal processing methods require stationarity of the signals. Although EEG signal may not be stationary, usually smaller windows or parts of those signals will exhibit stationarity [149]. An EEG signal is stationary for a small amount of time. That is the reason the recorded EEG signals of every class are split into several non-overlapping segments based on a particular time period to properly account for possible stationarities. Hence the EEG signals of each class are segmented into some fixed-size non-overlapping time windows (called ‘*Segm*’) to obtain representative values of a specific time period. Each *Segm* consists of EEG channel data within a time window. Figure 4.2 illustrates an example of determining the segments *Segms* in an EEG signal of a class. It is worthwhile to mention that the number of *Segms* (k) is determined empirically over time for any experiment design.

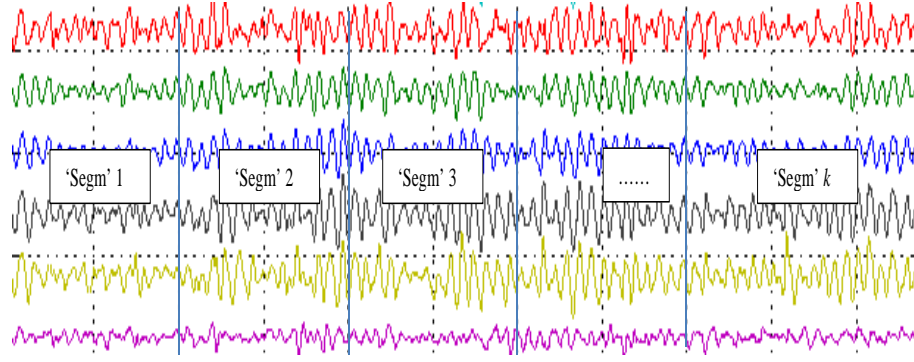


Figure 4.2: An example of determining *Segms* from an EEG signals of a class.

4.3.2 Douglas-Peucker Algorithm

The DP [251] is one of the most popular methods for line (trajectory) simplification. The algorithm simplifies a line by detecting and preserving the most significant points of a line while neglecting less important points. In this study, the DP technique is used to extract the representative samples from different '*Segms*'. Let the data series (trajectory) S be described by the set of N points $\langle P_1, P_2, \dots, P_N \rangle$. The main idea of DP algorithm is to determine a new data series with fewer and most significant points without deviating from the original data series by at most a simplification tolerance ϵ . As an initial step of DP, the algorithm approximates the data series S with a line segment $\overline{P_1 P_N}$ constructed from the first to the last data point. Then it calculates the perpendicular Euclidean distance between each intermediate data point and the line segment $\overline{P_1 P_N}$ and retains the point P_i which has the maximum distance D_{max} . The algorithm compares D_{max} with the given simplification tolerance ϵ . If the maximum distance D_{max} is less than the simplification tolerance ϵ , the algorithm removes all intermediate points in data series. Otherwise, it uses data point P_i to split the data series to two subseries $\langle P_1, P_2, \dots, P_i \rangle$ and $\langle P_i, P_{i+1}, \dots, P_N \rangle$ and recursively repeats the procedure for each subseries. The DP algorithm terminates when the D_{max} in a subseries is lower than

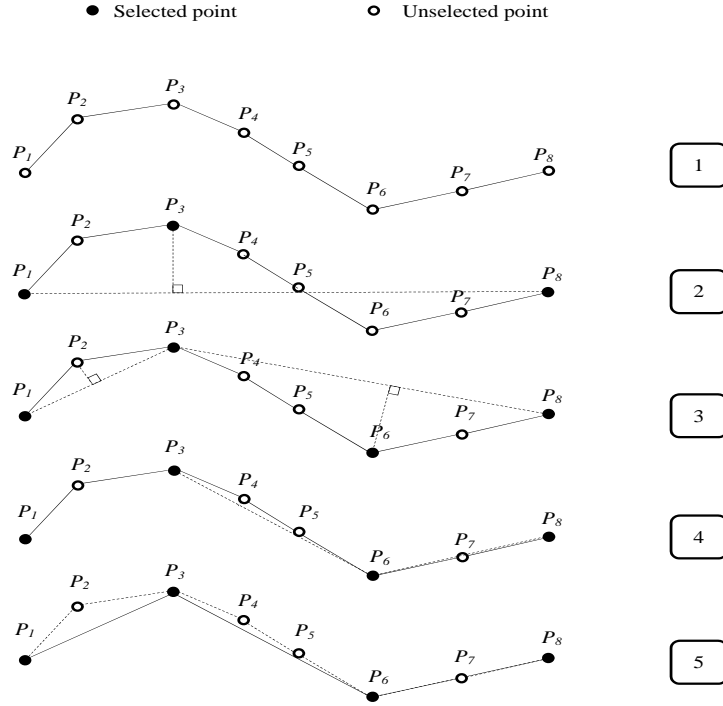


Figure 4.3: An example of DP sample point extraction.

the simplification tolerance ϵ or the subseries contains only two data points. Figure 4.3 illustrates an example of DP sample point extraction. The original data series contain eight points (P_1P_8). The distances from the points P_1P_8 to the line segment $\overline{P_1P_8}$ are first computed. Since the maximum distance D_{max} at point P_3 exceeds the given simplification tolerance ϵ , the data series are divided at this point into two subseries (step 2 in Figure 4.3). In the left subseries, the distance from P_2 to the line segment $\overline{P_1P_3}$ is lower than ϵ value, so the point P_2 is ignored. In the right subseries, the distance from the point P_6 to the line segment $\overline{P_3P_8}$ also exceeds the simplification tolerance ϵ , hence a new split is performed at the point P_6 and the process is repeated for each part, respectively. Figure 4.3 shows that, the original data series having 8 points finally becomes a 4-points data series after this process.

The value of simplification tolerance ϵ determines the degree of simplification.

Therefore, it is an important task in DP algorithm to determine the most significant ϵ value. Choosing a small value of ϵ will produce a minimally simplified data series (i.e. only a few redundant data points will be removed from data series) while selecting a large one will provide a highly simplified data series that might lead to losing some of the significant points from the data series. The following formula 4.1 is used to calculate the most significant ϵ value for each data series.

$$\epsilon = \frac{T}{100} \times \left(\sum_{i=1}^{N-1} D(P_{i+1}, P_i) \right) \quad (4.1)$$

where N is the number of points in data series; $D(P_{i+1}, P_i)$ is the Euclidean distance between two points p_{i+1} and p_i ; $\sum_{i=1}^{N-1} D(P_{i+1}, P_i)$ is the overall distance of the data series; and T is a real number which is determined empirically. The value of ϵ is changed as different percentages of the overall distance of the data series by setting different values for T .

As shown in Figure 4.1, the DP process consists of the following steps to extract the representative samples from various segms.

Step 1: Consider all the channels of the EEG data of a class.

Step 2: The EEG data of that class is split into k *Segms* considering a particular time period. Suppose the sizes of the *Segms* are N_1, N_2, \dots, N_k , respectively.

Step 3: The overall distance of each *Segm* is calculated. Then by setting the T value, the value of simplification tolerance ϵ for each *Segm* is calculated using Eq. 4.1.

Step 4: The representative samples from each *Segm* is extracted using DP algorithm.

Let n_1, n_2, \dots, n_k , be the sizes of samples obtained from the *Segms* whose sizes are

N_1, N_2, \dots, N_k , respectively. The representative samples selected from each *Segm* in a class make a vector set denoted as DP_Sample as shown in Figure 4.1.

Step 5: The vector sets of all classes construct a matrix (denoted as DP_samples set) that is used as input to the PCA, as discussed in the next section.

4.3.3 Dimension Reduction by PCA

PCA is a well-known statistical method for feature extraction and dimensionality reduction [182, 264]. It uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a smaller set of uncorrelated variables called principal components (PC). These components represent the most important linear characteristics of the data. The multi-channel EEG signals recorded from different scalp sites are highly correlated. They contain a large amount of redundant information. Therefore, it would be useful to remove this redundant information by converting the EEG signals into a set of new linearly uncorrelated variables (i.e., the PC space) and utilise these new variables as features for better classification of EEG signals. In this chapter, the PCA is used to reduce the dimensionality of the DP_Samples set and also to obtain EEG features for classification of epileptic EEG signals.

Let $DP_samples\ set = [x_1^T; x_2^T; \dots; x_n^T] \in R^{n \times p}$ where each row x_i represents a data point in a p dimensional space (considering a p -channel EEG signals as a p -dimensional space), and n is the number of the points selected by DP. PCA can be formulated as the following optimisation problem:

$$\max_{\mathbf{U} \in R^{p \times k}, \|\mathbf{U}\|=I} \sum_{i=1}^n \mathbf{U}^T (x_i - \mu) (x_i - \mu)^T \mathbf{U} \quad (4.2)$$

where \mathbf{U} is a matrix consisting of q dominant eigenvectors. This problem can be solved by deriving an eigenvalue decomposition problem of the covariance matrix.

$$\hat{\mathbf{C}}_{\mathbf{x}}\mathbf{U} = \mathbf{U}\Lambda \quad (4.3)$$

where

$$\hat{\mathbf{C}}_{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T \quad (4.4)$$

is the covariance matrix, μ is the global mean defined as $\mu = \frac{1}{n} \sum_{i=1}^n x_i$, Λ_i ($i = 1, 2, \dots, p$) are the eigenvalues and they are sorted in descending order, \mathbf{U}_i ($i=1,2,p$) are the corresponding eigenvectors. In order to reduce the dimensionality of the DP_Samples set, only the first q eigenvector ($q \leq p$) which correspond to the q largest eigenvalues are selected by the following Eq to represent the DP_Samples set.

$$\frac{\sum_{i=1}^q \Lambda_i}{\sum_{i=1}^p \Lambda_i} \geq \sigma \quad (4.5)$$

For a given precision parameter σ (considering $\sigma=90\%$ in this study), the matrix \mathbf{U} consisting of q dominant eigenvectors is constructed and the q -dimensional feature set denoted as DP_PCA feature set is computed as follows:

$$DP_PCA = (DP_Sampleset)\mathbf{U} \quad (4.6)$$

4.3.4 DP_PCA Feature Set

The new feature set denoted as DP_PCA feature set is generated by reducing the dimensionality of the DP_Samples set using PCA method as discussed in Section 4.3.3. This feature vector set is divided into a training set and a testing set using a ten-fold cross validation method, which is discussed in Section 4.4.2. As shown in Figure 4.1, this feature set is fed to each of the four classifiers discussed in the next section.

4.3.5 Classification by the RF, k -NN, SVM, and DT

This study considers four classifiers: RF, k -NN, DT, and SVM for testing the performance of the proposed feature extraction method. A brief explanation of these classification methods is provided in the following sections.

4.3.5.1 Random Forest

The RF is an ensemble learning technique developed by Breiman [265]. It consists of many individual classification trees, where each tree is constructed using a tree classification by selecting a random subset of input features and a different bootstrap sample from the training data. The RF aggregates the results of all classification trees to classify new samples. Each tree casts a unit vote at the input data and then the forest selects the class with the most votes for the input data. Figure 4.4 illustrates the structure of random forest classifier.

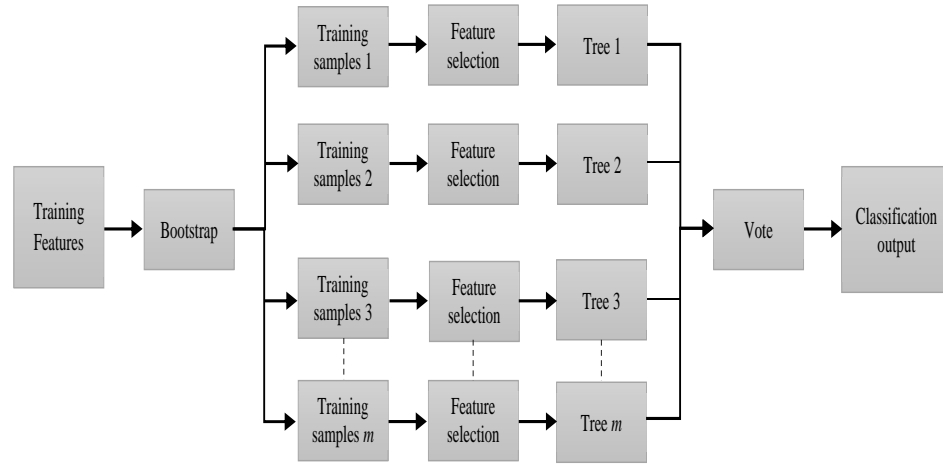


Figure 4.4: The structure of random forest classifier.

The RF algorithm proceeds as follows:

1. From the training data set, m training subsets are generated using the bootstrapping technique (randomly sampling with replacement). Each training subset has the same size as the training data set and contains approximately one-third of the samples of the training data set.
2. For each training subsets, a decision tree is built with the following criteria: at each node in building a decision tree, a random number of f features are selected from the F input features ($f \ll F$) and the best split (e.g., the largest Gini measure) among these f features is used to divide the node. The tree is grown to the maximum size with no pruning. The tree growing algorithm used in RF is random trees.
3. The m trees are combined into an RF ensemble and use a majority voting scheme to predict the class of new data by evaluating votes from each tree.

4.3.5.2 K-Nearest Neighbour Algorithm

The k -NN is a supervised learning algorithm for classifying objects based on closest training observations in the feature space [182, 266]. Although the k -NN is the simplest algorithm among all machine learning algorithms, it can still yield high performance, without priori assumptions about the distributions from which the training samples are drawn [182]. Given a query vector x_0 and a set of N labelled instances $\{x_i, y_i\}_1^N$, the aim of the classifier is to identify the class label of x_0 on the predefined P classes. The k -NN classification algorithm tries to find the K -nearest neighbors of x_0 , and uses a majority vote to determine the class label of x_0 . Without prior knowledge, the k -NN classifier usually applies Euclidean distances as the distance metric [267]. A detailed discussion of this method can be found in [182, 266, 268].

4.3.5.3 Support Vector Machine

The SVM is a machine learning algorithm based on statistical learning theory and structural risk minimisation principle presented by Vapnik [269]. The main idea of SVM is to map the input data into a higher dimensional space and then determines an optimal separating hyperplane between the two classes of data in the transformed space [270, 271]. For nonlinear classifier models, when the data are not linearly separable, SVMs map inseparable input data into a high-dimensional space by constructing a linear kernel function to make the input data linearly separable in new space and allows better fitting of the hyperplane to the input dataset. Although the SVM is originally designed as a two-class classifier, some methods have been proposed to extend the application of SVM to multi-class classifications. One common used procedure in practice is to employ a set of pairwise classifiers, based on one-against-one decomposition [271]. The decision function of binary SVM classifier can be expressed as follows:

$$f(x) = \text{sgn} \left(\sum_{i=1}^s y_i \alpha_i k(x_i, x) + b \right); 0 < \alpha_i < C \quad (4.7)$$

where sgn is the signum function, $k(x_i, x)$ is kernel function, and b is the bias of the training samples. There are several kernel functions such as Linear kernel, polynomial kernel, RBF kernel and sigmoid kernel. In this chapter, the polynomial kernel is considered as the best kernel function for identifying multi-categories EEG signals as it was found to give the best classification performance.

The regularisation parameter C is used to control the trade-off between training error and model complexity and can be calculated as follows:

$$C = N / \sum_{i=1}^N K(x_i, x), \text{ where } N \text{ is the size of the training set.} \quad (4.8)$$

In the multi-class classification problem, the SVMs work by using a collection of decision functions f_{kl} . The class decision can be obtained by the following formula [271]:

$$f_k(x) = \sum_{i=1}^n \text{sgn}(f_{kl}(x)) \quad (4.9)$$

where kl indicates each pair of classes selected from separated target classes and n is the number of separated target classes. The algorithm proceeds as follows: it assigns a label to the class: $\arg \max f_k(x)$, ($k = 1, 2, \dots, n$). The pairwise classification then converts the n -class classification problem into $n(n-1)/2$ two class problems which cover all pairs of classes. An overview of SVM classifier can be found in [269, 270, 271].

4.3.5.4 J48 Decision Tree

J48 decision tree is an implementation of the C4.5 algorithm [272] in the WEKA (The Waikato Environment for Knowledge Analysis) [273, 274]. C4.5 is an extension of the ID3 algorithm. It uses the top-down construction technique to recursively split the data set into smaller subsets based on the value of an attribute [272, 275]. This classifier builds a decision tree for the given dataset using the concept of information entropy. In a decision tree, each attribute can be used to make a decision by splitting the data into smaller subsets. At each node of the tree, the algorithm evaluates each attribute of the data for dividing the data into smaller subsets and chooses the attribute that gives the highest information gain. Once an attribute is selected, the data set is split into subsets, and the splitting process is repeated for each subset until further splitting is not gainful. In the resulting tree structure; each inner node in the tree corresponds to one of the input attributes, each branch represents a value or range of values of that attribute, and each leaf accounts for a classification.

4.4 Data and Implementation

4.4.1 The Epileptic EEG Data

The epileptic EEG data used in this work is obtained from publicly available EEG database of Department of Epileptology, University of Bonn, Germany [262, 263]. The whole database contains five subsets denoted as Sets A, B, C, D, and E. Each subset is containing 100 single-channel EEG signals with a duration of 23.6 s. The subsets A and B are recorded extracranially, whereas subsets C, D, and E are recorded intracranially. Set A and Set B were collected from five healthy volunteers with eyes open and eyes

closed, respectively. Sets C and D were collected from five epileptic patients during interictal periods. Set C was recorded from the hippocampal formation on the opposite side of the epileptogenic zone while set D was recorded from the epileptogenic zone. Set E was collected from all of the recording zones in Sets C and D during seizure activity (ictal periods). All EEG recordings were recorded using a 128-channel amplifier system with a sampling rate of 173.61 Hz and 12-bit A/D resolution. Signals were filtered using a 0.5340 Hz (12 dB/octave) band pass filter, and artifacts such as muscle and eye movements were removed by visual inspection. A summary description of the five set EEG data is provided in Table 4.1. Exemplary EEG time series from each of the five classes (Set A-Set E) are shown in Figure 4.5

Table 4.1: Summary of the epileptic EEG data.

	Set A	Set B	Set C	Set D	Set E
Subjects	Five healthy subjects	Five healthy subjects	Five epileptic subjects	Five epileptic subjects	Five epileptic subjects
Patient's state	Awake and eyes open (normal)	Awake and eyes closed (normal)	Seizure-free (interictal)	Seizure-free (interictal)	Seizure activity (ictal)
Electrode type	Surface	Surface	Intracranial	Intracranial	Intracranial
Electrode placement	Inernational 10-20 system	Inernational 10-20 system	Epileptogenic zone	Epileptogenic zone	Epileptogenic zone
Number of channels	100	100	100	100	100
Time duration (s)	23.6	23.6	23.6	23.6	23.6

4.4.2 Implementation

This section presents the implementation of the proposed method on the epileptic EEG data [262, 263]. As discussed in Section 4.4.1, the complete dataset contains five sets (denoted as A, B, C, D, and E), each containing 100 channels data of 23.6 s. Each channel consists of 4096 data samples. The implementation of the proposed method comprises five steps as follow:

1. Each class data is segmented into four segms ($K=4$), each containing the EEG data for 5.9 s. As each channel consists of 4096 data samples, the sizes of the four

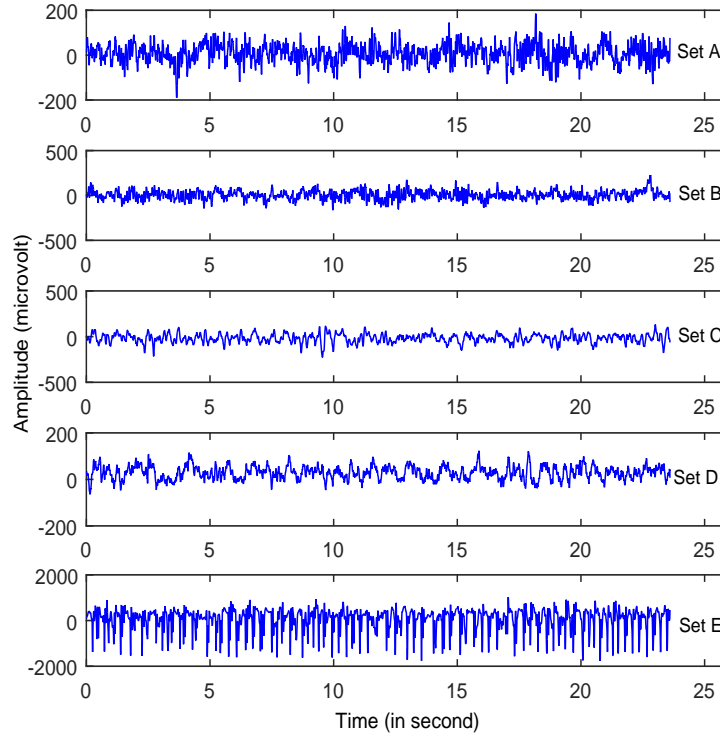


Figure 4.5: Example of five different sets of EEG signals taken from different subjects.

Segms are $N_1=1024$, $N_2=1024$, $N_3=1024$, $N_4=1025$, respectively.

