DEVELOPING ELECTROENCEPHALOGRAM SIGNAL PROCESSING TECHNIQUES FOR BRAIN COMPUTER INTERFACE APPLICATIONS

A

THESIS

Submitted in partial fulfilment of the Requirement for the Award of the

Degree of

DOCTOR OF PHILOSOPHY

IN

ELECTRONICS AND COMMUNICATION ENGINEERING

by

KOMAL JINDAL

(UID: 951706002)

Supervisor

Dr. Rahul Upadhyay (Assistant Professor, ECED)

Dr. Hari Shankar Singh (Assistant Professor, ECED)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

THAPAR INSTITUTE OF ENGINEERING & TECHNOLOGY

NOVEMBER, 2023

Dedicated to

The Grand Weaver of my life,

My Parents & Husband

DECLARATION

I, Komal Jindal hereby declare that all the written material submitted in the thesis entitled "Developing Electroencephalogram Signal Processing Techniques for Brain Computer Interface Applications" in fulfilment of the requirement for the award of the degree "Doctor of Philosophy" submitted to the Department of Electronics and Communication Engineering, Thapar Institute of Engineering and Technology, Patiala is an authentic record of my own research work carried out under the supervision and guidance of Dr. Rahul Upadhyay, Assistant Professor, Thapar Institute of Engineering and Technology, Patiala and Dr. Hari Shankar Singh, Assistant Professor, Thapar Institute of Engineering and Technology, Patiala. The results obtained in this thesis have not been submitted in part or in full to any other University or Institute for the award of any degree or diploma.

Komal Jindal

UID: 951706002

Date: November 2023

Place: Patiala, India

CERTIFICATE

Certified that the thesis entitled "Developing Electroencephalogram Signal Processing Techniques for Brain Computer Interface Applications" submitted by Mrs. Komal Jindal to the Thapar Institute of Engineering & Technology, Patiala for the award of "Doctor of Philosophy" has been thoroughly examined. I have reviewed the content and presentation style of this thesis and declare that it is free from plagiarism and grammatical errors. The research work stated in this thesis is the original work of the candidate. I confirm that the investigations have been conducted with the ethical policies of the institute and research is presented without prejudice.

03/11/2023

Dr. Rahul Upadhyay

Assistant Professor

Department of Electronics & Communication,

Thapar Institute of Engineering & Technology

Patiala - 147004

INDIA

Hari Shankar Singh 3/11/2023

Dr. Hari Shankar Singh

Assistant Professor

Department of Electronics & Communication,

Thapar Institute of Engineering & Technology

Patiala - 147004

INDIA

Acknowledgement

In this great endeavour of my life, I am always thankful to my godmother for guiding me in the right direction. She has given me the power to believe in myself.

I would like to express my sincere gratitude to my advisor Dr. Rahul Upadhyay for the continuous support of my PhD study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD study. I would also like to thank Dr. Hari Shankar Singh to provide motivation and support for presenting my IRB in front of the doctoral committee and successfully start my PhD journey.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Ravinder Aggarwal, Dr. Kulbir Singh, and Dr. Vinay Kumar, for their insightful comments and encouragement. Their suggestions have helped me in bringing this thesis in present shape.

My acknowledgement would be incomplete without thanking the biggest source of my strength, my family. I also express my sincere gratitude and deep love to my husband Mr. Mohit Bansal for enduring a seemingly endless ordeal for sacrificing some of his best years.

"...may the LORD reward everyone of you for your kindness... to me."

Last, but not least, I would like to thank my grandfather for inspiring me to start this wonderful journey of my PhD.

