A Multifeatured Method for Detection and Classification of Epileptic Seizure Based on Time Frequency Analysis of EEG Signals

by

Partha Pratim Acharjee

MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING

Department of Electrical and Electronic Engineering
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY

July 2012

The thesis entitled **"A Multifeatured Method for Detection and Classification of Epileptic Seizure Based on Time Frequency Analysis of EEG Signals"** submitted by Partha Pratim Acharjee, Student No.: 1009062042 F, Session: October, 2009 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING on July 12, 2012.

BOARD OF EXAMINERS

1. _____

(Dr. Celia Shahnaz) — **Chairman** (Supervisor)
Assistant Professor
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka - 1000, Bangladesh.

2. _____

(Dr. Pran Kanai Saha) — **Member** (Ex-officio)
Professor and Head
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka - 1000, Bangladesh.

3. _____

(Dr. Mohammed Imamul Hassan Bhuiyan) — **Member**
Associate Professor
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka - 1000, Bangladesh.

4. _____

(Dr. Farruk Ahmed) — **Member** (External)
Professor
School of Engineering and Computer Science (SECS)
Independent University, Bangladesh (IUB)
Block-B, Bashundhara R/A, Dhaka-1229.

CANDIDATE'S DECLARATION

I, do, hereby declare that neither this thesis nor any part of it has been submitted elsewhere for the award of any degree or diploma.

Signature of the Candidate

Partha Pratim Acharjee

Dedication

To my parents.

Acknowledgment

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study. First and foremost, I would like to express my deepest gratitude and indebtedness to my advisor, Dr. Celia Shahnaz, for her excellent guidance, caring, patience, encouragement, constructive suggestions and providing me with an excellent atmosphere for doing research. Dr. Shahnaz guide my research for the past one year and helping me to develop my background in digital signal processing, EEG and biomedical instrumentation. I also want to thank her for spending so many hours with me in exploring new areas of research and new ideas and improving the writing of this dissertation.

I would also like to thank the rest of the members of my thesis committee: Prof. Dr. Pran Kanai Saha, Dr. Mohammed Imamul Hassan Bhuiyan, and Prof. Dr. Farruk Ahmed, for their steadfast encouragement and insightful comments. I would like to thank the head of the department of Electrical and Electronic Engineering for allowing me to use the lab facilities and all sorts of the financial, academic and technical support, which contributed greatly in completing the work in time. I would like to thank Dr. Shaikh Anowarul Fattah, who as an academic advisor, was always willing to help and give his best suggestions and thoughtful comments. Special note of thanks goes to the research group especially Raju Sinha, Rajib Goswami, Asaduzzaman and Shamima Najnin for their continuous moral support, accompany and friendly cooperation. Thanks for the friendship and memories.

Last but not the least, and most importantly, my deepest gratitude goes to my beloved parents; Prof. Swapan Kumar Acharjee and Mrs. Joysree Acharjee and also to my brother for their endless love, prayers and encouragement. Without them I would never have come so far in pursuing my dream.

Abstract

Epileptic seizure describes a recurrent abnormal but synchronized surge of electrical activity in the brain, which is the second common neurological disease. Signal processing methods try to model visual information into few parameters, which can be easily detected thus decision making becomes more accurate compared to the method based on visual observation of EEG. For seizure detection, the features can be categorized as univariate/bivariate and linear/nonlinear types. But, the use of composite feature set has been limitedly reported. Since, EEG is a non-stationary signal and distribution of its energy demonstrates the seizure activities, time-frequency distribution can perform better than the conventional frequency analysis methods. Effectiveness of time-frequency based feature extraction depends on the choice of a kernel and its processing time. Moreover, development of a multifeatured set capable of detecting and classifying epileptic seizure originated from different parts and state of the brain is still a challenging problem. Prior to feature extraction, pre-processing involving Hilbert transform is performed. For the pre-processed EEG signal, time-frequency distributions (TFDs) are obtained by using twelve Cohen Class kernels capable of reducing influence of cross-terms. The TFDs are examined to find out nonuniform modules corresponding to dominant bands, namely $\delta, \theta, \alpha, \beta$ and γ components and to form a feature set containing modular cumulative energy at percentile frequencies and modular entropy. This feature set when fed to each of the decision tree, KNN and ANN classifier, can produce greater detection accuracy independent of the kernels and offers lesser processing time. But, it is unable to classify epileptic seizure originated from different parts and state of the brain. Therefore, a high number of uniform modules are formed to compute a mutifeatured set containing modular energy and entropy which is found effective in detecting as well as classifying epileptic seizure originated from five different parts and state of the brain. Simulations are carried out using standard EEG dataset to evaluate the performance of the proposed method in terms of selectivity, sensitivity and accuracy. It is shown that the proposed method outperforms a state-of-the-art method with superior efficacy.

Contents

Dedication iii

Acknowledgement iv

Abstract v

1 Introduction 1

1.0.1 Types of Seizure 1

1.1 Epilepsy 4

1.1.1 Prevalence of Epilepsy 4

1.1.2 Cause of Epilepsy 5

1.1.3 Diagnosis of Epilepsy 5

1.2 EEG 6

1.2.1 Source of EEG Signal 7

1.2.2 Types of EEG recording 8

1.2.3 10-20 Standard EEG System 9

1.3 Epilepsy Detection and Classification Methods 11

1.3.1 Conventional methods of seizure detection 12

1.3.2 Signal processing for seizure detection 13

1.4 Problem Definition 14

1.5 Objective of the Thesis 15

1.6 Organization of the Thesis 15

2 Literature Review 17

2.1 Introduction 17

2.2 Time Domain Approaches 17

2.3 Frequency Domain Approaches 22

2.4 Time-frequency Domain Approaches 24
2.5 Conclusion 25

3 Epileptic Seizure Detection from EEG Signals Based on Features Extracted from Non-uniform Modules in Time-frequency Domain 26
3.1 Introduction 26
3.2 Proposed Method 27
3.2.1 Pre-processing 27
3.2.2 Time-frequency Analysis 29
3.2.3 Feature Extraction 32
3.2.4 Cumulative Energy 39
3.2.5 Modular Entropy 41
3.2.6 Classification 42
3.3 Conclusion 44

4 Uniform Modular Feature Based Multiclass Epileptic Seizure Classification 45
4.1 Introduction 45
4.2 Proposed Method 45
4.2.1 Feature Extraction 46
4.2.2 Modular Energy 51
4.2.3 Modular Entropy 53
4.3 Conclusion 53

5 Simulation Results 54
5.1 Introduction 54
5.2 EEG Dataset 54
5.3 Performance Parameters 55
5.4 Simulation Results of Seizure Detection 57
5.4.1 Goodness of Feature 57
5.4.2 Performance Analysis 59
5.4.3 Performance comparisons 60
5.5 Simulation Results of Seizure Classification 64
5.5.1 Goodness of Feature 64

5.5.2 Selection of Number of Modules 67
5.5.3 Performance Analysis 68
5.5.4 Performance comparisons 70
5.6 Conclusion 71

6 Conclusion **80**
6.1 Concluding Remarks 80
6.2 Contributions of this Thesis 80
6.3 Scopes for Future Work 81

List of Tables

2.1 Statistical Moments 18

3.1 Time-Frequency Distribution 30

3.2 Implementation Equation of Distance Functions 43

5.1 Comparison of the average sensitivity obtained from the proposed method using 12 kernels and 6 classifiers 60

5.2 Performance comparison of the proposed method with that of a state-of-the-art method in terms of average sensitivity (Sen) and average selectivity (Sel) for all the kernels 62

5.3 Comparison of the average accuracy obtained from the proposed method using 12 kernels and 6 classifiers for two, three and five class problems 75

5.4 Performance comparison of the proposed method with that of a state-of-the-art method in terms of average sensitivity (Sen) and average selectivity (Sel) for two class problem 76

5.5 Performance comparison of the proposed method with that of a state-of-the-art method in terms of average sensitivity (Sen) and average selectivity (Sel) for three class problem 77

5.6 Performance comparison of the proposed method with that of a state-of-the-art method in terms of average sensitivity (Sen) and average selectivity (Sel) for five class problem 78

List of Figures

1.1 Different parts of the brain and functions of those parts 3
1.2 Disability-adjusted life year (DALY) rates from Epilepsy by country (per 100,000 inhabitants). 6
1.3 EEG electrodes position on the scalp in 10-20 EEG recording system 10
1.4 Multichannel EEG signal example with seizure 11
1.5 Time domain plot of different classes 13

3.1 Block diagram of the proposed method 28
3.2 Formation of an analytic signal from a real-valued signal 29
3.3 Time-frequency representation by using WV, PWV and SPWV kernel functions 32
3.4 Time domain (left) and time-frequency domain (right) plots of Z and S 33
3.5 Magnitude plot of 32 DFT coefficients of Z and S. 33
3.6 Examples of some EEG signals or waves at different state of the brain 36
3.7 Distribution of Energy at different frequency bands for seizure and non-seizure EEG signals. 36
3.8 Non-uniform modularization of time-frequency plane 37
3.9 Coefficients of 2D-DFT plot of all classes in both modular (Lower row) and non-modular (Upper row) cases. 38
3.10 Comparison of Euclidean distances for non-modular and modular cases in terms of 2D-DFT coefficients 39
3.11 Comparison of reduction in the length of feature vector in modularization case. 40
3.12 Modular cumulative energy for seizure and non-seizure data at each module. 41

4.1 Block diagram of the proposed method 46

4.2 Time domain (left) and time-frequency domain (right) plots of all class 47
4.3 Magnitude plot of 32 DFT coefficients of Z, O, N, F and S. 48
4.4 Coefficients of 2D-DFT plot of all classes in both modular (Lower row) and non-modular (Upper row) cases. 50
4.5 Comparison of all Euclidean distances for non-modular and modular cases in terms of 2D-DFT coefficients 50
4.6 Percentage improvement in Euclidean distances for all possible cases. 51
4.7 Comparison of reduction in the length of feature vector in modularization case. 52
4.8 Comparison of reduction in the length of feature vector in modularization case. 52

5.1 Confusion matrix for two, three and five class classification cases . . 56
5.2 Selection of proposed feature 58
5.3 Comparison of inter-class separability 59
5.4 Comparison of the average selectivity obtained from the proposed method using 12 kernels and 6 classifiers 60
5.5 Comparison of the average accuracy [%] obtained from the proposed method using 12 kernels and 6 classifiers 61
5.6 Performance comparison of the proposed method with that of a state-of-the-art method in terms of average accuracy [%] for all the kernels 63
5.7 Kernel processing time of a single sample for all kernels 64
5.8 Performance comparison of the proposed method with that of a state-of-the-art method in terms of average accuracy (%) using different classifier 65
5.9 Average accuracy with variance for different feature sets for two, three and five class problems 66
5.10 Comparison of non-modular and modular feature sets 67
5.11 Effect of varying the number of module on classification performance for two, three and five class problems 72
5.12 Comparison of the average selectivity obtained from the proposed method using 12 kernels and 6 classifiers for two, three and five class problems 73

5.13 Comparison of the average sensitivity obtained from the proposed method using 12 kernels and 6 classifiers for two, three and five class problems 74
5.14 Performance comparison of the proposed method with that of a state-of-the-art method in terms of average accuracy for two, three and five class problems. 79

List of abbreviations

FFT	Fast Fourier Transform
DFT	Discrete Fourier Transform
ANN	Artificial Neural Network
KNN	K-nearest neighbor
MH	Margenau-Hill
WV	Wigner-Ville
RIH	Rihaczak
PMH	Pseudo Margenau-Hill
PWV	Pseudo Wigner-Ville
BJ	Born-Jordan
BUT	Butterworth
CW	Choi-Williams
GRECT	Generalized rectangular
RI	Reduced Interference
SPWV	Smoothed Pseudo Wigner-Ville
ZAM	Zhao-Atlas-Marks

Chapter 1

Introduction

A seizure is an excessive abnormal but synchronized surge of electrical activity in the brain. It results from a sudden disturbance of brain function created by abnormal firing of cortical neurons causing neighboring cells to form into a critical mass which usually lasts from a few seconds up to a few minutes. The symptom of seizures vary from person to person. Some people may have simple staring spells, while others have violent shaking and loss of alertness. Seizure is uniformly distributed around the world without any racial, geographical or social class boundaries. It occurs in both genders at all ages, but especially affects neonates and old people. The symptom of seizures are dramatic and alarming and frequently elicit fear and misunderstanding. This type of physical and mental limitation led to profound social consequences for sufferers and has greatly added to the burden of this disease. So, seizure detection and classification methods utilizing the signal processing technique can make the diagnosis process more accurate and faster.

In this chapter, we describe about epilepsy and diagnosis methods, and motivation and contribution of the thesis. Finally outline of the thesis will be drawn for better clarification.

1.0.1 Types of Seizure

The International League Against Epilepsy (ILAE), a world-wide organization of epilepsy professionals, has compiled a list of the names of different seizure types. This is called the ILAE seizure classification. The names given to different types of seizures are based on this classification. Giving seizures the right names is important for doctors. This is because some drugs and treatments can help some seizure types but not others. Based on the type of behavior and brain activity, seizures are

divided into two broad categories: generalized and partial (also called local or focal). Generalized seizures are produced by electrical impulses from throughout the entire brain, whereas partial seizures are produced (at least initially) by electrical impulses in a relatively small part of the brain. The part of the brain generating the seizures is sometimes called the focus.

Generalized seizures: involve epileptic activity in both halves of the brain. The main types of generalized seizure are tonic-clonic, absence, myoclonic, tonic and atonic. The tonic-clonic seizure, also called the grand-mal seizure, is the most common and widely recognized generalized seizure. There are two phases to this type of seizure: the "tonic" phase, followed by the "clonic" phase. In this type of seizure, the patient loses consciousness and usually collapses. The loss of consciousness is followed by generalized body stiffening (called the "tonic" phase of the seizure) for 30 to 60 seconds, then by violent jerking (the "clonic" phase) for 30 to 60 seconds, after which the patient goes into a deep sleep (the "postictal" or after-seizure phase). During tonic-clonic seizures, injuries and accidents may occur, such as tongue biting and urinary incontinence. Absence seizures cause a short loss of consciousness (just a few seconds) with few or no symptoms. The patient, most often a child, typically interrupts an activity and stares blankly. These seizures begin and end abruptly and may occur several times a day. Patients are usually not aware that they are having a seizure, except that they may be aware of "losing time." Myoclonic seizures consist of sporadic jerks, usually on both sides of the body. Patients sometimes describe the jerks as brief electrical shocks. When violent, these seizures may result in dropping or involuntarily throwing objects. Tonic seizures are characterized by stiffening of the muscles. Atonic seizures consist of a sudden and general loss of muscle tone, particularly in the arms and legs, which often results in a fall.

Partial seizures: are mostly categorized according to the starting position of the seizure in the brain. Different parts of the brain are shown in Fig. 1.1. These parts of the brain are used to categorised the partial seizure.Focal seizures starting in the temporal lobes are common. The temporal lobes are responsible for many functions. Some examples of these functions are hearing, speech, memory, and emotions. So symptoms include these: flushing, sweating, going very pale, having a churning feeling in stomach, seeing things as smaller or bigger than they really are,

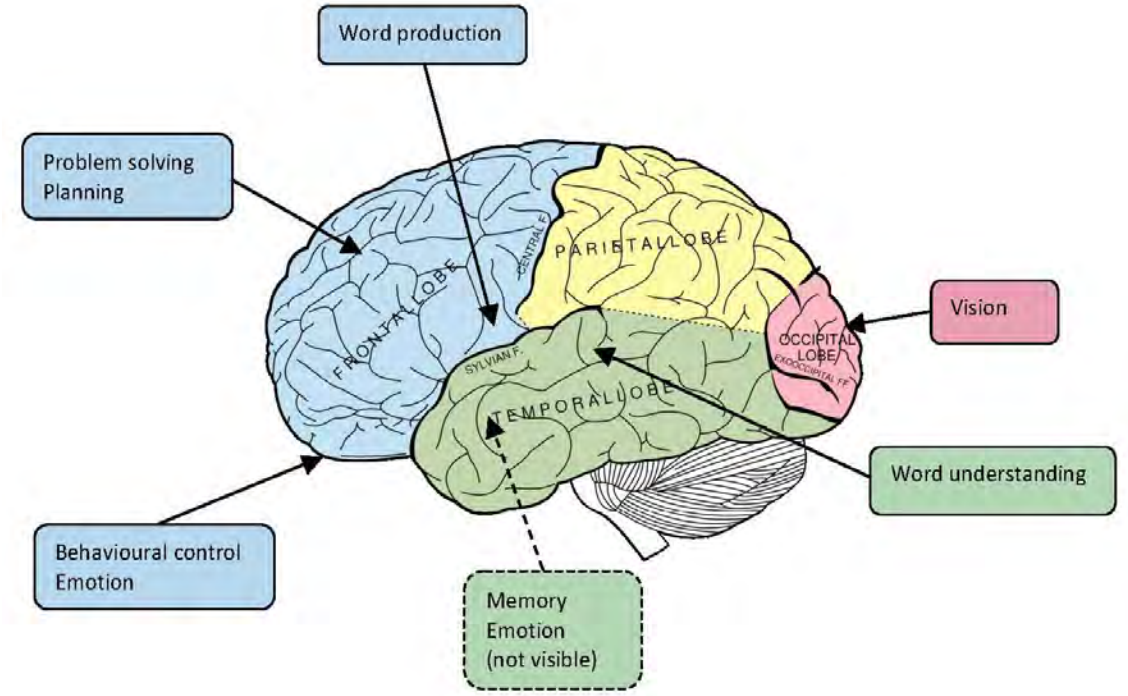


Fig. 1.1: Different parts of the brain and functions of those parts

seeing or hearing something that is not actually happening, smelling non-existent smells, tasting non-existent tastes etc. Frontal lobes also may be the starting point of seizure. Frontal lobes are responsible for many different functions. These include movement, emotions, memory, language, and social and sexual behavior. The frontal lobes are also considered to be home of human's personality. Seizure in frontal lobes may cause these symptoms: turning head to one side, arms or hands are becoming stiff and drawing upwards, cycling movements of legs, thrashing of arms, carrying out strange and complicated body movements, having problems speaking or understanding etc. Focal seizures starting in the parietal lobes are uncommon. The parietal lobes are responsible for bodily sensations. Focal seizures in this part of the brain cause strange physical feelings. A tingling or warm feeling down one side of the body is typical. These types of seizures are also known as 'sensory' seizures. Focal seizures starting in the occipital lobes are uncommon. The occipital lobes are responsible for vision. Focal seizures happening in this part of the brain affect the way of see things, seeing flashes, or balls of light, or having brief loss of vision, are typical symptoms.

1.1 Epilepsy

Epilepsy is caused by a synchronized electrical discharge of a group of neuron which is caused by recurrent seizure, producing a change in the sensation, awareness, and behavior. In reference to the world population, 0.6-0.8% is affected by Epilepsy, which is the second common neurological disease just after the stroke, of whom 30% have not been able to gain any control over their seizures using current pharmacological treatment measures [1], [2]. A person is diagnosed epileptic on the occurrence of two or more unprovoked seizures, and every year more than 2 million new cases of epilepsy are diagnosed. [3]. Seizure prevalence increases with age resulting in severe neurological damage that often becomes medically intractable, a condition in which seizure cannot be controlled by the administration of two or more anti-epileptic drugs (AEDs). Patients with medically intractable seizures are often candidates for surgical resection (removal of the epileptic foci in the brain), which requires accurate localization.

1.1.1 Prevalence of Epilepsy

Epilepsy seizures usually begin between ages 5 and 20, but they can happen at any age. There may be a family history of seizures or epilepsy. Epilepsy affects nearly 3 million Americans and 50 million people worldwide. In the U.S., it affects more than 300,000 children under the age of 15–more than 90,000 of whom have seizures that cannot be adequately be treated. More than 570,000 adults age 65 and above have the condition. Epilepsy is the third most common neurological disorder in the U.S. after Alzheimer's disease and stroke. Its prevalence is greater than cerebral palsy, multiple sclerosis and Parkinson's disease combined. Despite how common it is and major advances in diagnosis and treatment, epilepsy is among the least understood of major chronic medical conditions, even though one in three adults knows someone with the disorder. Disability-adjusted life year for epilepsy per 100,000 inhabitants in 2002 is shown in Fig. 1.2. The disability-adjusted life year (DALY) is a measure of overall disease burden, expressed as the number of years lost due to ill-health, disability or early death.

1.1.2 Cause of Epilepsy

Epilepsy occurs when permanent changes in brain tissue cause the brain to be too excitable or jumpy. The brain sends out abnormal signals. This results in repeated, unpredictable seizures. (A single seizure that does not happen again is not epilepsy). Epilepsy may be due to a medical condition or injury that affects the brain, or the cause may be unknown (idiopathic). Common causes of epilepsy include:

1. Stroke or transient ischemic attack (TIA)
2. Dementia, such as Alzheimer's disease
3. Traumatic brain injury
4. Infections, including brain abscess, meningitis, encephalitis, and AIDS
5. Brain problems that are present at birth (congenital brain defect)
6. Brain injury that occurs during or near birth
7. Metabolism disorders present at birth (such as phenylketonuria)
8. Brain tumor
9. Abnormal blood vessels in the brain
10. Other illness that damage or destroy brain tissue
11. Use of certain medications, including antidepressants, tramadol, cocaine, and amphetamines

1.1.3 Diagnosis of Epilepsy

The doctor's main tool in diagnosing epilepsy is a careful medical history with as much information as possible about what the seizures looked like and what happened just before they began. The doctor will also perform a thorough physical examination, especially of the nervous system, as well as analysis of blood and other bodily fluids. A second battery of diagnostic tools includes an electroencephalograph (EEG). An EEG test tells doctors about the electrical activity happening in the brain. An EEG only shows what is happening in the brain at the time the test

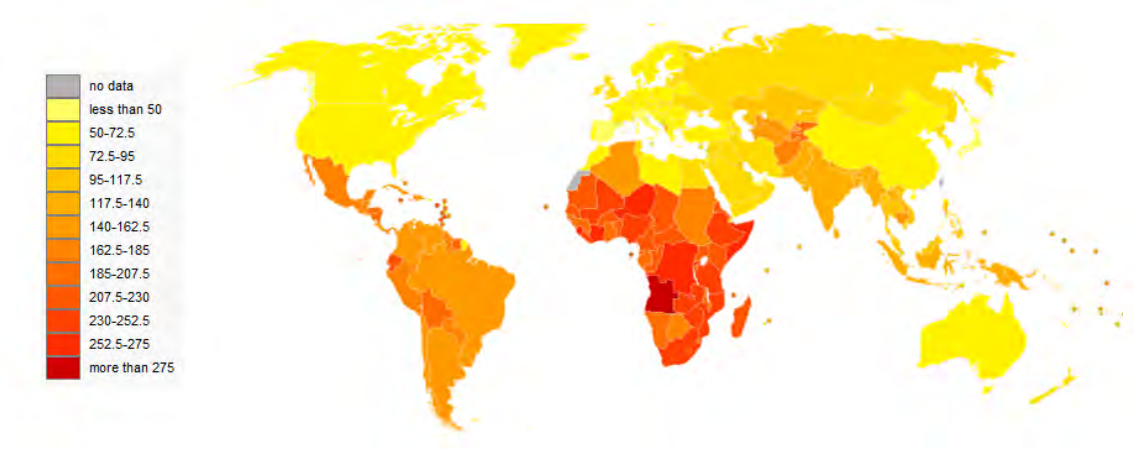


Fig. 1.2: Disability-adjusted life year (DALY) rates from Epilepsy by country (per 100,000 inhabitants).

is being done. It's not able to show what has already happened or what is going to happen in the future. Despite this, an EEG can sometimes be very helpful to doctors when they are diagnosing epilepsy. Imaging methods such as CT (computerized tomography) or MRI (magnetic resonance imaging) scans may be used to search for any growths, scars or other physical conditions in the brain that may be causing the seizures. In a few research centers, positron emission tomography (PET) imaging is used to identify areas of the brain which are producing seizures. Blood tests are used to check general health, and to look for any medical conditions that might be causing epilepsy. Those can also be used to find out if seizures are not caused by epilepsy, but another medical condition,such as diabetes.

1.2 EEG

An electroencephalogram (EEG) is a test that measures and records the electrical activity of the brain. EEG measures voltage fluctuations resulting from ionic current flows within the neurons of the brain. In clinical contexts, EEG refers to the recording of the brain's spontaneous electrical activity over a short period of time, usually 20-40 minutes, as recorded from multiple electrodes placed on the scalp. The electrode only picks up electric signal from the brain and doesn't affect the brain. So this process is totally painless and harmless. In neurology, the main diagnostic application of EEG is in the case of epilepsy, as epileptic activity can create clear abnormalities on a standard EEG study. A secondary clinical use of EEG is in the

diagnosis of coma, encephalopathy, and brain death. EEG used to be a first-line method for the diagnosis of tumors, stroke and other focal brain disorders, but this use has decreased with the advent of anatomical imaging techniques with high (¡1 mm) spatial resolution such as MRI and CT. Despite limited spatial resolution, EEG continues to be a valuable tool for research and diagnosis, especially when millisecond-range temporal resolution (not possible with CT or MRI) is required.

1.2.1 Source of EEG Signal

The brain's electrical charge is maintained by billions of neurons. Neurons are electrically charged (or "polarized") by membrane transport proteins that pump ions across their membranes. Neurons are constantly exchanging ions with the extracellular milieu, for example to maintain resting potential and to propagate action potentials. Ions of similar charge repel each other, and when many ions are pushed out of many neurons at the same time, they can push their neighbors, who push their neighbors, and so on, in a wave. This process is known as volume conduction. When the wave of ions reaches the electrodes on the scalp, they can push or pull electrons on the metal on the electrodes. Since metal conducts the push and pull of electrons easily, the difference in push or pull voltages between any two electrodes can be measured by a voltmeter. Recording these voltages over time gives us the EEG. The electric potential generated by single neuron is far too small to be picked up by EEG or MEG. EEG activity therefore always reflects the summation of the synchronous activity of thousands or millions of neurons that have similar spatial orientation. If the cells do not have similar spatial orientation, their ions do not line up and create waves to be detected. Pyramidal neurons of the cortex are thought to produce the most EEG signal because they are well-aligned and fire together. Because voltage fields fall off with the square of distance, activity from deep sources is more difficult to detect than currents near the skull. Scalp EEG activity shows oscillations at a variety of frequencies. Several of these oscillations have characteristic frequency ranges, spatial distributions and are associated with different states of brain functioning (e.g., waking and the various sleep stages). These oscillations represent synchronized activity over a network of neurons. The neuronal networks underlying some of these oscillations are understood (e.g., the

thalamocortical resonance underlying sleep spindles), while many others are not (e.g., the system that generates the posterior basic rhythm). Research that measures both EEG and neuron spiking finds the relationship between the two is complex with the power of surface EEG in only two bands (gamma and delta) relating to neuron spike activity.

1.2.2 Types of EEG recording

EEG test can be performed with different variation according to physician's need to fulfill the diagnostic requirement. The standard test duration will be 10 to 20 minutes which capture the brain activity snap for a particular time. As seizure is a random event this types of standard test may not fulfill the requirement. Ambulatory test is done to observe the brain activity more deeply. Ambulatory means designed for walking. So an ambulatory EEG can be used while you are moving around. An ambulatory EEG test can record the activity in your brain over a few hours, days or weeks. This allows more time for the test to pick up any unusual electrical activity in your brain, than during a standard EEG test. An ambulatory EEG uses electrodes similar to those used on a standard EEG test. However, the electrodes plug in to a small machine that records the results. Patient can wear the machine on a belt, so he or she is able to go about their daily business. Patient does not usually stay in hospital while the test is being done. Sometimes patients may be asked to have an EEG test while they are asleep. This could be because sometimes seizures happen when patients are asleep or when patients are tired. A sleep-deprived EEG test is done when patients have had less sleep than usual. When they are tired, there is more chance that there will be unusual electrical activity in the brain. A video-telemetry test involves wearing an ambulatory EEG. At the same time, all movements are recorded by a video camera. The test is usually carried out over a few days. This is to increase the chances that a seizure can be recorded. After the test, doctors can watch the video to see any occurrence of seizures. They can also look at the EEG results for the time patient was having the seizure. This will tell them about any changes to the brainwave patterns at the time of the seizure. Patient would usually only has a video-telemetry test if they have already been diagnosed with epilepsy.