2. To determine the value of simplification tolerance ϵ , the overall distance of each of the four *Segms* in each class are calculated, and then the value of T is changed from 0.01 to 0.1 with step size 0.01 in Eq. 4.1 to identify the most significant ϵ value for each *Segm*. From the experiment, it is considered that $T=0.06$, and then the value of ϵ for each *Segm* is calculated using Eq. 4.1. Table 4.2 presents the obtained value of ϵ for each *Segm* in each of the five classes. From Table 4.2, it is observed that the values of ϵ are not equal due to the differences in the overall distance of *Segms*.
3. Using the obtained value of ϵ shown in Table 4.2, The representative samples from each *Segm* are extracted using the DP algorithm. Figure 4.6 and 4.7 show

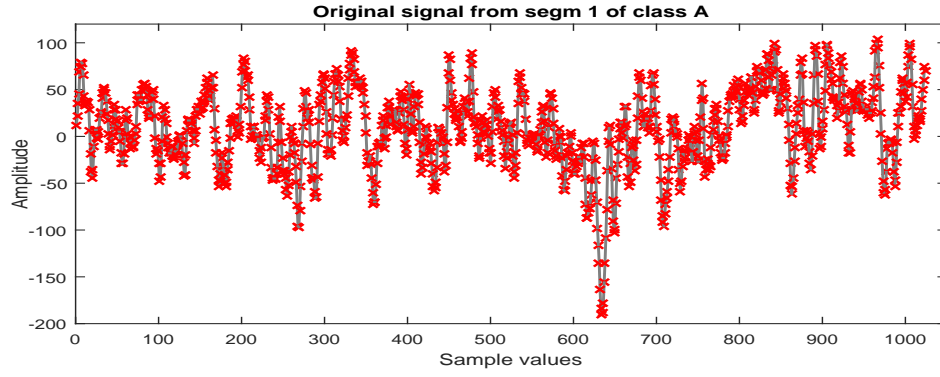
Table 4.2: The obtained value of ϵ for each *Segm* in each of the five classes

	Segm 1	Segm 2	Segm 3	Segm 4
Set A	96.80	84.45	88.31	97.38
Set B	137	139.87	179.87	167.87
Set C	66.37	99.77	62.75	68.71
Set D	52.14	75.96	63.53	54.37
Set E	908.46	821.18	747.53	727.49

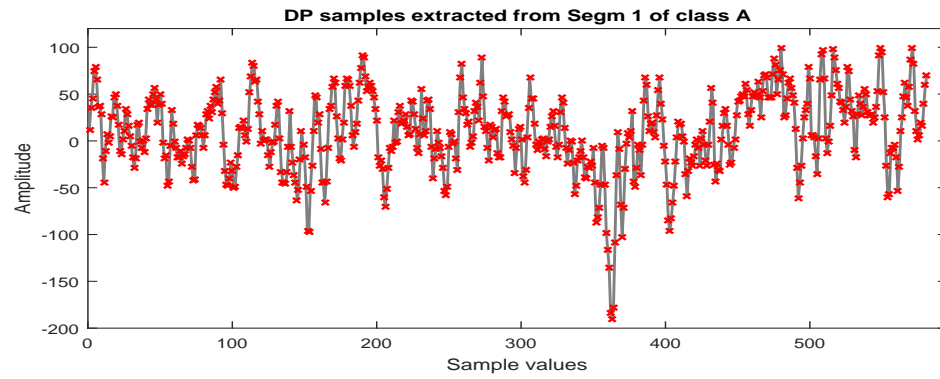
typical results of DP for the healthy subject (class A) and the epileptic patient (class E), respectively. In Figure 4.6 and 4.7, the first *Segm* of class A and class E are considered, respectively. It can be seen from Figure 4.6 and 4.7 that the DP samples can effectively represent the original signals with fewer points which indicate the ability of DP to select most significant points from each signal.

Table 4.3 provides the number of the representative samples chosen by DP for each *Segm* in each of the five classes. It can be seen from Table 4.3, the number of the representative samples for each *Segm* are not equal, e.g. in set A (Class 1), the total number of 581, 612, 599, and 562 samples are selected by DP from *Segm* 1, *Segm* 2, *Segm* 3, and *Segm* 4, respectively. The total number of the representative samples is 2354 for set A.

The representative samples selected from all *Segms* in a class create a vector set denoted as DP_Sample as shown in Figure 4.1. For example, the selected representative samples from each of the four *Segms* of class 1 create a DP_Sample 1 vector. The DP_Sample 1 is constructed as 581, 612, 599, 562 which contains all 2354 selected samples in class 1. The vector sets are created similarly: DP_Sample 2, DP_Sample 3, DP_Sample 4, DP_Sample 5 from class 2, class 3, class 4, and class 5, respectively. All DP_Samples from the five class EEG data



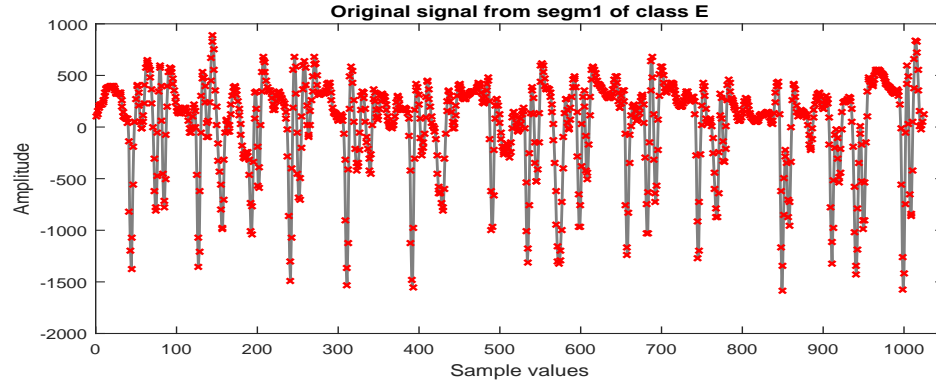
(a) Original signal from *Segm 1* of class A



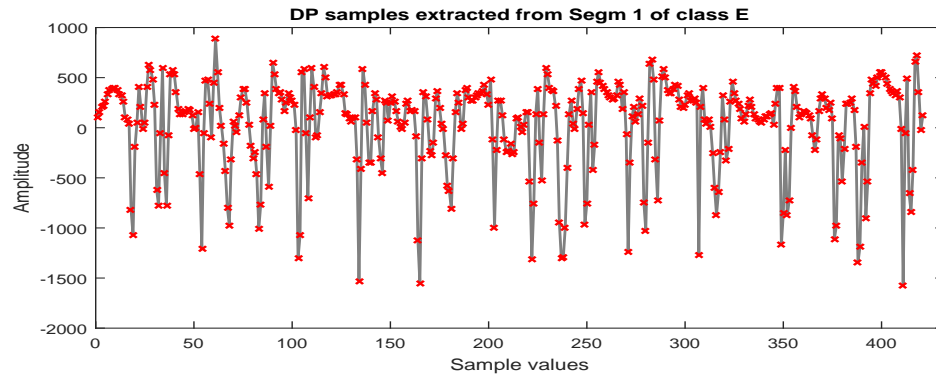
(b) DP samples extracted from *Segm 1* of class A

Figure 4.6: The typical results of DP for the healthy subject:

construct a matrix denoted as DP_sample set that is used as an input to the PCA. The DP_sample set contains all 10755 selected samples from five classes (2354 for class 1, 2008 for class 2, 2237 for class 3, 2398 for class 4, 1758 for class 5). It can be seen from Table 3, the DP reduces the data samples size of the five class from 20480 sample points to 10755 sample point (47.49% sample reduction). Here note that each sample has 100 dimensions as each class contains 100 channels of EEG data. Therefore, the DP_sample set consists of 10755 samples of 100 dimensions.



(a) Original signal from *Segm 1* of class E



(b) DP samples extracted from *Segm 1* of class E

Figure 4.7: The typical results of DP for the epileptic patient:

Table 4.3: The representative samples chosen by DP for each *Segm* in each of the five classes.

Sample size	Segm 1	Segm 2	Segm 3	Segm 4	Total
Set A	581 (1024)	612 (1024)	599 (1024)	562 (1025)	2354 (4096)
Set B	564 (1024)	583 (1024)	429 (1024)	432 (1025)	2008 (4096)
Set C	589 (1024)	408 (1024)	622 (1024)	618 (1025)	2237 (4096)
Set D	669 (1024)	489 (1024)	532 (1024)	708 (1025)	2398 (4096)
Set E	421 (1024)	438 (1024)	450 (1024)	449 (1025)	1758 (4096)

Note: In each cell, the number inside the parentheses is the *Segm* size (e.g. (1024) in the 1st cell) and the number outside of the parentheses is the sample size (e.g. 581 in the 1st cell) obtained by the DP.

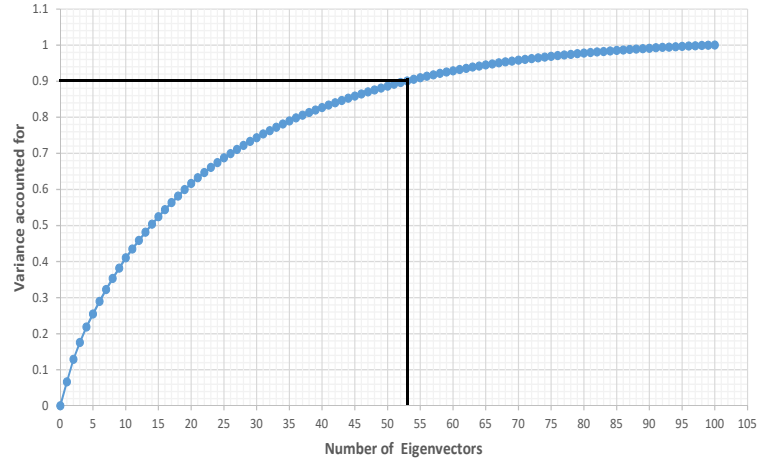


Figure 4.8: The cumulative eigenvalues for all 100 eigenvectors.

4. The PCA is applied to reduce the dimensionality of the DP_Sample set. Only the first q eigenvectors are selected to represent the DP_sample set based on the accumulation of their respective eigenvalues exceed 90% of total sum of eigenvalues (Eq4.5). Figure 4.8 illustrates the cumulative eigenvalues for all 100 eigenvectors. It is observed that the accumulation of the first 53 eigenvalues exceeds 90% of total sum of eigenvalues. Therefore, only the first 53 eigenvectors are considered for obtaining the DP_PCA feature set. The obtained DP_PCA feature set contains 10755 samples of 53 dimensions.
5. The DP_PCA feature set is divided into a training set and a testing set using a 10-fold cross validation method to evaluate the performances of the proposed methods. The DP_PCA feature set is split into ten mutually exclusive subsets (10-folds) of approximately equal sizes. Training and testing are performed ten times. Each time, one of the folds is used as a testing set and the remaining nine folds are combined into a set for training.

In this research, the performances of the proposed methods are evaluated based on different statistical measures, such as sensitivity (Se), overall classification accuracy

(OCA), false positive rate (FPR), kappa statistic, and receiver operating characteristic (ROC) curve area. Their formulas are given below:

$$Se(i) = \frac{\text{No. of true positive decisions in class } i}{\text{No. of actual positive cases in class } i} \times 100, i = 1, 2, 3, 4, 5 \quad (4.10)$$

$$OCA = \frac{\text{No. of correct decisions for all classes}}{\text{Total no. of cases for all classes}} \times 100 \quad (4.11)$$

$$FPR(i) = \frac{\text{No. of negative cases that are detected as positive cases in } i\text{th classes}}{\text{Total no. of negative cases}} \times 100 \quad (4.12)$$

$$Kappa = \frac{P_o - P_e}{1 - P_e} \quad (4.13)$$

where P_o denotes the overall observed agreement between the classifier and the true classes, and P_e represents the expected proportion of agreement. Besides, the Area Under the ROC Curve (AUC) is measured to compare the overall performance of the classifiers. The ROC curve is obtained by plotting the sensitivity versus false positive rates [19].

4.5 Results and Discussions

This section presents the experimental results of the proposed methods on the epileptic EEG datasets. First, the effectiveness of each of the four mentioned classifiers are evaluated on the DP_PCA feature set to select the most appropriate classifier as discussed

in section 4.5.1. Then a comparison between the proposed method and six existing methods is provided in section 4.5.2. All mathematical calculations are carried out in MATLAB R signal processing tool (version 7.11, R2010b). The classification executions for all four classifiers: RF, k -NN, SVM and J48 classifiers are executed in WEKA machine learning toolkit [273, 274]. The LIBSVM tools (version 3.2) [276] is used in WEKA for the SVM classification. It is worth mentioning that the default parameter values for each classifier in WEKA is used as there are no specific guidelines for selecting these parameters.

4.5.1 Classification Results for Each Classifier

As mentioned before, four machine learning methods: RF, k -NN, SVM and J48 classifiers are tested for detection of the multi-class EEG signals. Table 4.4 presents the classification results of all classifiers on the DP_PCA feature set. The performance results are given by averaging over the results of the 10-fold cross-validation test and expressed as the mean \pm standard deviation. As shown in Table 4.4, The RF classifier achieves the highest classification performance in terms of average classification accuracy which is 99.85%. The k -NN classifier stands at the second position and the SVM classifier achieves the third position with the average classification accuracy of 98.31% and 96.11%, respectively. The J48 classifier yields the lowest average classification accuracy among all tested methods. It can be seen from Table 4.4 that the RF classifier produces the best performance in terms of sensitivity among all classifiers and obtains the sensitivity rate of 99.79% for set A, 99.85% for set B, 99.96% for set C, 99.71% for set D, and 100% for set E. It is observed that both k -NN and SVM classifiers yield high sensitivity rates in sets A, B, C, and D, but they failed to correctly classify the epileptic patient during seizure activity class (Set E). Table 4.4 also shows that the standard de-

viation for every classifier is very low which indicates the consistency of the mentioned classifiers for the DP_PCA features set.

Table 4.4: Classification results on the epileptic EEG data.

Classifiers	<i>Se</i> (mean \pm standard deviation)					10-fold cross validation <i>OCA</i> (mean \pm standard deviation)
	Set A	Set B	Set C	Set D	Set E	
RF	99.79 \pm 0.36	99.85 \pm 0.24	99.96 \pm 0.14	99.71 \pm 0.34	100 \pm 0.00	99.85 \pm 0.13
<i>k</i> -NN	100 \pm 0.00	99.75 \pm 0.35	100 \pm 0.00	100 \pm 0.00	89.93 \pm 2.02	98.31 \pm 0.32
SVM	100 \pm 0.00	99.85 \pm 0.34	100 \pm 0.00	99.87 \pm 0.20	76.56 \pm 3.53	96.11 \pm 0.61
J48	95.80 \pm 1.54	91.68 \pm 1.31	93.52 \pm 1.75	94.83 \pm 1.33	95.39 \pm 1.43	94.27 \pm 0.78

To provide more detailed information about how the 10-fold cross-validation system produces the classification performance regarding sensitivity and accuracy in each of the ten folds for each of the four classifiers, the classification results in each of the ten folds are provided in Figures 4.9 and Figure 4.10. Figure 4.9 illustrates the classification results in terms of the sensitivity (*Se*) in each of the ten folds in each class (Set) for the reported classifiers and Figure 4.10 shows the classification results for all classifiers in terms of the overall classification accuracy (*OCA*) in each of the ten folds. The error bars in these figures represent the standard errors.

Figure 4.9 presents the patterns of the *Se* for each class. From Figure 4.9(a), little fluctuation is noted in the *Se* patterns among the ten folds in each of the five classes for the RF classifier. These results indicate the stability and robustness of the RF classifier. From Figure 4.9(b) and 4.9(c), it is seen that the *Se* patterns for Set A, Set B, Set C, Set D are almost similar but the patterns for Set E is different and dramatically lower than the other patterns for both *k*-NN and SVM classifiers. This indicates the weakness of the *k*-NN and SVM classifiers for detecting the epileptic signals during seizure activity class (Set E). Figure 4.9(d) shows that the J48 classifier produces similar *Se* patterns for all Sets. It can be seen from Figure 4.9 that the fluctuations in the *Se* patterns among the different folds are negligible in each class for all classifiers. These results demonstrate

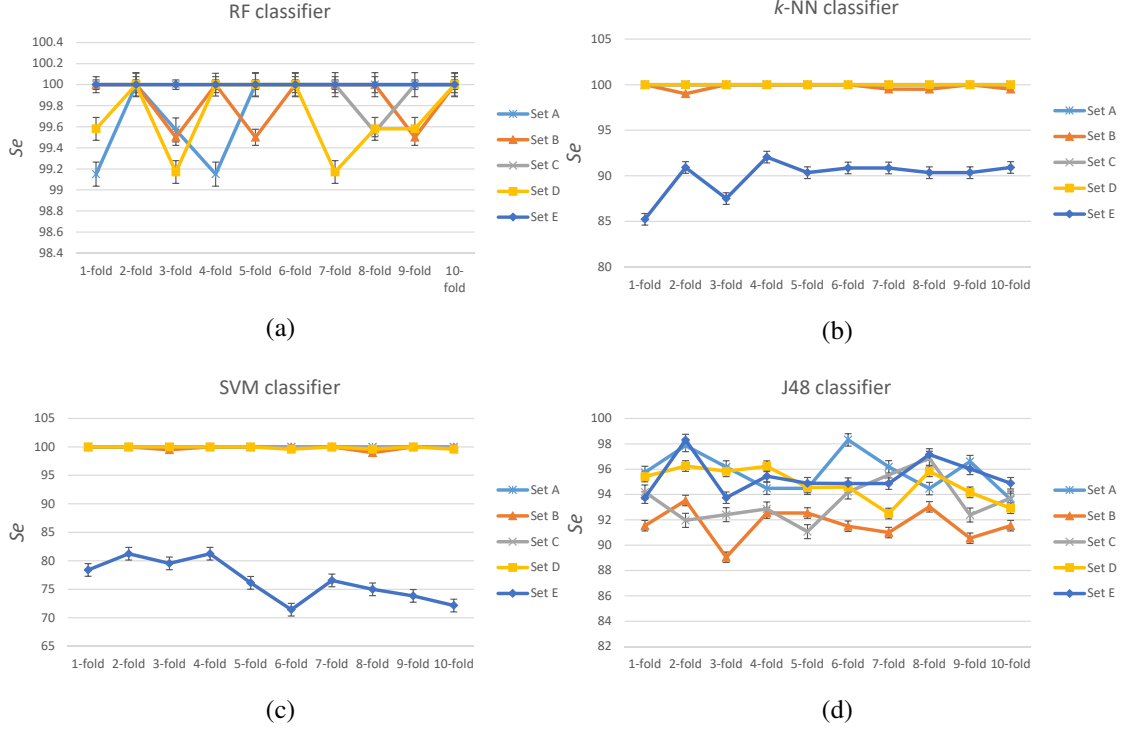


Figure 4.9: Individual classification performances of each of the ten folds in each class for the proposed classifiers: (a) RF, (b) k -NN, (c) SVM, (d) J48 .

the consistency of the classification methods.

Figure 4.10 shows the overall classification accuracies against each of the 10-folds for all classifiers. As can be seen from Figure 4.10, the RF classifier yields the best performance for each of the 10-folds compared to the k -NN, SVM and the J48 classifiers. It is observed that the fluctuations of the performance of the RF classifier are smaller among the different folds compared to other classifiers, indicating the stability of the RF classifier for the DP_PCA features Set. This figure also shows that the k -NN classifier produces a better performance than both SVM and J48 classifiers in each of the 10-folds. The lowest performance is obtained by the J48 classifier in each of the 10-folds.

Table 4.5 provides the false positive rate (FPR) for the four classifiers in each of the ten folds for Set A, Set B, Set C, Set D, and Set E. In Table 4.5, the overall results of

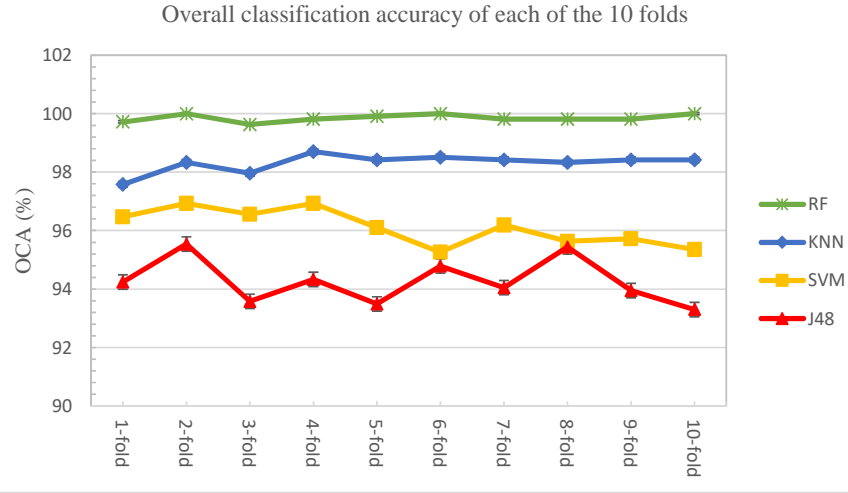


Figure 4.10: The overall classification accuracy (*OCA*) in each of the ten folds.

each class (Set) are also reported in terms of mean \pm standard deviation of the *FPR* over a 10-fold cross-validation. It is observed from Table 4.5 that the RF classifier produces an overall *FPR* of 0.04% for Set A, 0.03% for Set B, 0.11% for Set C, 0.01% for Set D, and 0.00% for Set E; while these values are 0.18%, 1.09%, 0.36%, 0.50%, 0.00%, respectively, for the *k*-NN classifier; and 1.06%, 1.63%, 1.14%, 1.07%, 0.00%, respectively, for the SVM classifier; and 1.18%, 1.81%, 2.20%, 1.75%, 0.29%, respectively, for the J48 classifier. The results show that in most cases the *FPRs* are zero in each of the folds in all classes for the RF classifier. It is also observed that the overall *FPRs* of the RF classifier are lower than those of the *k*-NN, the SVM, and the J48 classifiers in all classes.

Table 4.5: Obtained false positive rate (*FPR*) for each of the proposed classifiers.

Classifiers	Dataset	<i>FPR</i> (%)				
		Set A	Set B	Set C	Set D	Set E
RF	1-fold	0.00	0.11	0.23	0.00	0.00
	2-fold	0.00	0.00	0.00	0.00	0.00
	3-fold	0.00	0.23	0.12	0.12	0.00
	4-fold	0.00	0.00	0.23	0.00	0.00
	5-fold	0.00	0.00	0.12	0.00	0.00
	6-fold	0.00	0.00	0.00	0.00	0.00
	7-fold	0.12	0.00	0.12	0.00	0.00
	8-fold	0.24	0.00	0.00	0.00	0.00
	9-fold	0.00	0.00	0.23	0.00	0.00
	10-fold	0.00	0.00	0.00	0.00	0.00
	Overall	0.04±0.08	0.03±0.08	0.11±0.10	0.01±0.04	0.00±0.00
<i>k</i> -NN	1-fold	0.00	1.49	0.35	1.20	0.00
	2-fold	0.24	0.80	0.23	0.84	0.00
	3-fold	0.12	1.26	0.59	0.60	0.00
	4-fold	0.12	1.26	0.23	0.00	0.00
	5-fold	0.12	1.49	0.35	0.00	0.00
	6-fold	0.24	0.69	0.47	0.48	0.00
	7-fold	0.48	0.57	0.35	0.60	0.00
	8-fold	0.12	0.80	0.70	0.48	0.00
	9-fold	0.12	1.37	0.00	0.48	0.00
	10-fold	0.24	1.14	0.35	0.24	0.00
	Overall	0.18±0.13	1.09±0.34	0.36±0.20	0.50±0.36	0.00±0.00
SVM	1-fold	1.07	1.94	0.94	0.48	0.00
	2-fold	0.71	1.14	0.82	1.20	0.00
	3-fold	1.07	1.14	1.29	0.84	0.00
	4-fold	0.60	1.94	0.94	0.36	0.00
	5-fold	1.19	1.71	0.82	1.19	0.00
	6-fold	1.31	1.71	1.53	1.44	0.00
	7-fold	1.31	0.91	1.41	1.20	0.00
	8-fold	1.55	1.72	0.82	1.44	0.00
	9-fold	0.71	2.06	1.41	1.20	0.00
	10-fold	1.07	2.06	1.41	1.32	0.00
	Overall	1.06±0.31	1.63±0.42	1.14±0.29	1.07±0.38	0.00±0.00
J48	1-fold	0.59	2.17	2.00	2.03	0.44
	2-fold	1.43	1.60	1.06	1.56	0.00
	3-fold	1.66	1.60	2.00	2.63	0.22
	4-fold	1.19	0.69	2.82	2.03	0.44
	5-fold	1.19	2.40	2.46	1.91	0.22
	6-fold	1.19	1.14	2.47	0.96	0.78
	7-fold	1.07	2.51	2.12	1.68	0.11
	8-fold	0.83	1.60	2.00	1.20	0.11
	9-fold	1.79	1.83	2.35	1.20	0.44
	10-fold	0.83	2.52	2.70	2.28	0.11
	Overall	1.18 ±0.38	1.81 ±0.61	2.20 ±0.50	1.75 ±0.53	0.29 ±0.38

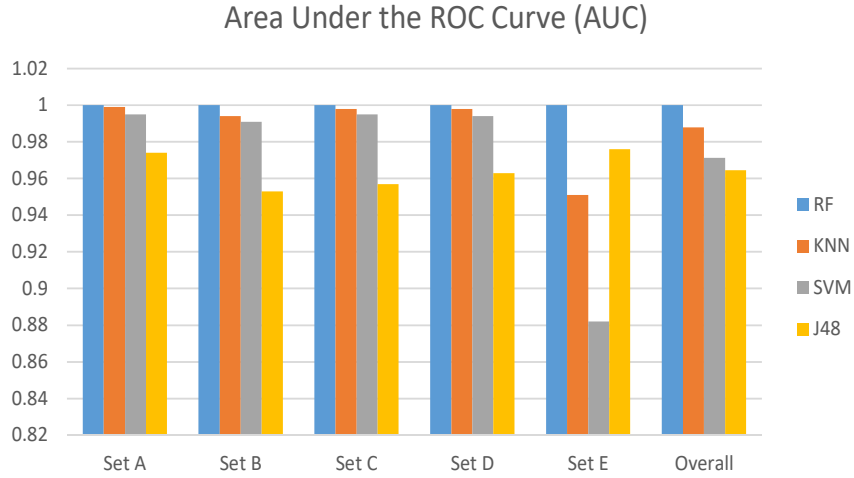


Figure 4.11: The AUC for the proposed classifiers.

In order to explore the best classifier for the DP_PCA features set, the performances of all four classifiers are compared in terms of kappa statistics and Area Under the ROC Curve (AUC). Figure 4.11 presents the AUC for the RF, k -NN, SVM, and J48 classifiers for the DP_PCA features set, separately for each of five classes and their overall AUC as well. The AUC is used as a measure for assessing the classifier performance (e.g., a higher value of the area indicates better performance of the classifier). As can be seen in Figure 4.11, each of the four classifiers achieve high AUC close to 1 for each class (Set), and the RF classifier produces slightly higher AUC in each class comparing to the other three classifiers. Figure 4.12 shows the performance of all reported classifiers in terms of kappa statistic. In this research, kappa statistics test is used to evaluate the consistency of the four classifiers: the RF, k -NN, SVM, and J48 on the DP_PCA features set. The kappa value (k) indicates the consistency of the classifier. The consistency is considered as mild if $k < 0.2$, fair if $0.21 < k < 0.40$, moderate if $0.41 < k < 0.60$, good if $0.61 < k < 0.80$, and excellent if $k > 0.81$. The maximum value of Kappa is one which defines total consistency. As can be seen in Figure 4.12, the kappa values are very high (close to 1) for all four classifiers, and the RF classifier achieves the highest kappa value

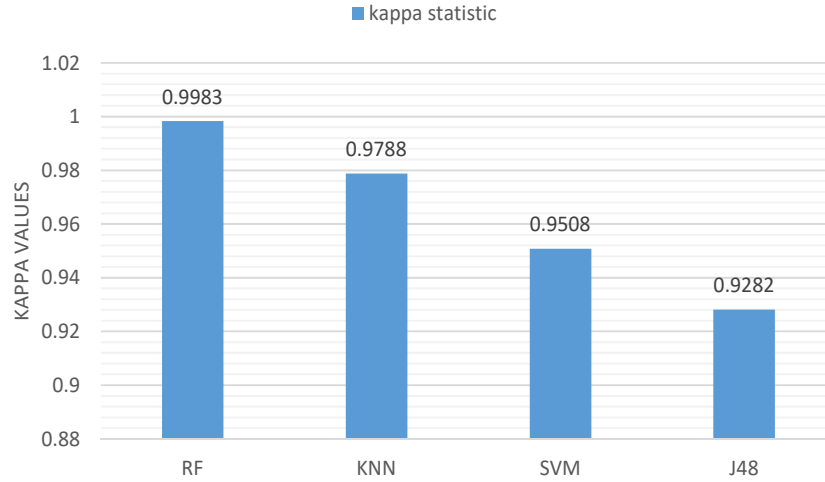


Figure 4.12: Kappa statistics values for the proposed classifiers

($K=0.998$). From Figure 4.11 and 4.12, it is clear that the RF classifier yields better performance with the DP_PCA features set in the EEG signals classification than the other three classifiers. Therefore, The RF classifier is selected as the best classifier for the DP_PCA features set in epileptic EEG signal classification.

4.5.2 Comparison

Although there are many studies in the literature for epileptic EEG classification, most are restricted to the two-class classification problems dealing with the benchmark epileptic EEG data [249, 138, 145, 144, 143, 250]. Few studies have focused on the multi-class EEG signal classification [17, 18, 4, 19, 20, 21](discussed in Section 4.2). To further evaluate the efficiency of the proposed method, a comparison of the proposed method with other six reported methods is provided. Table 4.6 provides a comparative study between the proposed method and the three reference algorithms for the same benchmark epileptic EEG dataset. This table reports the overall classification performance of the five categories of EEG signals in terms of sensitivity, specificity, and the classifi-

cation accuracy. The specificity is the complement of False positive rate (FPR) ($100 - FPR$). The highest overall classification performances among all reported methods are highlighted in bold. From Table 4.6, it is observed that the proposed method achieves the highest performance in each statistical parameter of each class compared to the six reference methods. The OCA of the proposed method is 99.85% while they are 99.30%, 99.28%, 99.20%, 98.05%, 97.60%, and 93.63% for methods reported in [18], [4], [17], [19], [20], and [21], respectively. These results indicate the proposed method outperforms all six referenced methods and improves the OCA by at least 0.55%.

4.6 Conclusion

This research investigates the effectiveness of DP in detection of epileptic seizures and presents an effective feature extraction technique for the classification of multi-class EEG signals. The DP and PCA are combinedly used to extract representative features named as DP_PCA feature set. To assess the performance of the proposed method, four machine learning techniques: RF, k -NN, SVM and DT classifiers are employed on the obtained features. The experimental results report that the proposed method is very effective and efficient for extracting features from epileptic EEG data. The high classification performance achieved by all classifiers confirm the consistency of the extracted features to detect epileptic EEG signals. The results show that the proposed RF classifier with the DP_PCA features obtains the best overall performance as compared to the other classifiers. The results also indicate that the proposed method outperforms the existing methods for the same epileptic EEG database. To conclude, the Douglas-Peucker algorithm is reliable for extracting the representative samples from the original EEG data and the RF with the DP_PCA feature set is an effective classifier for the classification of multi-class EEG signals.

Table 4.6: Comparison with the existing methods on epileptic EEG database.

Methods	Datasets	Statistical parameters (%)		
		Sensitivity (Se)	Specificity (Sp)	Overall classification accuracy (OCA)
Proposed Method	Set A	99.79	99.96	99.85
	Set B	99.85	99.97	
	Set C	99.96	99.90	
	Set D	99.71	99.99	
	Set E	100.00	100.00	
Emigdio et al. (2016) [20]	Set A	95.30	99.26	97.60
	Set B	96.83	98.69	
	Set C	97.23	99.53	
	Set D	98.40	99.42	
	Set E	99.28	99.80	
Murugavel and Ramakrishnan (2016) [21]	Set A	93.25	98.42	93.63
	Set B	93.63	98.36	
	Set C	94.00	98.16	
	Set D	94.13	97.17	
	Set E	93.13	99.54	
Ubeyli (2010) [19]	Set A	98.00	99.62	98.05
	Set B	98.25	99.53	
	Set C	98.13	99.37	
	Set D	98.00	99.15	
	Set E	97.88	99.87	
Ubeyli (2009) [17]	Set A	99.25	99.84	99.20
	Set B	99.13	99.81	
	Set C	99.25	99.72	
	Set D	99.38	99.62	
	Set E	99.00	100.00	
Ubeyli (2008) [18]	Set A	99.38	99.81	99.30
	Set B	99.25	99.87	
	Set C	99.13	99.78	
	Set D	99.50	99.65	
	Set E	99.25	100.00	
Guler and Ubeyli (2007) [4]	Set A	99.25	99.84	99.28
	Set B	99.38	99.84	
	Set C	99.25	99.75	
	Set D	99.38	99.65	
	Set E	99.13	100.00	

CHAPTER 5

A PCA AIDED CROSS-COVARIANCE SCHEME FOR CLASSIFICATION OF MI TASKS

Feature extraction of EEG signals plays a significant role in BCI as it can significantly affect the performance and the computational time of the system. The main aim of the current work is to introduce an innovative algorithm for acquiring reliable discriminating features from EEG signals to improve classification performances and to reduce the time complexity. This chapter develops a robust feature extraction method combining the PCA and the CCOV for the extraction of discriminatory information from the mental states based on EEG signals in BCI applications. The correlation based variable selection method with the best first search is applied on the extracted features to identify the best feature set for characterising the distribution of mental state signals. To verify the robustness of the proposed feature extraction method, three machine learning techniques including multilayer perceptron neural networks (MLP), least square support vector machine (LS-SVM), and logistic regression (LR) are employed on the obtained features. The proposed methods are evaluated on two publicly available datasets. The performances of the proposed methods are evaluated by comparing them with some recently reported algorithms. The experimental results show that the proposed feature extraction method is very effective for getting representative information from mental states EEG signals in BCI applications and is able to reduce the computational complexity of classifiers by reducing the number of extracted features.

5.1 Introduction

BCI is a communication system that provides a direct communication channel for transmitting messages from the human brain to computers by analyzing the brain's mental activities [12]. EEG is widely used for the acquisition of brain signals in BCI systems, as it is non-invasive and has excellent temporal resolution [59, 277]. Non-invasive BCI systems make the use of EEG signals to translate a subject's thought or intention into a control signal that allows a subject, such as a disabled person, to communicate with a device, such as a computer, a wheelchair or a neuroprosthesis [278]. MI is the mental rehearsal of a movement (such as a hand or foot movement) without executing it [279].

The detection of MI-based EEG signals is divided into mainly two parts: Feature extraction and classification. Feature extraction is the most important part of MI-based EEG signals classification system as it significantly affects the performance and the computational time of the classifier. The performance of the classifier will be degraded if the features extracted from EEGs contain irrelevant features. The computational time of the classifier may ascend by increasing the number of extracted features. Therefore, it is very important to extract optimal number of relevant features from the MI-based EEG signals to achieve a good classification rate and computational time for a classifier. However, the features extracted from MI-based EEG signals are noisy due to poor signal-to-noise ratio of the EEG signals and have high dimensionality [180, 212], which may result in increasing the computational time of the classifier and also reducing the classification accuracy.

Many methods for classification of MI-based EEG signals have been developed based on different feature extraction and classification methods. Features methods include common spatial patterns (CSP) [280], autoregressive (AR) [156], fast fourier

transform (FFT) [113], wavelet coefficients [281], spatio-spectral patterns [6], power spectral density (PSD) [161], correlation [25, 282], phase locking value (PLV) and spectral coherence [167]. Several classifiers have been used for MI-based EEG classification such as support vector machines (SVM) [187], linear discriminant analysis (LDA) [187], neural networks (NN) [283], LR [284], and k-nearest neighbour [285].

The existing feature extraction methods often extract several features from different EEG channels and different time segments and finally they concatenate all extracted features into final high dimensional feature vector set which is fed to a classifier for classification [180]. Not all of the features extracted from EEG signals are equally relevant to the target EEG signals. Therefore, some researchers have used feature selection techniques such as principal component analysis [286], and independent component analysis [287] to select a subset of features by performing a linear transformation of the feature vector set to a low-dimensional vector. In addition, some existing approaches apply different subset of features to the classifier to find the optimal feature set [284]. Although, these methods can reduce the dimensionality of the feature set, they still have low classification performance and require cumbersome experimental validations to identify the optimal feature sets.

To address the aforementioned problems, this study proposes a robust feature extraction method based on the PCA and CCOV for the extraction of discriminatory information from the MI-based EEG signals. The PCA is used to explore inter-channel information and extract a temporal pattern of MI tasks. Then, the CCOV is applied to identify the relationship between the EEG signals and the temporal pattern and provide more discriminative information about those signals. In order to characterise the distribution of EEG signals, eleven statistical features are extracted from each cross-covariance sequence. The correlation-based variable selection method (fsSubsetEval)

with the best first search is applied on the extracted features to find the optimal feature set. The extracted feature sets are then used as inputs to MLP, LS-SVM, and LR classifiers. The proposed method is evaluated on two benchmark datasets, IVa and IVb of BCI Competition III [288]. The performance of the proposed method is compared with ten most recent reported methods in dataset IVa. The results show that the MLP and LS-SVM classifiers achieve a better performance compared to the LR classifier for the same extracted feature set. The results also indicate that the proposed methods outperform all the recent reported existing methods in dataset IVa.

The remainder of this chapter is organised as follows. In Section 5.2, the experimental datasets used in this study are described, and the proposed methods are presented. Section 5.3 describes the performance evaluation procedure for the proposed methods. The experimental results and discussions are provided in Section 5.4. Finally, Section 5.5 concludes this chapter.

5.2 Materials and Method

In this section, first the EEG datasets used in this chapter are described, and then the proposed method and its implementation on the EEG datasets are presented.

5.2.1 EEG Datasets

Two publicly available EEG datasets, IVa and IVb from BCI Competition III [288] are used to assess the effectiveness of the proposed method. All EEG signals in both datasets were collected during MI tasks.

Dataset IVa [288] contains MI EEG signals recorded from five healthy persons (*aa*, *al*, *av*, *aw*, and *ay*) using 118 electrodes at the positions of the extended international 10/20 system. During each capturing trial, a visual cue representing one of three MI tasks was presented for 3.5 s to a subject who was required to perform the MI tasks according to the target cues: left hand, right hand, and right foot MI tasks. For subject to relax, the cue presentation was separated by relaxation time intervals with random length from 1.75 to 2.25 s. Only two classes of MI tasks (right hand and right foot classes) are provided for classification. Each subject provided a total number of 280 trials consisting of 140 trials for each class. For each subject, a training and testing data set containing different numbers of trials were formed. The numbers of training trials for subjects *aa*, *al*, *av*, *aw*, *ay* were 168, 224, 84, 56, and 28, respectively. The remaining trials were used as testing data. All EEG signals were down-sampled at 100 Hz, For succedent analysis.

Data IVb [288] was recorded from one healthy subject. It consists of two classes: left-hand and right-foot MI. In each trial, the subject performed left hand or right foot MI tasks according to visual cues displayed on a computer screen for 3.5 s (i.e., arrows pointing left and down were considered as visual cues indicating left hand and right foot imagery MI task, respectively). The cue presentation was separated by non-task intervals with random length from 1.75 to 2.25 s. The data consists of 210 trials and were collected using 118 EEG channels at the positions of the extended international 10/20 system. The data were down sampled to 100 Hz and bandpass filtered.

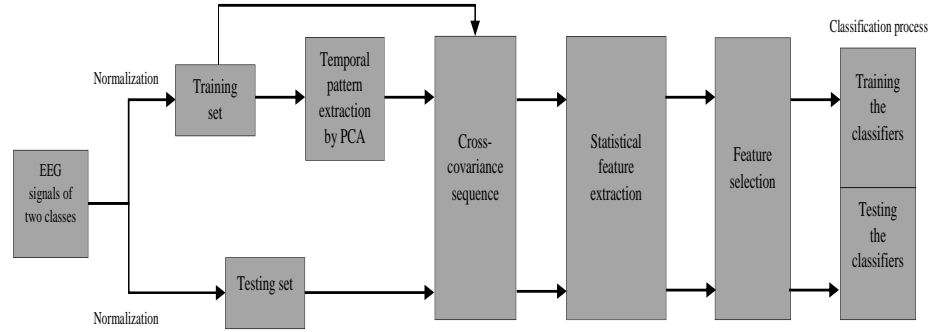


Figure 5.1: Block diagram of the proposed method for the MI tasks classification in BCIs.

5.2.2 Proposed Method

This study proposes an innovative scheme for classifying MI tasks in the development of BCI systems, which is shown in Figure 5.1. This scheme includes six main processing parts: normalisation, determination of dominant temporal pattern by PCA, computation of a cross-covariance sequence, statistical feature extraction, feature selection and classification by the MLP, LS-SVM, and LR. The descriptions of each part are provided below.

5.2.2.1 Normalisation

In this step, all EEG signals are preprocessed to have unit variance and zero mean to eliminate differences in power levels between the signals due to electrode position. The normalisation process facilitates the PCA calculations to ensure equal weighting. This process is done by subtracting the overall average from each EEG signal and dividing it by the standard deviation. The normalisation of each signal can be achieved from the

following equation.

$$y_i = \frac{x_i - \bar{x}}{\delta} \quad (5.1)$$

where y_i and x_i are the normalised and original signal values, respectively. \bar{x} and δ are the mean and standard deviation of signal x_i , respectively. As mentioned before, two datasets IVa and IVb from BCI Competition III are used in this study. Both dataset consist EEG signals involving a MI task with two classes. The dataset IVa contains the right-hand and the right-foot MI task while the dataset IVb contains the left-hand class and the right-foot class. Each subject dataset contains 118 EEG channel data in each class. Each dataset for a subject is divided into 50/50 fractions for training and testing sets for evaluation of the proposed methods. Therefore, each training set and testing set for a subject contains 59 channel data in each class. All channel data in both training and testing sets are normalised using Eq 5.1.