Inter-ictal findings from electroencephalography (EEG) offer the most specific test for diagnosing epileptic seizure. A short period EEG recording can be used to identify the inter-ictal indications of epilepsy. Generally inter-ictal events are characterized by isolated spikes, sharp waves and spike-wave-complex, whereas ictal period is manifested by rhythmic waveforms and poly-spikes. For the diagnosis of epileptic seizure, several types of EEG including Routine EEG, Ambulatory EEG, and Video-EEG are used. At the first stage, neurologist examines the patient's basic mental functions, such as the ability to remember words, calculate, and name objects, and then systematically test the functioning of the muscles and senses, along with reflexes, walking, and coordination. Routine EEG that mainly records 20 to 40 minutes post-ictal data is sufficient for general cases and neurologist generally detects epilepsy from visual observation. Ambulatory EEG can record up to 72 hours to observe both ictal and post-ictal brain activities thus allowing the characterization of epileptic seizures and seizure-like events even at home. However, due to the relatively infrequent nature of epileptic seizures, the long-term video-EEG monitoring is mandatory. Thus, ambulatory EEG can be integrated with video recording which correlates the patient behavior with EEG data. How-ever, visual seizure detection has not been proven very effective as visual observation suffers from misinterpretation frequently and needs highest level of expertise.Detecting dominance of different frequency components from visual observation does not correlate with few parameters only. For confident decision, involvements of few mathematical features are mandatory. Efficient automated seizure detection schemes facilitate the diagnosis of epilepsy and enhance the management of long-term EEG recordings.But for influent seizures long term monitoring is mandatory where automatic detection of seizure will be very helpful for the physician [4].

1.2.3 10-20 Standard EEG System

The 10-20 system or International 10-20 system is an internationally recognized method to describe and apply the location of scalp electrodes in the context of an EEG test or experiment. The position of the electrode of the 10-20 system are shown in Fig. 1.3 [5].This method was developed to ensure standardized reproducibility so that a subject's studies could be compared over time and subjects could be compared

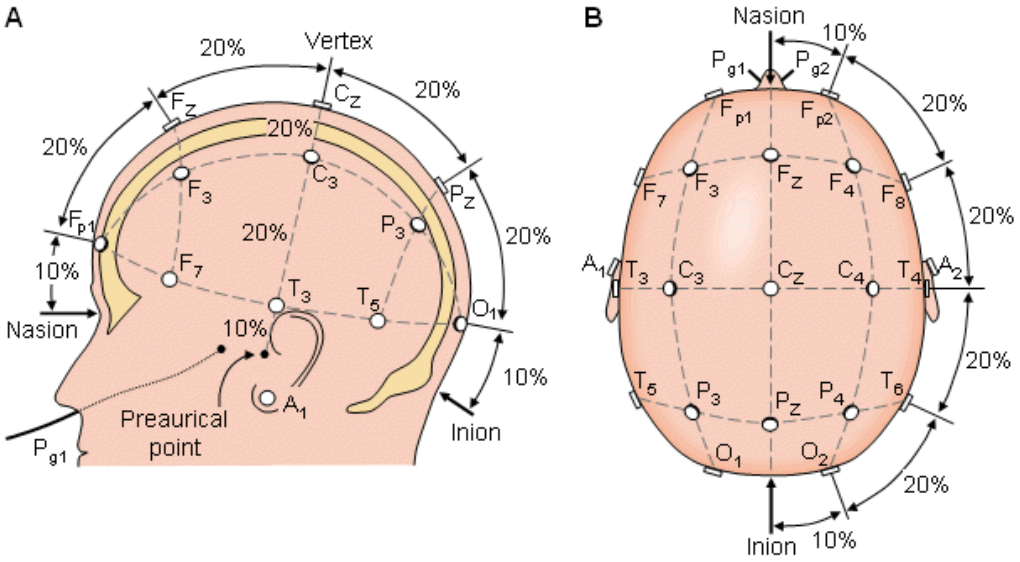


Fig. 1.3: EEG electrodes position on the scalp in 10-20 EEG recording system

to each other. This system is based on the relationship between the location of an electrode and the underlying area of cerebral cortex. The "10" and "20" refer to the fact that the actual distances between adjacent electrodes are either 10% or 20% of the total front-back or right-left distance of the skull. Each site has a letter to identify the lobe and a number to identify the hemisphere location. The letters F, T, C, P and O stand for frontal, temporal, central, parietal, and occipital lobes, respectively. Note that there exists no central lobe; the "C" letter is only used for identification purposes only. A "z" (zero) refers to an electrode placed on the midline. Even numbers (2,4,6,8) refer to electrode positions on the right hemisphere, whereas odd numbers (1,3,5,7) refer to those on the left hemisphere.

Two anatomical landmarks are used for the essential positioning of the EEG electrodes: first, the nasion which is the point between the forehead and the nose; second, the inion which is the lowest point of the skull from the back of the head and is normally indicated by a prominent bump. When recording a more detailed EEG with more electrodes, extra electrodes are added utilizing the spaces in-between the existing 10-20 system. This new electrode-naming-system is more complicated giving rise to the Modified Combinatorial Nomenclature (MCN). This MCN system uses 1, 3, 5, 7, 9 for the left hemisphere which represents 10%, 20%, 30%, 40% and 50% of the inion-to-nasion distance respectively. The introduction of extra letters allows the naming of extra electrode sites. Note that these new letters do not necessarily

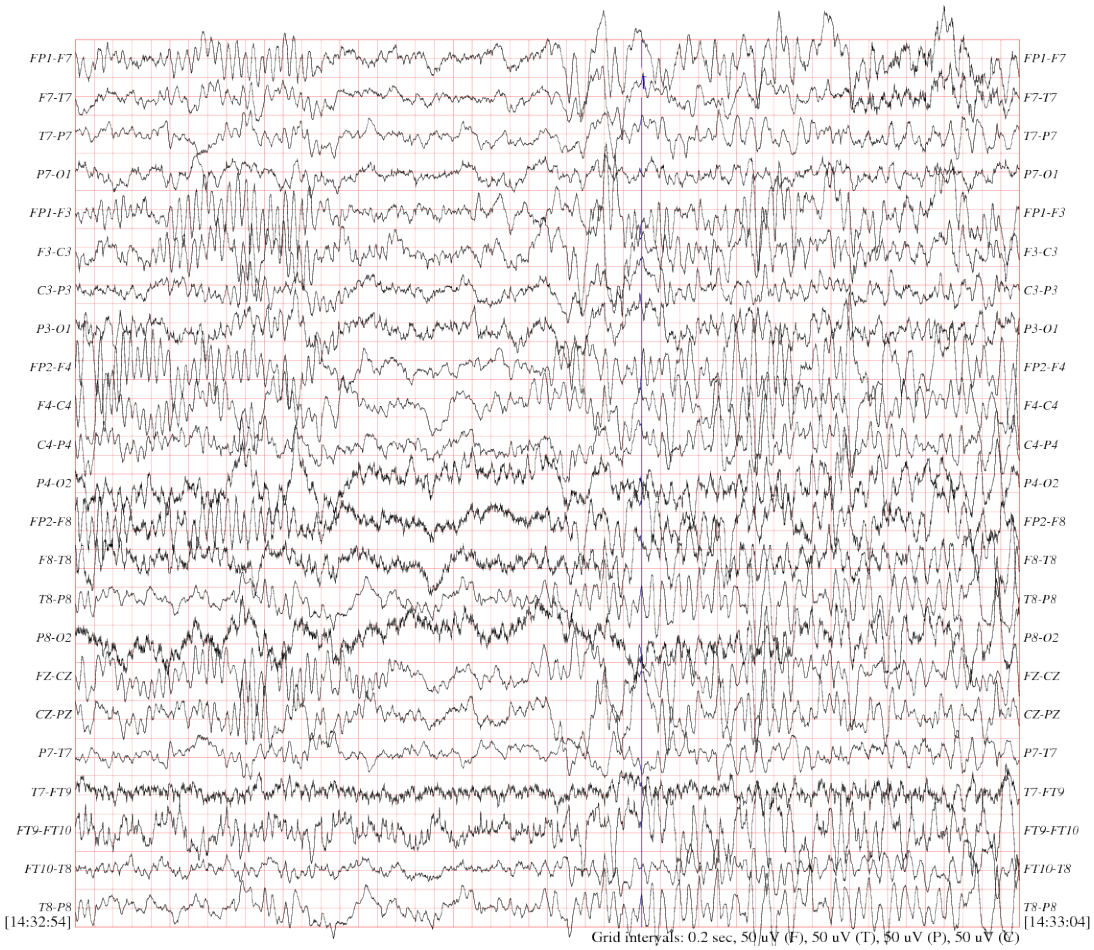


Fig. 1.4: Multichannel EEG signal example with seizure

refer to an area on the underlying cerebral cortex. Multichannel EEG measures the voltage difference in two different electrodes. A multichannel EEG signal example is shown in Fig. 1.4

1.3 Epilepsy Detection and Classification Methods

EEG can measure the abnormality of brain activities which are indicated by multiple channel data. Different techniques are utilized to detection and classification the the EEG signal from different channel. Conventionally, physicians use visual observation for decision making which needs superior expertise. Signal processing techniques introduce different methods to achieve expert like accuracy.

1.3.1 Conventional methods of seizure detection

Conventionally seizure is detected and classified by visual observation of EEG signal by experts. As EEG signal is varied with the age of the patient, classification of seizure from visual observation needs superior expertise. The classic EEG feature of LGS is a slow spike-wave complex, seen on both sides or over the entire head, repeating at 1-2 per second. The same pattern may be seen during an atypical absence seizure. The frequent presence of the slow spike-wave complexes and the waxing and waning presence can make it difficult to distinguish between the interictal (between seizure) and ictal (during seizure) pattern. The slow spike-wave complexes may not be present when the child is first diagnosed with seizures, and only becomes apparent on a subsequent EEG, making the initial diagnosis difficult in some children. The occurrence of slow spike-wave complexes may decrease in adolescents and adults. In patients whose seizure frequency decreases, the pattern can change to single spike-and-wave complexes during sleep and then their subsequent disappearance. Focal spike (epileptiform discharges) or focal slowing may be seen. Bursts of generalized polyspikes at 10 per second or more are seen frequently enough to be considered an additional criterion for LGS. These are seen best in sleep, and may be associated with a nocturnal tonic seizure (as the ictal pattern). As children reach early adult years, only 30% to 50% still have the characteristic EEG and clinical characteristics.

The EEG during a tonic seizure reveals generalized, fast (10-15 per second), low-amplitude activity, seen best over the anterior head regions. This slows in frequency and increases in amplitude as the seizure progresses. This may be preceded by a single generalized spike-and-wave complex. The clinical features typically begin within one second of the EEG manifestations. Diffuse, slow (less than 2-2.5 per second), spike-and-wave complexes characterize the EEG during an atypical absence seizure. This may be difficult to distinguish from the interictal bursts. Atonic, myoclonic and myoclonic-atonic seizures have an EEG characterized by bilateral slow spike-and-wave complexes, polyspike-and-wave complexes or rapid polyspikes.

However, visual seizure detection has not been proven very effective as visual observation suffers from misinterpretation frequently and needs highest level of expertise. Efficient automated seizure detection schemes facilitate the diagnosis of epilepsy and enhance the management of long-term EEG recordings. So different

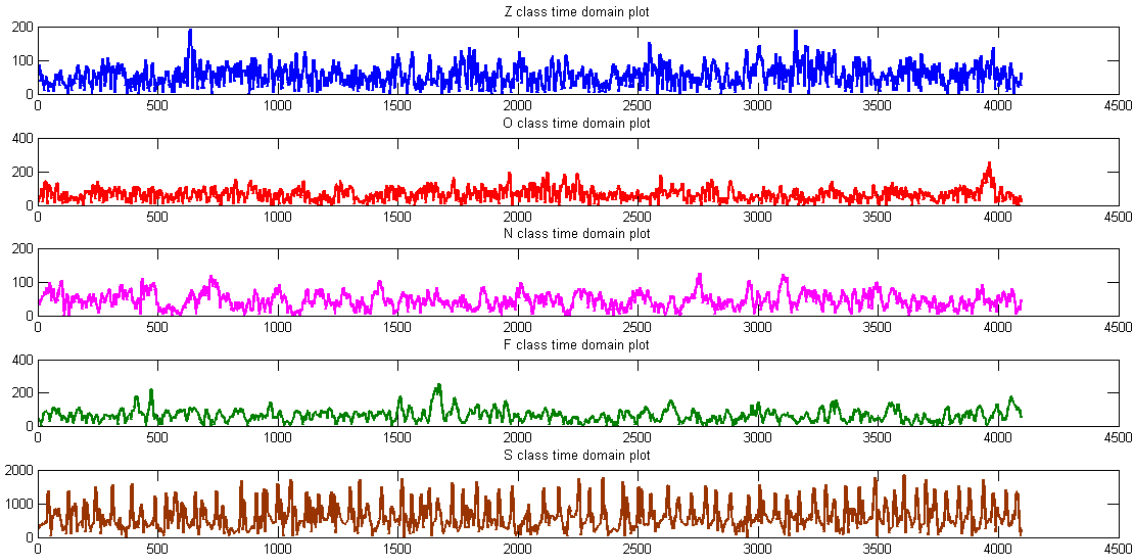


Fig. 1.5: Time domain plot of different classes

signal processing based EEG signal classification methods are utilized to facilitate expert decision.

1.3.2 Signal processing for seizure detection

Signal processing methods introduce the mathematical representations of scattered visual information by different feature set. These feature sets always try to model all visual information which can be detected from expert analysis. So signal processing methods precise all information and model those into few parameters from where decision making is easier and more accurate than conventional manual methods. Fig. 1.5 demonstrates time domain plot of different state of the brain and classification of different state of the brain from this visual observation is erroneous. As EEG is a non-stationary signal, taking perfect decision is mostly dependent on the accuracy of extracting feature in time and frequency domain. As distribution of energy at different frequency bands demonstrate the seizure activities, time frequency distribution performs better than conventional frequency analysis methods [6], [7].

At the very beginning stage of automatic EEG analysis, Viglione and Walsh (1975) try to detect absence seizure by linear approach [8]. Features used at previous works can be categorized as univariate/bivariate and linear/nonlinear types. Statistical moment, spectral band power, spectral edge frequency, accumulated energy, auto correlation, hjorth parameters and auto regressive modeling are linear

univariate measures [9–13]. On the other hand, correlation sum/density/ dimension/ entropy, marginal predictability, dynamic similarity, Lyapunov exponent etc are nonlinear univariate measures [14–20]. In order to find out similarity between two signals bivariate measured are used. In bivariate system, maximum linear cross-correlation and linear coherence are popular linear parameters [21], [22]. Non-linear interdependence, dynamic entrainment and phase synchronization are used as non-linear bivariate measures [23], [24]. Few works made composite to prepare the feature vector. These features are then classified by different classifier like different distance based classifier, neural network based classifier etc. Previous work on seizure can be categorize as seizure detection and seizure classification technique. Spike detection during inter-ictal period is investigated through several methods, like wavelet, frequency domain analysis, ICA, ANN, SVM, data mining and template matching etc [25–30].Many previous works used time frequency analysis to detect preseizure chirps and multi-resolution analysis of EEG [31], [32]. Effectiveness of these works depends on frequency or time domain smoothing. RI distribution is employed for smoothing purpose before feature extraction [33], [34]. Twelve Cohen class kernels are also employed for feature extraction [35].

1.4 Problem Definition

Long time EEG recording is needed to capture a seizure event which can be used for further diagnosis. From EEG recording types and source of the seizure can be identified. But random signal type, involvement of multiple node and long time of recording make it difficult to detect and classify EEG signals. EEG signal vary in time and frequency domain simultaneously, so only time and only frequency domain feature is not sufficient for higher class EEG signal classification problem. On the other hand, time-frequency analysis can extract time-frequency information more precisely than conventional frequency analysis method. Time-frequency analysis is recently used for seizure classification problem. As time-frequency analysis inherently introduce some interference, some time and frequency domain smoothing function are used for reducing those interference. So selection of smoothing function also affect the performance of the classification problem. Similarly, selection of feature which can extract necessary information from time-frequency plane is another

major criteria for good classification results. Moreover, development of a multi-featured set capable of detecting and classifying epileptic seizure originated from different parts and state of the brain is still a challenging problem.

1.5 Objective of the Thesis

The objective of this thesis are:

- To analyze the time-frequency distributions obtained by transforming the EEG signals using Cohen class kernels.
- To develop a multifeatured set from the time frequency representation to detect epileptic seizure from EEG signals.
- To derive another multifeatured set not only to detect epileptic seizure, but also to classify epileptic seizure originated from different parts and state of the brain.
- To investigate the performance of the proposed feature sets and that of different classifiers for the detection and classification of epileptic seizures using EEG signals available from the standard EEG database

The outcome of this thesis is the development of an EEG based method exploiting a composite feature set derived from the time frequency distributions involving the Cohen class kernels, which is able to detect as well as classify epileptic seizure with greater accuracy and lesser processing time.

1.6 Organization of the Thesis

The thesis is organized as follows.

- Chapter 1 gives the introduction of the overall thesis.
- Chapter 2 reviews popular seizure detection and classification methods.
- Chapter 3 describes a method of epileptic seizure detection from EEG signals based on non-uniform modules in time-frequency domain.

- Chapter 4 shows the formation of another uniform modular feature based multiclass epileptic seizure classification methods which can classify epileptic seizure originated from different parts and state of the brain. Here, three classification problems are considered, namely two,three and five class classification problem.

- Simulation results and quantitative performance analysis is discussed in Chapter 5 for both the methods described in Chapter 3 and 4. Performance of these methods are also compared with a state-of-the-art method.

- Finally, in Chapter 6, concluding remarks highlighting the contributions of the thesis and suggestions for further investigation are provided.

Chapter 2

Literature Review

2.1 Introduction

Epileptic seizures are caused by the transient and unexpected electrical disturbances of the brain. Roughly stated, one in every 100 persons is likely to experience a seizure at some time in their life. The recording of seizures by EEG is of prime importance in the assessment of epileptic patients. Seizures can be defined as the phenomenon of rhythmicity discharge from either a specific area or the whole brain and the individual behavior generally lasts from seconds to minutes. In general, as seizures are observed occasionally and unpredictably, automatic detection of seizures during long-term electroencephalograph (EEG) recordings is greatly recommended. As EEG involves multiple channel and frequency variation in different parts and state of the brain, making decision by observing the time domain plot is erroneous. Since EEG signals are non-stationary, the general methods of frequency analysis are not satisfactory for diagnostic purposes. A plentiful of researches is available in the literature concerned with automated detection and prediction of epileptic seizures using EEG signals. All those researches try to extract feature from EEG signal and then use different classifier for classification and detection purpose. Some of those use time domain feature, few use frequency domain feature and few also try to incorporate time-frequency base feature extraction.

2.2 Time Domain Approaches

Time series of EEG contain activities with different amplitudes and frequencies. Time domain feature extraction approaches measure the amplitude and phase information from the EEG time series. Different mathematical, statistical and empirical

measures are introduced to design time domain based the feature set.

Statistical moments provide information on the amplitude distribution of a time series signal $\{x_i\}$. The first and second statistical moment, the mean and the variance respectively, provide information on the location and variability of the amplitude distribution of the time series. The third (skewness) and fourth (kurtosis) moments also provide information on the shape of the distribution.

Table 2.1: Statistical Moments

Serial	Statistical Moment	Mathematical representation
1	First statistical moment, Mean	$\bar{x} = \frac{1}{N}\sum_{i=0}^{N-1} x_i$
2	Second statistical moment,Variance	$\sigma = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} x_i^2}$
3	Third statistical moment, Skewness	$\chi = \frac{1}{N}\sum_{i=0}^{N-1}(\frac{x_i}{\sigma})^3$
4	Fourth statistical moment, Kurtosis	$\kappa = \{\frac{1}{N}\sum_{i=0}^{N-1}(\frac{x_i}{\sigma})^4\} - 3$

The ability of these measures to distinguish between the interictal period and the preseizure period in the IEEG data have been compared in different literature. Using variance and kurtosis, a preictal period was found with significant changes (a decrease for variance and an increase for kurtosis) in comparison with the interictal period. Other attempts to extract seizure precursors from the EEG were carried out for seizure prediction using spectral analysis [12, 36]. But mean variation of amplitude and its variance can not extract all information from the EEG data. It can only distinguish those classes which are almost visually separable in time domain.

Autocorrelation is the correlation between values of the signal at different points in time. It is computed as a function of the two times or of the time difference. It is usually used to detect "whiteness" in data. The first zero-crossing of this function is defined as the decorrelation time [37]. Preictal period from the interictal period can be distinguishable by observing a decrease in the decorrelation time. It can not extract enough information for seizure classification rather than only detection.

In linear modeling of a time series, one assumes that each value of the series depends only on a weighted sum of the previous values of the same series plus "noise". The main assumption in linear modeling is the stationarity of the signal. So, for non-stationary signals like EEG, one needs to segment it into stationary parts. Using autoregressive modeling, preictal changes have been reported before

seizure onset in [12]. The most general linear (univariate) model for a time series is the autoregressive moving average (ARMA) model. It is composed of three linear model processes: a purely random process (white noise), an autoregressive (AR) process and a moving average (MA) process. An AR process is defined by

$$x_i = \sum_{l=1}^{p} a_l k_{i-l} + \sum_{k=1}^{q} b_k \varepsilon_{i-k} \tag{2.1}$$

where the coefficients $\{a_k\}$ and $\{b_k\}$ are to be determined by fitting the data, typically using a least-squares or an information-theoretic criterion. Identification of an appropriate ARMA model allows the design of special filters, forecasting time series or estimation of the power spectrum and derived measures such as the so-called transfer function.

The correlation dimension and the following two measures are centered on the concept of the correlation integral, which can be computed from the state space representation of EEG time series. The correlation integral is defined as the probability that any two randomly chosen points on the state space lie within a given distance of each other [38]. The correlation dimension gives a measure of the dimensionality of the state space. This measure is used to distinguish random signals from deterministic time series [39]. The ability of the correlation dimension for seizure prediction has been proved in [40]. In [40], it is demonstrated that significant drops in correlation dimension occurred prior to seizures. However, [41] suggest strongly that the correlation dimension has no predictive power for epileptic seizures.

The correlation density is calculated by computing the correlation integral for a fixed radius and using a combination of time delay and spatial embedding of EEG time series. In [14], it is demonstrated that in most cases, seizure onset could be anticipated well in advance. However, the ability of this measure for seizure prediction was questioned in [42].

The Kolmogorov entropy is a dynamic measure representing the rate at which information needs to be created as the dynamical system evolves over time [43]. It gives a measure of the level of uncertainty about the future state of the system. The feasibility of using trends in Kolmogorov entropy to anticipate seizures in pediatric patients with intractable epilepsy has been demonstrated in [44]. It has been concluded that the Kolmogorov entropy is as effective as the correlation dimension

in anticipating seizures. But effectiveness of Kolmogorov entropy and Correlation density is limited to few predefined restriction and assumptions.

Dynamical similarity index is composed of the phase space reconstruction of the EEG time series using time intervals between two positive zero-crossings and the measurement of dynamical similarity between a reference window and test windows using the cross-correlation integral. Le Van Quyen et al. In [45] showed that the method can track in real time spatio-temporal changes in brain dynamics several minutes prior to seizure. However, other studies [19] questioned the reliability of the optimistic results reported for the dynamical similarity index in [45].

The maximum linear cross correlation measure implies that two systems are linearly synchronized if their characteristic variables evolve identically over time [46]. In order to quantify the similarity of two signals $\{x_i\}$ and $\{y_i\}$ the maximum of a normalized cross-correlation function can be used as a measure for lag synchronization

$$C_{max} = max_{\tau}\{|\frac{C_{xy}(\tau)}{\sqrt{C_{xx}(0)C_{yy}(0)}}|\}, \tag{2.2}$$

here,$C_{\{}xx\}$=Auto-correlation,C_{xy}=Cross-correlation. A preictal loss in synchronization between EEG signals recorded simultaneously from different locations in the brain has been observed. In this method multichannel EEG data is required to observe the loss of synchronization. Effectiveness of the method is also dependent on the measurement accuracy of both channels. Selection of set a channels is vital for obtain reported accuracy.

The phase synchronization measures the degree to which two signals are phase locked during a short time period. Phase synchronization is traditionally defined as phase locking $[\phi_x(t) - \phi_y(t) = const]$ or, in the case of noisy and/or chaotic systems , as phase entrainment $[\phi_x(t) - \phi_y(t) < const]$, with $[\phi_x(t)$ and $[\phi_y(t)$ denoting the phase variables of two oscillating signals $x(t)$ and $y(t)$. Mean phase coherence is the most popular phase synchronization measure which can be expressed as

$$R =| \frac{1}{N}\sum_{j=1} N e^{i[\phi_x(t_j)-\phi_y(t_j)]} | \tag{2.3}$$

In intracranial EEG data, this measure has shown its power to discriminate transient synchronization [24]. Analysis of long EEG recordings has shown that the epileptogenic process during the interictal state can be characterized by a patholog-

ically increased level of synchronization as measured by the mean phase coherence [46]. In another study, a specific state of brain synchronization has been observed several hours before the actual seizure. The changes involved both increases and decreases of the synchronization levels often localized near the primary epileptogenic zone [47]. I this method multichannel EEG data is required and selection of pair of monitoring channel is crucial for accurate classification of the EEG signals.

The Lyapunov exponents quantify the exponential divergence of initially close state-space trajectories and determine the predictability of a dynamical system. The largest Lyapunov exponent gives a measure for detecting the presence of chaos in a dynamical system [48]. The largest Lyapunov exponent is used for characterizing intracranial EEG recordings and noted premonitory events several minutes prior to the onset of seizures in several recordings [20]. However, inability of Lyapunov exponents to predict epileptic seizures is highlighted in [49].

The efficacy of inter quartile range (IQR), a median based measure of statistical dispersion, as a discriminating feature that can be used for the classification of EEG signals into normal, interictal and ictal classes is shown in [50]. IQR along with variance and entropy are calculated for each frame of EEG. To reduce the feature vector size, standard statistical features such as mean, minimum, maximum and standard deviation were evaluated and were given as input to a linear classifier.

A novel seizure detection and analysis scheme based on the phase-slope index (PSI) is proposed in [51]. The PSI metric identifies increases in the spatio-temporal interactions between channels that clearly distinguish seizure from inter-ictal activity. We form a global metric of interaction between channels and compare this metric to a threshold to detect the presence of seizures. The threshold is chosen based on a moving average of recent activity to accommodate differences between patients and slow changes within each patient over time. A common threshold procedure is involve to determine the threshold value. So, calculating the threshold value is crucial for seizure detection in this case.

Time domain feature extracted effective information about magnitude and phase of the time series signal. Frequency is extracted from variation of magnitude with respect to time. So, minor and frequent change in different frequency band are overlooked during time domain feature extraction. So frequency domain analysis is

introduced to make a deeper look at the variation of frequency.

2.3 Frequency Domain Approaches

Observing variation in magnitude and phase in time domain are sufficient for classification of multiclass EEG data. Variation of different feature in frequency domain is useful to analyze the dominant frequencies. Characteristic of the frequency domain variation are also reported as useful feature for EEG classification.

The EEG signal has usually been described in terms of main frequency bands, δ (less than 4 Hz), θ (4-8 Hz), α (8-12 Hz), β (13-30 Hz), and γ (greater than 30 Hz). Relative power in any frequency band is defined as the area under the curve of the power spectrum within the bandwidth under consideration divided by total power for all bands. So The relative power contained in these bands can be defined as

$$\delta = \frac{1}{P}\sum_{f=0.5Hz}^{4Hz} P_f; \theta = \frac{1}{P}\sum_{f=4Hz}^{8Hz} P_f; \alpha = \frac{1}{P}\sum_{f=8Hz}^{12Hz} P_f; \beta = \frac{1}{P}\sum_{f=13Hz}^{30Hz} P_f; \gamma = \frac{1}{P}\sum_{f=30Hz}^{100Hz} P_f; \tag{2.4}$$

here, P is the total power of the signal. In [36], it have been shown that for the preictal period in comparison with the interictal period, there is a relative decrease of power in the delta band that is accompanied by a relative increase in the remaining bands. This variation can classify only preictal and interictal events but it is not capable of classifying EEG signal of different state in the interictal period.

The accumulated energy is computed for any moving observation window by averaging all successive values of energies calculated in that window. This can be considered as the running average of the energy [52]. Using this measure, promising results for seizure prediction have been reported [10]. However, the results are not reportedly good for all cases and without few predefined restrictions [53].

The loss of recurrence quantifies the degree of non-stationarity in a EEG signal [54]. The frequency distribution of time distances under stationary conditions with respect to each reference point is first computed. If the system is non-stationary, an increased deviation from this distribution is observed due to the absence of distant time indices in the neighborhood of the reference. That is considered as a loss of recurrence. The predictability of this measure for epileptic seizures has been shown in [36].

In a typical EEG signal, most of the power is contained within the frequency band from 0 Hz up to 40 Hz: $P_{40Hz} \approx P$. As a characterizing measure for the power distribution, the so-called spectral edge frequency can be used [9], which is defined as the minimum frequency up to which 50% the spectral power up to 40 Hz is contained in the signal:

$$f_{50} = min\{f^* | \sum_{f=0Hz}^{f^*} p_f > (P_{40Hz} \times 0.5) \tag{2.5}$$

Hjorth defined three timedomain parameters, activity, mobility and complexity,also called normalized slope descriptors [11]. The activity is the variance of the signal which gives a measure of mean power. The mobility is the ratio of the root mean square(RMS) of the slopes of the signal to the RMS of the amplitude. This parameter may be considered as an estimate of the mean frequency. The complexity gives a measure of the RMS of the rate of slope changes with reference to an ideal possible curve. This parameter gives an estimate of the bandwidth of the signal. In [36], it is found a preictal period with a significant increase in the Hjorth mobility and complexity with respect to the interictal period.In the frequency domain, the mobility and complexity can be estimated from the second and fourth statistical moment of the power spectrum which can be expressed as

$$HM = \sum k = 1N/2p_k K^2 j, \quad HC = \sum k = 1N/2p_k K^4 j \tag{2.6}$$

Variation of Hjorth parameters are not significant for minor variation of the EEG data.

The analysis shown in [55]is based upon extraction of relevant information and learning of object parameter values such as power distribution in various phases of EEG time series, seizure time and spectral power in various frequency bands of EEG. The averaged spectral power in EEG epochs of pre seizure, seizure, post seizure and non seizure is calculated. The power distribution, particularly in alpha band and delta band is computed, thereby alpha band delta band ratio (ADR) has been calculated to detect seizure. Only involving alpha and delta band frequencies limits the performance of the. Variation in other frequency band should be monitored also for classification of broader range EEG classification problem.

In this work [56], energy, variance, peaks, sharp and spike waves, duration, events

and covariance from the EEG signals are extracted based on wavelet transformation using Hard and Soft Thresholding Methods. They analyze the performance of Elman neural networks in optimization of code converter outputs for the classification of epilepsy risk levels from EEG (Electroencephalogram) signals. The efficacy of soft and hard Thresholding methods for four different wavelets are analyzed in the extraction of features also. As a heuristic based approach, selection of threshold value is vital for the expected performance of the method.

In [57], temporal characteristics of spectral subbands of epileptic EEG data associated with different pathological states of the brain including $\delta, \theta, \alpha, \beta$ and γ subbands are examined using two features of local minima and local maxima, referred to the number of local min-max and the variance of local min-max intervals. The computational results show the substantial differences between the temporal characteristics of the epileptic EEG signal during an epileptic seizure activity and during a non-seizure period in any subband. Furthermore, the epileptic EEG signal during an epileptic seizure activity exhibits different temporal characteristics between the low and high frequency subbands as compared to the epileptic EEG signal during a non-seizure period.

2.4 Time-frequency Domain Approaches

Time-frequency based feature extraction methods provides the variation of signals with respect to both time and frequency. So, periodicity and duration of any frequency band component can be observed from time domain segmentation. On the other hand, frequency domain sub-banding provides the information of existing frequencies from where frequency based feature can be extracted.

Time-frequency method is used for seizure detection in [58] where energy based features are extracted after time-frequency domain transformation. Teager's algorithm is used for emphasizing different frequencies nonuniformly. Teager's algorithm weights component of different frequency nonuniformly, emphasizing higher frequencies over lower frequencies by square law weighting method.

Modular energy based feature extraction from time-frequency plane is introduced in [35]. Twelve Cohen class kernels are used for time-frequency smoothing. For two and three class seizure classification problem are solved with 100% accuracy

when RI kernel is used. Dependency on kernel function and classifier increases the computational time and complexity. Although, classifying five class problem is still challenging.

2.5 Conclusion

In this chapter, a brief literature survey of the recent state-of-the-art seizure detection and classification methods are presented. All the methods have their pros and cons. In order to handle the practical situations of real life applications, a seizure detection and classification method, apart from providing simplicity in computation, is needed to be capable of producing optimal results with improved overall classification accuracy for multiclass problem, where EEG signals from different part and state of the brain are involved.

Chapter 3

Epileptic Seizure Detection from EEG Signals Based on Features Extracted from Non-uniform Modules in Time-frequency Domain

3.1 Introduction

Designing a feature set which is capable of extracting distinguishable information to detect seizure data from mixture of normal and seizure EEG signals is not an easy task. Although, time-frequency analysis itself is subjected to interference problem, selection of interference reduction technique is also crucial in this case. In this chapter, we introduce a feature which represents energy distribution on time-frequency plane perfectly. These features are extracted based on five popular frequency bands, namely $\alpha, \beta, \gamma, \theta$ and δ .This feature eliminates all dependence on kernel functions of time-frequency analysis which means any kind of time or frequency smoothing is optional for feature extraction. We incorporate twelve Cohen class kernels for time and/or frequency smoothing which reduces interference due to its quadratic properties. Time or frequency smoothing for feature extraction made other methods unsuitable for onset detection. Moreover previous work used principle component analysis (PCA) which made that process slower. So exclusion of PCA reduces processing time significantly. On the other hand, simple distance based classifier can distinguish the classification problem when improved feature is used. With all of these, training data is also reduced by 10%. So reduction of processing time and

improvement of accuracy made time frequency analysis suitable for real time seizure onset detection [59].

3.2 Proposed Method

The proposed EEG based epileptic seizure detection and classification Method consists of some major steps, namely, pre-processing, time-frequency analysis, feature extraction and classification. In the classification, we consider two classes of epileptic seizure data, namely Z and S. Pre-processing manipulates the signal to be ready for time-frequency analysis and feature extraction. For the purpose of detecting epileptic seizure and to classify epileptic seizure originated from different parts and state of the brain, a train-ing database is needed to be prepared consisting of template EEG signals of different classes as well as different persons. The detection and classification task is based on comparing a test EEG signal with template data. It is obvious that considering EEG signals themselves would require extensive computations for the purpose of comparison. Thus, instead of utilizing the EEG signals, some characteristic features are extracted for preparing the template. It is to be noted that the detection and classification accuracy strongly depends upon the quality of the extracted features. Therefore, the main focus of this paper is to develop an effective feature extraction algorithm. The block diagram of the proposed method is shown in Fig. 4.1.

3.2.1 Pre-processing

A signal with no negative-frequency is called analytical signal. Just by adding an imaginary part which is the phase-quadrature component of the sinusoid, any real sinusoidal can be converted to complex sinusoidal, which has only positive frequency. All real world signals can be expressed as a sum of sinusoidal. Thus, a filter which can shift each sinusoidal by a quarter cycle can be used for making a signal analytic. This filter is termed as Hilbert transformation.

The analytical signal is expressed as

$$C_z = C_0 + iC_q, \tag{3.1}$$

where, C_0 = Original signal and C_q =quadratic shift of the original signal. Using

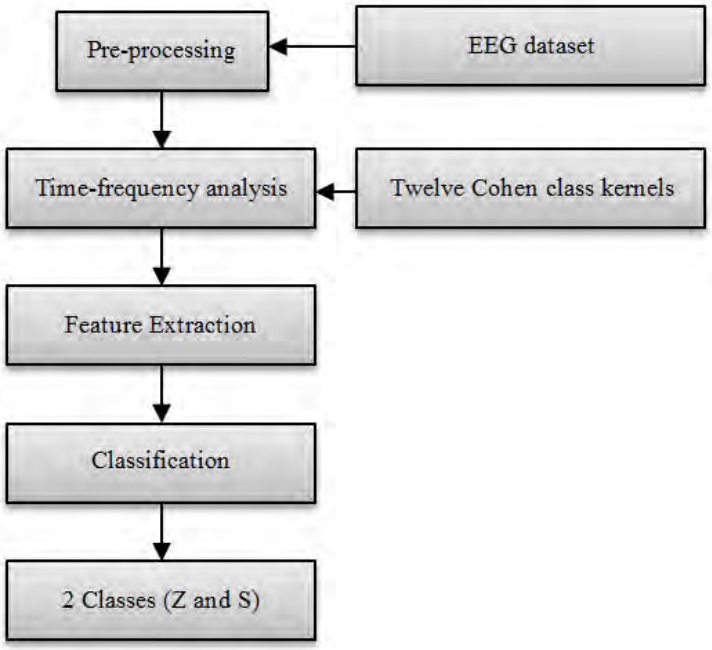


Fig. 3.1: Block diagram of the proposed method

the Cauchy Principal Value, here denoted as PV, Hilbert transform of a signal $c(t)$ can be written as

$$H[c(t)] = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{c(\tau)}{t-\tau} d\tau. \tag{3.2}$$

In Fig. 3.2, the formation of an analytical signal from a real valued signal is shown. It is seen from this figure that the negative frequency component in $c(t)$ is canceled out through Hilbert transformation thus keeping the positive frequency only in the Hilbert transformed signal, which is analytic. The transfer function of the discrete Hilbert transform is defined as

$$\mathcal{H}(\omega) = \begin{cases} j, & \text{for} \quad 0 < \omega < \pi \\ 0, & \text{for} \quad \omega = 0 \; and \; \omega = \pi \\ -j, & \text{for} \quad -\pi < \omega < 0. \end{cases} \tag{3.3}$$

The method of computing the discrete Hilbert transform is based upon its transfer function and utilizing the discrete Fourier transform (DFT) as a tool.

The Hilbert transformed signal has the same amplitude and frequency content as the original real data and includes phase information that depends on the phase of the original data. Thus, in this work, as a pre-processing, analytical representation of real EEG signal is obtained through Hilbert transformation. Such an analytical representation facilitates mathematical manipulation during time-frequency analy-

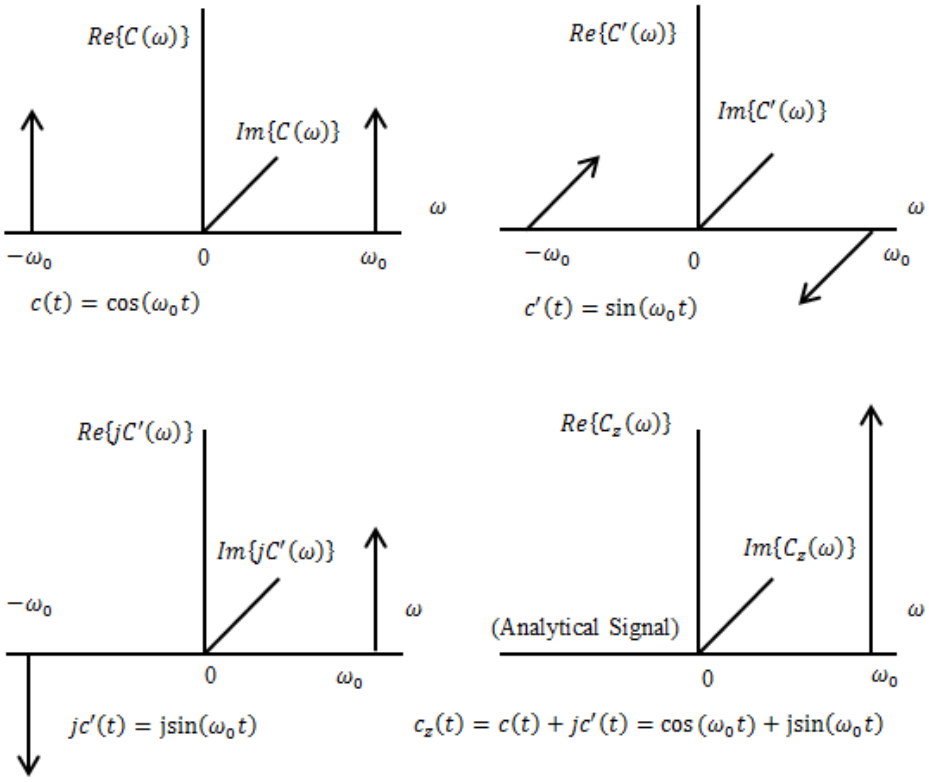


Fig. 3.2: Formation of an analytic signal from a real-valued signal

sis, which is performed prior to feature extraction.

3.2.2 Time-frequency Analysis

In order to find the dominance pattern of an EEG signal in a short time window, short time discrete Fourier transform is introduced, which is the basic frequency analysis technique decomposing the signal into individual frequency component. But, here, there is a limitation in the sense of making a tradeoff between time and frequency resolution. Since, EEG is a non-stationary signal and distribution of its energy demonstrates the seizure activities, time-frequency distribution can perform better than the conventional frequency analysis methods [60].

Gabor, Ville and Page try to find out an alternative to analyze non-stationary signals [61]. The primary motivation is to introduce a joint function of time and frequency that will describe energy density or intensity of a signal in time and frequency domain simultaneously. But simultaneous representation yields a distribution, which may satisfy many desirable properties. Since, such a distribution is quadratic, it introduces cross term interference, which reduces the reliability of time

frequency analysis.

For the pre-processed EEG signal, we obtain time-frequency distributions by using twelve Cohen Class kernels [35]. Cohen class distributions thus obtained utilize time and frequency co-variance properties. The general expression for such a distribution is as follows

$$\mathcal{C}(t,\omega) = \frac{1}{4\pi^2}\iiint e^{-j(\theta t+\tau\omega-\theta u)}\phi(\theta,\tau)c^*(u-\frac{\tau}{2})c(u+\frac{\tau}{2})dud\tau d\theta, \tag{3.4}$$

here, $\phi(\theta,\tau)$ is known as parameterization function or kernel function. For time and/or frequency smoothing, we employ twelve kernel functions that are capable of reducing interference. The twelve Cohen class time-frequency distributions with corresponding kernel function representations are summarized in Table 3.1, where, α symbolizes dissemmetry ratio, γ stands for scaling factor and $h(\tau)$ represents window function. Among all kernel functions, Margenau-Hill (MH), Wigner-Ville

Table 3.1: Time-Frequency Distribution

Serial	Name of Distribution	Kernel representation
1	Margenau-Hill (MH)	$\cos(\pi\theta\tau)$
2	Wigner-Ville (WV)	1
3	Rihaczak (RIH)	$e^{j\pi\theta\tau}$
4	Pseudo-MH (PMH)	$h(\tau)e^{j\pi\theta\tau}$
5	Pseudo-WV (PWV)	$h(\tau)$
6	Born-Jordan (BJ)	$\frac{\sin(\pi\theta\tau)}{\pi\theta\tau}$
7	Butterworth (BUT)	$(1+(\frac{\theta}{\theta_1})^{2N}+(\frac{\tau}{\tau_1})^{2M})^{-1}$ $\theta_1,\tau_1,N,M>0$
8	Choi-Williams (CW)	$e^{-(\frac{\pi\theta\tau}{\gamma})^2}$
9	Generalized rectangular (GRECT)	$sin(\frac{2\pi\gamma\theta}{\lvert\tau\rvert^\alpha})/(\pi\theta)$
10	Reduced Interference (RI)	$\int_{-\infty}^{+\infty} h(t)e^{-j2\pi\theta\tau t}dt$
11	Smoothed-PWV (SPWV))	$Q(\theta)h(\tau)$
12	Zhao-Atlas-Marks (ZAM)	$h(\tau)\sin(\pi\theta\tau)/(\pi\theta\tau)$

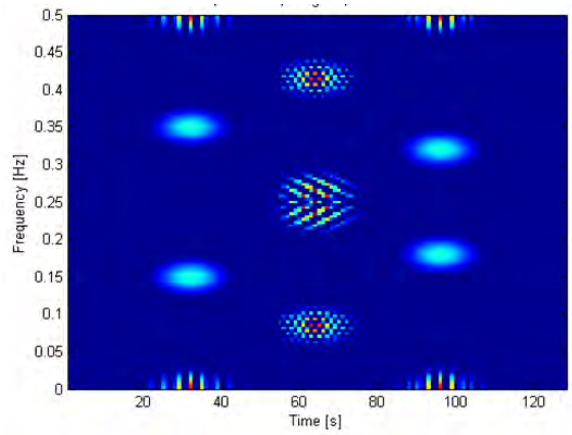
(WV) and Rihaczek (RIH) are basic kernel functions, which do not employ any of

kind of smoothing. Pseudo-Margenau-Hill (PMH) and Pseudo-Wigner-Ville(PWV) are the frequency domain smoothed version of the basic kernels as mentioned above, where $h(\tau)$ is used as windowing function in the mathematical representation. Born-Jordan (BJ), Butterworth (BUT), Choi-Williams (CW), Generalized rectangular (GRECT), Reduced Interference (RI), Smoothed Pseudo Wigner-Ville (SPWV) and Zhao-Atlas-Marks (ZAM) utilize both time and frequency domain smoothing [61].

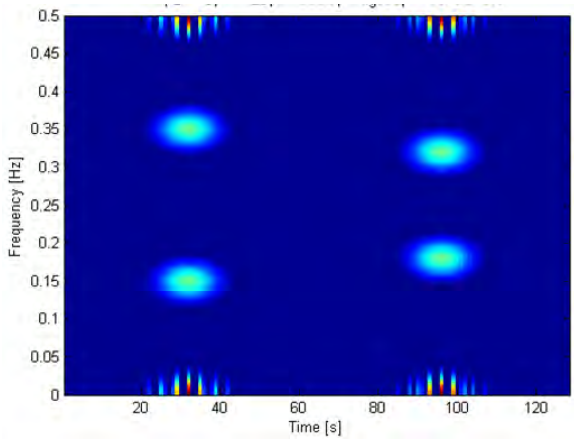
The effect of using a time and/or frequency smoothing kernel function on interference can be analyzed using a synthetic signal. The signal consists of two sinusoidal is generated such that one sinusoidal remains for the whole time under consideration and the another one appears in the middle for a while with a different frequency. We employ basic WV distribution using no smoothing, PWV distribution with frequency smoothing kernel function and SPWV distribution with both time and frequency smoothing to realize the effect on interference and resulting time frequency resolution of the distribution.

In Fig. 3.3 time-frequency representation by using WV, PWV and SPWV kernel functions are shown. From Fig. 3.3(a), it is seen that both the frequencies are spotted perfectly while using WV kernel. Comparing Figs. 3.3(a) through (c), it is found that time and frequency resolution degrades in an increasing order if we use no smoothing, frequency smoothing and time-frequency smoothing, respectively. From Fig. 3.3(c) representing SPWV distribution, it is clear that the horizontal frequency line seems thicker thus attesting the degradation in the time-frequency resolution. But the cross-term interference reduces gradually when moving from a no smoothing to a frequency smoothing kernel function and no interferences are observed while using the time and frequency smoothing kernel function.

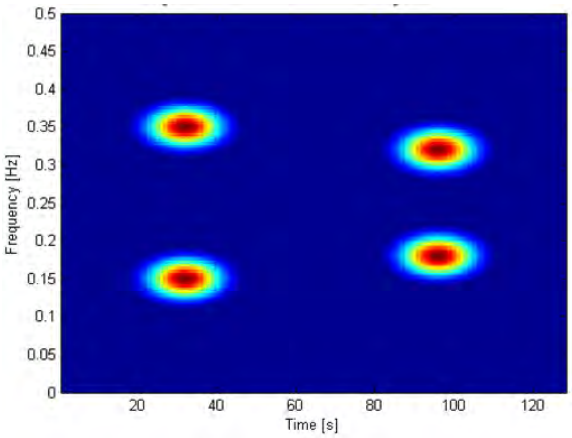
Thus, after performing time frequency analysis, distribution of EEG signal at time-frequency plane is found to be capable of extracting features that can characterize EEG data fully to distinguish different classes of epileptic seizure. However, effectiveness of time-frequency based feature extraction depends on the choice of a kernel function.



(a) WV kernel



(b) PWV kernel



(c) SPWV kernel

Fig. 3.3: Time-frequency representation by using WV, PWV and SPWV kernel functions

3.2.3 Feature Extraction

The most important task here is to extract distinguishing features from the template data, which directly dictates the detection and classification accuracy. For epileptic

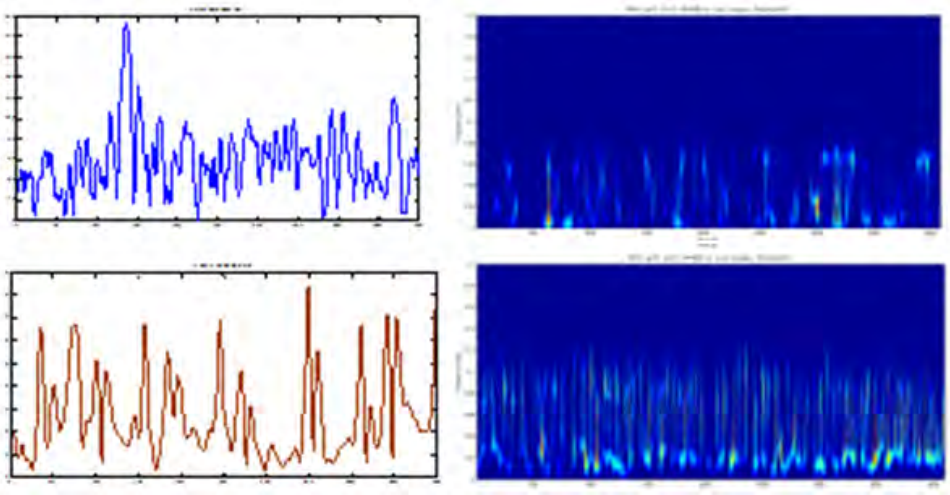


Fig. 3.4: Time domain (left) and time-frequency domain (right) plots of Z and S

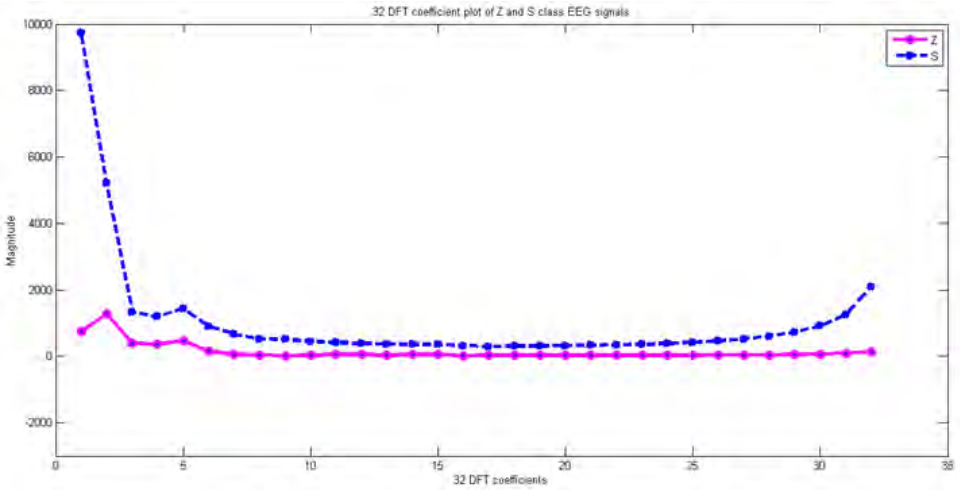


Fig. 3.5: Magnitude plot of 32 DFT coefficients of Z and S.

seizures from different classes, obtaining a significant feature space from the time domain variations as plotted in left column of Fig.4.2 is very challenging even for an expert. In Fig. 4.3, frequency domain representations in terms of Discrete Fourier Transform (DFT) coefficients for different classes of EEG signals are shown.

Comparing Figs. 4.2 and 4.3, it is clear that direct subjective correspondence between EEG features in the time domain and those in the frequency domain is not very apparent. In what follows, we are going to demonstrate the proposed feature extraction algorithm, where time-frequency domain local variation is extracted instead of considering only either time or frequency domain variation. After performing time frequency analysis, important information is extracted from the time-frequency plane as detailed as possible to resolve the detection problem.

In the right column of Fig. 4.2, epileptic seizures of different classes are represented in time-frequency plane using RI distribution. It is vivid from Fig. 4.2 that time-frequency representation has minor variation in magnitude at different frequencies that correspond to different time. Epileptic seizures from different classes are separable in terms of these minor variations in magnitude. These minor variations may not be properly enhanced or captured when considering the whole time-frequency plane and may not contribute to the feature vector. Hence, in that case, the feature vectors of the different EEG signals shown in Fig. 4.2 may be similar enough to be wrongfully classified as if they belong to the same class. Therefore, we propose to extract features from local zones of the time-frequency plane.

The proposed feature extraction algorithm is based on extracting minor variations precisely from a number of modules in the time-frequency plane of the EEG signals instead of utilizing the time-frequency plane as a whole. In view of this, a nonuniform modularization technique is employed first to segment the entire time-frequency plane. So, in case of modularization, the minor variations can be compared locally which should be more precise than that of the global comparison. Every module provides some local differences between class to class which can ensure that all distinguishable data are compared. The significance of modularization is that it helps extracting local information which may be lost in a global case.

We use a modularization approach based on performing a partitioning both in the time and the frequency axes. In the time domain, three equal-sized windows are selected, while in the frequency domain, the employed partition divided the frequency domain in five subbands, which are defined based on medical knowledge on EEG.

The dominant frequency bands in EEG signals correspond to δ (0-3 Hz), θ (4-7 Hz), α (8-13 Hz), β (14-30 Hz) and $\gamma(> 30Hz)$ components or waves [62]. The EEG signal is closely related to the level of consciousness of the person. As the activity increases, the EEG shifts to higher dominating frequency and lower amplitude. An example of each waveform is given in Fig. 3.6. δ wave has large and slow deflections, which is observed at frontal and at posterior in children and adults, respectively. δ waves occur in very deep sleep, in infancy, and in serious organic brain disease. In a certain phase of sleep, rapid eye movement called (REM) sleep, the

person dreams and has active movements of the eyes, which can be characterized from EEG signal. θ waves occur during emotional stress in some adults, particularly, during disappointment and frustration. α wave is dominant in the posterior region and in central region during rest. α waves are found in the EEGs of almost all normal adult people and can be measured from the occipital region when they are awake in a quiet, resting state of cerebration. When the eyes are closed, the α waves begin to dominate the EEG, whereas when the person falls asleep, the dominant EEG frequency decreases and α waves disappear. When the awake person's attention is directed to some specific type of mental activity, α waves are replaced by asynchronous, higher frequency but lower voltage β waves, which are detectable over the parietal and frontal lobes. γ wave is originated at somatosensory cortex. Study in [62] speculated that the function of γ band oscillation is to provide a reference clock to control the firing of the excitatory neurons during different mental activity. From EEG, no cerebral activity can be detected from a patient with complete cerebral death. Therefore, for analysis purpose, EEG specialists select five frequency bands, which vary at different parts and state of the brain.