5.2.2.2 Determination of Dominant Temporal Pattern by PCA

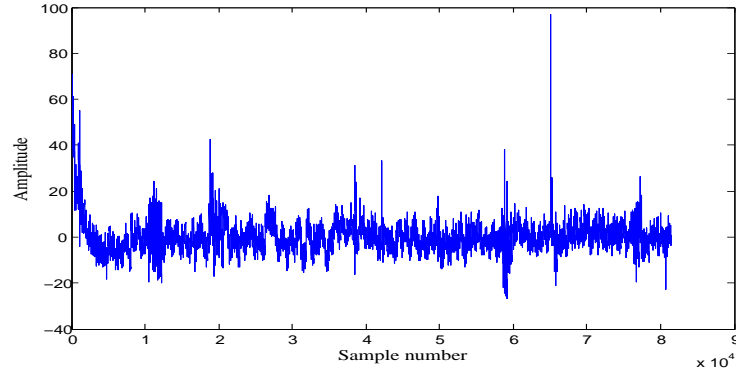
In this phase, a dominant temporal pattern is determined by PCA using only one of the two-class channel training data for each subject. The dominant temporal patterns are considered as a reference signal in each subject. Each dominant temporal pattern collects the information that is common to all EEG channels and characterises the EEG pattern across channels. The first principal component (PC1) captures the largest variability of the data, while the next components (PC2, PC3, etc.) represent, respectively, less variation. The PC1 is considered as a dominant temporal pattern from all EEG signals of every class. The multi-channel EEG signals recorded during MI tasks are highly correlated with each other. However, they contain different amount of discriminative

information regarding MI tasks due to different channel position on the scalp. For instance, channels positioned near to the active brain areas (responding to an MI task) may provide more information of MI task compared to channels positioned far from the active brain areas. Therefore, the PCA is used to explore inter-channel information and extract a temporal pattern of MI tasks. Detailed information about the PCA can be found in references [289, 182, 290].

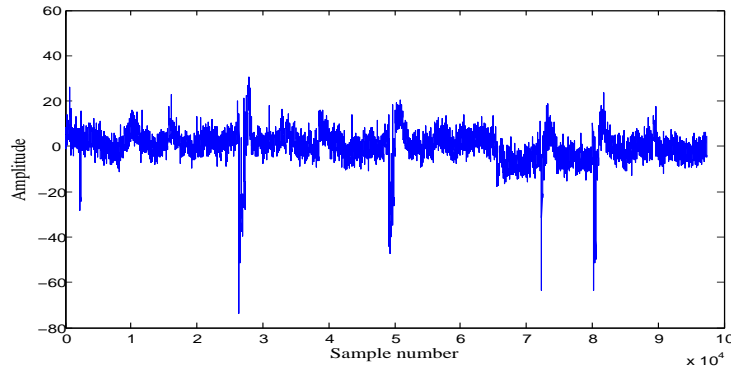
The dominant temporal pattern is extracted from only one of the two-class channels in each subject training set. In this study, For each subject in dataset IVa, all channel data with right-hand class of each training set are used to extract the dominant temporal pattern for each subject while all channel data with right-foot class are used to obtain the dominant temporal pattern for dataset IVb. Other class channel data are also tested to extract the dominant temporal pattern for each subject, and it is found that their performances are similar for both cases. Figures 5.2(a) and 5.2(b) illustrate the dominant temporal pattern of subject *aa* for dataset IVa and dataset IVb, respectively.

5.2.2.3 Computation of a Cross-Covariance Sequence

The CCOV [291, 292] is used to calculate a cross-covariance sequence between the dominant temporal pattern and any channel signal in any of the training and testing sets for each subject. The CCOV measures the similarity between two signals and it is commonly utilised to discover features from an unidentified signal by comparing it to a reference signal. The CCOV can be calculated by:



(a) Subject *aa* in Dataset IVa



(b) Dataset IVb.

Figure 5.2: The dominant temporal patterns extracted from:

$$CCOV_{xy}[m] = \sum_{n=0}^{N-|m|-1} (x[n+m] - \bar{x})(y[n] - \bar{y})^*; \quad (5.2)$$

$$m = -(N-1), -(N-2),$$

$$..., 0, 1, 2, 3, ..., (N-2), (N-1).$$

where $x[n]$ is considered as the dominant temporal, $y[n]$ is regarded as any channel signal in both training and testing sets for each subject, \bar{x} and \bar{y} are the mean of $x[n]$ and $y[n]$ respectively, N is the number of sample points, m represents time-shift parameters known as lag, and $C_{xy}[m]$ is the cross-covariance sequence.

In this study, 59 channels of each class are considered for a subject in each training

and testing set. Therefore, each training and testing dataset for each subject contains 118 cross-covariance sequences that are obtained from the dominant temporal pattern and any of the channels in each training and testing sets for a subject. For example, a total of 118 cross-covariance sequences are obtained from the dominant temporal pattern which are extracted from the channel data with right-hand class and all 118 channels in the training set for subject *aa* (59 right-hand channel signals and 59 right-foot channels signals).

Figure 5.3 illustrates typical cross-covariance sequences obtained for the right-hand and the right-foot MI signal for the subject *aa* in dataset IVa. To obtain each cross-covariance sequence, the dominant temporal pattern of subject *aa* (Figure 5.2(a)) is cross-correlated with either the right-hand signal or the right-foot signal.

Figure 5.4 shows typical signals of dataset IVb for the right-foot MI and the left-hand MI. This figure also presents typical results of the cross-covariance sequence for the right-foot MI signal and the left-hand MI signal. As shown in Figure 5.4, the cross-covariance sequence of the right-foot MI signal is obtained using the dominant temporal pattern and the right-foot MI signal, and the left-hand cross-covariance sequence is generated by the dominant temporal pattern and the left-hand MI signal.

Figure 5.4 illustrates typical the right-foot signal and the left-hand signal in dataset IVb. The typical results of the cross-covariance sequence for these signals are also shown in Figure 5.4. The dominant temporal pattern for dataset IVb (Figure 5.2(b)) is cross correlated with either the right-foot signals or the left-hand signals to obtain the right-foot and the left-hand cross-covariance sequences.

The cross-covariance sequences are less noisy than the original signals and they provide information about the similarity between the dominant temporal pattern and

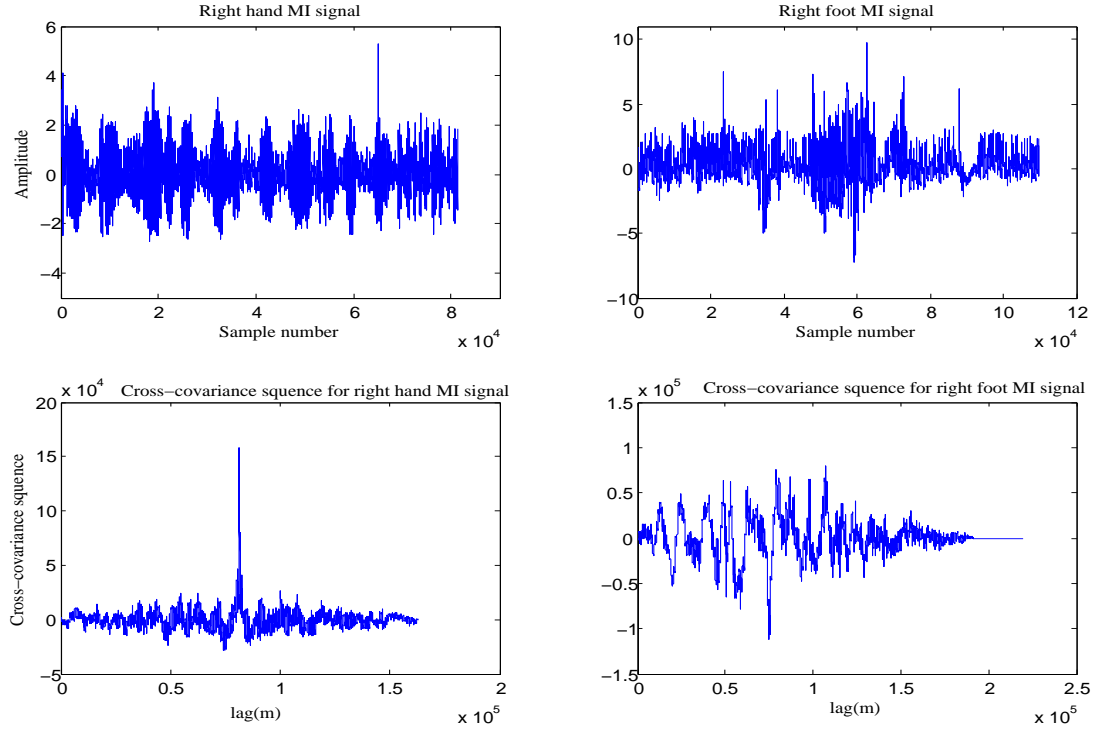


Figure 5.3: Typical cross-covariance sequences obtained for the right-hand and the right-foot MI signal for the subject *aa* in dataset IVa.

original signals. Therefore, they can provide more discriminative information for distinguishing MI tasks as compared to the original signal. As can be seen from Figures 5.3 and 5.4, the two cross-covariance sequences have different shapes, which may indicate the presence of statistical discrimination between the cross-covariance sequences of two class MI task.

5.2.2.4 Statistical Feature Extraction

To characterise the distribution of EEG signals, eleven statistical features (denoted by *F1-F11* as shown in Table 5.1) are extracted from each cross-covariance sequence in each training and testing set for a subject. The reason for considering the eleven features

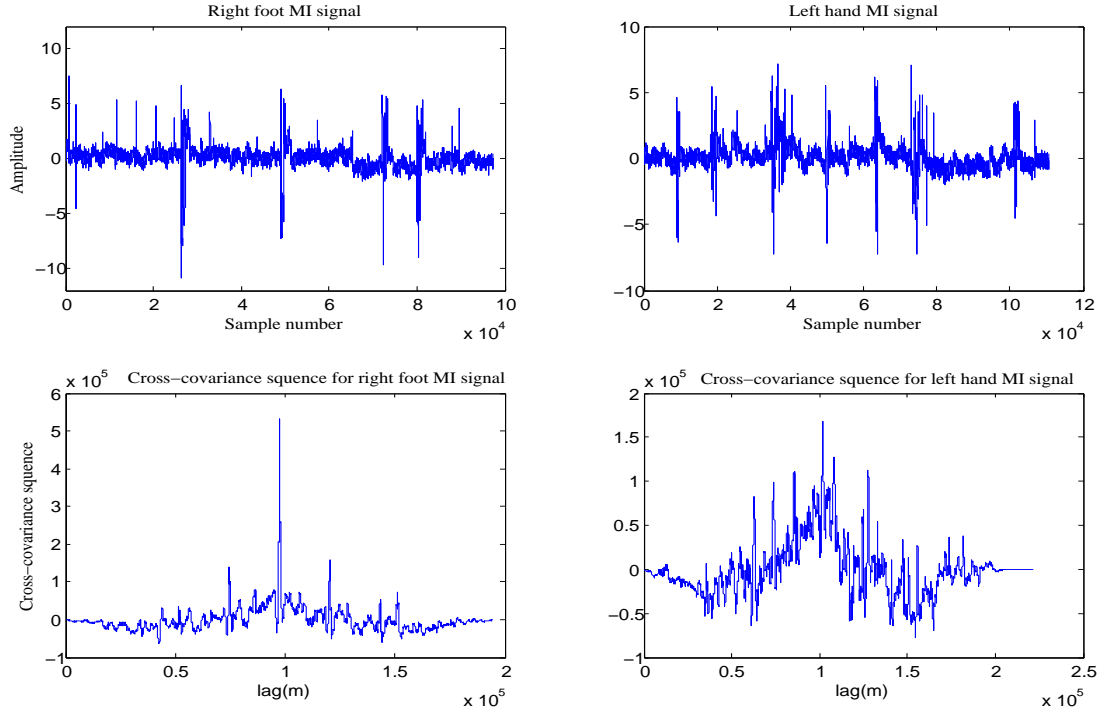


Figure 5.4: Typical cross-covariance sequences obtained for the right-foot and the left-hand MI signal for dataset IVb.

in this study are described here. *Mean* corresponds to the centre of a set of values while *Median* is the middle most observation. *Mode* is the value in the data set that occurs most often. In a tabular form, the *Mode* is the value with the highest frequency. *Mean* and *Median* are the measures irrespective of data are discrete or continuous. However, the *Mode* is most suitable for discrete data but is tricky for continuous case. The *mode* for a continuous probability distribution is defined as the peak of its histogram or density function. *Mean*, *median* and *Mode* are the most used features that can describe almost all distributions with a reasonable degree of accuracy [25, 293, 294, 23] and provide a fairly good idea about the nature of the data. *Standard deviation* gives information about the spread of data on how close the entire set of data is to the average value in the distribution. *Maximum* and *Minimum* values are used to describe the range of

Table 5.1: Set of features extracted from each cross-covariance sequence.

Feature #	Feature description
<i>F1</i>	Minimum
<i>F2</i>	Maximum
<i>F3</i>	Mean
<i>F4</i>	Median
<i>F5</i>	Mode
<i>F6</i>	Standard deviation
<i>F7</i>	First quartile
<i>F8</i>	Third quartile
<i>F9</i>	Inter quartile range
<i>F10</i>	skewness
<i>F11</i>	kurtosis

observations in the distribution. *First quartile* and *Third quartile*, measure how the data is distributed in the two sides of the *Median*. *Inter quartile range* is difference between *Third quartile* and *First quartile* that is used in measuring the spread of a data set that excludes most outliers. *Skewness* describes the shape of a distribution that characterises the degree of asymmetry of a distribution around its mean [284]. *Kurtosis* measures of whether the data are peaked or flat relative to a normal distribution. These features create feature vector set. Not all features extracted are relevant to MI tasks. Therefore feature selection method is used to find an optimal subset of features, rather than using all eleven features. Feature selection algorithms require the use of two main components namely, a feature evaluator and a search method. The search method determines the best feature set and the evaluator method evaluates the worth of each subset of features. This process not only reduces the size of feature set by removing irrelevant features but also improves the performance of the classifiers.

Table 5.2: The selected features by BestFirst+CfsSubsetEval for each subject in datasets (IVa and IVb).

Subject	Selected features
IVa-aa	F2, F4, F5, F6, F7, F8, F9, F11
IVa-al	F1, F2, F4, F5, F6, F9, F10, F11
IVa-av	F2, F3, F5, F6, F7, F8, F9, F10
IVa-aw	F2, F4, F5, F6, F8, F11
IVa-ay	F1, F2, F4, F5, F6, F11
IVb	F1, F2, F4, F5, F8, F10, F11, F11

5.2.2.5 Feature Selection

In this research, The Best First (Best-First [295]) is used to search through the features' subsets by greedy hill-climbing augmented with a backtracking facility, and then the Correlation-based variable selection (CfsSubsetEval [296]) is used to assess the worth of a subset of features, by measuring the predictive ability of each feature individually and the degree of redundancy among them. It chooses the features sets that are highly correlated with the class but have low inter-correlation with each other. Table 5.2 shows the selected features by Best-First+CfsSubsetEval for each subject in both datasets (IVa and IVb).

To explore which set of features can produce the best performance of the proposed classifier, two feature sets (Fs1 and Fs2) are constructed. The Fs1 includes all eleven features extracted from each cross-covariance sequences while the Fs2 includes the two most frequent features (*F2 and F5*) selected by BestFirst+CfsSubsetEval for each subject in datasets (IVa and IVb).

In this study, 118 cross-covariance sequences in each training and testing datasets are obtained for a subject (59 cross-covariance sequences for each class). The mentioned

eleven features are extracted from each cross-covariance sequence and then two feature sets are constructed. For Fs1, a total of 118 feature vectors of eleven dimension sequence in each training and testing datasets are acquired for each subject while it is a total of 118 feature vectors of two dimensions for Fs2. The training feature sets are used to train the three classifiers that are discussed in the next section while the testing feature sets are used to evaluate those classifiers.

5.2.2.6 Classification by the MLP, LS-SVM, and LR

This study employs three classifiers: MLP, LS-SVM, and LR classifier for the classification of EEG signals. A brief explanation of these classification methods is provided in the following sections.

Multilayer Perceptron Neural Networks: the MLP with Backpropagation is employed as a classifier to classify MI tasks based on the obtained features. In this study, the MLP is chosen as a potential classifier because the MLP is widely used classification technique in BCI application which has ability of producing nonlinear decision boundaries [199, 297]. The MLP consists of three layers: input, hidden and output layers. The hidden layer process and transmits the input information to the output layer. Each layer is made up of neurons. The number of neurons in the input layer is the same as the number of features present in the feature set and the number of neurons in the output layer is the same as the number of classes in the dataset. The number of neurons in the hidden layer can be found by trial and error [298, 299].

Least Square Support Vector Machine: the LS-SVM is originally proposed by Suykens and Vandewalle [300]. It is a modified version of a SVM [301]. There are some differences between SVMs and LS-SVMs trainings. A set of linear equations is

used for LS-SVM training while a quadratic optimisation formulation is used for SVM training. The decision function of LS-SVM is given by [300].

$$y(x) = \text{sign} \left(\sum_{i=1}^N y_i \alpha_i k(x, x_i) + b \right) \quad (5.3)$$

where $y(x)$ is the predicted class on the basis of the input index x , b is the bias term, α_i denotes Lagrange multipliers called support values, and $k(x, x_i)$ is the kernel function. The radial basis function (RBF) is used as the kernel function in this study, which is defined [300, 302, 303] as

$$K(x, x_i) = \exp \left(\frac{-(\|x - x_i\|)^2}{2\sigma^2} \right) \quad (5.4)$$

where σ is the kernel parameter or width. The values chosen for two (\mathcal{C} , σ^2) important parameters of the LS-SVM affect the classification accuracy of the LS-SVM classifier. Therefore, a two-step grid search technique is used to find the optimum values of these parameters for the LS-SVM, that is discussed in section 5.4.1. The details of the LS-SVM algorithm could be found in reference [300].

Logistic Regression: the LR is a statistical technique for modeling the probability of the dichotomous outcome event in terms of suitable explanatory variables x . The logit model of the LR is defined by the following equation [244, 304, 305]:

$$\begin{aligned} \text{logit}(P) &= \ln \left(\frac{P}{1-P} \right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \\ &= \beta_0 + \sum_{i=1}^n \beta_i x_i \end{aligned} \quad (5.5)$$

where β_0 is the intercept, $\beta_1, \beta_2, \dots, \beta_n$ are the regression coefficient related to the inde-

pendent variables x_1, x_2, \dots, x_n , and P is the probability of occurrence of an event($y=1$) and is given by [244, 304, 305, 306]:

$$P(y = 1 \mid x_1, x_2, \dots, x_n) = \left(\frac{e^{B_0 + \sum_{i=1}^n \beta_i x_i}}{1 + e^{B_0 + \sum_{i=1}^n \beta_i x_i}} \right) \quad (5.6)$$

where x_1, x_2, \dots, x_n are vectors of input features and y is its class label that can be either 0 or 1. The complementary event probability $P(y = 0 \mid x_1, x_2, \dots, x_n)$ can therefore be calculated as $1 - P(y = 1 \mid x_1, x_2, \dots, x_n)$. The maximum likelihood estimation (MLE) [307] is used in the LR equation to estimate the coefficients $\beta_1, \beta_2, \dots, \beta_n$.

5.3 Performance Evaluation

In order to evaluate the performance of the proposed methods, this study uses training data sets for each subject to train the classifiers and testing datasets to evaluate the performance of the classifiers for the classification of the two class MI data. The performance of the proposed classifiers are evaluated by sensitivity (S_e), specificity (S_p), and overall Accuracy (Acc). Their formulas are given below:

$$\begin{aligned} S_e &= \frac{TP}{P} \\ S_p &= \frac{TN}{N} \\ Acc &= \frac{TP + TN}{P + N} \end{aligned} \quad (5.7)$$

where N and P represent the total number of cases in each class. TP and TN are the number of correctly classified cases in each class. In dataset IVa, N and P denote the total number of the right-hand class and the right-foot class, respectively. Correctly predicted

the right-hand class and the right-foot class are defined as TP and TN , respectively. In dataset IVb, N and P are the total number of the right-foot class and the left-hand class, respectively. Correctly predicted the right-foot class and the left-hand class are defined as TP and TN , respectively. In addition, execution time of each classifier for training and testing is measured to compare the computational cost of each classifier based on different feature sets. The proposed methods were tested on Intel R Core™ i5 CPU 3.2GHz computer with 8.00GB RAM using 32-bit OS.

5.4 Experiments and Results

In this section, first the parameters selection for the three mentioned classifiers namely, MLP, LS-SVM, and LR are explained in section 5.4.1. Then the performance of each classifier is assessed on both feature sets (Fs1 and Fs2) separately for datasets IVa and IVb discussed in section 5.4.2. Finally, a comparative study of the proposed methods with some existing methods is presented in section 5.4.3. The LS-SVM classification is carried out in MATLAB R signal processing tool (version 7.11, R2010b) using the LS-SVMlab toolbox (version 1.8) [308]. The classification executions for the MLP and LR classifiers are executed in WEKA [309].

5.4.1 Parameter Selection

This study employs three classifiers, MLP, LS-SVM, and LR for classification of MI task. The performance of each classifier is evaluated on both features sets (Fs1 and Fs2). To select the number of neurons in the hidden layer of the MLP classifier, The WEKA implementation of the MLP classifier with default settings is used. The number

of neurons n in the hidden layer can be calculated as:

$$n = \frac{\text{number of attributes} + \text{number of classes}}{2} \quad (5.8)$$

For Fs1, the MLP model with a single hidden layer of 6 hidden neurons is used while the MLP model with a single hidden layer of 2 hidden neurons is used for the Fs2 feature set.

In LS-SVM, the RBF kernel function is used as an optimal kernel function after testing different kernels such as Linear kernel, polynomial kernel, and sigmoid kernel. The LS-SVM with RBF has two important parameters γ and σ^2 . The regularisation parameter γ determines the trade-off between minimising the training error and minimising the model complexity. The parameter σ^2 is the bandwidth and implicitly defines the nonlinear mapping from the input space to a high dimensional feature space. These two parameters play an important role in the classification performance of the LS-SVM classifier. In this study, a two-step grid search technique is used to select the optimum values of these two parameters. The two-step grid search method is applied in each of the training feature sets (Fs1 and Fs2) for a subject in both datasets IVa and IVb. Table 5.3 and 5.4 present the obtained values of γ and σ^2 for the LS-SVM classifier in dataset IVa and dataset IVb, respectively. In the LR classifier, The LR parameters are selected automatically using maximum likelihood estimation (MLE) method.