In particular, for EEG signals in seizure and non-seizure cases, distribution of energy at different frequency bands are shown in Fig. 3.7. In EEG signal during epileptic seizure, among δ, θ, α, β or γ frequency bands, one or several bands may have dominance and it is possible to differentiate α, β, δ, γ and θ waves as well as spikes associated with epilepsy. Thus, from different EEG channels, experts can detect epileptic seizure using spatial distribution of dominant frequency bands. The dominance of the frequency bands vary among EEG signals of different classes of epileptic seizure originating from different parts and state of the brain. So, we segment the frequency domain according to these well established five frequency bands. The purpose of the time window is to obtain three segments, where the variation in time among different classes of epileptic seizure EEG signals is analyzed. The size of the time window is within the range of windows selected in [63],where it is defined using expert neurologist knowledge. Therefore, based on the five popular EEG frequency bands and three time windows as discussed above, we modularize the whole time-frequency plane in 5×3 nonuniform modules. Each module thus obtained can be used to extract specific features, which can ensure some local differences

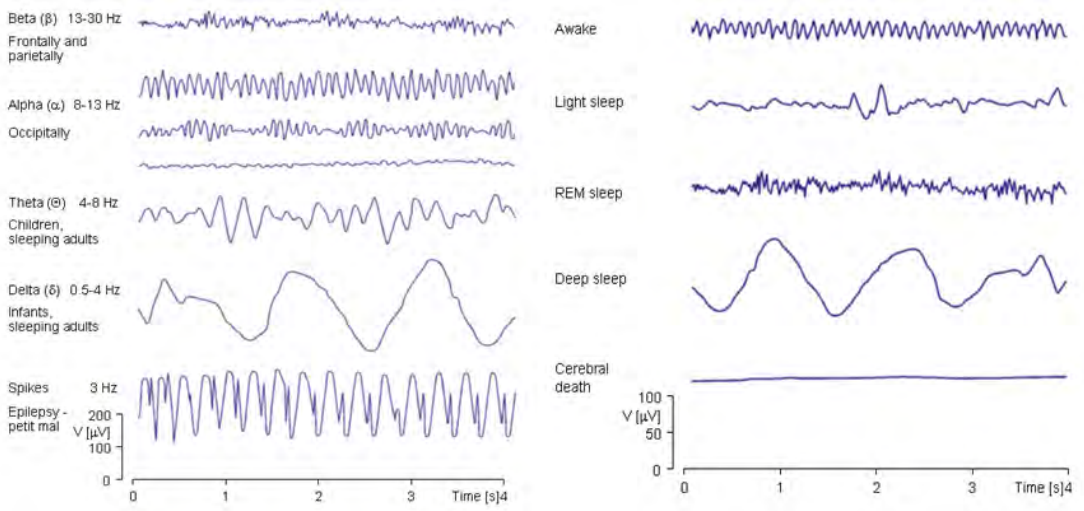
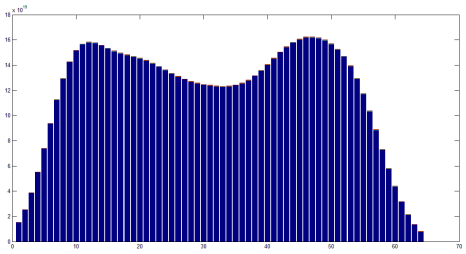
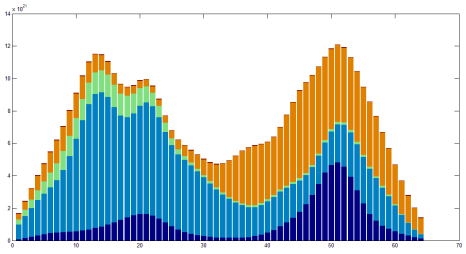


Fig. 3.6: Examples of some EEG signals or waves at different state of the brain

between class to class. Fig. 3.8 presents the nonuniform modules used for feature extraction.



(a) Non-seizure EEG signal



(b) Seizure EEG signal

Fig. 3.7: Distribution of Energy at different frequency bands for seizure and non-seizure EEG signals.

In order to observe the impact of nonuniform modularization as discussed above, we choose well-known DFT based algorithms that offer ease of implementation in practical applications. For a function $g(x, y)$ of size $m \times n$ with two-dimensional (2D) variation, the 2D discrete Fourier transform (2D-DFT) is given by [64]

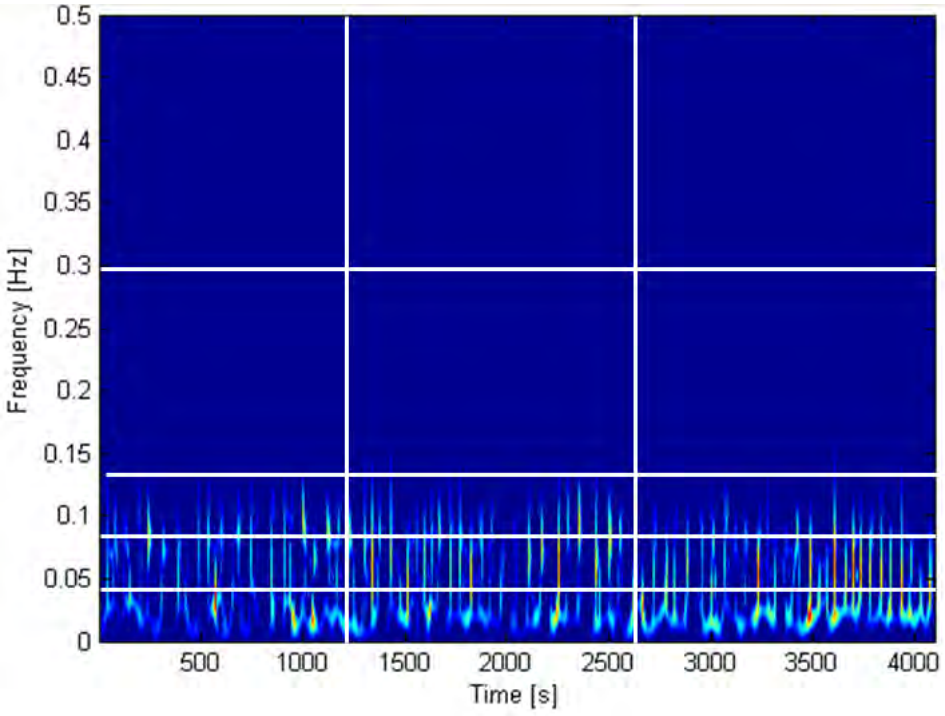


Fig. 3.8: Non-uniform modularization of time-frequency plane

$$G(u,v) = \frac{1}{m \times n} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} g(x,y) e^{-j2\pi(\frac{ux}{m} + \frac{uy}{n})} \tag{3.5}$$

where $u = 1, 2, 3, \ldots\ldots, m-1$ and $v = 1, 2, 3 \ldots\ldots, n-1$. By considering one random sample from each class and obtaining time frequency representations using RI distribution, the coefficients of 2D-DFT plots of Z and S in non-modular case of whole time-frequency plane is shown in the upper row of Fig. 4.4 shows From these figures, it is evident that, in case of the non-modular time-frequency plane, there exists no significant difference in terms of 2D-DFT coefficients among the five classes of EEG signals. Hence, they are difficult to distinguish from each other.

On the other hand, the lower row of Fig. 4.4 shows the coefficients of 2D-DFT plots of Z and S EEG signals in modular case. Here, for each sample of a particular class, the coefficients of 2D-DFT are calculated for all the modules belonging to the time-frequency plane and average value of these coefficients are considered to form the modular coefficient matrix. The figures in the lower row attest that the undistinguishable difference in time signals as well as in non-modular time-frequency plane are clearly signified and enhanced in terms of the 2D-DFT coefficients obtained employing modularization technique. For both the non-modular and modular cases,

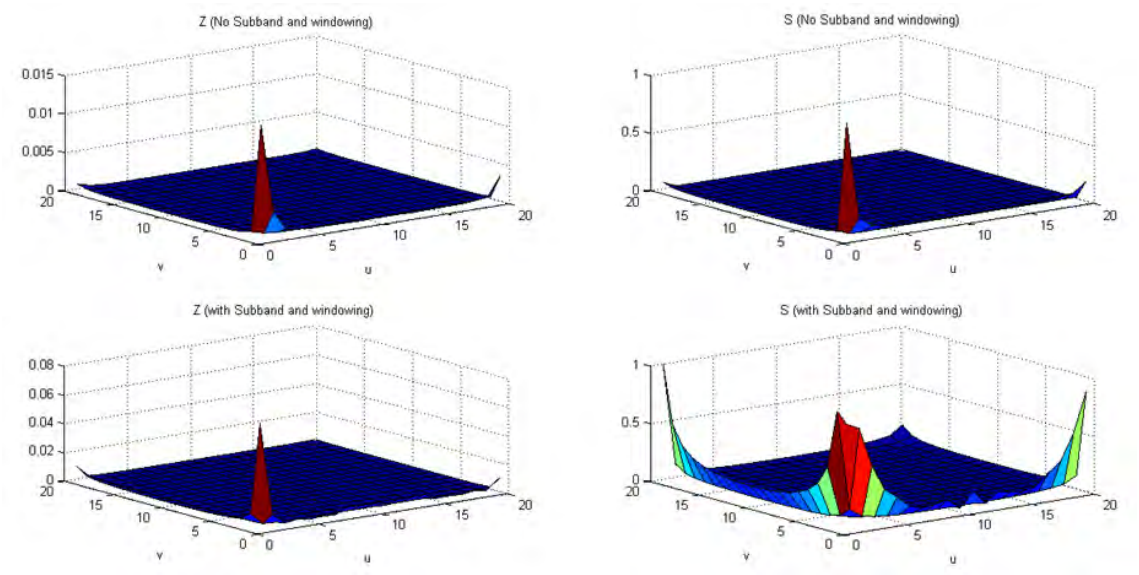


Fig. 3.9: Coefficients of 2D-DFT plot of all classes in both modular (Lower row) and non-modular (Upper row) cases.

2D-DFT coefficients of both classes are normalized with respect to the maximum coefficient value of the S class.

To investigate the similarity among different classes, we find the Euclidean distances between the 2D-DFT coefficients of these two classes as shown in Fig. 4.5. From the figure, it is vivid that Euclidean distance is increased significantly after modularization.

It is observed from Fig. 4.5 that the Euclidean distance is increased significantly, in particular, by 144% in case of modularization compared to the distances obtained in non-modular cases. The analysis as performed above clearly indicates that better distinguishable features are extracted from nonuniform smaller modules than that from the entire time-frequency plane as a whole. Change of class position are clearly shown in Fig. 4.7 in a virtual plane, where distance between Z and S are increased after modularization.

However, instead of taking the 2D-DFT coefficients of the entire time-frequency plane, if all the coefficients obtained from each module are considered to be used as a feature, it would definitely result in a feature vector with a very large dimension. In view of reducing the feature dimension, we propose to utilize one or two feature from each module thus reducing the length of the feature vector. The length of the feature vector becomes 199% higher if all 2D-DFT coefficients from each module are used as features instead of using one or two feature per module. The complexity

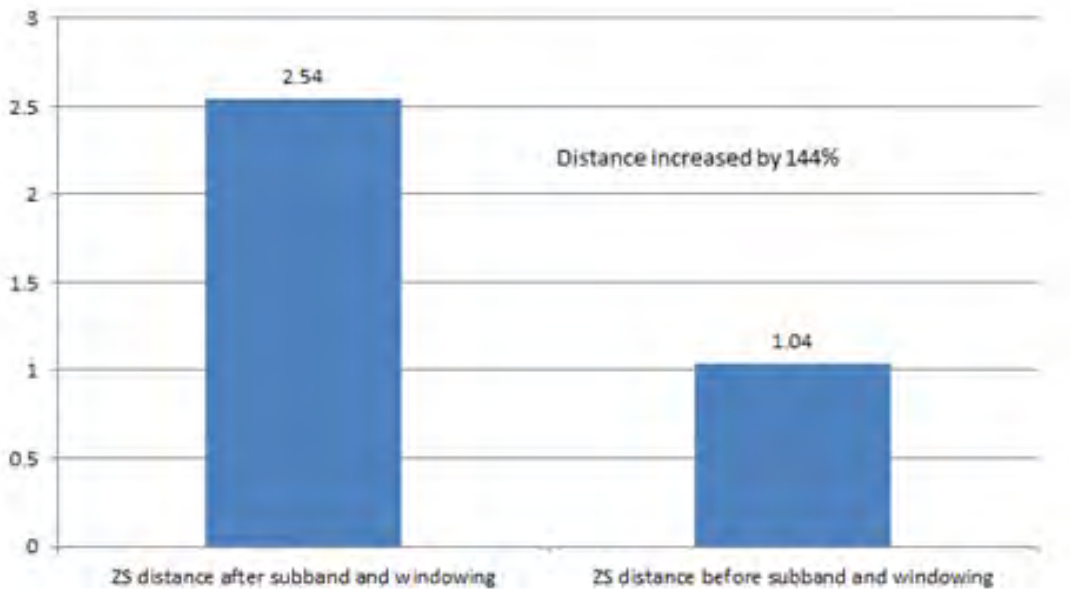


Fig. 3.10: Comparison of Euclidean distances for non-modular and modular cases in terms of 2D-DFT coefficients

of feature extraction and ease of its implementation is another important criterion while developing a feature vector for multiclass epileptic seizure classification. Since computation of 2D-DFT coefficients is expensive, we propose to extract cumulative energy at marked percentile frequency and modular entropy from each module

3.2.4 Cumulative Energy

Time-frequency distribution is used to calculate power spectrum density (PSD) which represents energy distribution at time-frequency domain. The energy at a time-frequency module is the sum of PSD H at each time and frequency range corresponding to the module, which is expressed as

$$E_{energy}(l,r) = \int_{t_{l(start)}}^{t_{l(end)}} \int_{\omega_{r(start)}}^{\omega_{r(end)}} H(t,\omega) d\omega dt. \tag{3.6}$$

We propose the cumulative energy at some predefined marked frequency percentiles. The percentiles are selected with respect to the maximum frequency (f_{max}). For a (l,r)-th module, the cumulative energy at the k-th percentile of frequency is expressed as

$$E^{(l,r)}_{energy}(k) = E^{(l,r)}_{energy}(k-1) + \int_{\omega_{k(start)}}^{\omega_{k(end)}} H^{(l,r)}(\omega) d\omega. \tag{3.7}$$

Thus, instead of providing a single total energy, a distribution of energy pattern per module is obtained. The cumulative energy at marked frequency percentiles

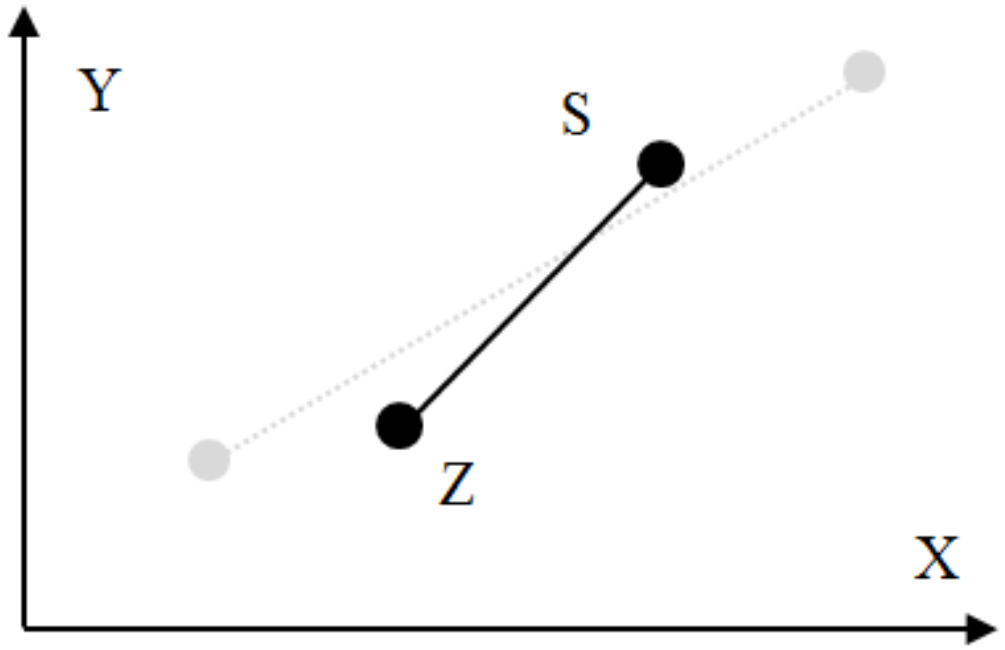


Fig. 3.11: Comparison of reduction in the length of feature vector in modularization case.

characterizes the distribution of energy in the frequency range of the module. We use fourteen predefined percentile marks to find the energy distribution. It is to be noted that at each module, the cumulative energy at f_{max} representing the total energy is also included in the formation of the cumulative energy. Although, seizure activities contain high frequency variation, most of the energy of EEG signal is concentrated in low frequency region, where the difference in variation of energy from class to class is prominent. Thus, the lower frequency is prioritized to define those percentile marks in order to capture the variation between classes. In Fig. 3.12, cumulative energy plots of seizure(S) and non-seizure(Z) data are shown for fourteen predefined frequencies at each module. Value of each cumulative energy is normalized with respect to total energy at f_{max} to compare the pattern between Z and S. It is seen from this figure that since the non-seizure signal(Z) involves low frequencies, its cumulative energy becomes saturated at a faster rate than that obtained from the seizure(S) data.

By using equation.(3.7), the cumulative energy in the whole time-frequency plane can be obtained as the sum of cumulative energy calculated at all the modules.

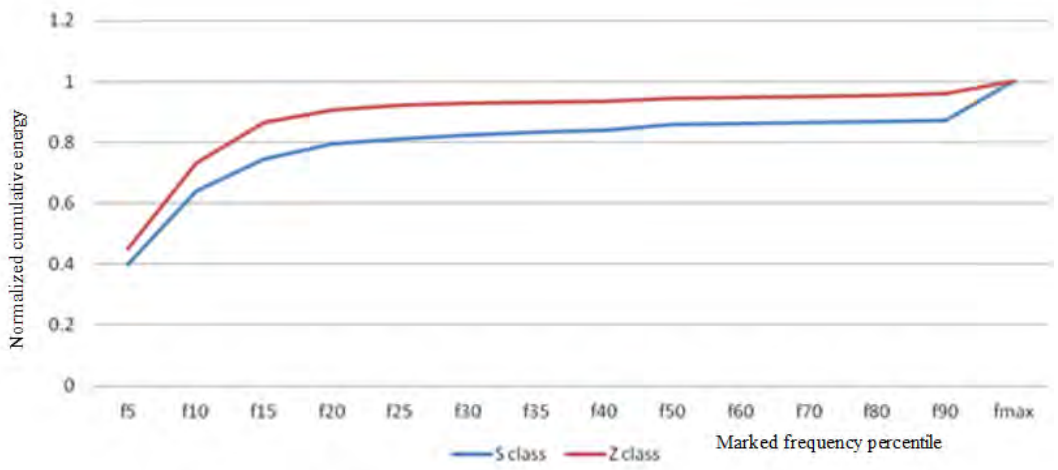


Fig. 3.12: Modular cumulative energy for seizure and non-seizure data at each module.

$$E_{energy}(\bar{k}) = \frac{1}{M+N} \sum_{l=0}^{M-1} \sum_{r=0}^{N-1} E_{energy}^{(l,r)}(k). \tag{3.8}$$

3.2.5 Modular Entropy

Entropy is first introduced in physics to describe the system disorganization, which is later adopted in signal processing for the same purpose. The definition of entropy in signal processing is based on the hypothesis that noise is totally disorganized thus carrying the highest entropy. Entropy of a signal can be defined as

$$E_{entropy}(\chi) = -\sum_{i=1}^{N} p(\chi_i) \log_2 p(\chi_i), \tag{3.9}$$

where $\chi = \{\chi_1, \chi_2, \chi_3, \chi_N\}$ is a set of random data, such as EEG, and $p(\chi_i)$ is the probability of a random data χ_i. Since, entropy represents the information content in a signal, entropy of each module $E_{modular-entropy}(l, r)$ can extract local information content.

Therefore, for classifying different EEG data, the proposed multi-featured set can be formed as

$$F = \{E_{energy}(\bar{k}); E_{modular-entropy}(l, r)\}, \tag{3.10}$$

l and r are the module number in the time and frequency domain, respectively where $\bar{k} = 0, 1, ..., 13;\ l = 0, 1..., 4\ and\ r = 0, 1, 2.$

3.2.6 Classification

In this paper, we employ six different classifiers to deter-mine the efficacy of the feature vector in classifying epileptic seizure originating from different parts and states of the brain. Some of these classifiers are linear and few are non-linear.

k-nearest neighbor (KNN) classifier: The k-nearest neigh-bor (KNN) is simplest linear classifier [65]. Here, an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors, where k is a typically small positive integer. If k = 1, then the object is simply assigned to the class of its nearest neighbor. In a KNN classifier, different types of mathematical distances is used to rate all neighbors. Among them, KNN classifier with Euclidian distance is attractive in the sense of reducing the processing time.

In the Euclidean distance based classifier, the classification task is carried out based on the Euclidean distances between the feature vectors of the training EEG and the feature vectors of the testing EEG. Given the q-dimensional feature vector for the f-th training EEG belonging to class ψ be $\{\alpha_{\psi f}(1), \alpha_{\psi f}(2), ..., \alpha_{\psi f}(q), \}$ and a ν-th test EEG with a feature vector $\{\beta_\nu(1), \beta_\nu(2), ..., \beta_\nu(q)\}$, Euclidian distance is measured between the test EEG ν and the training EEG f belonging to class ψ as

$$ED^\nu_{\psi f} = \sqrt{\sum_{i=1}^{q} |\alpha_{\psi f}(f) - \beta_\nu(i)|^2} \tag{3.11}$$

Considering total Γ training EEG data belonging to class ψ minimum Euclidean distance is obtained from

$$ED^\nu_\psi = min^\Gamma_{f=1} ED^\nu_{\psi f} \tag{3.12}$$

Therefore, test EEG beat will be classified as ψ class among Ψ number of classes if it satisfies the condition

$$ED^\nu_\psi < ED^\nu_a, \ \forall \psi \neq a, \ \forall a \in 1, 2, ..., \Psi \tag{3.13}$$

In this paper, we are interested to handle a five-class problem ($\Psi = 5$). We also used cityblock, cosine and correlation as distance function in the KNN classifier. The implementation equations for all the distance functions are summarized in Table 3.2.

Generally, linear classifiers are easy to compute but have lesser accuracy rate than non-linear ones. Among non-linear classifiers, namely decision trees and artificial

Table 3.2: Implementation Equation of Distance Functions

Distance	Implementing Equation
Euclidean	$\sum_{i=1}^{L}(W_i - V_i)^2$
City-block	$\sum_{i=1}^{L}\lvert W_i - V_i\rvert$
Cosine	$1 - \frac{\sum_{i=1}^{L} W_i V_i}{\sqrt{\sum_{i=1}^{L} W_i^2}\sqrt{\sum_{i=1}^{L} V_i^2}}$
Correlation	$1 - \frac{\sum_{i=1}^{L}(W_i - W_{avg})-(V_i - V_{avg})}{\sqrt{\sum_{i=1}^{L}(W_i - W_{avg})^2}\sqrt{\sum_{i=1}^{L}(V_i - V_{avg})^2}}$

neural network (ANN), ANN is more popular in approaches for epileptic seizure classification.

Decision trees: Based on search heuristics, decision trees find explicit and understandable rules-like relationships among the input and output variables [66]. The original data is split into finer subsets using recursive partitioning algorithms in the search heuristics. In order to maximize the in-formation gain, the algorithm finds the optimum number of splits and determines where to partition the data. Since there are fewer rules to understand in case of fewer the splits, the output becomes more explainable. For indicating the variables, conditions, and outcomes, decision trees are built of nodes, branches, and leaves, respectively. The most predictive variable is placed at the top node of the tree. In this work, the C4.5 decision tree induction algorithm is employed.

Artificial neural network (ANN): Inspired by the structure and/or functional aspects of biological neural networks, ANN is developed as a mathematical or computational model [67, 68]. A neural network processes information using a connectionist approach to computation as it consists of an interconnected group of artificial neurons. In most cases, based on external or internal information that flows through the network during the learning phase, an ANN changes its structure thus it can be treated as an adaptive system. Modern neural networks are non-linear statistical data modeling tools that are usually used to model complex relationships between inputs and outputs or to find patterns in data, such as EEG.

3.3 Conclusion

Conventional time or frequency domain analysis is found inadequate to describe the characteristics of a non-stationary signal such as, EEG. In this chapter, we propose to transform the EEG data using twelve Cohen class kernels in order to facilitate the time-frequency analysis. The transformed data thus obtained is exploited to formulate modular entropy and cumulative frequency-related energy that can better model the time-frequency behavior of the EEG data. These features contribute to develop a multifeatured vector that is separately fed to each of KNN with four different distance calculation method, ANN and Decision tree classifiers in order to classify epileptic seizure and normal EEG data. It will be shown that even with a less complex fast processing kernel, a simple classifier and a smaller training data set, the proposed method is capable of producing 100 percent accuracy in epileptic seizure/normal EEG data detection. But, this non-uniform feature extraction method has less classification accuracy for higher class classification problem. So, for three and five class problems, more effective information should be extracted by the feature set to achieve higher classification performance. Therefore, we propose another uniform modularization technique for feature extraction which can perform better in case of higher classification problem.

Chapter 4

Uniform Modular Feature Based Multiclass Epileptic Seizure Classification

4.1 Introduction

Classification of EEG data needs better information extraction from time-frequency plane. If classification problem involves EEG data from different part and state of the brain from the same person, more precise and distinguishable feature is needed to be extracted from time-frequency plane. So, extraction of local information rather than global one is required for better perfection. In this chapter, we introduce a highly modularized feature, which represents all relevant information in time-frequency plane perfectly. Uniform modularization up to a certain level is proved effective for multiclass EEG classification [69]. We incorporate twelve Cohen class kernels for time and/or frequency smoothing, which reduces interference due to its quadratic properties. By using this method, we can get the highest classification performance for multiclass epileptic seizure classification problem. We consider three different classes, namely two, three and five class problem to verify the performance of the proposed method.

4.2 Proposed Method

The proposed EEG based epileptic seizure detection and classification Method consists of some major steps, namely, pre-processing, time-frequency analysis, feature extraction and classification as describe at previous chapter. In this classification, we consider three different classification problem, namely two class, three class and five

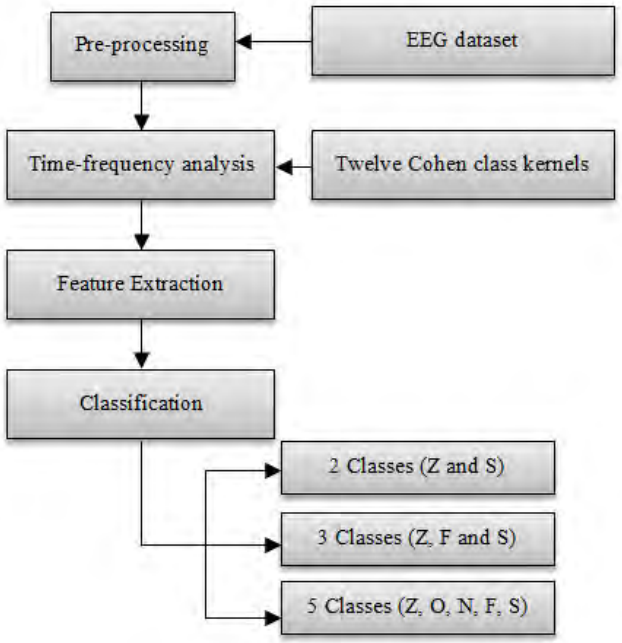


Fig. 4.1: Block diagram of the proposed method

class problem. Pre-processing manipulates the signal to be ready for time-frequency analysis and feature extraction. We incorporate Hilbert transform in this method also as stated in previous chapter. The Hilbert transformed signal is fed for time-frequency decomposition using twelve Cohen class kernels as previous. Extracted feature set is fed to six different classifiers which are introduced in the previous chapter.Then uniform modularization is employed for feature extraction. So, the proposed method for seizure classification can be shown as Fig. 4.1.

4.2.1 Feature Extraction

The most important task here is to extract distinguishing features from the template data, which directly dictates the detection and classification accuracy. For epileptic seizures from different classes, obtaining a significant feature space from the time domain variations as plotted in left column of Fig.4.2 is very challenging even for an expert. In Fig. 4.3, frequency domain representations in terms of Discrete Fourier Transform (DFT) coefficients for different classes of EEG signals are shown.

Comparing Figs. 4.2 and 4.3, it is clear that direct subjective correspondence between EEG features in the time domain and those in the frequency domain is not very apparent. In what follows, we are going to demonstrate the proposed feature extraction algorithm, where time-frequency domain local variation is ex-

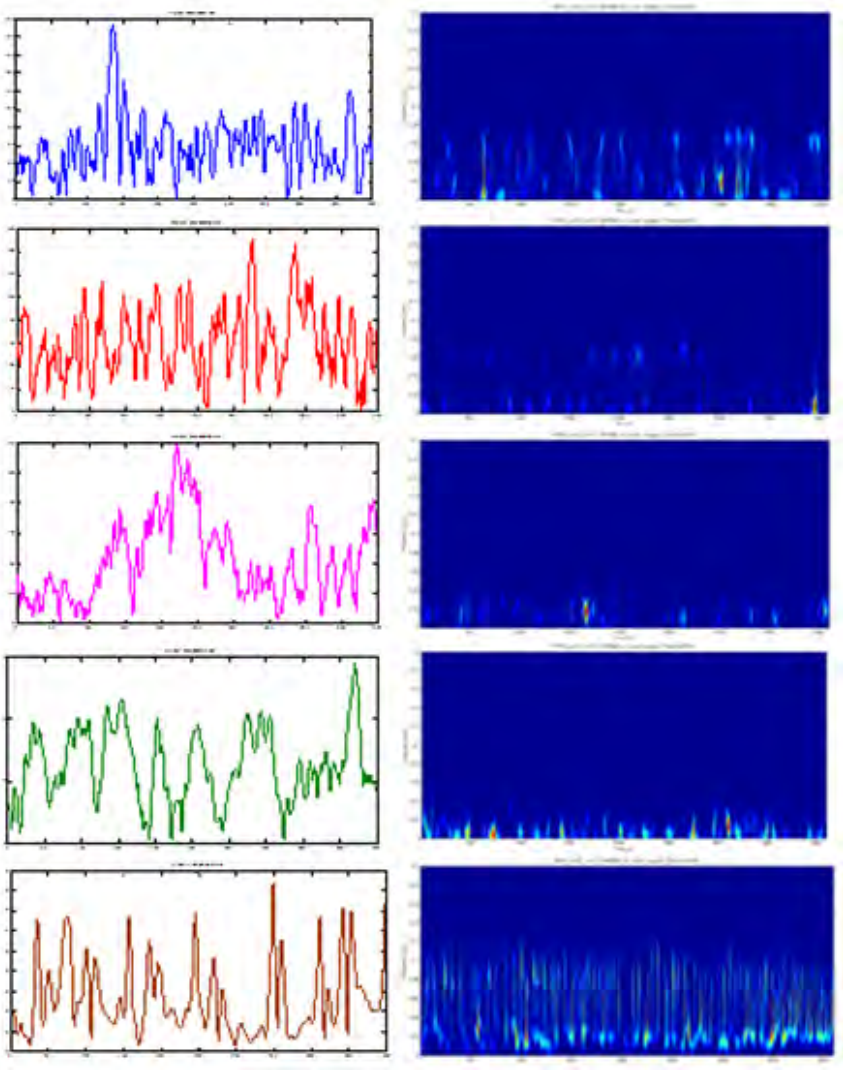


Fig. 4.2: Time domain (left) and time-frequency domain (right) plots of all class

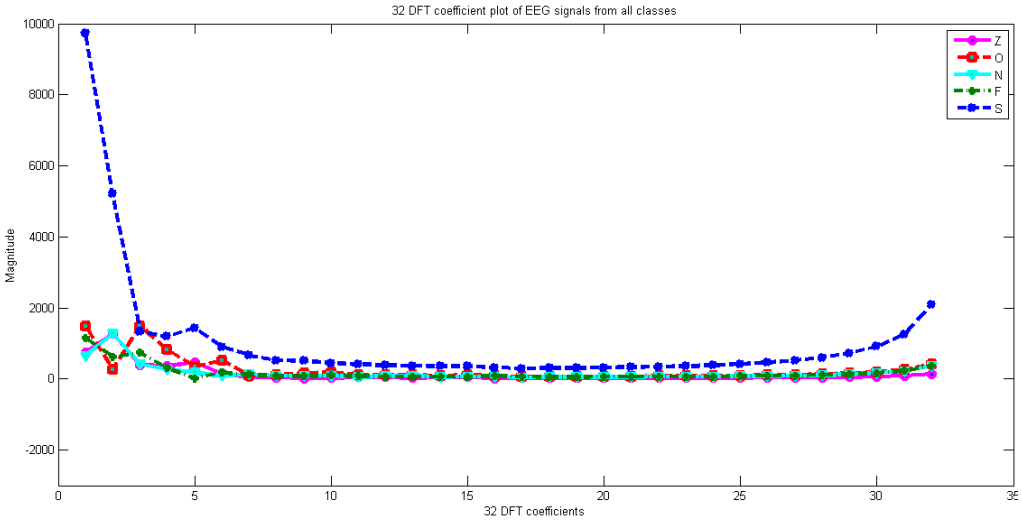


Fig. 4.3: Magnitude plot of 32 DFT coefficients of Z, O, N, F and S.

tracted instead of considering only either time or frequency domain variation. After performing time frequency analysis, important information is extracted from the time-frequency plane as detailed as possible to resolve the detection and multiclass classification problem.

In the right column of Fig. 4.2, epileptic seizures of five different classes are represented in time-frequency plane using RI distribution. It is vivid from Fig. 4.2 that time-frequency representation has minor variation in magnitude at different frequencies that correspond to different time. Epileptic seizures from five different classes are separable in terms of these minor variations in magnitude. These minor variations may not be properly enhanced or captured when considering the whole time-frequency plane and may not contribute to the feature vector. Hence, in that case, the feature vectors of the different EEG signals shown in Fig. 4.2 may be similar enough to be wrongfully classified as if they belong to the same class. Therefore, we propose to extract features from local zones of the time-frequency plane.

The proposed feature extraction algorithm is based on extracting minor variations precisely from a number of modules in the time-frequency plane of the EEG signals instead of utilizing the time-frequency plane as a whole. In view of this, a uniform modularization technique is employed first to segment the entire time-frequency plane. So, in case of modularization, the minor variations can be compared locally which should be more precise than that of the global comparison. Every module provides some local differences between class to class which can ensure that all dis-

tinguishable data are compared. The significance of modularization is that it helps extracting local information which may be lost in a global case.

It is well-known that DFT based algorithms offer ease of implementation in practical applications. DFT coefficient of modular and non-modular cases are calculated as equation. 3.5. In order to observe the modularization impact, we choose one random sample from each class and obtain time frequency representations using RI distribution. We modularize the whole time-frequency plane in uniform 32×32 modules. The upper row of Fig. 4.4 shows the coefficients of 2D-DFT plots of all classes of epileptic EEG signals in non-modular case of whole time-frequency plane. From these figures, it is evident that, in case of the non-modular time-frequency plane, there exists no significant difference in terms of 2D-DFT coefficients among the five classes of EEG signals. Hence, they are difficult to distinguish from each other.

On the other hand, the lower row of Fig. 4.4 shows the coefficients of 2D-DFT plots of all classes of epileptic EEG signals in modular case. Here, for each sample of a particular class, the coefficients of 2D-DFT are calculated for all the modules belonging to the time-frequency plane and average value of these coefficients are considered to form the modular coefficient matrix. The figures in the lower row attest that the undistinguishable difference in time signals as well as in non-modular time-frequency plane are clearly signified and enhanced in terms of the 2D-DFT coefficients obtained employing modularization technique. For both the non-modular and modular cases, 2D-DFT coefficients of every class are normalized with respect to the maximum coefficient value of the S class.

To investigate the similarity among different classes, we find the Euclidean distances between the 2D-DFT coefficients of every two classes as shown in Fig. 4.5. For a five class problem, there exist ten possible distances. Distances: ZO, ZN, ZF, ZS, ON, OF, OS, NF, NS and FS are plotted and numbered sequentially from one to ten in Fig. 4.5. Thus Fig. 4.5 shows a comparison between these Euclidean distances determined for both non-modular and modular cases. In the former case, where the whole time-frequency plane is considered, the value of the Euclidean distance is smaller than that obtained in the latter case, where only the 2D-DFT coefficients of the modules are considered. For both non-modular and modular cases, it may infer

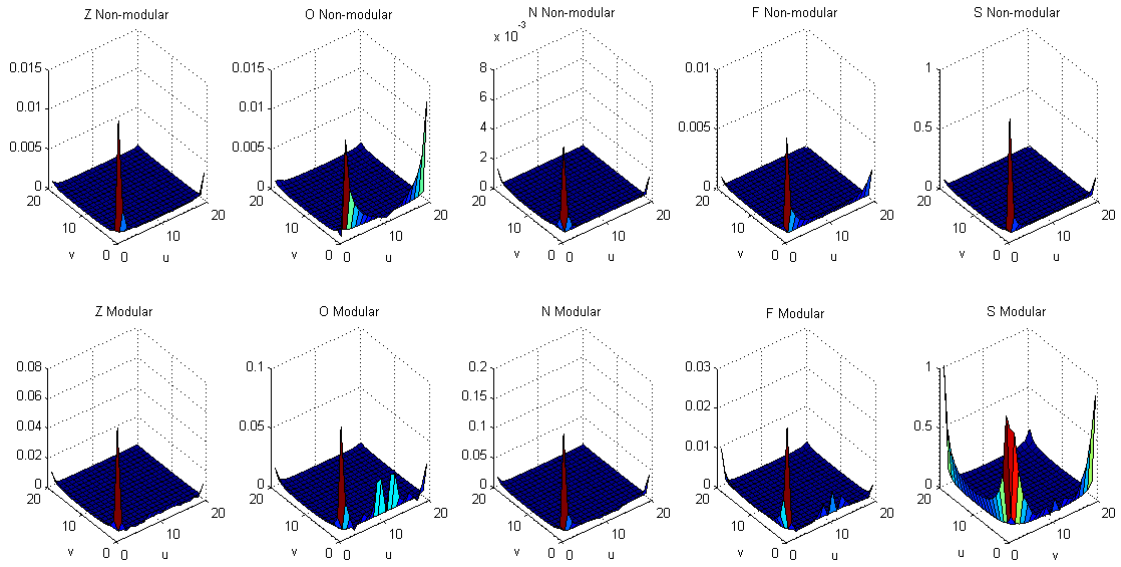


Fig. 4.4: Coefficients of 2D-DFT plot of all classes in both modular (Lower row) and non-modular (Upper row) cases.

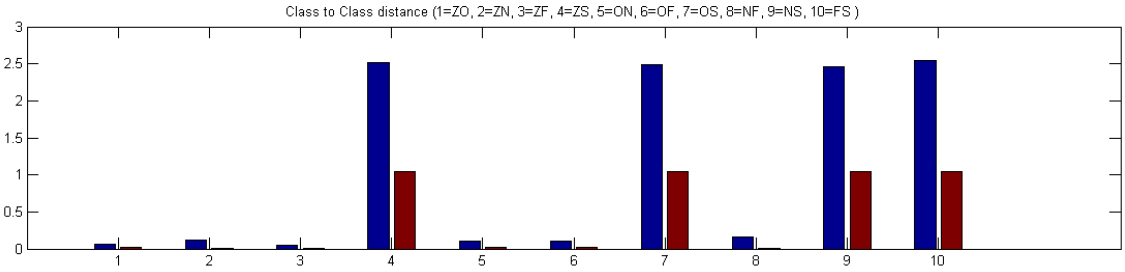


Fig. 4.5: Comparison of all Euclidean distances for non-modular and modular cases in terms of 2D-DFT coefficients

from Fig. 4.5 that some of the Euclidean distances, namely ZS, OS, NS and FS are similar, which may mislead towards an idea that Z, O, N and F are non-separable. Since, all Euclidean distances ZO, ZN, ZF, ZS, ON, OF, OS, NF, NS and FS are found non-zero, each of Z,O, N and F surely has mutual distance between each other thus indicating their separability. In Fig. 4.6, percentage improvement in Euclidean distance for all possible cases is presented. It is observed from Fig. 4.6 that all distances are increased significantly, in particular, approximately by at least 60% in case of modularization compared to the distances obtained in non-modular cases. For a better clarification, in a virtual plane, comparison of distances among different classes in non-modular (dark) and modular (shaded) cases are plotted in Fig. 4.7 as pentagons. Fig. 4.7 demonstrates that distances between each class is increased for employing modularization which is indicated by lines in a shaded pentagon. Similar

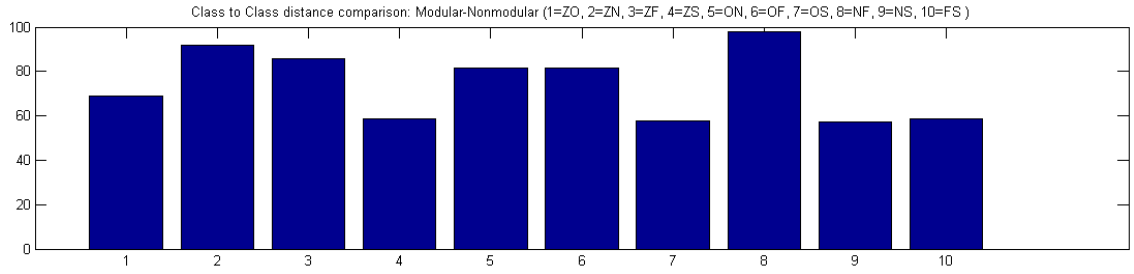


Fig. 4.6: Percentage improvement in Euclidean distances for all possible cases.

to Fig. 4.5, Fig. 4.7 also illustrates that although virtual distances ZS (4) and OS (7) seem similar, Z and O remain separable due to mutual distance ZO (1) be-tween them. The analysis as performed above clearly indicates that better distinguishable features are extracted from uniform smaller modules than that from the entire time-frequency plane as a whole. However, instead of taking the 2D-DFT coefficients of the entire time-frequency plane, if all the coefficients obtained from each module are considered to be used as a feature, it would definitely result in a feature vector with a very large dimension. In view of reducing the feature dimension, we propose to utilize one or two feature from each module thus reducing the length of the feature vector to $32 \times 32 \times 2$. A comparison of reduction in the length of feature vector in the case of modularization is shown in Fig.4.8. It is inferred from Fig.4.8 that the length of the feature vector be-comes 199% higher if all 2D-DFT coefficients from each module are used as features instead of using one or two feature per module. The complexity of feature extraction and ease of its implementation is another important criterion while developing a feature vector for multiclass epileptic seizure classification. Since computation of 2D-DFT coefficients is expensive, we propose to extract energy and entropy from each module

4.2.2 Modular Energy

Time frequency distribution is used to calculate power spectrum density $H(PSD)$ at each module, which represents energy distribution at time-frequency domain. The modular energy is the sum of PSD at each module, which is expressed as

$$E_{energy}(l,r) = \int_{t_{l(start)}}^{t_{l(end)}} \int_{\omega_{r(start)}}^{\omega_{r(end)}} H_{PSD}(t,\omega) d\omega dt \tag{4.1}$$

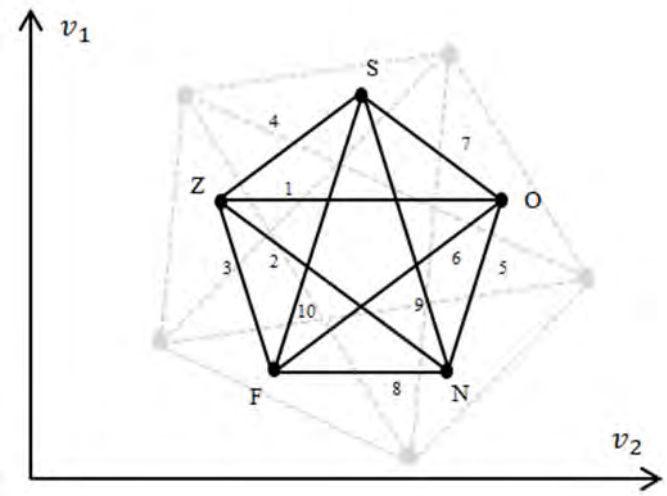


Fig. 4.7: Comparison of reduction in the length of feature vector in modularization case.

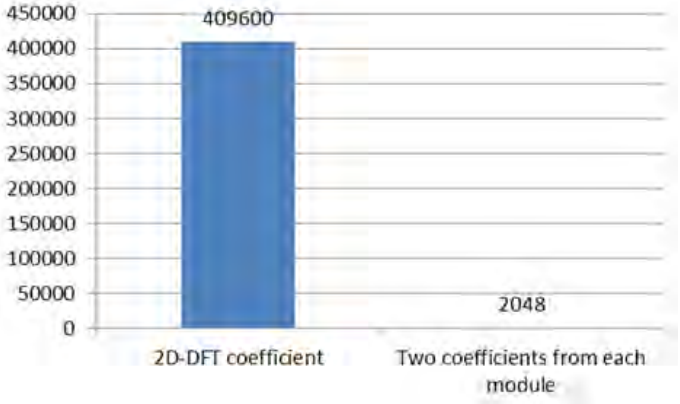


Fig. 4.8: Comparison of reduction in the length of feature vector in modularization case.

4.2.3 Modular Entropy

Entropy is first introduced in physics to describe the system disorganization, which is later adopted in signal processing for the same purpose. The definition of entropy in signal processing is based on the hypothesis that noise is totally disorganized thus carrying the highest entropy. Entropy of a signal can be defined as equation. (3.9). Since, modular entropy represents the information content in a module, entropy of each module can serve as a feature for classifying different EEG data. Therefore, the proposed multi-featured set can be formed as

$$F = \{E_{energy}(l,r); E_{entropy}(l,r)\}, \;\; l,r = 0,1,..31 \tag{4.2}$$

here l and r are the module number in the time and frequency domain, respectively.

4.3 Conclusion

Classification of seizure from the Electroencephalogram (EEG) data is the first step towards the automation of diagnosis and treatment of seizure. Since conventional time or frequency domain analysis is found inadequate to describe the characteristics of a non-stationary signal such as, EEG, we propose to transform the EEG data using twelve Cohen class kernels in order to facilitate the time-frequency analysis. The transformed data thus obtained is exploited to formulate a feature set involving modular energy and modular entropy that can better model the time-frequency behavior of the EEG data originated from different part and state of the brain. This uniform modular multifeature set is separately fed to each of KNN, ANN and Decision tree classifiers in order to classify multiclass epileptic seizure data.

Chapter 5

Simulation Results

5.1 Introduction

A number of simulations is carried out to evaluate the performance of the proposed methods. Performance is analyzed for both seizure detection and classification cases. Performance of proposed method is compared with a state-of-art method for evaluation purpose. A popular EEG database which consists five class EEG data is utilized for simulation purpose for both detection and classification methods.

5.2 EEG Dataset

For training and testing of the proposed method, the EEG dataset described in [70] is used, where all EEG signals are recorded with the same 128-channel amplifier system using an average common reference. The data are sampled at 173.61 samples per second using 12 bit resolution. The spectral bandwidth of the data is same as the acquisition system, which varies from 0.5 to 85 Hz. The dataset includes five subsets (de-noted as Z, O, N, F, and S) each containing 100 single channel EEG segments of 23.6 s duration. Subsets Z and O include the EEG data from five non-seizure volunteers, with eyes open and closed, respectively. F and N subsets consist of EEG data of five seizure patients during seizure free interval, where F is collected from the epileptogenic zone and on the other hand, N is collected from the opposite hemisphere of the brain in hippocampal formation. Subset S is the recording of the EEG data from patients during seizure activities. Therefore, subsets Z and O are two types of EEG data of seizure free healthy people and EEG subsets F, N and S are three types of data originated from different part and state of the brain of seizure patients. In all of our simulations, we employ 64-point Hamming window for both

time and frequency domain smoothing. On extracting the modularized feature set and using the set in classifying different EEG signals based on different classifiers, performance is evaluated through ten iterations by random selection of training and testing dataset of EEG signals at each step of iteration. At each iteration, from 500 EEG segments, we have randomly selected 50%(for classification case) and 40%(for detection case) EEG data as training dataset consisting of equal number of data for each class and then the rest of the EEG segments are employed for validation purpose at the testing phase. Thus, training and testing data are completely independent. Finally, average classification performance over all iterations is considered for detail analysis and comparison. Therefore, the simulation study is not dependent on particular training sets.

- Two class problem is designed with Z and S type EEG data. Here, Z consists of EEG recording from healthy person and S is collected from seizure patient during seizure activities. Distinguishing these classes means the detection of seizure from non-seizure dataset. Two hundred EEG samples are used in this classification problem.

- Second problem is designed as three class problem where Z,F and S class data are used. As usual, Z is normal data, S is seizure patients' data during seizure and F is seizure patients' data in non-seizure state. This problem handles with 300 EEG sample from one hundred EEG data from each class

- The final problem is a five class classification problem where all data are used as mentioned above. Here, Z and O is normal, F and N is seizure free interval data of seizure patient and S is seizure data.

5.3 Performance Parameters

For the performance evaluation of the proposed method, criteria considered in our simulation study are: 1) Sensitivity 2) Selectivity 3) Accuracy.

All the criteria as mentioned above can be derived from the confusion matrix, which is a form of representing the result from a classification exercise. The rows in the matrix stand for the actual classes to be tested and columns provide the class classified by a method. In particular, any $[row, column]$ entry in the confusion

matrix indicates the number of cases from the test database that belongs to the class corresponding to the row but classified as the class corresponding to the column.

In Fig. 5.1, a general confusion matrix for a two, three and five class problem is shown, where TP, FP, FN and TN are represented for Z class. In general, TP_i, true positive for any class i, measures the number of testing cases, which are correctly classified as class i. FP_i, false positive for any class i, denotes the number of testing cases i, which are incorrectly classified as class i. FN_i, false negative for any class i, indicates the number of testing cases, which are incorrectly classified as other than class i. TN_i, true negative for any class i, means the number of testing cases, which are correctly classified as other than class i.

		Classified by the classifier	
		Z	S
Actual Class	Z	TP_z	FN_z
	S	FP_z	TN_z

(a) Two class with respect to Z

		Detected by the classifier		
		Z	F	S
Actual Class	Z	TN_S		FP_S
	F			
	S	FN_S		TP_S

(b) Three class with respect to S

		Detected by the classifier				
		Z	O	N	F	S
Actual Class	Z	TN_N		FP_N	TN_N	
	O					
	N	FN_N		TP_N	FN_N	
	F	TN_N		FP_N	TN_N	
	S					

(c) Five class with respect to N

Fig. 5.1: Confusion matrix for two, three and five class classification cases

Sensitivity of a class relates the number of positive testing cases which are correctly classified to the number of testing cases of that particular class. Thus, sensitivity, for a particular class i, can be defined as

$$Sensitivity_i = \frac{TP_i}{TP_i + FN_i} \tag{5.1}$$

A classifier, which always indicates positive, regardless of the class of the testing case, provides 100% sensitivity for that class. Therefore the sensitivity alone cannot be used to determine the usefulness of the classifier in practice. Selectivity of a class relates the number of positive testing cases which are correctly classified to the number of classified cases in that particular class. Therefore, selectivity, for a particular class i, can be expressed as

$$Selectivity_i = \frac{TP_i}{TP_i + FP_i} \tag{5.2}$$

Accuracy of a class relates the number of testing cases which are correctly classified to the number of total testing cases. Therefore, accuracy, for a particular class i, can be written as

$$Accuracy_i = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{5.3}$$

5.4 Simulation Results of Seizure Detection

Performance of the epilectical seizure detection method based on non-modular time-frequency modules of EEG signals, described in Chapter.3,are analyzed and compared with a state-of-art method.

5.4.1 Goodness of Feature

In order to explain the rationale behind proposing the combination of modular cumulative energy and entropy as a feature set, a bar graph plotting average classification accuracy is shown in Fig. 5.2 for non-modular and modular feature sets considering all EEG segments and a simple Euclidean distance based KNN classifier. In non-modular case, only total energy or only total entropy is considered as a feature set. It is seen that although total energy provides an average classification accuracy of 98.8%, cumulative energy produces an average classification accuracy as high as 99.4%. As modular entropy can extract local information more precisely than total

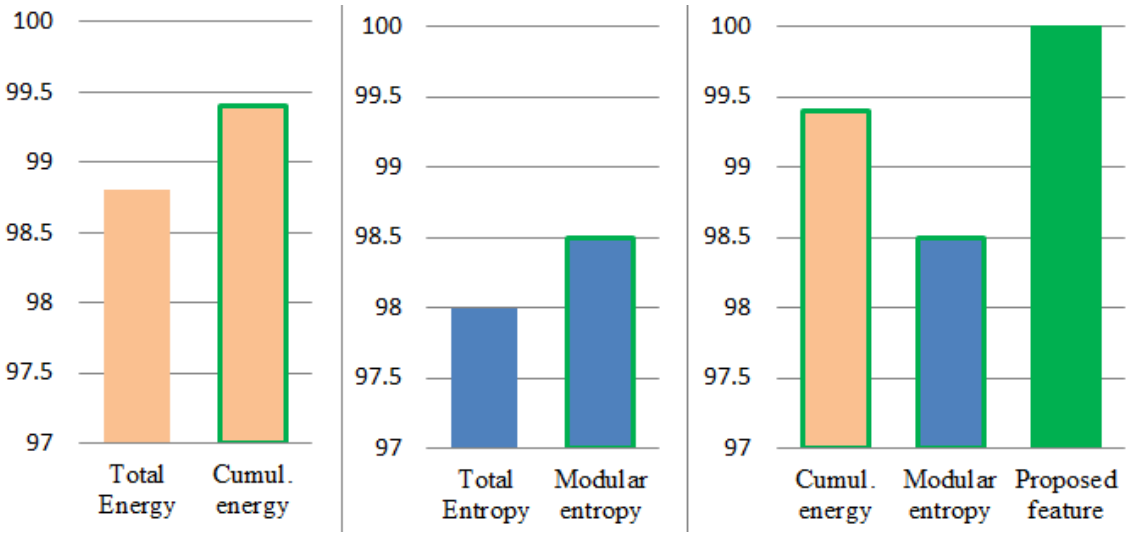


Fig. 5.2: Selection of proposed feature

entropy, modular entropy yields an average accuracy of 98.5%, which is higher than that obtained from the total entropy. It is to be noted that performing modularization prior to extracting energy or entropy can enhance the accuracy significantly compared to that obtained using non-modular energy and entropy. Therefore, the combination of modular cumulative energy and entropy expectedly offers a very high average classification accuracy, such as 100%.

The effectiveness of the proposed modular multi-featured set in classifying different types of EEG signals is justified by inter-class separability of the feature. For this purpose, a clustering analysis is performed for a feature set consisting of non-modular energy and that of entropy. The similar analysis is carried out for a feature set containing the proposed modular cumulative energy and entropy. In Fig. 5.3, a clustering plot of non-modular and modular feature sets is shown. In case of either non-modular or modular feature set, for a particular class, we form the cluster using the mean of the feature vectors obtained over all the EEG segments and two principle components are deduced to plot the cluster position in the virtual plane. From Fig. 5.3, it is clear that modular cumulative energy is more separable than modular entropy and our proposed feature set has the highest inter-class separability. Therefore, in this paper, we are motivated to employ a combination of modular cumulative energy and entropy for classifying EEG signals originating from different part and state of the brain. In both the Figs. 5.2 and 5.3, time-frequency distribution smoothed by WV kernel is used prior to feature extraction.