Table 5.3: The obtained values for the LS-SVM parameters (γ and σ^2) for each subject in dataset IVa.

Obtained optimal parameter values of γ and σ^2										
Subject	<i>aa</i>		<i>al</i>		<i>av</i>		<i>aw</i>		<i>ay</i>	
parameters	γ	σ^2	γ	σ^2	γ	σ^2	γ	σ^2	γ	σ^2
Training dataset Fs1	5.46	153.5	1.70	2.58	1.47	1.01	81.25	439.28	0.62	10.78
Training dataset Fs2	1.75	1.99	1.37	1.8	1.52	2.16	1.06	1.29	1.76	2.64

Table 5.4: The obtained values for the LS-SVM parameters (γ and σ^2) for the subject in dataset IVb.

Obtained optimal parameter values of γ and σ^2		
subject	IVb	
parameters	γ	σ^2
Training dataset Fs1	1.25	8.20
Training dataset Fs2	1.71	1.07

5.4.2 The MI Classification Results for Each Classifier Testing Different Features Fs1 and Fs2

In this study, three classifiers are used to identify the MI EEG signals where the two training feature sets (Fs1 and Fs2) are used separately to train each classifier and the testing feature sets are used to evaluate the general efficacy of each classifier.

The classification results of each classifier, the MLP, the LS-SVM, and the LR for

Table 5.5: The classification results of each classifier based on the eleven-feature set (Fs1) in dataset IVa.

Subject	MLP			LS-SVM			LR		
	$S_e(\%)$	$S_p(\%)$	Acc(%)	$S_e(\%)$	$S_p(\%)$	Acc(%)	$S_e(\%)$	$S_p(\%)$	Acc(%)
<i>aa</i>	98.31	100	99.15	98.31	100	99.15	98.31	100	99.15
<i>al</i>	100	100	100	100	100	100	100	100	100
<i>av</i>	100	100	100	100	96.61	98.31	100	100	100
<i>aw</i>	100	100	100	100	100	100	96.61	100	98.31
<i>ay</i>	98.31	100	99.15	98.31	100	99.15	98.31	100	99.15
Average	99.32	100	99.66	99.32	99.32	99.32	98.65	100	99.32

Table 5.6: The classification results of each classifier based on the two-features set (Fs2) in dataset IVa.

Subject	MLP			LS-SVM			LR		
	$S_e(\%)$	$S_p(\%)$	Acc(%)	$S_e(\%)$	$S_p(\%)$	Acc(%)	$S_e(\%)$	$S_p(\%)$	Acc(%)
<i>aa</i>	98.31	100	99.15	98.31	100	99.15	98.31	100	99.15
<i>al</i>	100	100	100	100	100	100	100	100	100
<i>av</i>	100	100	100	100	100	100	100	100	100
<i>aw</i>	100	100	100	100	100	100	98.31	100	99.15
<i>ay</i>	98.31	100	99.15	98.31	100	99.15	98.31	100	99.15
Average	99.32	100	99.66	99.32	100	99.66	98.99	100	99.49

the eleven-feature set (Fs1) and the two-feature set (Fs2) are presented in Table 5.5 and Table 5.6, respectively, for each subject in dataset IVa. In both Table 5.5 and 5.6, the results of each classifier are shown in terms of sensitivity (S_e), specificity (S_p), and overall Accuracy(Acc).

It can be seen from Table 5.5 and 5.6 that the MLP classifier achieves the same performance in both feature sets. It produces the accuracy of 99.15% for Subject *aa*,

100% for subjects *al*, *av* and *aw*, and 99.15% for Subject *ay* for both feature sets in dataset IVa. The average classification accuracy, sensitivity, and specificity of the MLP classifier for both Fs1 and Fs2 feature sets are 99.66%, 99.32%, and 100%, respectively.

For the same dataset, the LS-SVM for the eleven-feature set (Fs1) obtains 99.15%, 100%, 98.31%, 100%, and 99.15% classification accuracy for subjects *aa*, *al*, *av*, *aw*, *ay*, respectively, whilst those values are 99.15%, 100%, 100%, 100%, and 99.15%, respectively, for the two-feature set Fs2. As shown in Table 5.5 and 5.6, The overall classification accuracy, sensitivity, and specificity of the LS-SVM classifier for the feature set Fs1 are 99.32%, 99.32%, and 99.32%, respectively, while these values are 99.66%, 99.32%, 100%, respectively, for the feature set Fs2. The results show that the LS-SVM classifier for the feature set Fs2 performed slightly better in terms of Acc and S_p on dataset IVa.

On the other hand, it is observed from Table 5.5 and 5.6 that the LR classifier achieves the classification accuracy of 99.15%, 100%, 100%, 98.31%, and 99.15% for subjects *aa*, *al*, *av*, *aw*, *ay*, respectively, for the eleven-feature set Fs1; whereas these values are 99.15%, 100%, 100%, 99.15%, and 99.15%, respectively, for the two-feature set Fs2. The average accuracy, sensitivity, and specificity of this classifier are 99.32%, 98.65%, and 100%, respectively, for the feature set Fs1; where the values are 99.49%, 98.99%, and 100%, respectively, for the feature set Fs2. The results show that this classifier for the feature set Fs2 performed slightly better in terms of Acc and S_e on dataset IVa.

From Table 5.5, it is also seen that the highest average classification accuracy is achieved for the MLP classifier for the feature set Fs1, which is 99.66% as compared to the LS-SVM and LR classifiers. While for the feature set Fs2 (Table 5.6), both MLP and LS-SVM achieves the overall accuracy of 99.66%.

Table 5.7: The classification results of each classifier based on the eleven-feature set (Fs1) and the two-feature set (Fs2) in dataset IVb.

Feature set	MLP			LS-SVM			LR		
	$S_e(\%)$	$S_p(\%)$	Acc(%)	$S_e(\%)$	$S_p(\%)$	Acc(%)	$S_e(\%)$	$S_p(\%)$	Acc(%)
Fs1	100	100	100	100	98.31	99.15	100	100	100
Fs2	100	100	100	100	100	100	100	100	100

Table 5.7 presents the classification results of each classifier for the eleven-feature set Fs1 comparing with the results of the two-features set Fs2 for dataset IVb. As shown in Table 5.7, both MLP and LR classifiers obtain the same performance in both feature sets. They achieve an accuracy of 100%, the sensitivity of 100% and the specificity of 100% for dataset IVb. On the other hand, The LS-SVM classifier for the feature set Fs2 performed better in terms of Acc and S_p on dataset IVb as compared to the results for the feature set Fs1. It produces the accuracy of 99.15% the sensitivity of 100%, and specificity of 98.31% for the feature set Fs1, while those values are 100%, 100%, and 100%, respectively, for the feature set Fs2.

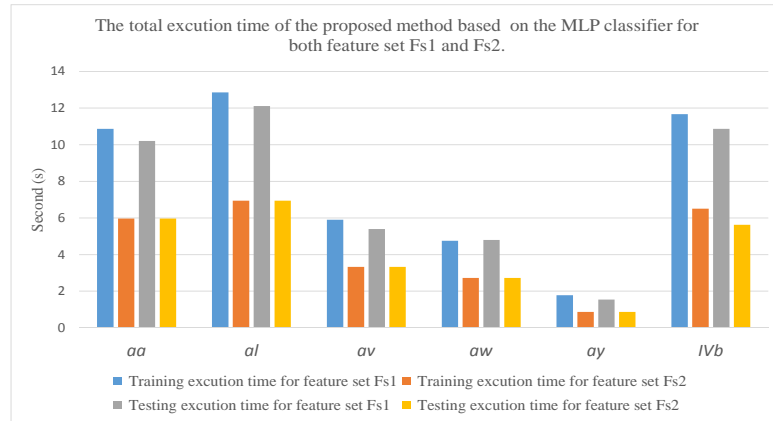
Figure 5.5 illustrates the total training and testing execution time for each classifier based on both feature set Fs1 and Fs2. The training execution time is total time of training phase of the proposed method including time for normalisation, temporal pattern extraction, cross-covariance sequence computation, statistical features extraction, and time for training the classifier. The testing execution time includes all processing times in testing part of the proposed method including time for normalisation, cross-covariance sequence computation, statistical features extraction, and time for testing the classifier. From Figure 5.5, it can be seen that the total training and testing execution time for each classifier with feature set Fs2 is much less than the total training and testing execution time of same classifier with feature set Fs1. These results show the efficiency of the feature selection method in reducing the execution time.

From the classification results and execution time results, one can see that the two features: *Maximum*, and *Mode* selected by Best-First+CfsSubsetEval method; are the best features to characterise the distribution of MI tasks signal. In addition, they provide shorter execution time as compared to the eleven-feature set. Thus, the feature set Fs2 is considered as the best feature set for MI task classification process in this study. It is worth to mention that it is not possible to provide a comparison report with existing methods [22, 23, 24, 25, 6, 26, 27, 28, 29, 30] in terms of computational complexity since the computational complexity of the existing methods are not reported in their paper.

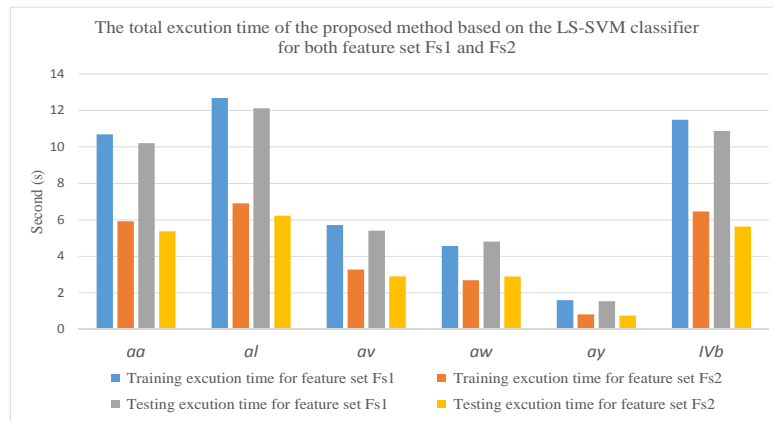
Figure 5.6 presents ROC areas for the MLP, LS-SVM, and LR classifiers for feature set Fs2, separately for all subjects in both dataset and their overall ROC area as well. The area of the ROC curve is used as a measure for assessing the classifier performance (e.g., a higher value of the area indicates better performance of the classifier). As seen in Figure 5.6, each of the three classifiers achieves high ROC area close to 1 for each subject and the MLP classifier produces slightly higher overall ROC area than the other two classifiers. From the experimental results for both data sets, it is obvious that the proposed feature extraction approach based on PCA and CCOV is capable of extracting robust features for MI tasks data and both MLP and LS-SVM classifiers have the ability to accurately classify MI tasks in BCI application.

5.4.3 Performance Comparisons Between the Proposed Methods and the Existing Techniques

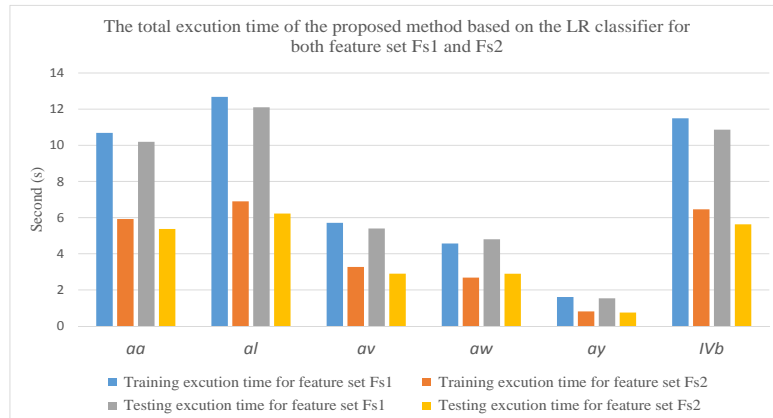
To further evaluate the efficiency of the proposed methods, the comparisons of the proposed methods with other ten recently reported methods for dataset IVa are provided.



(a) The MLP classifier



(b) The LS-SVM classifier



(c) The LR classifier

Figure 5.5: The total execution time of the proposed method based on:

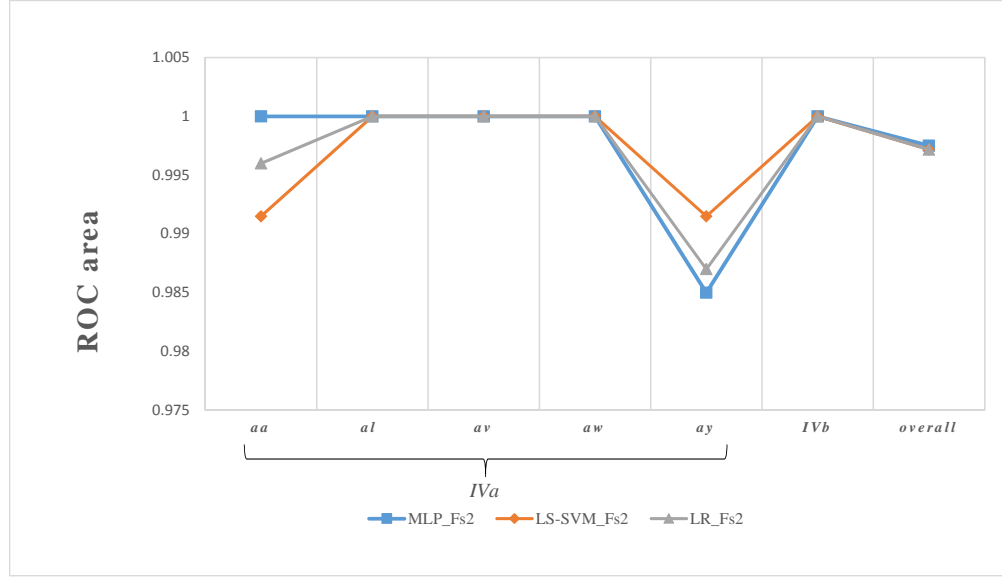


Figure 5.6: ROC area for the MLP, LS-SVM and LR classifiers with feature set Fs2 in dataset IVa and IVb.

Table 5.8 shows the comparison results in terms of classification accuracy for the proposed methods and the ten algorithms for dataset IVa. The classification performance for each subject as well as the overall accuracy are reported in Table 5.8. From literature, it is clear that different evaluation methods are used in different studies for accessing the performance of the classifiers. the experimental training and testing configurations of each method are provided in Table 5.8. It is worth to mention that, no methods were found that use 50/50 fractions for training and testing datasets for evaluation of the algorithm. Not enough literature is found for a comparison on dataset IVb.

From Table 5.8, it is observed that both proposed PCA-CCOV based MLP and LS-SVM provide superior performance compared to the ten algorithms. The overall performance of the proposed methods is 99.66% which is the highest performance among the other methods. Subject specific rates are 99.15% for subject *aa*, 100% for subject *al*, 100% for subject *av*, 100% for subject *aw* and 99.15% for subject *ay* achieved by the proposed approaches. The results indicate that the proposed methods achieve at least

Table 5.8: Performance comparisons with the existing methods for dataset IVa

Authors	Methods	Training/testing partition	Classification accuracy rate (%)					
			aa	al	av	aw	ay	Average
The Proposed	PCA-CCOV based MLP	50/50 random	99.15	100	100	100	99.15	99.66
	PCA-CCOV based LS-SVM		99.15	100	100	100	99.15	99.66
Bhattacharyya et al. [22]	LA-DE-SVM	10-fold cross validation	97.06	100	100	100	100	99.41
Siuly et al. [23]	OA-NB	10-fold cross validation	97.92	97.88	98.26	94.47	93.26	96.36
Ince et al. [24]	CS-SVM	10-fold cross validation	95.60	99.70	90.50	98.40	95.70	96.00
Siuly and Li [25]	CC-Based LS-SVM	10-fold cross validation	97.88	99.17	98.75	93.43	89.36	95.72
Wu et al. [6]	ISSPL	10-fold cross validation, 10 times	93.57	100	79.29	99.64	98.57	94.21
Siuly et al. [26]	CT-based LS-SVM	10-fold cross validation	92.63	84.99	90.77	86.5	86.73	88.32
Song et al. [27]	CSP-SVM	50/50 cross validation, 50 times	87.42	97.35	69.70	96.84	88.57	87.98
Lu and Yin [28]	MALS-GNMF	5-fold cross validation	86.41	98.32	71.39	83.08	82.18	84.28
Lu et al. [29]	R-CSP with aggregation	Different testing and training set for each patient	76.8	98.2	74.5	92.2	77	83.90
	R-CSP with aggregation	71.43/28.59 random, 20 times	79	88.7	69.7	83.9	89.6	82.2
	R-CSP with aggregation	57.14/42.86 random, 20 times	77.60	89.00	68.5	81.90	88.30	81.1
Lotte et al [30]	SSRCSP	40/60	70.54	96.43	53.57	71.88	75.39	73.56
	TRCSP		71.43	96.43	63.27	71.88	86.9	77.98
	WTRCSP		69.64	98.21	54.59	71.88	85.32	75.93
	SRCSP		72.32	96.43	60.2	77.68	86.51	78.63

0.25% improvement in terms of average accuracy over the recently reported methods in the MI tasks classification. It is also noted from Table 5.8, that the classification performance of LS-SVM classifier are improved by using the proposed features extraction compared to other features extraction methods in [25, 26]. These results indicate the effectiveness of the proposed PCA-CCOV feature extraction method.

5.5 Conclusions

This chapter presents a robust feature extraction method for extracting discriminatory information from the mental states based EEG signals. The PCA and CCOV are combinedly used for feature extraction from EEG signals. The correlation based variable selection method with the best first search is utilised to select the optimal feature set from the extracted features. To assess the performance of the proposed method, three

machine learning techniques; MLP, LS-SVM, and LR classifiers are employed on the obtained features. In addition, the proposed method is compared with some existing methods. The experimental results report that the proposed PCA-CCOV is very effective for feature extraction for MI-based EEG signals in BCI systems and the two features, *Maximum* and *Mode*, are the best features to characterise the distribution of MI tasks signal. The results show that both proposed PCA-CCOV based MLP and LS-SVM classifiers achieve the highest overall performance as compared to the LR classifier. The results also indicate that the proposed methods outperform the aforementioned ten reported methods by at least 0.25% average accuracy improvement in dataset IVa.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Summary

ECG and EEG signals are the two most vital biomedical time series that provide considerable information about the health and functioning of the heart and brain, respectively. Automatic analysis and classification of ECG and EEG signals are important in various clinical applications as they can aid physicians to diagnose abnormalities and diseases based on these signals. However, accurate classification of these biomedical signals is challenging due to the chaotic nature and complexity of these signals.

This thesis has focused on developing classification methods for accurate detection of different types of ECG and EEG signals with four main objectives:

- To develop a low complexity algorithm for online premature ventricular contractions detection from ECG signals.
- To propose a method for accurate epileptic seizure detection in multi-category EEG signal.
- To introduce a method for classification of motor imagery based EEG signals in BCI system.
- To empirically evaluate the classification performance of the proposed methods on different benchmark ECG and EEG databases with some state-of-the-art methods presented in the literature.

To achieve these objectives, this study proposed three methods for the classification of different types of ECG and EEG signals. All three proposed methods were evaluated

on publicly available databases and compared to their counterparts found in the literature to verify the efficiency of the methods.

A PVC classification method was developed to effectively and efficiently detect PVC beats in an online manner (see Chapter 3). This algorithm utilises a new replacing strategy to check the effects of different types of heartbeats on the variation of principal directions to classify heartbeats into two classes: Non-PVC and PVC heartbeat. This study showed that the variation of principal directions in replacing strategy can be used as a measure to detect PVC heartbeats. As the proposed method requires to extract the principal directions for each heartbeat, an efficient computation for estimating principal directions based on the power method algorithm is used to reduce the computational complexity of the proposed method. The PVC detection method was experimentally validated using two non-overlapping data sets from a public ECG database through comparison with some methods reported in the literature. The experimental results demonstrate that the proposed method achieves not only high classification performance but is also capable of detecting PVC beats in an online manner. The results also showed that the proposed method outperforms the existing state-of-the-art methods. This technique can aid clinicians to diagnose PVC arrhythmia disease more accurately.

In order to accurately classify epileptic seizure in multi-category EEG signal, a novel feature extraction technique based on the DP and the PCA was proposed to extract discriminatory information and features from epileptic EEG data (see Chapter 4). In this method, the DP algorithm is used to extract the most representative samples from EEG signals. Then the PCA is applied to produce uncorrelated variables and to reduce the dimensionality of the DP samples for better recognition. To verify the robustness of the proposed method, four machine learning techniques: RF, k -NN, SVM, and DT classifiers are employed on the obtained features for classification of multi-category EEG sig-

nals. The performance of the proposed method was experimentally validated using the 10-fold cross-validation procedure on a benchmark epileptic EEG dataset. Moreover, the performance of the proposed method was accessed by comparing it with the six most recently reported methods in the literature. The experimental results showed that the RF classifier achieved the best performance on the proposed features set compared to other three classifiers tested. The experimental results also indicated that the proposed method outperforms the other six recently reported methods in terms of sensitivity, specificity and classification accuracy. The research findings indicate that this proposed method is very effective for accurate epileptic seizure detection in multi-category EEG signal which can be used to provide clinical information about patients who have epilepsy.