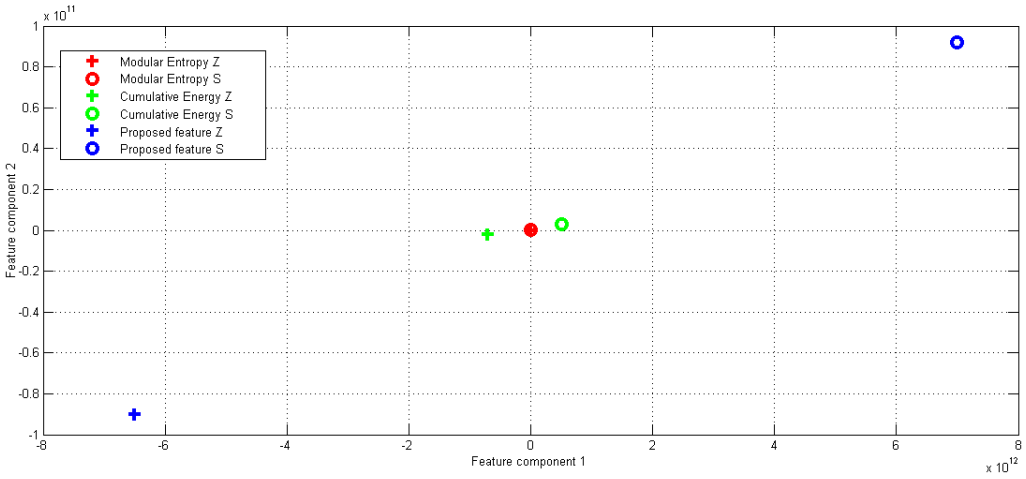


Fig. 5.3: Comparison of inter-class separability

5.4.2 Performance Analysis

The classification performance of the proposed method using six different classifiers is investigated with twelve different kernels in terms of average selectivity, sensitivity and accuracy and results are depicted in Figs. 5.4 and5.5 and Table.5.1.

Fig. 5.4 shows the comparison of the average selectivity obtained from the proposed method using different kernels, namely MH, WV, RIH, PWV, PMH, BJ, BUT, CW, GRECT, RI, SPWV and ZAM for a particular classifier, namely decision tree. For each kernel, the results of average selectivity are also obtained and plotted for KNN classifier with different distance types, such as Euclidean, City-block, Cosine and Correlation and ANN classifier. It is evident from this figure that the average selectivity of the proposed method while using KNN-cosine and KNN-correlation classifiers are lower than 85% for all kernels. On the other hand, for KNN-City block and KNN-Euclidean classifiers, the proposed method produces high average selectivity, such as 100% with all kernel types. It is to be mentioned that for all the kernels, the proposed method using ANN classifier also results in 100% average selectivity.

The average sensitivity and accuracy obtained from the proposed method using twelve kernels and six classifiers are plotted in Table. 5.1 and Fig. 5.5, respectively. From Table. 5.1, it is vivid that the proposed method yields 100% average sensitivity for all kernels except ZAM while using KNN-Euclidean, KNN-Cityblock and ANN classifiers. It is attested from Fig. 5.5 that the proposed feature is capable of

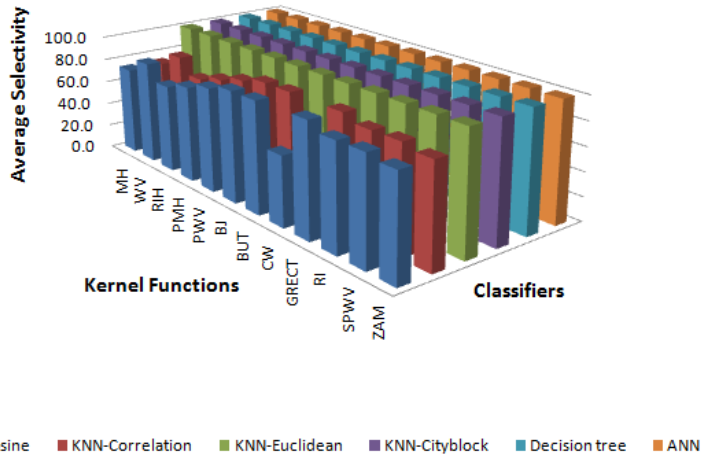


Fig. 5.4: Comparison of the average selectivity obtained from the proposed method using 12 kernels and 6 classifiers

Table 5.1: Comparison of the average sensitivity obtained from the proposed method using 12 kernels and 6 classifiers

Kernel function	KNN-correlation	KNN-cosine	Decision tree	KNN-Euclidean	KNN-cityblock	ANN
MH	73.7	73.5	99.2	100	100	100
WV	86.1	86	99.2	100	100	100
RIH	73.6	73.2	99.2	100	100	100
PMH	80.2	79.8	99.2	100	100	100
PWV	86.7	86.6	99.2	100	100	100
BJ	92.3	92.2	99.2	100	100	100
BUT	92.5	92.3	99.2	100	100	100
CW	57.8	57.9	99.2	100	100	100
GRECT	92.4	92.5	99.2	100	100	100
RI	87.2	87	99.2	100	100	100
SPWV	87.3	87.3	99.2	100	100	100
ZAM	83.5	84.1	99.7	99.7	99.7	99.7

producing 100% average accuracy while using all classifiers except the KNN-cosine and KNN-correlation.

5.4.3 Performance comparisons

Such a classification problem using the same dataset [70] has been limitedly reported [35]. In [35], energy distribution in time-frequency plane corresponding to three time windows and five frequency subbands, such as $0-2.5$ Hz, $2.5-5.5$ Hz, $5.5-10.5$ Hz, $10.5-21.5$ Hz, and $21.5-43.5$ Hz are considered as feature and ANN classifier is used to classify the different classes of EEG data as mentioned above.

Table. 5.2 presents a detail performance comparison of the proposed method with that of a state-of-the-art method [35] in terms of average sensitivity (Sen) and average selectivity (Sel) using the same EEG dataset and the same ANN classifier

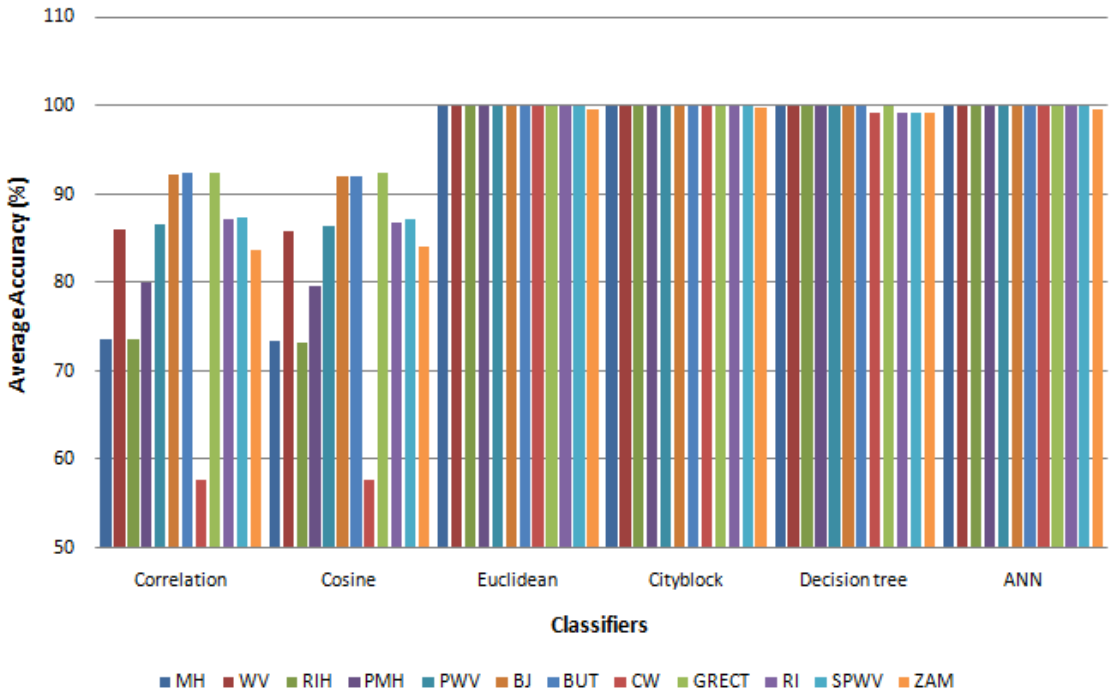


Fig. 5.5: Comparison of the average accuracy [%] obtained from the proposed method using 12 kernels and 6 classifiers

for all the kernels. It is observed from Table. 5.2 that by using a kernel function with no smoothing, while the other method achieves a low average selectivity and sensitivity of 69.6% and 69.7%, respectively, the proposed method yields much higher average selectivity and sensitivity, such as 100%. While using the frequency or time-frequency smoothing kernel functions, the other method continues to give less average selectivity with a maximum value of 100% only with the RI and SPWV kernel functions, whereas the proposed method remain better in performance with 100% average selectivity and sensitivity for all the kernels except ZAM.

For all the kernels, Fig. 5.6 presents another detail performance comparison of the proposed method with that of a state-of-the-art method [35] in terms of average accuracy using the same EEG dataset the same ANN classifier as used in Table. 5.2. It is clear from Fig. 5.6 that the highest average accuracy is obtained using the RI and SPWV kernel functions for the methods described in [35]. It is to be noted that the proposed method is capable of producing better average accuracy compared to that of the method in [35] for all the kernel functions.

In general, effectiveness of a classification method depends an the choice of kernel function as well as kernel processing time. Kernel processing time is needed to be

Table 5.2: Performance comparison of the proposed method with that of a state-of-the-art method in terms of average sensitivity (Sen) and average selectivity (Sel) for all the kernels

Serial No	Kernel Short Name	Indicator	Proposed Method			Comparison Method			Kernel Processing Time
			Z	S	Average	Z	S	Average	
1	MH	Sel	100	100	100	71.6	67.6	69.6	Low
		Sen	100	100	100	68.9	70.4	69.7	
2	WV	Sel	100	100	100	98.4	94.2	96.3	Low
		Sen	100	100	100	94.4	98.3	96.4	
3	RIH	Sel	100	100	100	86.8	80.6	83.7	Low
		Sen	100	100	100	68.8	82.1	75.5	
4	PMH	Sel	100	100	100	94.2	96	95.1	Very Low
		Sen	100	100	100	95.9	94.3	95.1	
5	PWV	Sel	100	100	100	98	100	99.0	Very Low
		Sen	100	100	100	100	98	99.0	
6	BJ	Sel	100	100	100	96.4	99.8	98.1	High
		Sen	100	100	100	99.8	96.5	98.2	
7	BUT	Sel	100	100	100	98	100	99.0	Very high
		Sen	100	100	100	100	98	99.0	
8	CW	Sel	100	100	100	96.4	100	98.2	High
		Sen	100	100	100	100	96.5	98.3	
9	GRECT	Sel	100	100	100	96.2	100	98.1	Very high
		Sen	100	100	100	100	96.3	98.2	
10	RI	Sel	100	100	100	100	100	100.0	Very high
		Sen	100	100	100	100	100	100.0	
11	SPWV	Sel	100	100	100	100	100	100.0	Very high
		Sen	100	100	100	100	100	100.0	
12	ZAM	Sel	100	99.2	99.6	99.8	97.4	98.6	High
		Sen	99.2	100	99.6	97.5	99.8	98.7	

considered while calculating the time for classification. In Fig. 5.7, time required for processing an EEG segment for each kernel function are shown. It is found from this figure and the right most column of Table.5.2 that kernel processing time for frequency smoothing by PMH and PWV kernels are very low, namely 0.75 and 0.96 respectively. In Table. 5.2, Figs. 5.6 and 5.7, it is shown that 100% average sensitivity, selectivity and accuracy are obtained by the comparison method in [35] while using time and frequency domain smoothing, such as RI and SPWV kernels. It is clear from Fig. 5.7 and Table.5.2 that time required for processing RI and SPWV kernel functions are almost 8 to 10 percent higher than that for PMH and PWV kernels. It is demonstrated through Table. 5.2, Figs. 5.6 and 5.7 that the proposed method produces 100% average selectivity, sensitivity and accuracy for kernels with very lower processing time.

PCA can not be employed to a single feature vector, rather it is applied to the whole training matrix by appending the test feature vector during every test. Since,

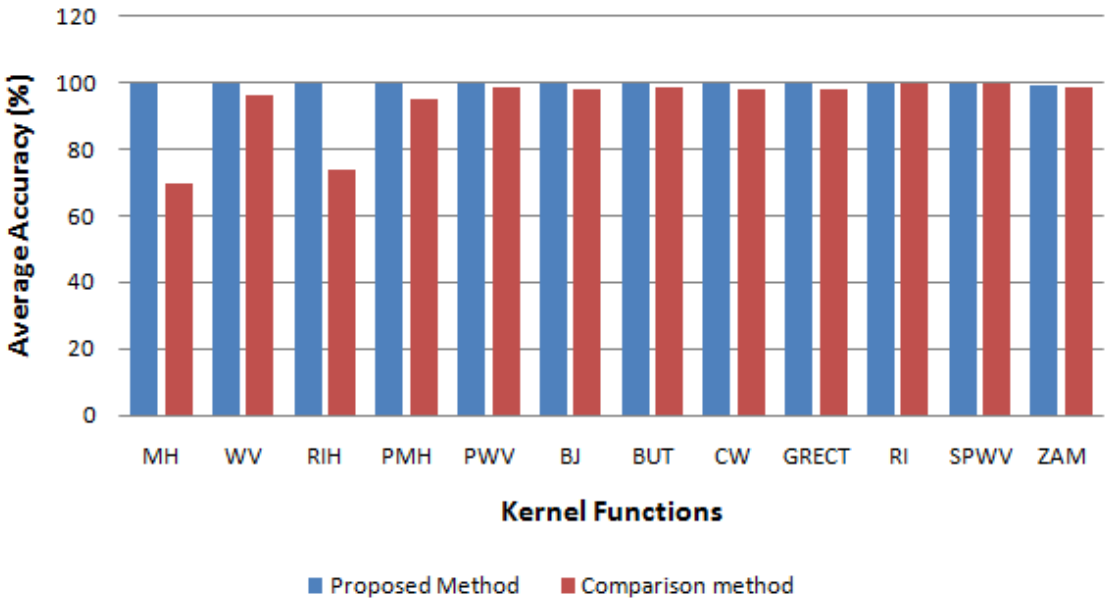


Fig. 5.6: Performance comparison of the proposed method with that of a state-of-the-art method in terms of average accuracy [%] for all the kernels

PCA involves the handling of complex matrix operation, which increases processing time for higher dimension matrix thus increases the total time for classification. PCA are used in the comparison method in [35] to reduce feature dimension, high average selectivity, sensitivity and accuracy are achieved at the expense of higher kernel processing time. On the other hand, in proposed method, using a feature set without employing PCA a higher average selectivity, sensitivity and accuracy are obtained with almost all kernels, especially even with a kernel of with lower processing time.

In Fig.5.8, the performance of the proposed method with that of a state-of-the-art method in [35] is compared in terms of average accuracy(%) using different classifier. Since the method in [35] produces best performance with time-frequency distribution smoothed by RI kernel, in Fig.5.8 RI kernel is used for a fair comparison.

It is observed from Fig.5.8 that the method in [35] produces lower average accuracy(%) compared to the proposed method while using KNN-Euclidean and Decision tree classifier and both the method yields 100% average accuracy(%) using ANN classifier. It is known that computational complexity of an ANN classifier is significantly high in comparison to the simple Euclidean distance based KNN classifier. Convergence time of the ANN classifier can be varied according to the classification problem, whereas KNN only calculate simple Euclidean distances which are

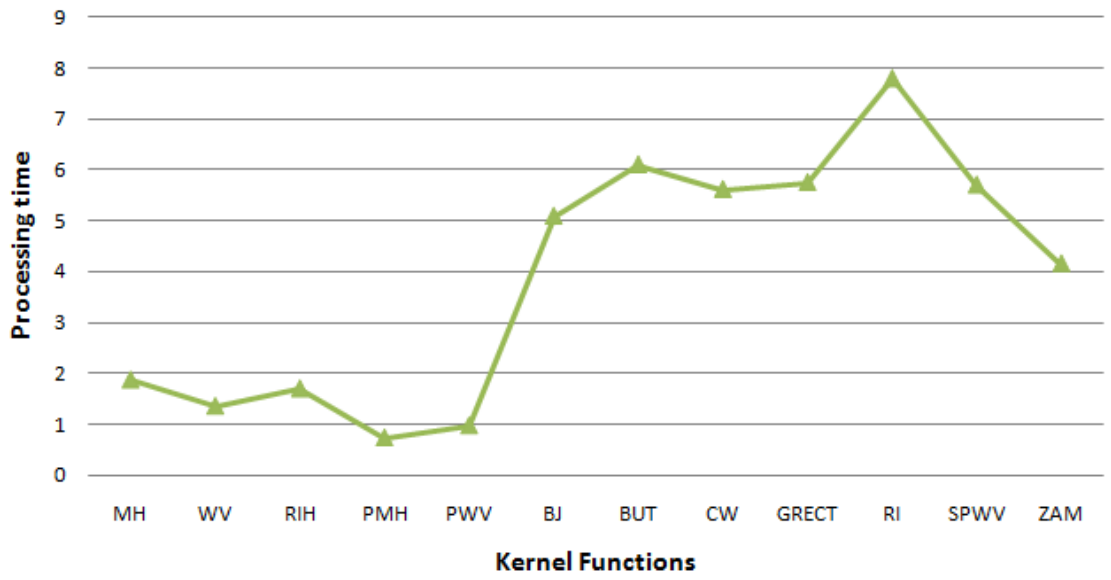


Fig. 5.7: Kernel processing time of a single sample for all kernels

not affected noticeable with the classification problem. It is clear from Fig.5.8 that in contrast to the method in [35], the proposed method is able to produce 100% average accuracy even while using a simple classifier.

Therefore, by using kernels with very lower processing time, without employing PCA for feature reduction and by exploiting simple classifier, the proposed method achieves lower implementation time thus making it more suitable for real time seizure detection.

In comparison to the method in [35], it is to be noted that the proposed method gains the highest average accuracy(%) using 10% less training data, which also signifies the effectiveness of the proposed feature set in detecting epileptic seizure.

5.5 Simulation Results of Seizure Classification

Performance of the uniform modular feature based multiclass epileptic seizure classification method, described in Chapter. 4,are analyzed and compared with a state-of-art method.

5.5.1 Goodness of Feature

In order to explain the rationale behind proposing the combination of modular energy and modular entropy as a feature set, bar graph plotting average accuracy with variance is shown in Fig. 5.9 for non-modular and modular feature sets considering

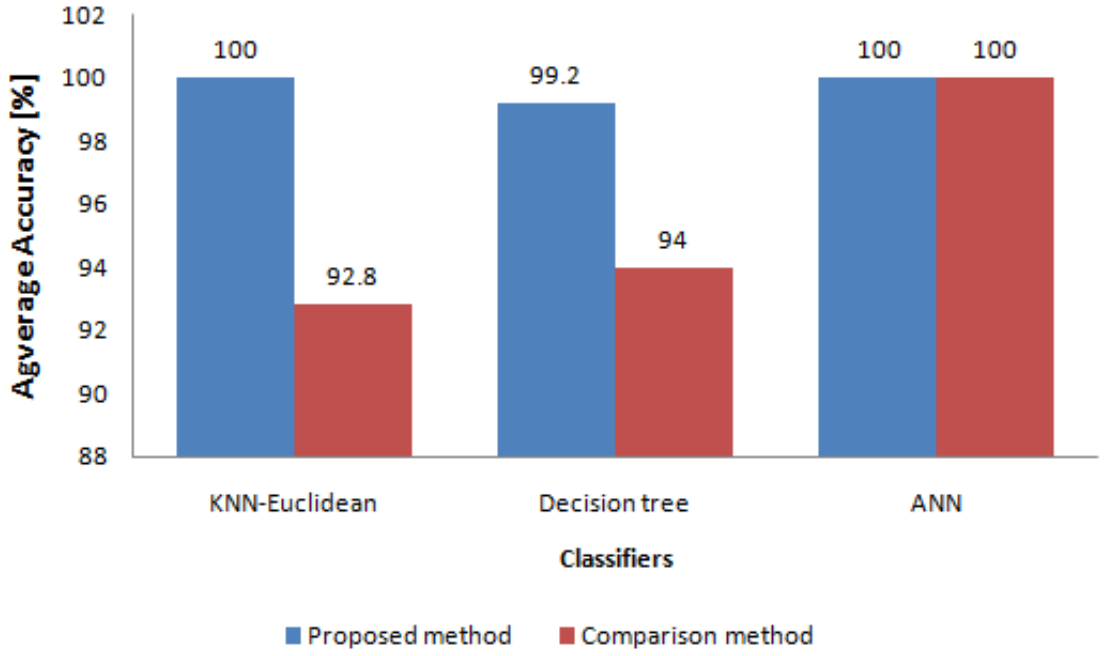
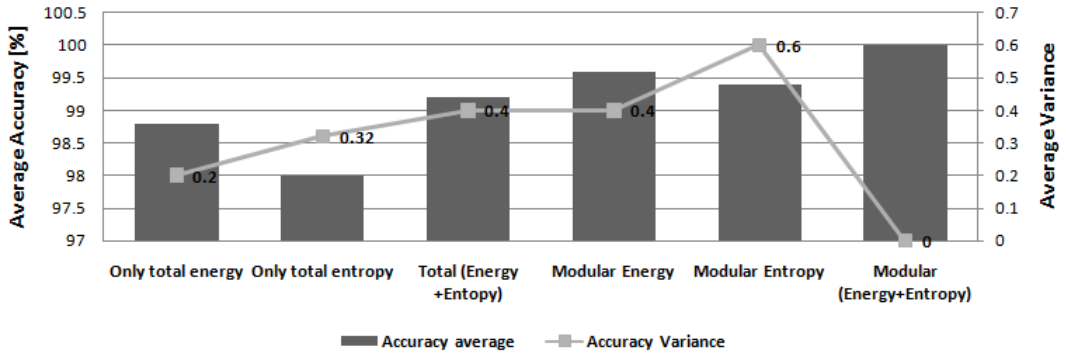


Fig. 5.8: Performance comparison of the proposed method with that of a state-of-the-art method in terms of average accuracy (%) using different classifier

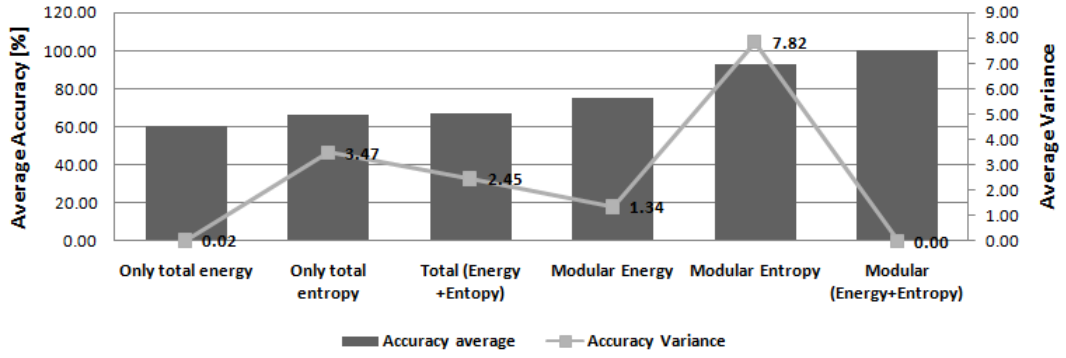
all EEG segments. In non-modular case, only total energy or only total entropy or combination of total energy and total entropy is considered as a feature set.

From Figs.5.9(a) and 5.9(b), we observed that when only total energy and that of entropy is utilized as feature set, an average classification accuracy is less than 99% and 65% in case of two class and three class classification problem, respectively. For five class problem, it is also seen from Fig.5.9(c) that total energy or total entropy provides an average classification accuracy less than 60%. Although, only total energy and entropy shows less accuracy in all classification problem, the combination of total energy and total entropy as a feature increases the average accuracy. It is to be noted that performing modularization prior to extracting energy or entropy can enhance the accuracy significantly compared to that obtained using non-modular energy and entropy. Therefore, the combination of modular energy and entropy expectedly produces the highest average accuracy, such as approximately 100% for two and three class and 98% for five class classification problems. The variance of average accuracy is zero for two and three class problems and as low as 1.22 in five class classification results.

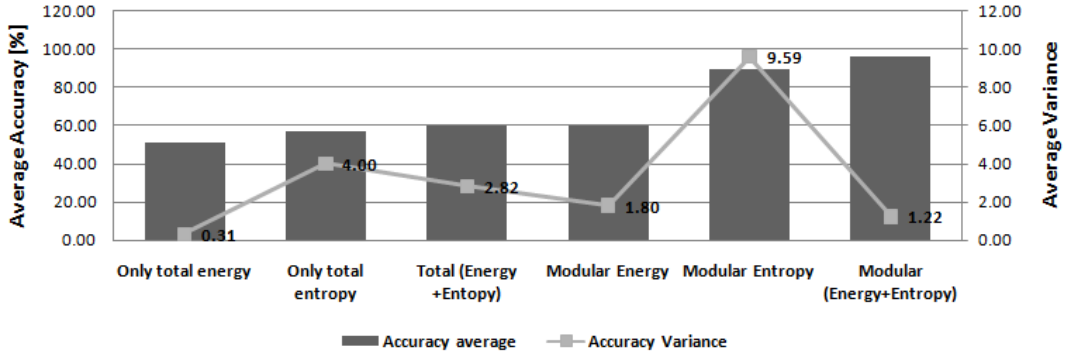
The effectiveness of the proposed modular multi-featured set in classifying different types of EEG signals is justified by inter-class separability of the feature. For



(a) Two class problem



(b) Three class problem



(c) Five class problem

Fig. 5.9: Average accuracy with variance for different feature sets for two, three and five class problems

this purpose, a clustering analysis is performed for a feature set consisting of non-modular energy and that of entropy. The similar analysis is carried out for a feature set containing the proposed modular energy and that of entropy. In both the cases, time-frequency distribution smoothed by RI kernel is used prior to feature extraction. In Fig. 5.10, clustering plot of non-modular and modular feature sets is shown for a five class problem. In case of either non-modular or modular feature set, for a particular class, we form the cluster using the mean of the feature vectors obtained

over all the EEG segments. Since, Z and O data are from two different state of normal patients, F and N data are from seizure free state of the seizure patients and S data represents the seizure state of the seizure patients, Z and O should be less separable from each other, N and F should be closer to each other and S should be far enough from all the classes. Therefore, in this paper, we are motivated to

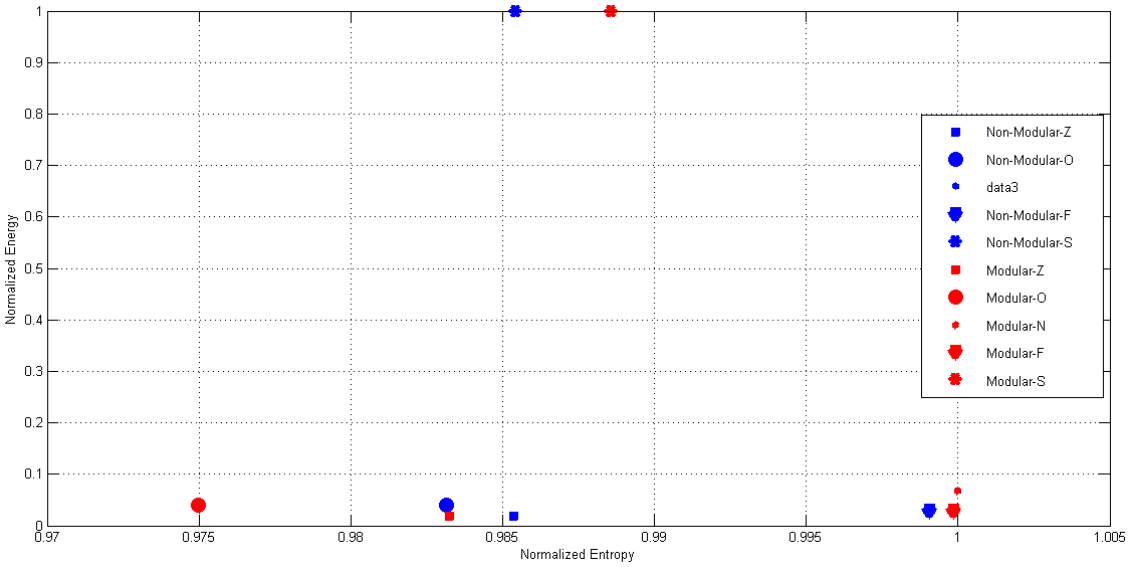


Fig. 5.10: Comparison of non-modular and modular feature sets

employ a combination of modular energy and modular entropy for classifying EEG signals originating from different part and state of the brain.