A robust feature extraction technique based on the PCA and CCOV was introduced for the extraction of discriminatory features from MI-based EEG signals to improve classification performances and to reduce the time complexity (see Chapter 5). In this method, a temporal pattern of MI task is extracted by discovering the inter-channel information in EEG signals using PCA. Then, the CCOV method is utilised to calculate cross-covariance sequences between the EEG signals and the temporal pattern to provide more discriminative information about those signals. From each cross-covariance sequence, eleven statistical features are calculated to characterise the distribution of EEG signals. Then, the correlation-based variable selection method is applied to find optimal features from the extracted features. To verify the effectiveness of the proposed method, three machine learning techniques: MLP, LS-SVM, and LR are employed on the obtained optimal features to classify the MI-based EEG signals. The proposed method was tested on two benchmark datasets (IVa and IVb of BCI Competition III) and the effectiveness of proposed method was evaluated by comparing it with ten most recently reported algorithms in the literature. The experimental results demonstrated that all three classifiers achieved high classification performance and outperform the other ten recently reported

methods in terms of average accuracy. The findings of this study also indicated that the proposed feature extraction technique can effectively extract characteristic information from MI-based EEG signals and also reduce the computational complexity of classifiers by reducing the number of extracted features.

classifiers

6.2 Future Work

The ECG and EEG classification methods developed as part of this research project have shown high performances in detecting PVC beats, epileptic seizure and MI tasks signals. However, there are still some recommendations for future research, as mentioned below:

- The ECG classification method proposed in this thesis is only focused on detecting PVC arrhythmia. However, the proposed algorithm has the potential to be extended to classify other heartbeat arrhythmias. To achieve this aim, further study is needed to explore the effects of different types of arrhythmias on the variation of principal directions in the proposed replacing strategy.
- Concerning the proposed algorithm for MI task classification, it can be extended for the classification of multi-class EEG signals.
- All proposed methods in this thesis are based either on ECG or EEG signal analysis. However, the combination of ECG and EEG signals analysis can be used for some medical diagnostics such as seizures detection and sleep staging. Concerning the proposed algorithm based on EEG signals for seizure detection, it could be improved by using fused features extracted from EEG and ECG signals. To reach this aim, further study is needed to extract different features from ECG signals

and to investigate the effects of these additional features on the performance of the proposed seizures detection method. In the future, the proposed method will be extended for the detection of seizures using the combination of ECG and EEG features.

- In practical cases, ECG and EEG signals are highly contaminated with various noises and artefacts both from the subject and from equipment interferences. In the methods proposed in this dissertation, removing artefacts and noises were not considered. For future works, implementing noise analysis for all the methods can be considered in classifying noisy EEG and ECG signals.

BIBLIOGRAPHY

- [1] P. De Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2535–2543, 2006.
- [2] A. H. Khandoker, M. Palaniswami, and C. K. Karmakar, "Support vector machines for automated recognition of obstructive sleep apnea syndrome from ECG recordings," *IEEE transactions on information technology in biomedicine*, vol. 13, no. 1, pp. 37–48, 2009.
- [3] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: a new approach in human identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 3, pp. 808–812, 2001.
- [4] I. Guler and E. D. Ubeyli, "Multiclass support vector machines for EEG-signals classification," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 2, pp. 117–126, 2007.
- [5] K.-Q. Shen, C.-J. Ong, X.-P. Li, Z. Hui, and E. P. Wilder-Smith, "A feature selection method for multilevel mental fatigue EEG classification," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 7, pp. 1231–1237, 2007.
- [6] W. Wu, X. Gao, B. Hong, and S. Gao, "Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL)," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 6, pp. 1733–1743, 2008.
- [7] A. Iwasa, M. Hwa, A. Hassankhani, T. Liu, and S. M. Narayan, "Abnormal heart rate turbulence predicts the initiation of ventricular arrhythmias," *Pacing and clinical electrophysiology*, vol. 28, no. 11, pp. 1189–1197, 2005.
- [8] F. Santoro, L. D. BIASE, P. Hranitzky, J. E. Sanchez, P. Santangeli, A. P. Perini, J. D. Burkhardt, and A. Natale, "Ventricular fibrillation triggered by PVCs from papillary muscles: clinical features and ablation," *Journal of cardiovascular electrophysiology*, vol. 25, no. 11, pp. 1158–1164, 2014.
- [9] H. Watanabe, N. Tanabe, Y. Makiyama, S. S. Chopra, Y. Okura, H. Suzuki, K. Matsui, T. Watanabe, Y. Kurashina, and Y. Aizawa, "ST-segment abnormalities and premature complexes are predictors of new-onset atrial fibrillation: the niigata preventive medicine study," *American heart journal*, vol. 152, no. 4, pp. 731–735, 2006.

- [10] D. J. Thurman, E. Beghi, C. E. Begley, A. T. Berg, J. R. Buchhalter, D. Ding, D. C. Hesdorffer, W. A. Hauser, L. Kazis, R. Kobau *et al.*, “Standards for epidemiologic studies and surveillance of epilepsy,” *Epilepsia*, vol. 52, no. s7, pp. 2–26, 2011.
- [11] M. Z. Parvez and M. Paul, “Epileptic seizure detection by analyzing EEG signals using different transformation techniques,” *Neurocomputing*, vol. 145, pp. 190–200, 2014.
- [12] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain–computer interfaces for communication and control,” *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [13] L.-Y. Shyu, Y.-H. Wu, and W. Hu, “Using wavelet transform and fuzzy neural network for VPC detection from the holter ECG,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1269–1273, 2004.
- [14] O. T. Inan, L. Giovangrandi, and G. T. Kovacs, “Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2507–2515, 2006.
- [15] J. S. Lim, “Finding features for real-time premature ventricular contraction detection using a fuzzy neural network system,” *IEEE Transactions on Neural Networks*, vol. 20, no. 3, pp. 522–527, 2009.
- [16] P. Li, C. Liu, X. Wang, D. Zheng, Y. Li, and C. Liu, “A low-complexity data-adaptive approach for premature ventricular contraction recognition,” *Signal, Image and Video Processing*, vol. 8, no. 1, pp. 111–120, 2014.
- [17] E. D. Übeyli, “Decision support systems for time-varying biomedical signals: EEG signals classification,” *Expert Systems with Applications*, vol. 36, no. 2, pp. 2275–2284, 2009.
- [18] E. D. Übeyli, “Analysis of EEG signals by combining eigenvector methods and multiclass support vector machines,” *Computers in Biology and Medicine*, vol. 38, no. 1, pp. 14–22, 2008.
- [19] E. D. Übeyli, “Lyapunov exponents/probabilistic neural networks for analysis of EEG signals,” *Expert Systems with Applications*, vol. 37, no. 2, pp. 985–992, 2010.

- [20] Z. Emigdio, L. Trujillo, A. Sotelo, P. Legrand, L. N. Coria *et al.*, “Regularity and matching pursuit feature extraction for the detection of epileptic seizures,” *Journal of neuroscience methods*, vol. 266, pp. 107–125, 2016.
- [21] A. M. Murugavel and S. Ramakrishnan, “Hierarchical multi-class SVM with ELM kernel for epileptic EEG signal classification,” *Medical & biological engineering & computing*, vol. 54, no. 1, pp. 149–161, 2016.
- [22] S. Bhattacharyya, A. Sengupta, T. Chakraborti, A. Konar, and D. Tibarewala, “Automatic feature selection of motor imagery EEG signals using differential evolution and learning automata,” *Medical & biological engineering & computing*, vol. 52, no. 2, pp. 131–139, 2014.
- [23] S. Siuly, H. Wang, and Y. Zhang, “Detection of motor imagery EEG signals employing naïve bayes based learning process,” *Measurement*, vol. 86, pp. 148–158, 2016.
- [24] N. F. Ince, F. Goksu, A. H. Tewfik, and S. Arica, “Adapting subject specific motor imagery EEG patterns in space–time–frequency for a brain computer interface,” *Biomedical signal processing and control*, vol. 4, no. 3, pp. 236–246, 2009.
- [25] S. Siuly and Y. Li, “Improving the separability of motor imagery EEG signals using a cross correlation-based least square support vector machine for brain–computer interface,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 4, pp. 526–538, 2012.
- [26] S. Siuly, Y. Li, and P. P. Wen, “Clustering technique-based least square support vector machine for EEG signal classification,” *Computer methods and programs in biomedicine*, vol. 104, no. 3, pp. 358–372, 2011.
- [27] L. Song, E. Gordon, and E. Gysels, “Phase synchrony rate for the recognition of motor imagery in brain-computer interface,” in *Advances in neural information processing systems*, 2005, pp. 1265–1272.
- [28] N. Lu and T. Yin, “Motor imagery classification via combinatory decomposition of ERP and ERSP using sparse nonnegative matrix factorization,” *Journal of neuroscience methods*, vol. 249, pp. 41–49, 2015.
- [29] H. Lu, H.-L. Eng, C. Guan, K. N. Plataniotis, and A. N. Venetsanopoulos, “Regularized common spatial pattern with aggregation for EEG classification in small-sample setting,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 12, pp. 2936–2946, 2010.

- [30] F. Lotte and C. Guan, “Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms,” *IEEE Transactions on biomedical Engineering*, vol. 58, no. 2, pp. 355–362, 2011.
- [31] D. Venes, *Taber’s cyclopedic medical dictionary*. FA Davis, 2013.
- [32] H. Gray, *Anatomy of the human body*. Lea & Febiger, 1918.
- [33] C. Starr, C. Evers, and L. Starr, *Biology today and tomorrow with physiology*. Cengage Learning, 2015.
- [34] S. Standring, *Gray’s anatomy: the anatomical basis of clinical practice*. Elsevier Health Sciences, 2015.
- [35] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, USA, 1995.
- [36] J. G. Betts, “Anatomy & physiology,” *Open Stax College*, pp. 787–846, 2013.
- [37] J. P. Keener and J. Sneyd, *Mathematical physiology*. Springer, 2009, vol. 1.
- [38] M. Endo, “Calcium release from the sarcoplasmic reticulum,” *Physiological Reviews*, vol. 57, no. 1, pp. 71–108, 1977.
- [39] E. Marbán, “Cardiac channelopathies,” *Nature*, vol. 415, no. 6868, pp. 213–218, 2002.
- [40] W. J. Germann, C. L. Stanfield, and J. G. Cannon, *Principles of human physiology*. Benjamin Cummings San Francisco CA, 2002, vol. 1.
- [41] D. G. Zill and J. M. Dewar, *Precalculus with calculus previews*. Jones & Bartlett Publishers, 2009.
- [42] R. Gupta, M. Mitra, and J. Bera, *ECG acquisition and automated remote processing*. Springer, 2014.
- [43] T. Baas, *ECG based analysis of the ventricular repolarisation in the human heart*. KIT Scientific Publishing, 2012, vol. 18.
- [44] W. Einthoven, G. Fahr, and A. De Waart, “Über die richtung und die manifeste grösse der potentialschwankungen im menschlichen herzen und über den einfluss

der herzlage auf die form des elektrokardiogramms,” *Pflügers Archiv European Journal of Physiology*, vol. 150, no. 6, pp. 275–315, 1913.

- [45] E. Goldberger, “A simple, indifferent, electrocardiographic electrode of zero potential and a technique of obtaining augmented, unipolar, extremity leads,” *American Heart Journal*, vol. 23, no. 4, pp. 483–492, 1942.
- [46] K. Wesley, *Huszar’s ECG and 12-Lead Interpretation*. Elsevier Health Sciences, 2016.
- [47] A. S. Go, D. Mozaffarian, V. L. Roger, E. J. Benjamin, J. D. Berry, M. J. Blaha, S. Dai, E. S. Ford, C. S. Fox, S. Franco *et al.*, “Heart disease and stroke statistics-2014 update,” *Circulation*, vol. 129, no. 3, 2014.
- [48] “Causes of death 2015 - australian bureau of statistics,” <http://www.abs.gov.au/ausstats/abs@.nsf/mf/3303.0>, accessed: 2017-03-19.
- [49] H. L. Kennedy, J. A. Whitlock, M. K. Sprague, L. J. Kennedy, T. A. Buckingham, and R. J. Goldberg, “Long-term follow-up of asymptomatic healthy subjects with frequent and complex ventricular ectopy,” *New England Journal of Medicine*, vol. 312, no. 4, pp. 193–197, 1985.
- [50] F. Gaita, C. Giustetto, P. Di Donna, E. Richiardi, L. Libero, M. C. R. Brusin, G. Molinari, and G. Trevis, “Long-term follow-up of right ventricular monomorphic extrasystoles,” *Journal of the American College of Cardiology*, vol. 38, no. 2, pp. 364–370, 2001.
- [51] S. K. Agarwal, G. Heiss, P. M. Rautaharju, E. Shahar, M. W. Massing, and R. J. Simpson, “Premature ventricular complexes and the risk of incident stroke,” *Stroke*, vol. 41, no. 4, pp. 588–593, 2010.
- [52] X. Jouven, M. Zureik, M. Desnos, D. Courbon, and P. Ducimetière, “Long-term outcome in asymptomatic men with exercise-induced premature ventricular depolarizations,” *New England Journal of Medicine*, vol. 343, no. 12, pp. 826–833, 2000.
- [53] J. P. Frolkis, C. E. Pothier, E. H. Blackstone, and M. S. Lauer, “Frequent ventricular ectopy after exercise as a predictor of death,” *New England Journal of Medicine*, vol. 348, no. 9, pp. 781–790, 2003.
- [54] M. Reed and A. Gray, “Collapse query cause: the management of adult syncope

- in the emergency department,” *Emergency medicine journal*, vol. 23, no. 8, pp. 589–594, 2006.
- [55] A. Garcia-Touchard, V. K. Somers, T. Kara, J. Nykodym, A. Shamsuzzaman, P. Lanfranchi, and M. J. Ackerman, “Ventricular ectopy during REM sleep: implications for nocturnal sudden cardiac death,” *Nature Clinical Practice Cardiovascular Medicine*, vol. 4, no. 5, pp. 284–288, 2007.
 - [56] D. Purves, G. J. Augustine, D. Fitzpatrick, L. C. Katz, A.-S. LaMantia, J. O. McNamara, and S. M. Williams, *Neuroscience*. Sinauer associates, third edition, Sunderland, Massachusetts, USA., 2004.
 - [57] D. J. Cech and S. T. Martin, *Functional movement development across the life span*. Elsevier Health Sciences, 2002.
 - [58] S. B. Rutkove and A. Blum, *The clinical neurophysiology primer*. Springer, 2007.
 - [59] S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2013.
 - [60] C. Burger, “A novel method of improving EEG signals for BCI classification,” Ph.D. dissertation, Stellenbosch: Stellenbosch University, 2014.
 - [61] H. Jasper, “Report of the committee on methods of clinical examination in electroencephalography,” *Electroencephalography and Clinical Neurophysiology*, vol. 10, pp. 370–375, 1958.
 - [62] R. Cooper, J. W. Osselton, and J. C. Shaw, *EEG technology*. Butterworth-Heinemann, 2014.
 - [63] J. B. Ochoa, “EEG signal classification for brain computer interface applications,” *Ecole Polytechnique Federale de Lausanne*, vol. 7, pp. 1–72, 2002.
 - [64] V. K. Manchala, “Human computer interface using electroencephalography,” Ph.D. dissertation, ARIZONA STATE UNIVERSITY, 2015.
 - [65] S. Fazel, A. Wolf, N. Långström, C. R. Newton, and P. Lichtenstein, “Premature mortality in epilepsy and the role of psychiatric comorbidity: a total population study,” *The Lancet*, vol. 382, no. 9905, pp. 1646–1654, 2013.
 - [66] L. Roller, J. Gowan *et al.*, “Disease state management: Epilepsy,” *AJP: The Australian Journal of Pharmacy*, vol. 97, no. 1155, p. 62, 2016.

- [67] M. Le Van Quyen, J. Martinerie, V. Navarro, P. Boon, M. D'Havé, C. Adam, B. Renault, F. Varela, and M. Baulac, "Anticipation of epileptic seizures from standard EEG recordings," *The Lancet*, vol. 357, no. 9251, pp. 183–188, 2001.
- [68] B. Koo, H.-G. Lee, Y. Nam, H. Kang, C. S. Koh, H.-C. Shin, and S. Choi, "A hybrid NIRS-EEG system for self-paced brain computer interface with online motor imagery," *Journal of neuroscience methods*, vol. 244, pp. 26–32, 2015.
- [69] S. G. Mason and G. E. Birch, "A general framework for brain-computer interface design," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 11, no. 1, pp. 70–85, 2003.
- [70] S. Siuly, Y. Li, and Y. Zhang, *EEG Signal Analysis and Classification: Techniques and Applications*. Springer, 2017.
- [71] P. B. LUKÁŠ, "Wheelchair control using eeg signal classification," Master's thesis, BRNO UNIVERSITY OF TECHNOLOGY, 2015.
- [72] 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, p. R32, 2007.
- [73] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor, "A spelling device for the paralysed," *Nature*, vol. 398, no. 6725, pp. 297–298, 1999.
- [74] B. Graimann, B. Allison, C. Mandel, T. Lüth, D. Valbuena, and A. Gräser, "Non-invasive brain-computer interfaces for semi-autonomous assistive devices," in *Robust intelligent systems*. Springer, 2008, pp. 113–138.
- [75] C. Escolano, A. R. Murguialday, T. Matuz, N. Birbaumer, and J. Minguez, "A telepresence robotic system operated with a P300-based brain-computer interface: initial tests with ALS patients," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2010, pp. 4476–4480.
- [76] B. Blankertz, M. Tangermann, C. Vidaurre, S. Fazli, C. Sannelli, S. Haufe, C. Maeder, L. E. Ramsey, I. Sturm, G. Curio *et al.*, "The Berlin brain-computer interface: non-medical uses of BCI technology," *Frontiers in neuroscience*, vol. 4, p. 198, 2010.
- [77] S. Haufe, M. S. Treder, M. F. Gugler, M. Sagebaum, G. Curio, and B. Blankertz,

- “EEG potentials predict upcoming emergency brakings during simulated driving,” *Journal of neural engineering*, vol. 8, no. 5, p. 056001, 2011.
- [78] R. N. Roy, S. Bonnet, S. Charbonnier, and A. Campagne, “Mental fatigue and working memory load estimation: interaction and implications for EEG-based passive BCI,” in *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2013, pp. 6607–6610.
 - [79] S. J. Johnston, S. G. Boehm, D. Healy, R. Goebel, and D. E. Linden, “Neurofeedback: A promising tool for the self-regulation of emotion networks,” *Neuroimage*, vol. 49, no. 1, pp. 1066–1072, 2010.
 - [80] V. Zotev, R. Phillips, H. Yuan, M. Misaki, and J. Bodurka, “Self-regulation of human brain activity using simultaneous real-time fMRI and EEG neurofeedback,” *NeuroImage*, vol. 85, pp. 985–995, 2014.
 - [81] L. Bonnet, F. Lotte, and A. Lécuyer, “Two brains, one game: design and evaluation of a multiuser BCI video game based on motor imagery,” *IEEE Transactions on Computational Intelligence and AI in games*, vol. 5, no. 2, pp. 185–198, 2013.
 - [82] D. Tan and A. Nijholt, *Brain-computer interfaces and human-computer interaction*. Springer, 2010.
 - [83] P. De Chazal, M. O’Dwyer, and R. B. Reilly, “Automatic classification of heartbeats using ECG morphology and heartbeat interval features,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1196–1206, 2004.
 - [84] J. Moraes, M. Seixas, F. Vilani, and E. Costa, “A real time QRS complex classification method using Mahalanobis distance,” in *Computers in Cardiology, 2002*. IEEE, 2002, pp. 201–204.
 - [85] O. Wieben, V. Afonso, and W. Tompkins, “Classification of premature ventricular complexes using filter bank features, induction of decision trees and a fuzzy rule-based system,” *Medical and Biological Engineering and Computing*, vol. 37, no. 5, pp. 560–565, 1999.
 - [86] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, and L. Sörnmo, “Clustering ECG complexes using hermite functions and self-organizing maps,” *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 7, pp. 838–848, 2000.
 - [87] İ. Güler and E. D. Übeyli, “ECG beat classifier designed by combined neural network model,” *Pattern recognition*, vol. 38, no. 2, pp. 199–208, 2005.

- [88] L. Senhadji, G. Carrault, J. Bellanger, and G. Passariello, "Comparing wavelet transforms for recognizing cardiac patterns," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 2, pp. 167–173, 1995.
- [89] Y. H. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach," *IEEE transactions on biomedical engineering*, vol. 44, no. 9, pp. 891–900, 1997.
- [90] M. I. Owis, A. H. Abou-Zied, A.-B. Youssef, and Y. M. Kadah, "Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification," *IEEE transactions on Biomedical Engineering*, vol. 49, no. 7, pp. 733–736, 2002.
- [91] E. D. Übeyli, "Adaptive neuro-fuzzy inference system for classification of ECG signals using Lyapunov exponents," *Computer methods and programs in biomedicine*, vol. 93, no. 3, pp. 313–321, 2009.
- [92] A. Lanata, G. Valenza, C. Mancuso, and E. P. Scilingo, "Robust multiple cardiac arrhythmia detection through bispectrum analysis," *Expert Systems with Applications*, vol. 38, no. 6, pp. 6798–6804, 2011.
- [93] M. L. Talbi and A. Charef, "PVC discrimination using the QRS power spectrum and self-organizing maps," *Computer methods and programs in biomedicine*, vol. 94, no. 3, pp. 223–231, 2009.
- [94] H. Al-Nashash, "Cardiac arrhythmia classification using neural networks," *Technology and Health Care*, vol. 8, no. 6, pp. 363–372, 2000.
- [95] I. Christov and G. Bortolan, "Ranking of pattern recognition parameters for premature ventricular contractions classification by neural networks," *Physiological Measurement*, vol. 25, no. 5, p. 1281, 2004.
- [96] F. Melgani and Y. Bazi, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," *IEEE transactions on information technology in biomedicine*, vol. 12, no. 5, pp. 667–677, 2008.
- [97] E. J. D. S. Luz, T. M. Nunes, V. H. C. De Albuquerque, J. P. Papa, and D. Menotti, "ECG arrhythmia classification based on optimum-path forest," *Expert Systems with Applications*, vol. 40, no. 9, pp. 3561–3573, 2013.
- [98] R. Mark and G. Moody, "MIT-BIH arrhythmia database 1997," URL <http://ecgmit.edu/dbinfo.html>, 1997.

- [99] M. G. Tsipouras, D. I. Fotiadis, and D. Sideris, "An arrhythmia classification system based on the RR-interval signal," *Artificial intelligence in medicine*, vol. 33, no. 3, pp. 237–250, 2005.
- [100] S.-W. Chen, "A nonlinear trimmed moving averaging-based system with its application to real-time QRS beat classification," *Journal of medical engineering & technology*, vol. 31, no. 6, pp. 443–449, 2007.
- [101] V. Krasteva and I. Jekova, "QRS template matching for recognition of ventricular ectopic beats," *Annals of Biomedical Engineering*, vol. 35, no. 12, pp. 2065–2076, 2007.
- [102] S. Hu, Z. Shao, and J. Tan, "A real-time cardiac arrhythmia classification system with wearable electrocardiogram," in *International Conference on Body Sensor Networks (BSN)*. IEEE, 2011, pp. 119–124.
- [103] A. Daamouche, L. Hamami, N. Alajlan, and F. Melgani, "A wavelet optimization approach for ECG signal classification," *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 342–349, 2012.
- [104] X. Liu, H. Du, G. Wang, S. Zhou, and H. Zhang, "Automatic diagnosis of premature ventricular contraction based on lyapunov exponents and LVQ neural network," *Computer methods and programs in biomedicine*, vol. 122, no. 1, pp. 47–55, 2015.
- [105] M. L. Talbi and P. Ravier, "Detection of PVC in ECG signals using fractional linear prediction," *Biomedical Signal Processing and Control*, vol. 23, pp. 42–51, 2016.
- [106] Ö. F. Alçın, S. Siuly, V. Bajaj, Y. Guo, A. Şengu, Y. Zhang *et al.*, "Multi-category EEG signal classification developing time-frequency texture features based Fisher Vector encoding method," *Neurocomputing*, vol. 218, pp. 251–258, 2016.
- [107] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated EEG analysis of epilepsy: a review," *Knowledge-Based Systems*, vol. 45, pp. 147–165, 2013.
- [108] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 2, pp. 512–518, 2008.

- [109] A. Subasi and M. I. Gursay, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8659–8666, 2010.
- [110] U. R. Acharya, S. V. Sree, A. P. C. Alvin, and J. S. Suri, "Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9072–9078, 2012.
- [111] S. Xie and S. Krishnan, "Dynamic principal component analysis with nonoverlapping moving window and its applications to epileptic EEG classification," *The Scientific World Journal*, vol. 2014, 2014.
- [112] T.-P. Jung, S. Makeig, M. J. McKeown, A. J. Bell, T.-W. Lee, and T. J. Sejnowski, "Imaging brain dynamics using independent component analysis," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1107–1122, 2001.
- [113] K. Polat and S. Güneş, "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast fourier transform," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 1017–1026, 2007.
- [114] E. D. Übeyli, "Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals," *Expert Systems with Applications*, vol. 37, no. 1, pp. 233–239, 2010.
- [115] V. Srinivasan, C. Eswaran, and N. Sriraam, "Artificial neural network based epileptic detection using time-domain and frequency-domain features," *Journal of Medical Systems*, vol. 29, no. 6, pp. 647–660, 2005.
- [116] R. B. Pachori and P. Sircar, "EEG signal analysis using fb expansion and second-order linear TVAR process," *Signal Processing*, vol. 88, no. 2, pp. 415–420, 2008.
- [117] H. Ocak, "Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2027–2036, 2009.
- [118] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection," *IEEE transactions on biomedical engineering*, vol. 54, no. 9, pp. 1545–1551, 2007.
- [119] A. Subasi, "EEG signal classification using wavelet feature extraction and a mix-

- ture of expert model,” *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084–1093, 2007.
- [120] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, “Epileptic seizure detection in EEGs using time–frequency analysis,” *IEEE transactions on information technology in biomedicine*, vol. 13, no. 5, pp. 703–710, 2009.
 - [121] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, “Automatic seizure detection based on time-frequency analysis and artificial neural networks,” *Computational Intelligence and Neuroscience*, vol. 2007, 2007.
 - [122] L. Guo, D. Rivero, and A. Pazos, “Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks,” *Journal of neuroscience methods*, vol. 193, no. 1, pp. 156–163, 2010.
 - [123] R. J. Martis, U. R. Acharya, J. H. Tan, A. Petznick, R. Yanti, C. K. Chua, E. K. Ng, and L. Tong, “Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals,” *International journal of neural systems*, vol. 22, no. 06, p. 1250027, 2012.
 - [124] R. B. Pachori, “Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition,” *Research Letters in Signal Processing*, vol. 2008, p. 14, 2008.
 - [125] A. Subasi, “Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction,” *Computers in Biology and Medicine*, vol. 37, no. 2, pp. 227–244, 2007.
 - [126] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, “Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis,” *Seizure*, vol. 26, pp. 56–64, 2015.
 - [127] S.-H. Lee, J. S. Lim, J.-K. Kim, J. Yang, and Y. Lee, “Classification of normal and epileptic seizure EEG signals using wavelet transform, phase-space reconstruction, and Euclidean distance,” *Computer methods and programs in biomedicine*, vol. 116, no. 1, pp. 10–25, 2014.
 - [128] R. B. Pachori and S. Patidar, “Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions,” *Computer methods and programs in biomedicine*, vol. 113, no. 2, pp. 494–502, 2014.
 - [129] K. Samiee, P. Kovacs, and M. Gabbouj, “Epileptic seizure classification of EEG

- time-series using rational discrete short-time Fourier transform,” *IEEE transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 541–552, 2015.
- [130] U. R. Acharya, C. K. Chua, T.-C. Lim, Dorithy, and J. S. Suri, “Automatic identification of epileptic EEG signals using nonlinear parameters,” *Journal of Mechanics in Medicine and Biology*, vol. 9, no. 04, pp. 539–553, 2009.
 - [131] K. C. Chua, V. Chandran, U. R. Acharya, and C. M. Lim, “Application of higher order spectra to identify epileptic EEG,” *Journal of Medical Systems*, vol. 35, no. 6, pp. 1563–1571, 2011.
 - [132] U. R. Acharya, S. V. Sree, and J. S. Suri, “Automatic detection of epileptic EEG signals using higher order cumulant features,” *International journal of neural systems*, vol. 21, no. 05, pp. 403–414, 2011.
 - [133] K. Lehnertz and C. Elger, “Spatio-temporal dynamics of the primary epileptogenic area in temporal lobe epilepsy characterized by neuronal complexity loss,” *Electroencephalography and clinical Neurophysiology*, vol. 95, no. 2, pp. 108–117, 1995.
 - [134] R. G. Andrzejak, G. Widman, K. Lehnertz, C. Rieke, P. David, and C. Elger, “The epileptic process as nonlinear deterministic dynamics in a stochastic environment: an evaluation on mesial temporal lobe epilepsy,” *Epilepsy research*, vol. 44, no. 2, pp. 129–140, 2001.
 - [135] N. F. Güler, E. D. Übeyli, and I. Güler, “Recurrent neural networks employing lyapunov exponents for EEG signals classification,” *Expert systems with applications*, vol. 29, no. 3, pp. 506–514, 2005.
 - [136] A. Accardo, M. Affinito, M. Carrozzi, and F. Bouquet, “Use of the fractal dimension for the analysis of electroencephalographic time series,” *Biological cybernetics*, vol. 77, no. 5, pp. 339–350, 1997.
 - [137] N. Kannathal, M. L. Choo, U. R. Acharya, and P. Sadasivan, “Entropies for detection of epilepsy in EEG,” *Computer methods and programs in biomedicine*, vol. 80, no. 3, pp. 187–194, 2005.
 - [138] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K.-H. Ng, and J. S. Suri, “Automated diagnosis of epileptic EEG using entropies,” *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 401–408, 2012.
 - [139] V. Srinivasan, C. Eswaran, and N. Sriraam, “Approximate entropy-based epileptic

- EEG detection using artificial neural networks,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 3, pp. 288–295, 2007.
- [140] Y. Song, P. Liò *et al.*, “A new approach for epileptic seizure detection: sample entropy based feature extraction and extreme learning machine,” *Journal of Biomedical Science and Engineering*, vol. 3, no. 06, p. 556, 2010.
 - [141] Y. Song, J. Crowcroft, and J. Zhang, “Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine,” *Journal of neuroscience methods*, vol. 210, no. 2, pp. 132–146, 2012.
 - [142] U. R. Acharya, S. V. Sree, S. Chattopadhyay, W. Yu, and P. C. A. Ang, “Application of recurrence quantification analysis for the automated identification of epileptic EEG signals,” *International journal of neural systems*, vol. 21, no. 03, pp. 199–211, 2011.
 - [143] M. Niknazar, S. Mousavi, B. V. Vahdat, and M. Sayyah, “A new framework based on recurrence quantification analysis for epileptic seizure detection,” *IEEE journal of biomedical and health informatics*, vol. 17, no. 3, pp. 572–578, 2013.
 - [144] 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. 13 475–13 481, 2011.
 - [145] U. R. Acharya, S. V. Sree, P. C. A. Ang, R. Yanti, and J. S. Suri, “Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals,” *International journal of neural systems*, vol. 22, no. 02, p. 1250002, 2012.
 - [146] O. Faust, U. R. Acharya, L. C. Min, and B. H. Spath, “Automatic identification of epileptic and background EEG signals using frequency domain parameters,” *International journal of neural systems*, vol. 20, no. 02, pp. 159–176, 2010.
 - [147] A. UR, “Automated diagnosis of epileptic electroencephalogram using independent component analysis and discrete wavelet transform for different electroencephalogram durations,” *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 227, no. 3, pp. 234–244, 2013.
 - [148] C. A. Lima, A. L. Coelho, and M. Eisenkraft, “Tackling EEG signal classification with least squares support vector machines: a sensitivity analysis study,” *Computers in Biology and Medicine*, vol. 40, no. 8, pp. 705–714, 2010.

- [149] S. Siuly and Y. Li, “Designing a robust feature extraction method based on optimum allocation and principal component analysis for epileptic EEG signal classification,” *Computer methods and programs in biomedicine*, vol. 119, no. 1, pp. 29–42, 2015.
- [150] K. Polat and S. Güneş, “A novel data reduction method: distance based data reduction and its application to classification of epileptiform EEG signals,” *Applied Mathematics and Computation*, vol. 200, no. 1, pp. 10–27, 2008.
- [151] S. Siuly, E. Kabir, H. Wang, and Y. Zhang, “Exploring sampling in the detection of multicategory EEG signals,” *Computational and mathematical methods in medicine*, vol. 2015, 2015.
- [152] D. Wang, D. Miao, and C. Xie, “Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection,” *Expert Systems with Applications*, vol. 38, no. 11, pp. 14 314–14 320, 2011.
- [153] B. Obermaier, C. Guger, C. Neuper, and G. Pfurtscheller, “Hidden markov models for online classification of single trial EEG data,” *Pattern recognition letters*, vol. 22, no. 12, pp. 1299–1309, 2001.
- [154] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer, “EEG-based discrimination between imagination of right and left hand movement,” *Electroencephalography and clinical Neurophysiology*, vol. 103, no. 6, pp. 642–651, 1997.
- [155] W. D. Penny, S. J. Roberts, E. A. Curran, and M. J. Stokes, “EEG-based communication: a pattern recognition approach,” *IEEE transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 214–215, 2000.
- [156] A. Schlögl, C. Neuper, and G. Pfurtscheller, “Estimating the mutual information of an EEG-based brain-computer interface,” *Biomedizinische Technik/Biomedical Engineering*, vol. 47, no. 1-2, pp. 3–8, 2002.
- [157] G. Pfurtscheller, C. Neuper, A. Schlogl, and K. Lugger, “Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters,” *IEEE transactions on Rehabilitation Engineering*, vol. 6, no. 3, pp. 316–325, 1998.
- [158] D. P. Burke, S. P. Kelly, P. de Chazal, R. B. Reilly, and C. Finucane, “A parametric feature extraction and classification strategy for brain-computer interfacing,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 1, pp. 12–17, 2005.

- [159] B. H. Jansen, J. R. Bourne, and J. W. Ward, "Autoregressive estimation of short segment spectra for computerized EEG analysis," *IEEE Transactions on Biomedical Engineering*, vol. BME-28, no. 9, pp. 630–638, 1981.
- [160] D. J. Krusienski, D. J. McFarland, and J. R. Wolpaw, "Value of amplitude, phase, and coherence features for a sensorimotor rhythm-based brain–computer interface," *Brain research bulletin*, vol. 87, no. 1, pp. 130–134, 2012.
- [161] S.-A. Park, H.-J. Hwang, J.-H. Lim, J.-H. Choi, H.-K. Jung, and C.-H. Im, "Evaluation of feature extraction methods for EEG-based brain–computer interfaces in terms of robustness to slight changes in electrode locations," *Medical & biological engineering & computing*, vol. 51, no. 5, pp. 571–579, 2013.
- [162] G. Blanchard and B. Blankertz, "BCI competition 2003-data set IIa: spatial patterns of self-controlled brain rhythm modulations," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1062–1066, 2004.
- [163] H. Ramoser, J. Muller-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.
- [164] M. Grosse-Wentrup and M. Buss, "Multiclass common spatial patterns and information theoretic feature extraction," *IEEE transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 1991–2000, 2008.
- [165] T. Demiralp, J. Yordanova, V. Kolev, A. Ademoglu, M. Devrim, and V. J. Samar, "Time–frequency analysis of single-sweep event-related potentials by means of fast wavelet transform," *Brain and language*, vol. 66, no. 1, pp. 129–145, 1999.
- [166] D. Farina, O. F. Do Nascimento, M.-F. Lucas, and C. Doncarli, "Optimization of wavelets for classification of movement-related cortical potentials generated by variation of force-related parameters," *Journal of neuroscience methods*, vol. 162, no. 1, pp. 357–363, 2007.
- [167] E. Gysels and P. Celka, "Phase synchronization for the recognition of mental tasks in a brain–computer interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 12, no. 4, pp. 406–415, 2004.
- [168] J.-P. Lachaux, E. Rodriguez, J. Martinerie, F. J. Varela *et al.*, "Measuring phase synchrony in brain signals," *Human brain mapping*, vol. 8, no. 4, pp. 194–208, 1999.

- [169] R. Q. Quiroga, “Bivariable and multivariable analysis of EEG signals,” 2009.
- [170] T. Wang and B. He, “An efficient rhythmic component expression and weighting synthesis strategy for classifying motor imagery EEG in a brain–computer interface,” *Journal of Neural Engineering*, vol. 1, no. 1, p. 1, 2004.
- [171] R. Palaniappan, “Brain computer interface design using band powers extracted during mental tasks,” in *Conference Proceedings of 2nd International IEEE EMBS Conference on Neural Engineering*. IEEE, 2005, pp. 321–324.
- [172] B. Obermaier, C. Neuper, C. Guger, and G. Pfurtscheller, “Information transfer rate in a five-classes brain-computer interface,” *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 9, no. 3, pp. 283–288, 2001.
- [173] K. Tavakolian, F. Vasefi, K. Naziripour, and S. Rezaei, “Mental task classification for brain computer interface applications,” in *Canadian Student Conference on Biomedical Computing*, 2006.
- [174] O. Dressler, G. Schneider, G. Stockmanns, and E. Kochs, “Awareness and the EEG power spectrum: analysis of frequencies,” *British journal of anaesthesia*, vol. 93, no. 6, pp. 806–809, 2004.
- [175] T. Madan, “Compression of long-term EEG using power spectral density,” Ph.D. dissertation, Concordia University, 2005.
- [176] F. Lotte, C. Guan, and K. K. Ang, “Comparison of designs towards a subject-independent brain-computer interface based on motor imagery,” in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2009)*. IEEE, 2009, pp. 4543–4546.
- [177] Q. Wei, Y. Wang, X. Gao, and S. Gao, “Amplitude and phase coupling measures for feature extraction in an EEG-based brain–computer interface,” *Journal of Neural Engineering*, vol. 4, no. 2, p. 120, 2007.
- [178] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller, “Optimizing spatial filters for robust EEG single-trial analysis,” *IEEE Signal processing magazine*, vol. 25, no. 1, pp. 41–56, 2008.
- [179] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio, and K.-R. Muller, “Combined optimization of spatial and temporal filters for improving brain-computer interfacing,” *IEEE transactions on biomedical engineering*, vol. 53, no. 11, pp. 2274–2281, 2006.

- [180] 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, p. R1, 2007.
- [181] M. M. Fouad, K. M. Amin, N. El-Bendary, and A. E. Hassanien, “Brain computer interface: A review,” in *Brain-Computer Interfaces*. Springer, 2015, pp. 3–30.
- [182] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.
- [183] G. Pfurtscheller and F. L. Da Silva, “Event-related EEG/MEG synchronization and desynchronization: basic principles,” *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [184] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, “Comparison of linear, nonlinear, and feature selection methods for EEG signal classification,” *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 11, no. 2, pp. 141–144, 2003.
- [185] G. Pfurtscheller and C. Neuper, “Motor imagery and direct brain-computer communication,” *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.
- [186] G. N. Garcia, T. Ebrahimi, and J.-M. Vesin, “Support vector EEG classification in the fourier and time-frequency correlation domains,” in *Conference Proceedings of the First International IEEE EMBS Conference on Neural Engineering*. IEEE, 2003, pp. 591–594.
- [187] A. Schlögl, F. Lee, H. Bischof, and G. Pfurtscheller, “Characterization of four-class motor imagery EEG data for the BCI-competition 2005,” *Journal of neural engineering*, vol. 2, no. 4, p. L14, 2005.
- [188] V. Gandhi, G. Prasad, D. Coyle, L. Behera, and T. M. McGinnity, “Quantum neural network-based EEG filtering for a brain–computer interface,” *IEEE transactions on neural networks and learning systems*, vol. 25, no. 2, pp. 278–288, 2014.
- [189] J. d. R. Millán, J. Mourino, F. Babiloni, F. Cincotti, M. Varsta, and J. Heikkonen, “Local neural classifier for EEG-based recognition of mental tasks,” in *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN 2000)*, vol. 3. IEEE, 2000, pp. 632–636.
- [190] K. Nakayama and K. Inagaki, “A brain computer interface based on neural net-

- work with efficient pre-processing,” in *International Symposium on Intelligent Signal Processing and Communications (ISPACS'06)*. IEEE, 2006, pp. 673–676.
- [191] S. Solhjoo and M. Moradi, “Mental task recognition: A comparison between some of classification methods,” in *BIOSIGNAL 2004 International EURASIP Conference*, 2004, pp. 24–26.
 - [192] D. Balakrishnan and S. Puthusserypady, “Multilayer perceptrons for the classification of brain computer interface data,” in *Proceedings of the IEEE 31st Annual Northeast Bioengineering Conference*. IEEE, 2005, pp. 118–119.
 - [193] A. K. 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.
 - [194] M. K. Hazrati and A. Erfanian, “An online EEG-based brain–computer interface for controlling hand grasp using an adaptive probabilistic neural network,” *Medical engineering & physics*, vol. 32, no. 7, pp. 730–739, 2010.
 - [195] W. Ting, Y. Guo-zheng, Y. Bang-hua, and S. Hong, “EEG feature extraction based on wavelet packet decomposition for brain computer interface,” *Measurement*, vol. 41, no. 6, pp. 618–625, 2008.
 - [196] T. Felzer and B. Freisieben, “Analyzing EEG signals using the probability estimating guarded neural classifier,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 4, pp. 361–371, 2003.
 - [197] T. Hoya, G. Hori, H. Bakardjian, T. Nishimura, T. Suzuki, Y. Miyawaki, A. Funase, and J. Cao, “Classification of single trial EEG signals by a combined principal+ independent component analysis and probabilistic neural network approach,” in *Proc. ICA2003*, vol. 197, 2003.
 - [198] R. Palaniappan, R. Paramesran, S. Nishida, and N. Saiwaki, “A new brain-computer interface design using fuzzy ARTMAP,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 10, no. 3, pp. 140–148, 2002.
 - [199] E. Haselsteiner and G. Pfurtscheller, “Using time-dependent neural networks for EEG classification,” *IEEE transactions on rehabilitation engineering*, vol. 8, no. 4, pp. 457–463, 2000.
 - [200] A. O. Barbosa, D. R. Achancaray, and M. A. Meggiolaro, “Activation of a mo-

- bile robot through a brain computer interface,” in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2010, pp. 4815–4821.
- [201] G. A. Barreto, R. A. Frota, and F. N. de Medeiros, “On the classification of mental tasks: a performance comparison of neural and statistical approaches,” in *Proceedings of the IEEE Workshop on Machine Learning for Signal Processing*. IEEE, 2004, pp. 529–538.
 - [202] S. Lemm, C. Schafer, and G. Curio, “BCI competition 2003-data set III: probabilistic modeling of sensorimotor/spl mu/rhythms for classification of imaginary hand movements,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1077–1080, 2004.
 - [203] P.-E. Danielsson, “Euclidean distance mapping,” *Computer Graphics and image processing*, vol. 14, no. 3, pp. 227–248, 1980.
 - [204] B. Blankertz, G. Curio, and K.-R. Müller, “Classifying single trial EEG: Towards brain computer interfacing,” *Advances in neural information processing systems*, vol. 1, pp. 157–164, 2002.
 - [205] J. F. Borisoff, S. G. Mason, A. Bashashati, and G. E. Birch, “Brain-computer interface design for asynchronous control applications: improvements to the LF-ASD asynchronous brain switch,” *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 985–992, 2004.
 - [206] F. Cincotti, A. Scipione, A. Timperi, D. Mattia, A. Marciani, J. Millan, S. Salinari, L. Bianchi, and F. Babilioni, “Comparison of different feature classifiers for brain computer interfaces,” in *In Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering*. IEEE, 2003, pp. 645–647.
 - [207] D. Ming, Y. Zhu, H. Qi, B. Wan, Y. Hu, and K. D. Luk, “Study on EEG-based mouse system by using brain-computer interface,” in *IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurements Systems (VECIMS’09)*. IEEE, 2009, pp. 236–239.
 - [208] R. Boostani and M. H. Moradi, “A new approach in the BCI research based on fractal dimension as feature and Adaboost as classifier,” *Journal of Neural Engineering*, vol. 1, no. 4, p. 212, 2004.
 - [209] U. Hoffmann, G. Garcia, J.-M. Vesin, K. Diserens, and T. Ebrahimi, “A boosting approach to P300 detection with application to brain-computer interfaces,” in *Conference Proceedings of the 2nd International IEEE EMBS Conference on Neural Engineering*. IEEE, 2005, pp. 97–100.