5.5.2 Selection of Number of Modules

In modularization, selection of number of module in the time-frequency plane is an important task for feature extraction. Increasing the module number overloads the classifier with a larger feature set, which increases the classification time as well as affect the overall accuracy. Therefore, the module number is needed to be optimized since the number of the module has a significant impact on classification performance. We investigate the number of the module for feature extraction in terms of average selectivity and sensitivity. The average selectivity and that of sensitivity computed over all classes are plotted in Fig. 5.11 using different number of modules, namely $4 \times 4, 8 \times 8, 16 \times 16, 24 \times 24, 32 \times 32$ and 36×36. In Fig. 5.11(a), it is shown that 100% accuracy is obtained at 16×16 modularization for two class problem. Thus, in case of two class problem, 16×16 modularization is selected

as an optimum module number. On the other hand, it is found from Figs. 5.11(b) and 5.11(c) that the average selectivity and sensitivity of the proposed method generally increases with increasing module number and reaches to the highest with a module number of 32×32 for three class and five class problems. Since, in case of five class problem, the average selectivity and sensitivity declines for a further modularization with a module number of 36×36, in the proposed method, we choose 32×32 as an optimum module number for solving the intended multiclass EEG signal classification problems.

5.5.3 Performance Analysis

The classification performance of the proposed method using six different classifiers is investigated with twelve different kernels in terms of average selectivity, sensitivity and accuracy and results are depicted in Figs. 5.12 and 5.13 and in Table.5.3,.

Fig. 5.12 shows the comparison of the average selectivity obtained from the proposed method using different kernels, namely MH, WV, RIH, PWV, PMH, BJ, BUT, CW, GRECT, RI, SPWV and ZAM for a particular classifier, namely decision tree. For each kernel, the results of average selectivity are also obtained and plotted in Fig. 5.12 for KNN classifier with different distance types, such as Euclidean, City-block, Cosine and Correlation, and ANN classifier. It is evident from these figures that the average selectivity of the proposed method while using KNN-cosine and KNN-correlation classifier is lower for all kernels. From Fig. 5.12(a) plotted for a two class problem, it is shown that 100% selectivity is obtained while using other four classifiers, namely decision tree, KNN-Euclidean, KNN-City-block and ANN classifier. In Fig. 5.12(b) depicted for a three class problem, 100% average accuracy is found when RI kernel with ANN classifier is used. Similarly, average accuracy is approximately 100% for the same kernel and classifier for a five class problem represented in Fig. 5.12(c). However, in order to make a tradeoff between the choice of a kernel and that of a classifier for implementing the proposed method, we investigate the performance of the proposed method for all kernels in terms of average selectivity as well as average sensitivity and accuracy and compare the corresponding performance with respect to all classifiers.

Comparison of the average sensitivity is shown in Fig. 5.13, where the proposed

method using twelve kernels and six classifiers is considered. Fig.5.13(a) shows the comparison of average sensitivity in case of two class classification problem. Here, it is observed that 100% sensitivity is obtained by using all classifiers except KNN-cosine and KNN-correlation. From Figs.5.13(b) and 5.13(c), it is vivid that for three and five class problems, four classifiers, namely KNN-cosine, KNN-cityblock, Decision tree and ANN, perform better and average sensitivity increases when time and frequency domain smoothing kernels are used. However, for all the kernels, the proposed method using ANN classifier is shown to yield the highest average sensitivity for the multi-class problems. In particular, the average sensitivity is found to reach a very high value of 100% and 97.3% in case of three class and five class problems, respectively, while the proposed method employs RI kernel and ANN classifier. Thus attesting the significance of using a kernel employing both time and frequency smoothing for reducing interference greatly. Therefore, we found it advantageous to choose RI for time-frequency distribution prior to feature extraction and then employ ANN classifier to classify three and five class EEG signals. On the other hand, faster kernels, namely PMH, PWV etc, and Euclidean distance based KNN classifier are found suitable for a two class problem.

Table. 5.3 shows the comparison of the average accuracy (%) obtained from the proposed method using 12 kernels and 6 classifiers. From Table. 5.3(a), it is clear that KNN-euclidian and ANN have 100% accuracy for all kernels in case of two class classification problem. From Table. 5.3(b)summarized for a three class problem and Table. 5.3(c)represented for a five class problem, it is demonstrated that the average accuracy of the proposed method while using ANN classifier is the highest for all kernels compared to that obtained using all other classifiers. So, in basis of average selectivity, sensitivity and accuracy, we can consider two types of optimum choice using the proposed method:

- *For two class:* KNN-euclidian with low processing time kernels, such as PMH, PWV etc

- *For three and five class:* ANN with RI kernel

5.5.4 Performance comparisons

Such a classification problem using the same dataset [70] has been limitedly reported [35]. In [35], energy distribution in time-frequency plane corresponding to the dominant frequencies, such as delta, theta, alpha, beta and gamma components of EEG is considered as feature and ANN classifier is used to classify the different classes of EEG data as mentioned above.

Tables. 5.4 through 5.6 present a detail performance comparison of the proposed method with that of a state-of-the-art method in [35] in terms of average selectivity and sensitivity using the same ANN classifier and the same EEG dataset utilizing all the kernels for two, three and five class problems, respectively. For two class problem, it is observed from Table. 5.4 that 100% average selectivity and sensitivity is obtained by using the proposed method for all the kernels. On the other hand, 100% average selectivity and sensitivity are obtained by using the method in [35] only for RI and SPWV kernels which are slower in terms of processing time.

Table. 5.5 shows the comparison of the proposed method with the method in [35] in terms of average selectivity and sensitivity for a three class problem. From this figure, it is vivid that 100% average accuracy is obtained for both the methods when RI kernel, involving time and frequency domain smoothing, is used. But the method described in [35] has lower average selectivity and sensitivity in comparison to that of the proposed method and the lowest value is approximately 75%, when no time-frequency domain smoothing is used. On the other hand, in the proposed method, average accuracy is higher than 98% for any types of kernels.

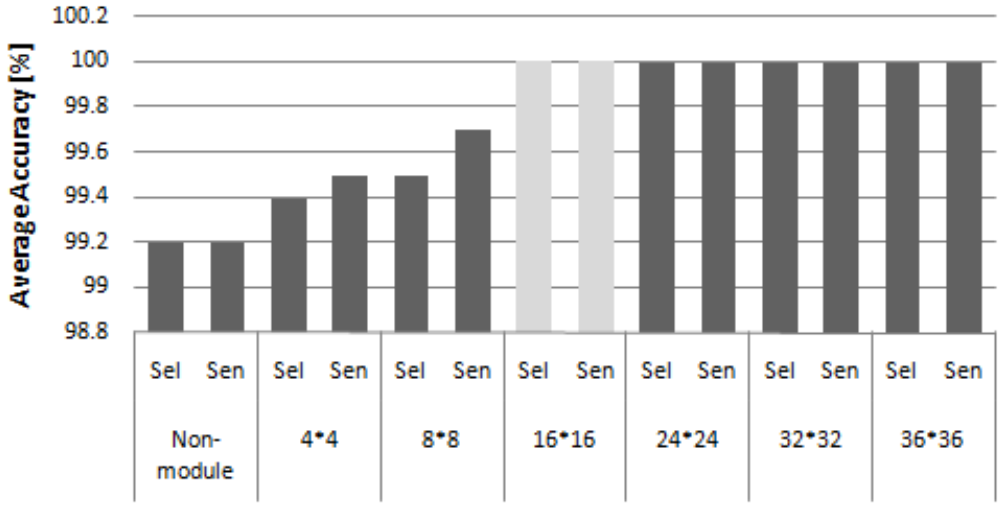
For a five class problem, it is observed from Table. 5.6 that by using a kernel function with no smoothing, while the other method achieves a low average selectivity of 54.8%, the proposed method yields much higher average selectivity of 90.9%. While using the frequency or time-frequency smoothing kernel functions, the other method continues to give less average selectivity with a maximum value of 89.2% with the RI kernel function, whereas the proposed method remain better in performance with all the kernels and achieves a maximum and very high average selectivity of 97.3% using the RI kernel. It can be found from Table. 5.6 that the proposed method maintains much higher average sensitivity while using all the kernels compare to that of the other method. Particularly, the average sensitivity of

the proposed method also shows the highest value for the RI kernel function.

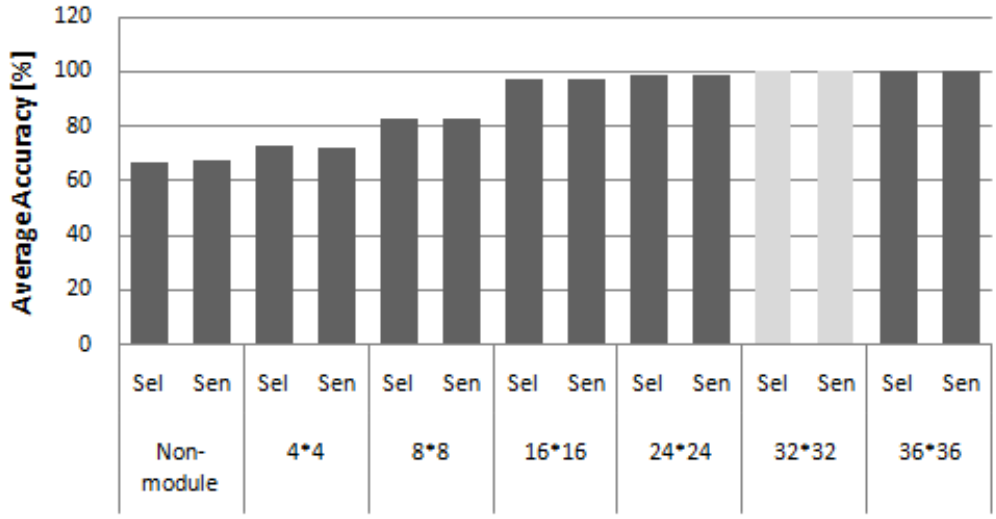
For all the kernels, Figs. 5.14(a) through (c) present another detail performance comparison of the proposed method with that of a state-of-the-art method in [35] in terms of average accuracy using the same ANN classifier and the same EEG dataset as used in average selectivity and sensitivity comparison for two, three and five class problem, respectively. It is clear from Fig. 5.14 that the highest average accuracy is obtained with the RI kernel function for all classes while using both the methods. It is to be noted that the proposed method is capable of producing better average accuracy compared to that of the method in [35] for all the kernel functions in case of two, three and five class problems.

5.6 Conclusion

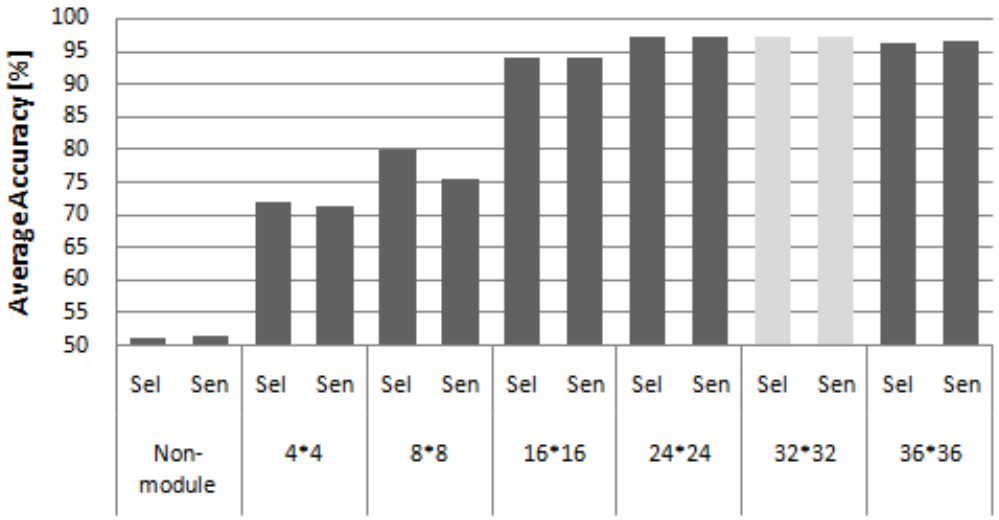
Cumulative energy and modular entropy based feature set is found suitable for seizure detection from EEG dataset. This feature set is faster and 100% accurate in seizure detection without any dependency on kernel functions. But, this feature can not extract enough information for seizure classification. Therefore, a uniformly modularize feature set is introduced for seizure classification problems. Using RI kernel function with ANN classifier, this uniform modularized feature set achieve 100% and 97.3% average accuracy for three class and five class problem, respectively.



(a) Two class problem

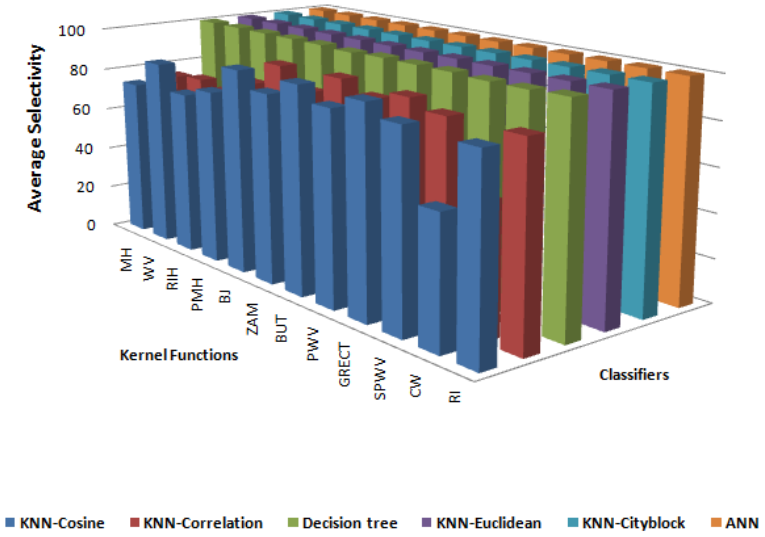


(b) Three class problem

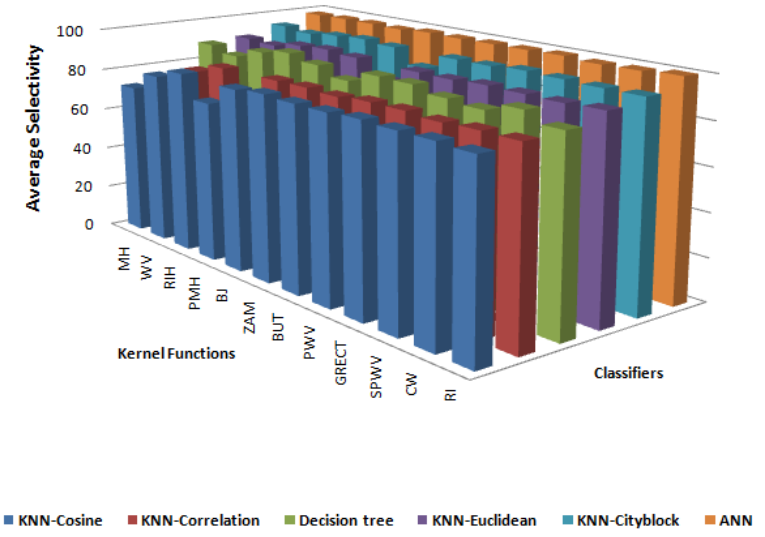


(c) Five class problem

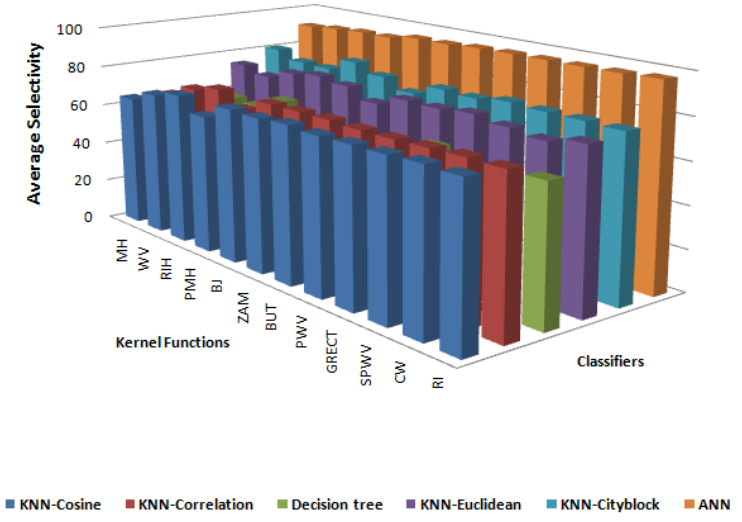
Fig. 5.11: Effect of varying the number of module on classification performance for two, three and five class problems



(a) Two class problem

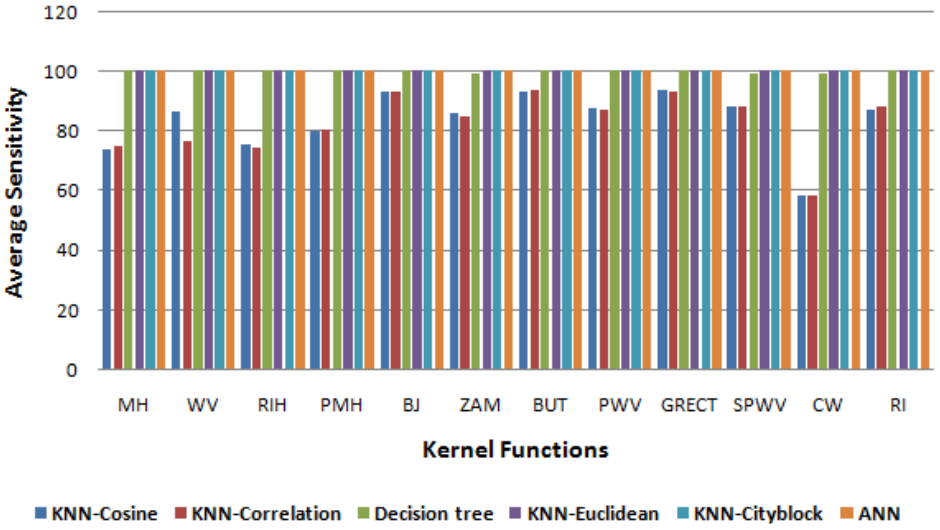


(b) Three class problem

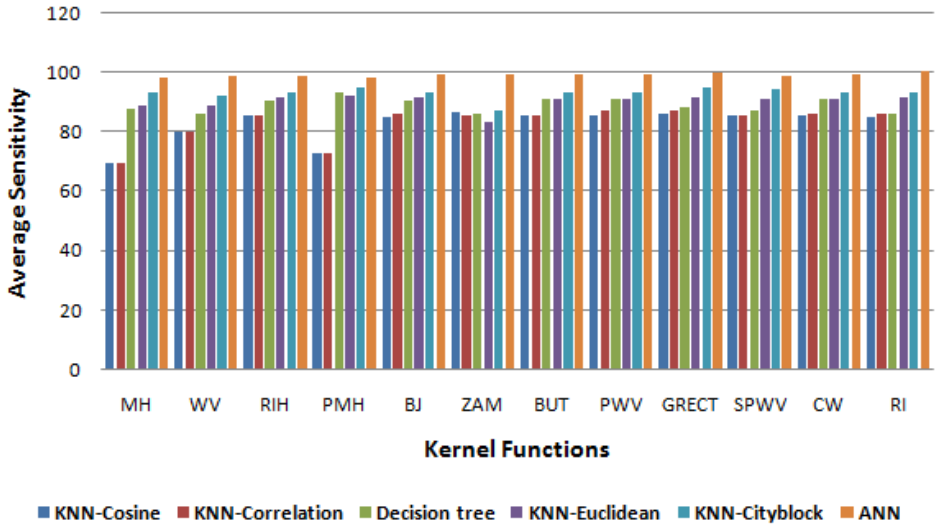


(c) Five class problem

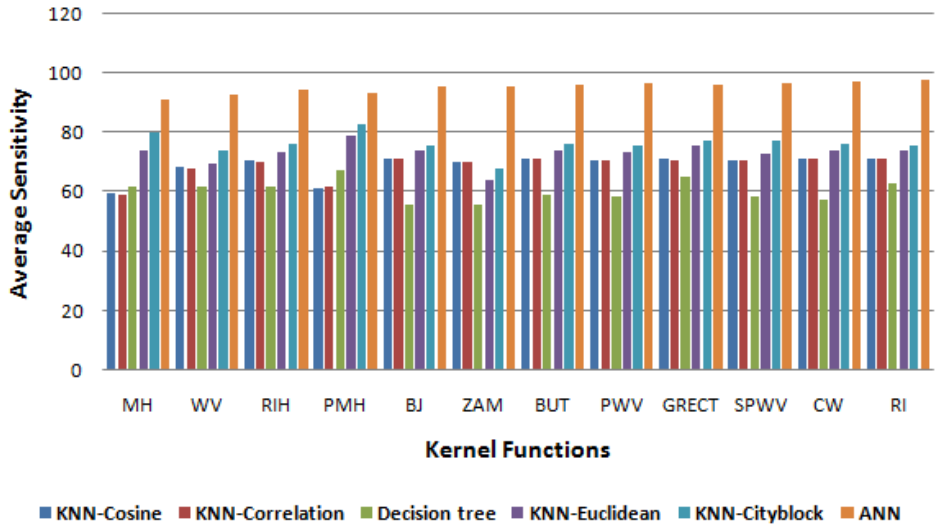
Fig. 5.12: Comparison of the average selectivity obtained from the proposed method using 12 kernels and 6 classifiers for two, three and five class problems



(a) Two class problem



(b) Three class problem



(c) Five class problem

Fig. 5.13: Comparison of the average sensitivity obtained from the proposed method using 12 kernels and 6 classifiers for two, three and five class problems

Table 5.3: Comparison of the average accuracy obtained from the proposed method using 12 kernels and 6 classifiers for two, three and five class problems

Kernel	KNN-Cosine	KNN-Correlation	Decision tree	KNN-Euclidean	KNN-Cityblock	ANN
MH	73.7	74.8	100	100	100	100.0
WV	86.4	76.7	100	100	100	100.0
RIH	75.2	74.3	100	100	100	100.0
PMH	79.8	80.2	100	100	100	100.0
BJ	93.2	92.9	100	100	100	100.0
ZAM	86.1	84.6	99.2	100	100	100.0
BUT	93.2	93.5	100	100	100	100.0
PWV	87.3	87.2	100	100	100	100.0
GRECT	93.6	92.8	100	100	100	100.0
SPWV	88.2	88.3	99.2	100	100	100.0
CW	58.1	58.3	99.2	100	100	100.0
RI	87.2	88.2	100	100	100	100.0

a) Two Class Problem

Kernel	KNN-Cosine	KNN-Correlation	Decision tree	KNN-Euclidean	KNN-Cityblock	ANN
MH	69.3	69.3	87.3	88.7	93.3	98.0
WV	80	80	86	88.7	92	98.4
RIH	85.3	85.3	90	91.3	93.3	98.4
PMH	72.7	72.7	93.3	92	94.7	98.1
BJ	84.7	86	90	91.3	93.3	99.1
ZAM	86.3	85.2	86	83.3	86.7	98.9
BUT	85.3	85.3	90.7	90.7	93.3	99.1
PWV	85.3	86.7	90.7	90.7	93.3	99.2
GRECT	86	86.7	88	91.3	94.7	99.6
SPWV	85.3	85.3	86.7	90.7	94	98.7
CW	85.3	86	90.7	90.7	93.3	99.0
RI	84.7	86	86	91.3	93.3	100.0

b) Three Class Problem

Kernel	KNN-Cosine	KNN-Correlation	Decision tree	KNN-Euclidean	KNN-Cityblock	ANN
MH	59.2	58.8	61.6	73.6	80	90.9
WV	68	67.6	61.6	69.2	74	92.3
RIH	70.4	70	61.6	73.2	76	93.9
PMH	61.2	61.6	67.2	78.8	82.8	93.3
BJ	70.8	70.8	55.6	73.6	75.6	95.1
ZAM	70	70	55.6	64	67.6	95.1
BUT	71.2	70.8	58.8	73.6	76	96.0
PWV	70.4	70.4	58.4	73.2	75.2	96.4
GRECT	70.8	70.4	65.2	75.2	76.8	95.7
SPWV	70.4	70.4	58.4	72.8	76.8	96.4
CW	71.2	70.8	57.2	73.6	76	96.6
RI	70.8	70.8	62.8	73.6	75.6	97.3

c) Five Class Problem

Table 5.4: Performance comparison of the proposed method with that of a state-of-the-art method in terms of average sensitivity (Sen) and average selectivity (Sel) for two class problem

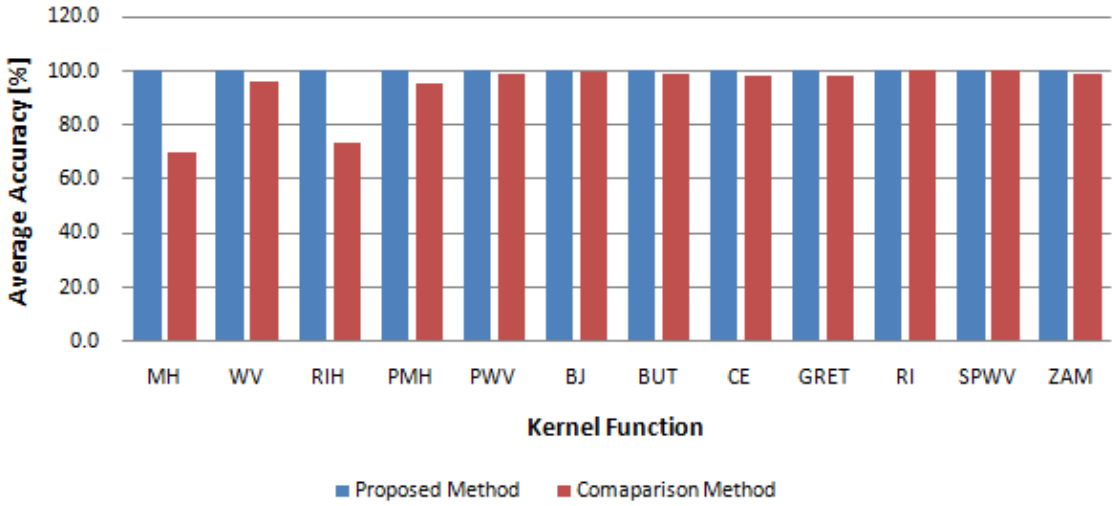
Kernel	Indicator	Proposed Method			Comparison Method		
		Z	S	Avg	Z	S	Avg
MH	Sel	100.0	100.0	**100.0**	68.9	70.4	**69.7**
	Sen	100.0	100.0	**100.0**	71.6	67.6	**69.6**
WV	Sel	100.0	100.0	**100.0**	94.4	98.3	**96.4**
	Sen	100.0	100.0	**100.0**	98.4	94.2	**96.3**
RIH	Sel	100.0	100.0	**100.0**	68.8	82.1	**75.5**
	Sen	100.0	100.0	**100.0**	86.8	60.6	**73.7**
PMH	Sel	100.0	100.0	**100.0**	95.9	94.3	**95.1**
	Sen	100.0	100.0	**100.0**	94.2	96.0	**95.1**
PWV	Sel	100.0	100.0	**100.0**	100.0	98.0	**99.0**
	Sen	100.0	100.0	**100.0**	98.0	100.0	**99.0**
BJ	Sel	100.0	100.0	**100.0**	99.8	96,5	**99.8**
	Sen	100.0	100.0	**100.0**	96,4	99.8	**99.8**
BUT	Sel	100.0	100.0	**100.0**	100.0	98.0	**99.0**
	Sen	100.0	100.0	**100.0**	98.0	100.0	**99.0**
CW	Sel	100.0	100.0	**100.0**	100.0	96.5	**98.3**
	Sen	100.0	100.0	**100.0**	96.4	100.0	**98.2**
GRECT	Sel	100.0	100.0	**100.0**	100.0	96.3	**98.2**
	Sen	100.0	100.0	**100.0**	96.2	100.0	**98.1**
RI	Sel	100.0	100.0	**100.0**	100.0	100.0	**100.0**
	Sen	100.0	100.0	**100.0**	100.0	100.0	**100.0**
SPWV	Sel	100.0	100.0	**100.0**	100.0	100.0	**100.0**
	Sen	100.0	100.0	**100.0**	100.0	100.0	**100.0**
ZAM	Sel	100.0	100.0	**100.0**	97.4	99.8	**98.6**
	Sen	100.0	100.0	**100.0**	99.8	97.5	**98.7**

Table 5.5: Performance comparison of the proposed method with that of a state-of-the-art method in terms of average sensitivity (Sen) and average selectivity (Sel) for three class problem

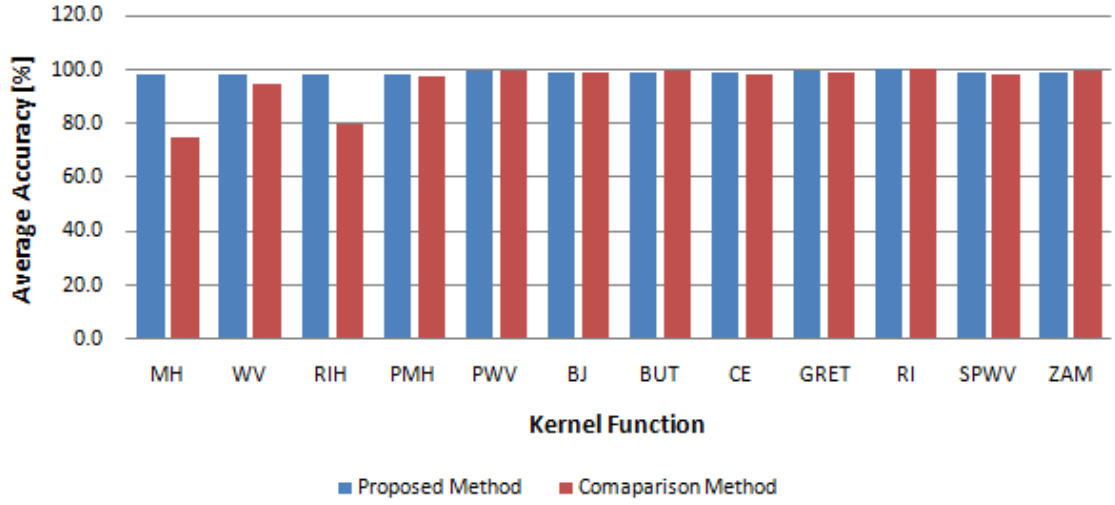
Kernel	Indicator	Proposed Method				Comparison Method			
		Z	F	S	Avg	Z	F	S	Avg
MH	Sel	98.0	98.0	98.0	**98.0**	65.4	97.2	64.2	**75.6**
	Sen	98.7	96.7	98.7	**98.0**	68.8	90.0	65.6	**74.8**
WV	Sel	97.7	99.3	98.0	**98.3**	96.0	93.1	93.8	**94.3**
	Sen	98.7	97.4	99.0	**98.4**	95.6	93.8	93.4	**94.3**
RIH	Sel	98.3	97.7	99.3	**98.4**	66.5	97.9	80.4	**81.6**
	Sen	98.7	98.0	98.7	**98.4**	91.2	92.6	54.8	**79.5**
PMH	Sel	99.0	96.0	99.3	**98.1**	97.9	100.0	95.2	**97.7**
	Sen	97.4	98.3	98.7	**98.1**	95.0	100.0	98.0	**97.7**
PWV	Sel	100.0	99.0	98.7	**99.2**	99.8	100.0	98.0	**99.3**
	Sen	98.4	99.3	100.0	**99.2**	98.0	100.0	99.8	**99.3**
BJ	Sel	100.0	99.0	98.3	**99.1**	99.4	100.0	97.6	**99.0**
	Sen	99.3	99.0	99.0	**99.1**	97.6	100.0	99.4	**99.0**
BUT	Sel	99.3	99.0	99.0	**99.1**	100.0	100.0	98.0	**99.3**
	Sen	99.3	98.0	100.0	**99.1**	98.0	100.0	100.0	**99.3**
CW	Sel	100.0	98.0	99.0	**99.0**	98.6	100.0	96.1	**98.2**
	Sen	98.1	99.3	99.7	**99.0**	96.0	100.0	98.6	**98.2**
GRECT	Sel	100.0	98.7	100.0	**99.6**	100.0	100.0	96.5	**98.8**
	Sen	98.7	100.0	100.0	**99.6**	96.4	100.0	100.0	**98.8**
RI	Sel	100.0	100.0	100.0	**100.0**	100.0	100.0	100.0	**100.0**
	Sen	100.0	100.0	100.0	**100.0**	100.0	100.0	100.0	**100.0**
SPWV	Sel	99.3	98.3	98.3	**98.7**	100.0	100.0	95.2	**98.4**
	Sen	97.1	99.3	99.7	**98.7**	95.0	100.0	100.0	**98.3**
ZAM	Sel	99.7	97.0	100.0	**98.9**	99.6	100.0	100.0	**99.9**
	Sen	98.7	99.7	98.4	**98.9**	100.0	100.0	99.6	**99.9**

Table 5.6: Performance comparison of the proposed method with that of a state-of-the-art method in terms of average sensitivity (Sen) and average selectivity (Sel) for five class problem

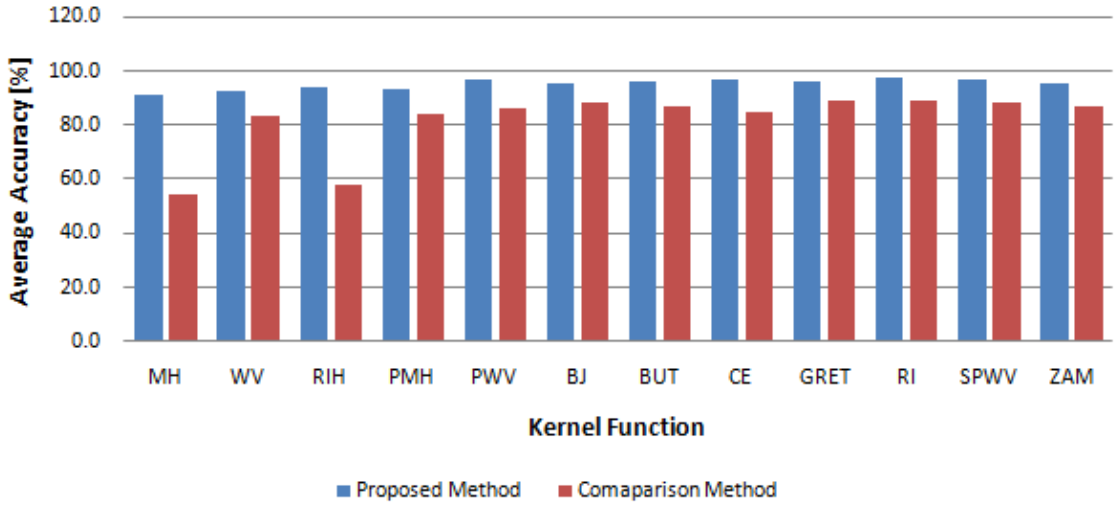
Kernel	Indicator	Proposed Method						Comparison Method					
		Z	O	F	N	S	Avg	Z	O	F	N	S	Avg
MH	Sel	90.7	93.7	87.0	84.3	98.7	**90.9**	39.4	40.4	63.6	90.9	39.8	**54.8**
	Sen	90.1	92.8	85.9	87.1	98.7	**90.9**	59.6	20.6	70.2	85.6	37.0	**54.6**
WV	Sel	94.7	95.0	81.3	90.7	98.7	**92.1**	76.0	72.5	86.7	93.0	90.4	**83.7**
	Sen	92.6	96.6	90.1	83.1	99.0	**92.3**	67.2	81.0	87.2	89.8	92.2	**83.5**
RIH	Sel	92.7	88.7	92.3	93.3	97.7	**92.9**	40.2	41.9	71.7	93.3	46.6	**58.7**
	Sen	95.9	94.1	88.5	92.9	98.0	**93.9**	73.4	10.8	71.0	89.4	45.2	**58.0**
PMH	Sel	92.3	94.3	89.0	90.7	98.7	**93.0**	77.0	70.8	93.9	99.8	81.8	**84.7**
	Sen	92.6	93.8	93.6	88.4	98.0	**93.3**	56.8	86.0	86.0	100.0	92.4	**84.2**
PWV	Sel	100.0	95.7	91.0	94.7	100.0	**96.3**	76.8	72.7	97.8	99.0	86.4	**86.5**
	Sen	95.0	100.0	95.8	91.3	99.7	**96.4**	67.6	77.4	89.4	99.8	98.0	**86.4**
BJ	Sel	97.7	97.7	92.0	89.7	98.3	**95.1**	84.3	76.3	95.8	99.8	86.7	**88.6**
	Sen	95.8	97.7	90.9	91.5	99.7	**95.1**	70.8	83.2	90.4	100.0	97.4	**88.4**
BUT	Sel	98.3	96.0	94.0	91.7	99.7	**95.9**	79.8	75.9	97.6	98.6	84.9	**87.4**
	Sen	94.6	99.3	93.1	93.2	99.7	**96.0**	72.0	78.6	87.8	99.8	97.6	**87.2**
CW	Sel	98.3	97.3	93.3	94.7	99.3	**96.6**	71.0	71.7	98.7	100.0	82.9	**84.9**
	Sen	97.1	98.4	95.0	93.7	99.0	**96.6**	63.0	72.8	88.8	100.0	99.2	**84.8**
GRECT	Sel	98.7	98.3	91.7	93.7	99.3	**96.3**	79.4	79.5	97.7	98.6	88.4	**88.7**
	Sen	96.7	99.7	91.2	91.3	99.7	**95.7**	76.2	77.6	93.0	99.8	97.2	**88.8**
RI	Sel	97.0	98.0	94.7	97.3	99.3	**97.3**	83.2	82.6	97.3	99.4	83.3	**89.2**
	Sen	97.7	99.3	95.6	93.9	100.0	**97.3**	70.4	85.2	92.0	99.8	97.6	**89.0**
SPWV	Sel	97.3	96.7	93.0	97.0	97.7	**96.3**	80.2	77.2	100.0	100.0	83.8	**88.2**
	Sen	95.8	97.3	96.6	92.7	99.7	**96.4**	66.2	82.0	92.0	99.8	100.0	**88.0**
ZAM	Sel	96.7	94.0	93.0	92.0	99.7	**95.1**	74.3	74.1	100.0	100.0	87.5	**87.2**
	Sen	94.2	97.3	92.4	93.0	98.7	**95.1**	69.8	74.8	93.0	100.0	98.0	**87.1**



(a) Two class problem



(b) Three class problem



(c) Five class problem

Fig. 5.14: Performance comparison of the proposed method with that of a state-of-the-art method in terms of average accuracy for two, three and five class problems.

Chapter 6

Conclusion

6.1 Concluding Remarks

In this thesis, a time-frequency based approach to solve the seizure detection and classification problem using the energy and entropy of time-frequency module has been presented. We incorporated a uniform and nonuniform modularization technique to extract necessary local information from time-frequency plane. Twelve Cohen class kernels are used for interference reduction before feature extraction. Introducing Hilbert transform to make the signal analytic is also facilitate time-frequency decomposition. Usefulness of six different classifiers with all Cohen class kernels are examined to get the suitable combination which can serve the seizure detection and classification problem better .Seizure detection and classification methods using the proposed feature has the highest reported accuracy with faster processing time.

6.2 Contributions of this Thesis

The major contribution of the thesis are,

- Introducing Hilbert transform to make the real EEG signal to analytic. Hilbert transform add a imaginary part in a real signal from the phase information of that signal. This imaginary signal has no negative frequency and suitable for complex decomposition. Time-frequency decomposition technique can handle the Hilbert transformed signal more easily.

- Twelve Cohen class kernels are utilized to reduce the inherent interference in time-frequency analysis. Effect of those Cohen class kernel functions on

interference reduction and time-frequency resolution are investigated to select the suitable kernel for both seizure detection and classification cases.

- Effect of modularization for feature extraction is investigated and usefulness of the modularization technique are shown also. Suitable number of module size are also determined from details analysis.

- A nonuniform modularized feature involving cumulative energy at some predefined marked percentiles frequencies and modular entropy is deduced, which has 100% accuracy in seizure detection for all Cohen class kernel. Previous reported work only get this type of accuracy for RI kernel which processing time is significantly large. So, reducing processing time with obtaining the highest accuracy is the key point of the feature.

- A uniform modularization feature involving modular energy and that of entropy is introduced, which can classify higher class classification problem. We form three classification problems, same as previous work, namely two, three and five class, to prove the excellence of the proposed method. The proposed method has 100% accuracy for two class problem for all kernel. For three and five class problems, accuracy is the highest, 100% and 97.3%, respectively.

- Six types of classifier are used to select the suitable one. For seizure detection case, we found Euclidean distance based KNN is suitable for its faster processing time. On the other hand, for seizure classification problem, ANN is better in terms of all performance criteria.

6.3 Scopes for Future Work

In this thesis, effective time-frequency based methods for seizure detection and classification is developed. However, there are some scopes for future research, as mentioned below:

- In this research, we use a popular EEG database which consists of five class EEG data. The proposed method can classify those with highest accuracy using time-frequency analysis. In future, effectiveness of the proposed method using different EEG databases can be verified.

- For seizure detection technique, whole sample of EEG data is used as per given database. So, in future, online seizure detection can be performed using different databases.

- As modularization impact is significant on seizure classification accuracy, performance of other feature set utilizing the modularization technique can be investigated.

Bibliography

[1] P. Rajdev, M. Ward, J. L. Rickus, R. M. Worth, and P. Irazoqui, "Real-time seizure prediction from local field potentials using an adaptive wiener algorithm," *Comp. in Bio. and Med.*, vol. 40, no. 1, pp. 97–108, 2010.

[2] L. Iasemidis, D. Shiau, W. Chaovalitwongse, J. Sackellares, P. Pardalos, J. Principe, P. Carney, A. Prasad, B. Veeramani, and K. Tsakalis, "Adaptive epileptic seizure prediction system," *Biomedical Engineering, IEEE Transactions on*, vol. 50, no. 5, pp. 616–627, 2003.

[3] R. Fisher, G. Harding, G. Erba, G. Barkley, and A. Wilkins, "Photic-and pattern-induced seizures: A review for the epilepsy foundation of america working group," *Epilepsia*, vol. 46, no. 9, pp. 1426–1441, 2005.

[4] R. A. U. D. Puthankattil Subha, Paul K. Joseph and C. M. Lim, "EEG signal analysis: A survey," *Journal of Medical Systems*, vol. 34, no. 2, pp. 195–212, 2010.

[5] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, USA, 1995.

[6] B. Boashash, *Time-frequency signal analysis and processing: A comprehensive reference*, 1st ed. Oxford,UK: Elsvier, 2003.

[7] A. Temko, C. Nadeu, W. Marnane, G. Boylan, and G. Lightbody, "EEG signal description with spectral-envelope-based speech recognition features for detection of neonatal seizures," *Information Technology in Biomedicine, IEEE Transactions on*, no. 99, pp. 839–847, 2011.

[8] S. Viglione and G. Walsh, "Proceedings: Epileptic seizure prediction." *Electroencephalography and clinical neurophysiology*, vol. 39, no. 4, p. 435, 1975.

[9] D. Stanski, R. Hudson, T. Homer, L. Saidman, and E. Meathe, "Pharmacodynamic modeling of thiopental anesthesia," *Journal of Pharmacokinetics and Pharmacodynamics*, vol. 12, no. 2, pp. 223–240, 1984.

[10] B. Litt, R. Esteller, J. Echauz, M. D'Alessandro, R. Shor, T. Henry, P. Pennell, C. Epstein, R. Bakay, M. Dichter *et al.*, "Epileptic seizures may begin hours in advance of clinical onset: a report of five patients," *Neuron*, vol. 30, no. 1, pp. 51–64, 2001.

[11] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalography and Clinical Neurophysiology*, vol. 29, no. 3, pp. 306–310, 1970.

[12] Z. Rogowski, I. Gath, and E. Bental, "On the prediction of epileptic seizures," *Biological cybernetics*, vol. 42, no. 1, pp. 9–15, 1981.

[13] Y. Salant, I. Gath, and O. Henriksen, "Prediction of epileptic seizures from two-channel EEG," *Medical and Biological Engineering and Computing*, vol. 36, no. 5, pp. 549–556, 1998.

[14] J. Martinerie, C. Adam, M. Le Van Quyen, M. Baulac, S. Clemenceau, B. Renault, F. Varela *et al.*, "Epileptic seizures can be anticipated by non-linear analysis," *Nature Medicine*, vol. 4, no. 10, pp. 1173–1176, 1998.

[15] K. Lehnertz and C. Elger, "Neuronal complexity loss in temporal lobe epilepsy: effects of carbamazepine on the dynamics of the epileptogenic focus," *Electroencephalography and clinical neurophysiology*, vol. 103, no. 3, pp. 376–380, 1997.

[16] R. Aschenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H. Voss, J. Timmer, and A. Schulze-Bonhage, "How well can epileptic seizures be predicted? an evaluation of a nonlinear method," *Brain*, vol. 126, no. 12, pp. 2616–2626, 2003.

[17] D. Li, W. Zhou, I. Drury, and R. Savit, "Non-linear, non-invasive method for seizure anticipation in focal epilepsy," *Mathematical biosciences*, vol. 186, no. 1, pp. 63–77, 2003.

[18] I. Drury, B. Smith, D. Li, and R. Savit, "Seizure prediction using scalp electroencephalogram," *Experimental neurology*, vol. 184, pp. 9–18, 2003.

[19] 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:Elsvier*, vol. 357, no. 9251, pp. 183–188, 2001.

[20] L. Iasemidis, J. Chris Sackellares, H. Zaveri, and W. Williams, "Phase space topography and the lyapunov exponent of electrocorticograms in partial seizures," *Brain Topography*, vol. 2, no. 3, pp. 187–201, 1990.

[21] M. Rosenblum, A. Pikovsky, and J. Kurths, "From phase to lag synchronization in coupled chaotic oscillators," *Physical Review Letters*, vol. 78, no. 22, pp. 4193–4196, 1997.

[22] R. Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, "Performance of different synchronization measures in real data: a case study on electroencephalographic signals," *Physical Review E*, vol. 65, no. 4, p. 041903, 2002.

[23] L. Iasemidis, P. Pardalos, J. Sackellares, and D. Shiau, "Quadratic binary programming and dynamical system approach to determine the predictability of epileptic seizures," *Journal of combinatorial optimization*, vol. 5, no. 1, pp. 9–26, 2001.

[24] F. Mormann, K. Lehnertz, P. David, and C. E Elger, "Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients," *Physica D: Nonlinear Phenomena*, vol. 144, no. 3-4, pp. 358–369, 2000.

[25] S. Wilson and R. Emerson, "Spike detection: a review and comparison of algorithms," *Clinical Neurophysiology*, vol. 113, no. 12, pp. 1873–1881, 2002.

[26] J. Gotman and P. Gloor, "Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG," *Electroencephalography and clinical neurophysiology*, vol. 41, no. 5, pp. 513–529, 1976.

[27] M. Dümpelmann and C. Elger, "Automatic detection of epileptiform spikes in the electrocorticogram: a comparison of two algorithms," *Seizure*, vol. 7, no. 2, pp. 145–152, 1998.

[28] S. Wilson, C. Turner, R. Emerson, and M. Scheuer, "Spike detection ii: automatic, perception-based detection and clustering," *Clinical neurophysiology*, vol. 110, no. 3, pp. 404–411, 1999.

[29] A. Tzallas, P. Karvelis, C. Katsis, D. Fotiadis, S. Giannopoulos, S. Konitsiotis *et al.*, "A method for classification of transient events in EEG recordings: application to epilepsy diagnosis." *Methods of information in medicine*, vol. 45, no. 6, p. 610, 2006.

[30] N. Acır and C. Güzeliş, "Automatic spike detection in EEG by a two-stage procedure based on support vector machines," *Computers in Biology and Medicine*, vol. 34, no. 7, pp. 561–575, 2004.

[31] J. Benedetto and D. Colella, "Wavelet analysis of spectrogram seizure chirps," in *Proceedings of SPIE*, vol. 2569, 1995, p. 512.

[32] S. Schiff, J. Heller, S. Weinstein, and J. Milton, "Controlled wavelet transforms for EEG spike and seizure localization," in *Proceedings of SPIE*, vol. 2242, 1994, p. 762.

[33] W. Williams, H. Zaveri, and J. Sackellares, "Time-frequency analysis of electrophysiology signals in epilepsy," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 14, no. 2, pp. 133–143, 1995.

[34] H. Zaveri, R. Duckrow, N. De Lanerolle, and S. Spencer, "Distinguishing subtypes of temporal lobe epilepsy with background hippocampal activity," *Epilepsia*, vol. 42, no. 6, pp. 725–730, 2001.

[35] A. Tzallas, M. Tsipouras, and D. Fotiadis, "Epileptic seizure detection in EEGs using time–frequency analysis," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, no. 5, pp. 703–710, 2009.

[36] F. Mormann, T. Kreuz, C. Rieke, R. Andrzejak, A. Kraskov, P. David, C. Elger, and K. Lehnertz, "On the predictability of epileptic seizures," *Clinical neurophysiology*, vol. 116, no. 3, pp. 569–587, 2005.

[37] E. George, *Time series analysis: Forecasting and control*. Holden-D., 1970.

[38] C. Stam, "Nonlinear dynamical analysis of EEG and MEG: review of an emerging field," *Clinical Neurophysiology*, vol. 116, no. 10, pp. 2266–2301, 2005.

[39] P. Grassberger and I. Procaccia, "Characterization of strange attractors," *Physical review letters*, vol. 50, no. 5, pp. 346–349, 1983.

[40] K. Lehnertz and C. Elger, "Can epileptic seizures be predicted? evidence from nonlinear time series analysis of brain electrical activity," *Physical Review Letters*, vol. 80, no. 22, pp. 5019–5022, 1998.

[41] M. Harrison, I. Osorio, M. Frei, S. Asuri, and Y. Lai, "Correlation dimension and integral do not predict epileptic seizures," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 15, no. 3, pp. 033 106–033 106, 2005.

[42] P. McSharry, L. Smith, and L. Tarassenko, "Prediction of epileptic seizures: are nonlinear methods relevant?" *Nature medicine*, vol. 9, no. 3, pp. 241–242, 2003.

[43] E. Ott, *Chaos in dynamical systems*. Cambridge Univ Pr, 2002.

[44] W. Drongelen, S. Nayak, D. Frim, M. Kohrman, V. Towle, H. Lee, A. McGee, M. Chico, and K. Hecox, "Seizure anticipation in pediatric epilepsy: use of kolmogorov entropy," *Pediatric neurology*, vol. 29, no. 3, pp. 207–213, 2003.

[45] M. Van Quyen, J. Martinerie, M. Baulac, and F. Varela, "Anticipating epileptic seizures in real time by a non-linear analysis of similarity between EEG recordings," *Neuroreport*, vol. 10, no. 10, p. 2149, 1999.

[46] F. Mormann, R. Andrzejak, T. Kreuz, C. Rieke, P. David, C. Elger, and K. Lehnertz, "Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients," *Physical Review E*, vol. 67, no. 2, p. 021912, 2003.

[47] M. Le Van Quyen, J. Soss, V. Navarro, R. Robertson, M. Chavez, M. Baulac, and J. Martinerie, "Preictal state identification by synchronization changes in long-term intracranial EEG recordings," *Clinical neurophysiology*, vol. 116, no. 3, pp. 559–568, 2005.

[48] M. Rosenstein, J. Collins, and C. De Luca, "A practical method for calculating largest lyapunov exponents from small data sets," *Physica D: Nonlinear Phenomena*, vol. 65, no. 1-2, pp. 117–134, 1993.

[49] Y. Lai, M. Harrison, M. Frei, and I. Osorio, "Inability of lyapunov exponents to predict epileptic seizures," *Physical review letters*, vol. 91, no. 6, p. 68102, 2003.

[50] M. Bedeeuzzaman, O. Farooq, Y. Khan, R. Perumal, W. Sultana, S. Sahoo, R. Rajak, N. Pasricha, A. Arora, R. Cheema *et al.*, "Automatic seizure detection using inter quartile range," *International Journal of Computer Applications*, vol. 44, no. 11, pp. 1–5, 2012.

[51] P. Rana, J. Lipor, H. Lee, W. van Drongelen, M. Kohrman, and B. Van-Veen, "Seizure detection using the phase-slope index and multichannel ECoG," *Biomedical Engineering, IEEE Transactions on*, no. 99, pp. 1–1, 2012.

[52] R. Esteller, J. Echauz, M. D'Alessandro, G. Worrell, S. Cranstoun, G. Vachtsevanos, and B. Litt, "Continuous energy variation during the seizure cycle: towards an on-line accumulated energy," *Clinical neurophysiology*, vol. 116, no. 3, pp. 517–526, 2005.

[53] T. Maiwald, M. Winterhalder, R. Aschenbrenner-Scheibe, H. Voss, A. Schulze-Bonhage, and J. Timmer, "Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic," *Physica D: Nonlinear Phenomena*, vol. 194, no. 3, pp. 357–368, 2004.

[54] C. Rieke, K. Sternickel, R. Andrzejak, C. Elger, P. David, and K. Lehncrtz, "Measuring nonstationarity by analyzing the loss of recurrence in dynamical systems," *Physical review letters*, vol. 88, no. 24, pp. 244 102–244 102, 2002.

[55] D. Ravish and S. Devi, "Automated seizure detection and spectral analysis of EEG seizure time series," *European Journal of Scientific Research*, vol. 68, no. 1, pp. 72–82, 2012.

[56] R. Harikumar, C. Babu, and T. Vijayakumar, "Performance analysis of elman neural networks as post classifiers for wavelet transforms based feature extraction using hard and soft thresholding methods in the classification of epilepsy risk levels from eeg signals," *European Journal of Scientific Research*, vol. 71, no. 2, pp. 221–232, 2012.

[57] S. Janjarasjitt and K. Loparo, "Examination of temporal characteristics of epileptic EEG subbands based on the local min-max," *Proceedings of the International MultiConference of Engineers and Computer Scientists*, vol. 2, 2012.

[58] H. Zaveri, W. Williams, and J. Sackellares, "Energy based detection of seizures," in *Pro Annu Conf Enh Med Biol, IEEE*, vol. 15, 1993, pp. 363–364.

[59] P. P. Acharjee and C. Shahnaz, "Epileptic seizure detection from EEG signals based on features extracted from non-uniform modules in time-frequency domain," *Submitted to Digital Signal Processing, Elsevier.*

[60] H. Hassanpour, M. Mesbah, and B. Boashash, "EEG spike detection using time-frequency signal analysis," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5. IEEE, 2004, pp. V–421.

[61] L. Cohen, "Time-frequency distributions-a review," *Proceedings of the IEEE*, vol. 77, no. 7, pp. 941–981, 1989.

[62] R. D. T. J. G. R. Jefferys and M. Whittington, "Neuronal networks for induced "40 hz" rhythms," *Trends in Neurosciences*, vol. 19, no. 5, pp. 202–208, 1996.

[63] H. Qu and J. Gotman, "A patient-specific algorithm for the detection of seizure onset in long-term EEG monitoring: possible use as a warning device," *Biomedical Engineering, IEEE Transactions on*, vol. 44, no. 2, pp. 115–122, 1997.

[64] R. González and R. Woods, *Tratamiento digital de imágenes*. Addison-Wesley Longman, 1996.

[65] D. Bremner, E. Demaine, J. Erickson, J. Iacono, S. Langerman, P. Morin, and G. Toussaint, "Output-sensitive algorithms for computing nearest-neighbour decision boundaries," *Discrete & Computational Geometry*, vol. 33, no. 4, pp. 593–604, 2005.

[66] T. Hastie, R. Tibshirani, and J. Friedman, "The elements of statistical learning: Data mining, inference, and prediction," *BeiJing: Publishing House of Electronics Industry*, 2004.

[67] A. Nigrin, *Neural networks for pattern recognition*. The MIT press, 1993.

[68] O. Omidvar and J. Dayhoff, *Neural networks and pattern recognition*. Academic Pr, 1998.

[69] P. P. Acharjee and C. Shahnaz, "Modular feature based multiclass epileptic seizure classification," *Submitted to IEEE Transactions on Neural Networks and Learning Systems*.

[70] R. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C. Elger *et al.*, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review-Series E*, vol. 64, no. 6, pp. 61 907–61 907, 2001.