- [210] G. Pfurtscheller, D. Flotzinger, and J. Kalcher, "Brain-computer interface: a new communication device for handicapped persons," *Journal of Microcomputer Applications*, vol. 16, no. 3, pp. 293–299, 1993.
- [211] J. Qin, Y. Li, and A. Cichocki, "ICA and committee machine-based algorithm for cursor control in a BCI system," in *International Symposium on Neural Networks*. Springer, 2005, pp. 973–978.
- [212] A. Rakotomamonjy, V. Guigue, G. Mallet, and V. Alvarado, "Ensemble of SVMs for improving brain computer interface P300 speller performances," in *International conference on artificial neural networks*. Springer, 2005, pp. 45–50.
- [213] D. H. Wolpert, "Stacked generalization," *Neural networks*, vol. 5, no. 2, pp. 241–259, 1992.
- [214] H. Lee and S. Choi, "Pca+ hmm+ svm for EEG pattern classification," in *Proceedings of Seventh International Symposium on Signal Processing and Its Applications*, vol. 1. IEEE, 2003, pp. 541–544.
- [215] G. D. Clifford and G. B. Moody, "Signal quality in cardiorespiratory monitoring," *Physiological Measurement*, vol. 33, no. 9, p. E01, 2012.
- [216] Y. H. Hu, W. J. Tompkins, J. L. Urrusti, V. X. Afonso *et al.*, "Applications of artificial neural networks for ECG signal detection and classification," *Journal of electrocardiology*, vol. 26, pp. 66–73, 1993.
- [217] S. Osowski and T. H. Linh, "ECG beat recognition using fuzzy hybrid neural network," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 11, pp. 1265–1271, 2001.
- [218] M. Llamedo and J. P. Martínez, "Heartbeat classification using feature selection driven by database generalization criteria," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 616–625, 2011.
- [219] R. Mark and G. Moody, "MIT-BIH arrhythmia database directory," *Cambridge: Massachusetts Institute of Technology*, 1988.
- [220] G. H. Golub and C. F. Van Loan, *Matrix computations*. JHU Press, 2012, vol. 3.
- [221] R. Mark and R. Wallen, "AAMI-recommended practice: Testing and reporting performance results of ventricular arrhythmia detection algorithms," *Association*

for the Advancement of Medical Instrumentation, Arrhythmia Monitoring Subcommittee, AAMI ECAR, 1987.

- [222] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, 1985.
- [223] A. for the Advancement of Medical Instrumentation *et al.*, *Testing and Reporting Performance Results of Cardiac Rhythm and ST-segment Measurement Algorithms*. The Association, 1999.
- [224] M. Escalona-Moran, M. Soriano, I. Fischer, and C. Mirasso, "Electrocardiogram classification using reservoir computing with logistic regression." *IEEE journal of biomedical and health informatics*, vol. 19, no. 3, pp. 892–898, 2015.
- [225] R. J. Martis, U. R. Acharya, and L. C. Min, "ECG beat classification using PCA, LDA, ICA and discrete wavelet transform," *Biomedical Signal Processing and Control*, vol. 8, no. 5, pp. 437–448, 2013.
- [226] M. Nazarahari, S. G. Namin, A. H. D. Markazi, and A. K. Anaraki, "A multi-wavelet optimization approach using similarity measures for electrocardiogram signal classification," *Biomedical Signal Processing and Control*, vol. 20, pp. 142–151, 2015.
- [227] R. S. Fisher, C. Acevedo, A. Arzimanoglou, A. Bogacz, J. H. Cross, C. E. Elger, J. Engel, L. Forsgren, J. A. French, M. Glynn *et al.*, "ILAE official report: a practical clinical definition of epilepsy," *Epilepsia*, vol. 55, no. 4, pp. 475–482, 2014.
- [228] S. Siuly and Y. Zhang, "Medical big data: Neurological diseases diagnosis through medical data analysis," *Data Science and Engineering*, vol. 1, no. 2, pp. 54–64, 2016.
- [229] N. Moghim and D. W. Corne, "Predicting epileptic seizures in advance," *PloS one*, vol. 9, no. 6, p. e99334, 2014.
- [230] H. Wang, G. Zhuo, Y. Zhang *et al.*, "Analyzing EEG signal data for detection of epileptic seizure: Introducing weight on visibility graph with complex network feature," in *Australasian Database Conference*. Springer, 2016, pp. 56–66.
- [231] G. Ray, "An algorithm to separate nonstationary part of a signal using mid-prediction filter," *IEEE transactions on signal processing*, vol. 42, no. 9, pp. 2276–2279, 1994.

- [232] S. Mukhopadhyay and G. Ray, "A new interpretation of nonlinear energy operator and its efficacy in spike detection," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 2, pp. 180–187, 1998.
- [233] S. Chandaka, A. Chatterjee, and S. Munshi, "Cross-correlation aided support vector machine classifier for classification of EEG signals," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1329–1336, 2009.
- [234] S. Altunay, Z. Telatar, and O. Eroglu, "Epileptic EEG detection using the linear prediction error energy," *Expert Systems with Applications*, vol. 37, no. 8, pp. 5661–5665, 2010.
- [235] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 2, pp. 205–211, 2007.
- [236] A. R. Hassan, S. Siuly, and Y. Zhang, "Epileptic seizure detection in EEG signals using tunable-q factor wavelet transform and bootstrap aggregating," *Computer Methods and Programs in Biomedicine*, vol. 137, pp. 247–259, 2016.
- [237] S. Li, W. Zhou, Q. Yuan, S. Geng, and D. Cai, "Feature extraction and recognition of ictal EEG using EMD and SVM," *Computers in biology and medicine*, vol. 43, no. 7, pp. 807–816, 2013.
- [238] C.-P. Shen, S.-T. Liu, W.-Z. Zhou, F.-S. Lin, A. Y.-Y. Lam, H.-Y. Sung, W. Chen, J.-W. Lin, M.-J. Chiu, M.-K. Pan *et al.*, "A physiology-based seizure detection system for multichannel EEG," *PloS one*, vol. 8, no. 6, p. e65862, 2013.
- [239] Y. Li, P. P. Wen *et al.*, "Clustering technique-based least square support vector machine for EEG signal classification," *Computer methods and programs in biomedicine*, vol. 104, no. 3, pp. 358–372, 2011.
- [240] Siuly, Y. Li, and P. Wen, "EEG signal classification based on simple random sampling technique with least square support vector machine," *International Journal of Biomedical Engineering and Technology*, vol. 7, no. 4, pp. 390–409, 2011.
- [241] S. Supriya, S. Siuly, H. Wang, J. Cao, and Y. Zhang, "Weighted visibility graph with complex network features in the detection of epilepsy," *IEEE Access*, vol. 4, pp. 6554–6566, 2016.
- [242] S. Supriya, S. Siuly, and Y. Zhang, "Automatic epilepsy detection from EEG

introducing a new edge weight method in the complex network,” *Electronics Letters*, vol. 52, no. 17, pp. 1430–1432, 2016.

- [243] E. Kabir, Y. Zhang *et al.*, “Epileptic seizure detection from EEG signals using logistic model trees,” *Brain Informatics*, vol. 3, no. 2, pp. 93–100, 2016.
- [244] A. Subasi and E. Ercelebi, “Classification of EEG signals using neural network and logistic regression,” *Computer methods and programs in biomedicine*, vol. 78, no. 2, pp. 87–99, 2005.
- [245] Y. Kumar, M. Dewal, and R. Anand, “Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine,” *Neurocomputing*, vol. 133, pp. 271–279, 2014.
- [246] N. Nicolaou and J. Georgiou, “Detection of epileptic electroencephalogram based on permutation entropy and support vector machines,” *Expert Systems with Applications*, vol. 39, no. 1, pp. 202–209, 2012.
- [247] Z. Iscan, Z. Dokur, and T. Demiralp, “Classification of electroencephalogram signals with combined time and frequency features,” *Expert Systems with Applications*, vol. 38, no. 8, pp. 10 499–10 505, 2011.
- [248] S. Siuly and Y. Li, “Discriminating the brain activities for brain–computer interface applications through the optimal allocation-based approach,” *Neural Computing and Applications*, vol. 26, no. 4, pp. 799–811, 2015.
- [249] K. Natarajan, R. Acharya, F. Alias, T. Tiboleng, and S. K. Puthusserypady, “Non-linear analysis of EEG signals at different mental states,” *BioMedical Engineering OnLine*, vol. 3, no. 1, p. 1, 2004.
- [250] G. Chen, “Automatic EEG seizure detection using dual-tree complex wavelet-fourier features,” *Expert Systems with Applications*, vol. 41, no. 5, pp. 2391–2394, 2014.
- [251] D. H. Douglas and T. K. Peucker, “Algorithms for the reduction of the number of points required to represent a digitized line or its caricature,” *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 10, no. 2, pp. 112–122, 1973.
- [252] U. C. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen, “Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready

- information,” *ISPRS Journal of photogrammetry and remote sensing*, vol. 58, no. 3, pp. 239–258, 2004.
- [253] A. Orzan, A. Bousseau, P. Barla, H. Winnemöller, J. Thollot, and D. Salesin, “Diffusion curves: a vector representation for smooth-shaded images,” *Communications of the ACM*, vol. 56, no. 7, pp. 101–108, 2013.
 - [254] K. Zhang, J. Yan, and S.-C. Chen, “Automatic construction of building footprints from airborne LIDAR data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 9, pp. 2523–2533, 2006.
 - [255] J. García and J. Fdez-Valdivia, “Boundary simplification in cartography preserving the characteristics of the shape features,” *Computers & Geosciences*, vol. 20, no. 3, pp. 349–368, 1994.
 - [256] B. Boucheham, Y. Ferdi, and M. C. Batouche, “Recursive versus sequential multiple error measures reduction: a curve simplification approach to ECG data compression,” *Computer methods and programs in biomedicine*, vol. 81, no. 2, pp. 162–173, 2006.
 - [257] B. Boucheham, Y. Ferdi, and M. C. Batouche, “Piecewise linear correction of ECG baseline wander: a curve simplification approach,” *Computer methods and programs in biomedicine*, vol. 78, no. 1, pp. 1–10, 2005.
 - [258] B. Boucheham, “ShaLTERR: A contribution to short and long-term redundancy reduction in digital signals,” *Signal Processing*, vol. 87, no. 10, pp. 2336–2347, 2007.
 - [259] N. Hazarika, J. Z. Chen, A. C. Tsoi, and A. Sergejew, “Classification of EEG signals using the wavelet transform,” *Signal processing*, vol. 59, no. 1, pp. 61–72, 1997.
 - [260] H. Adeli, Z. Zhou, and N. Dadmehr, “Analysis of EEG records in an epileptic patient using wavelet transform,” *Journal of neuroscience methods*, vol. 123, no. 1, pp. 69–87, 2003.
 - [261] R. Sharma and R. B. Pachori, “Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions,” *Expert Systems with Applications*, vol. 42, no. 3, pp. 1106–1117, 2015.
 - [262] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, “Indications of nonlinear deterministic and finite-dimensional structures in

- time series of brain electrical activity: Dependence on recording region and brain state,” *Physical Review E*, vol. 64, no. 6, p. 061907, 2001.
- [263] “EEG time series (epileptic EEG data),” <http://www.meb.uni-bonn.de/epileptologie/science/physik/eegdata.html>, accessed: 2017-03-19.
- [264] L. I. Smith, “A tutorial on principal components analysis,” *Cornell University, USA*, vol. 51, p. 52, 2002.
- [265] L. Breiman, “Random forests,” *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [266] T. Cover and P. Hart, “Nearest neighbor pattern classification,” *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [267] Y. Song, J. Huang, D. Zhou, H. Zha, and C. L. Giles, “Iknn: Informative k-nearest neighbor pattern classification,” in *European Conference on Principles of Data Mining and Knowledge Discovery*. Springer, 2007, pp. 248–264.
- [268] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*. Elsevier, 2011.
- [269] V. N. Vapnik and V. Vapnik, *Statistical learning theory*. Wiley New York, 1998, vol. 1.
- [270] R. K. Begg, M. Palaniswami, and B. Owen, “Support vector machines for automated gait classification,” *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 5, pp. 828–838, 2005.
- [271] X. Yin, B. W.-H. Ng, B. M. Fischer, B. Ferguson, and D. Abbott, “Support vector machine applications in terahertz pulsed signals feature sets,” *IEEE Sensors Journal*, vol. 7, no. 12, pp. 1597–1608, 2007.
- [272] J. R. Quinlan, *C4. 5: programs for machine learning*. Elsevier, 2014.
- [273] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.
- [274] E. Frank, M. Hall, G. Holmes, R. Kirkby, B. Pfahringer, I. H. Witten, and L. Trigg, “Weka-a machine learning workbench for data mining,” in *Data mining and knowledge discovery handbook*. Springer, 2009, pp. 1269–1277.

- [275] G. Ilczuk, R. Mlynarski, A. Wakulicz-Deja, A. Drzewiecka, and W. Kargul, "Rough set techniques for medical diagnosis systems," in *Computers in Cardiology, 2005.* IEEE, 2005, pp. 837–840.
- [276] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.
- [277] M. Grosse-Wentrup, C. Liefhold, K. Gramann, and M. Buss, "Beamforming in noninvasive brain–computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1209–1219, 2009.
- [278] T. Ebrahimi, J. Vesin, and G. Garcia, "Brain-computer interface in multimedia communication," *IEEE Signal Processing Magazine*, vol. 20, no. LTS-ARTICLE-2003-011, pp. 14–24, 2003.
- [279] K. P. Thomas, C. Guan, C. T. Lau, A. P. Vinod, and K. K. Ang, "A new discriminative common spatial pattern method for motor imagery brain–computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 11, pp. 2730–2733, 2009.
- [280] S. Lemm, B. Blankertz, G. Curio, and K.-R. Müller, "Spatio-spectral filters for improving the classification of single trial EEG," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 9, pp. 1541–1548, 2005.
- [281] L. Qin and B. He, "A wavelet-based time–frequency analysis approach for classification of motor imagery for brain–computer interface applications," *Journal of neural engineering*, vol. 2, no. 4, p. 65, 2005.
- [282] D. H. Krishna, I. Pasha, and T. S. Savithri, "Classification of EEG motor imagery multi class signals based on cross correlation," *Procedia Computer Science*, vol. 85, pp. 490–495, 2016.
- [283] T. Wang, J. Deng, and B. He, "Classifying EEG-based motor imagery tasks by means of time–frequency synthesized spatial patterns," *Clinical Neurophysiology*, vol. 115, no. 12, pp. 2744–2753, 2004.
- [284] S. Siuly, Y. Li, and P. P. Wen, "Modified CC-LR algorithm with three diverse feature sets for motor imagery tasks classification in EEG based brain–computer interface," *Computer methods and programs in biomedicine*, vol. 113, no. 3, pp. 767–780, 2014.

- [285] K.-R. Müller, M. Krauledat, G. Dornhege, G. Curio, and B. Blankertz, "Machine learning techniques for brain-computer interfaces," *Biomed. Tech.*, vol. 49, no. 1, pp. 11–22, 2004.
- [286] X. Yu, P. Chum, and K.-B. Sim, "Analysis the effect of PCA for feature reduction in non-stationary EEG based motor imagery of BCI system," *Optik-International Journal for Light and Electron Optics*, vol. 125, no. 3, pp. 1498–1502, 2014.
- [287] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, and F. Yang, "BCI competition 2003-data set Iib: enhancing P300 wave detection using ICA-based subspace projections for BCI applications," *IEEE transactions on biomedical engineering*, vol. 51, no. 6, pp. 1067–1072, 2004.
- [288] B. Blankertz, K.-R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. R. Millan, M. Schröder, and N. Birbaumer, "The BCI competition III: Validating alternative approaches to actual BCI problems," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153–159, 2006.
- [289] L. Cao, K. Chua, W. Chong, H. Lee, and Q. Gu, "A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine," *Neurocomputing*, vol. 55, no. 1, pp. 321–336, 2003.
- [290] M. Chawla, H. Verma, and V. Kumar, "ECG modeling and QRS detection using principal component analysis. paper no. 04," *Proceedings of the IET 3rd International Conference on Advances in Medical, Signal and Information Processing (MEDSIP)*, 2006.
- [291] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing (3rd Ed.): Principles, Algorithms, and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.
- [292] R. Mark and G. Moody, "Signal processing toolbox for use with MATLAB, user guide, version 5," *The Mathworks, Natwick, MA (2001)*, 2001.
- [293] M. Ghosh-Dastidar, R. D. De Veaux, P. F. Velleman, and D. E. Bock, *Intro Stats*. Boston: Pearson Addison Wesley, 2008.
- [294] M. N. Islam, "An introduction to statistics and probability," *Mullick & Brothers, Dhaka New Market, Dhaka*, 2004.

- [295] R. E. Korf, "Linear-space best-first search," *Artificial Intelligence*, vol. 62, no. 1, pp. 41–78, 1993.
- [296] M. A. Hall, "Correlation-based feature selection for machine learning," Ph.D. dissertation, The University of Waikato, 1999.
- [297] B. Mahmoudi and A. Erfanian, "Electro-encephalogram based brain–computer interface: improved performance by mental practice and concentration skills," *Medical and Biological Engineering and Computing*, vol. 44, no. 11, pp. 959–969, 2006.
- [298] S. N. Oğulata, C. Şahin, and R. Erol, "Neural network-based computer-aided diagnosis in classification of primary generalized epilepsy by EEG signals," *Journal of medical systems*, vol. 33, no. 2, pp. 107–112, 2009.
- [299] S. Haykin, *Neural networks and learning machines*. Pearson Education Upper Saddle River, 2009, vol. 3.
- [300] J. A. Suykens, T. Van Gestel, and J. De Brabanter, *Least squares support vector machines*. World Scientific, 2002.
- [301] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 2013.
- [302] X. Guo, Y. Liang, C. Wu, and C. Wang, "PSO-based hyper-parameters selection for LS-SVM classifiers," in *International Conference on Neural Information Processing*. Springer, 2006, pp. 1138–1147.
- [303] H. Esen, F. Ozgen, M. Esen, and A. Sengur, "Modelling of a new solar air heater through least-squares support vector machines," *Expert Systems with Applications*, vol. 36, no. 7, pp. 10 673–10 682, 2009.
- [304] D. W. Hosmer Jr and S. Lemeshow, *Applied logistic regression*. John Wiley & Sons, 1989.
- [305] W. Caesarendra, A. Widodo, and B.-S. Yang, "Application of relevance vector machine and logistic regression for machine degradation assessment," *Mechanical Systems and Signal Processing*, vol. 24, no. 4, pp. 1161–1171, 2010.
- [306] B. Hosseinifard, M. H. Moradi, and R. Rostami, "Classifying depression patients and normal subjects using machine learning techniques and nonlinear features

from EEG signal,” *Computer methods and programs in biomedicine*, vol. 109, no. 3, pp. 339–345, 2013.

- [307] A. Subasi, A. Alkan, E. Koklukaya, and M. K. Kiymik, “Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing,” *Neural Networks*, vol. 18, no. 7, pp. 985–997, 2005.
- [308] K. De Brabanter, P. Karsmakers, F. Ojeda, C. Alzate, J. De Brabanter, K. Pelckmans, B. De Moor, J. Vandewalle, and J. Suykens, “LS-SVMlab toolbox (version 1.8),” <http://www.esat.kuleuven.ac.be/sista/lssvmlab/>, accessed: 2017-03-19.
- [309] S. R. Garner *et al.*, “Weka: The waikato environment for knowledge analysis,” in *Proceedings of the New Zealand computer science research students conference*. Citeseer, 1995, pp. 57–64.