Komal Jindal

Abstract

Multi-class motor imagery EEG activity decoding has always been challenging for the development of Brain Computer Interface (BCI) system. Deep learning has recently emerged as a powerful approach for BCI system development using EEG activity. However, the EEG activity analysis and classification should be robust, automated and accurate. Currently, available BCI systems perform well for binary task identification whereas, multi-class classification of EEG activity for BCI applications is still a challenging task. The frequent occurrence of multiple artifactual origins makes it imperative to adopt an adequate artifacts removal methodology prior to feature estimation in Brain-Computer Interface applications. This study proposes a novel artifact removal methodology using a joint application of Fast-Power Independent Component Analysis and General Linear Chirplet Transform for automatic identification and rejection of artifactual origins. After segregating ongoing Electroencephalogram activity into Independent Components, Katz-Fractal Sparsity criterion is employed to identify artifactual components. The identified artifactual components are treated by General Linear Chirplet Transform-based EEG de-noising method to recover useful cerebral information leaked with artifactual origins. Thereafter, Inverse Independent Component Analysis yields artifact corrected clean Electroencephalogram activity for further analysis. The effectiveness of the developed methodology is validated with simulated as well as empirical Electroencephalogram datasets. The experimental results establish the proposed method as a potential candidate for non-cerebral artifacts correction and noise suppression from Electroencephalogram records.

Further, to enhance the multi-class BCI system's performance, a novel channel selection, and features optimization methodology have been proposed. First, the multi-channel EEG dataset is reduced to an optimum no. of channel subsets using the developed MDA-SOGWO-based EEG channel selection criterion. After that, the discriminable feature-set is generated from the time, frequency, and FAWT-based time-frequency domain of the EEG dataset. Then, an informative feature-set is constructed from extracted features through the CCA-RFE-based feature selection criterion. Finally, training and validation of the selected feature-set are carried out using ELM, LDA, RF, and MLP classifiers. The proposed methodology is evaluated on multi-class MI datasets of BCI Competition IV 2a and BCI Competition III 3a, yielding classification accuracy of *85.11%* and *92.46%*, respectively. The classification results pose the proposed methodology as an efficient method for future BCI system design.

Afterwards, for efficient feature extraction and classification methodology development, a hybrid AttentionNet ensemble voting classifier model is developed. The Time-Frequency Representation (TFR) of the multi-class EEG activity is generated using Transient Extracting Transform (TET). The TFR spectrogram images are input to the designed AttentionNet ensemble voting classifier model for training and classification purposes. The model is evaluated using different fusion strategies viz. feature-level and score-level fusion of layers. The proposed model is evaluated on two MI-BCI datasets, BCI competition IV 2a and BCI competition III 3a, yielding the highest classification accuracies of *88.14 %* and *93.13 %*, respectively. The results obtained on a large multi-class MI-BCI dataset confirm that the proposed hybrid AttentionNet ensemble voting classifier model significantly outperforms the conventional algorithm and achieves significantly high classification accuracy for the feature-level fusion of layers. The developed framework aids in identifying different tasks for multi-class MI-BCI EEG activity.

LIST OF PUBLICATIONS

Journal Publications:

1. **K. Jindal**, R. Upadhyay, H. S. Singh, "A Novel EEG Channel Selection and Classification Methodology for Multi-Class Motor Imagery-based BCI System Design", *International Journal of Imaging Systems & Technology*, Vol. 32(4), pp. 1318-1337,2022.

2. **K. Jindal**, R. Upadhyay, H. S. Singh, "Application of hybrid GLCT-PICA de-noising method in automated EEG artefact removal", *Biomedical Signal Processing and Control*, vol. 60, no. pp. 101977, 2020.

3. **K. Jindal**, R. Upadhyay, H. S. Singh, "A Hybrid Ensemble Voting based AttentionNet Model for Motor Imagery EEG Classification", *Analog Integrated Circuits and Signal Processing*, 2022. (Under Review)

Conference Publication:

1. **K. Jindal**, R. Upadhyay, H. Singh, "EEG Artifact Removal and Noise Suppression using Hybrid GLCT-ICA Technique", IEEE *The 10th International Congress on Ultra-Modern Telecommunications and Control Systems*, Moscow(Russia), 05-09 November, 2018.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

Preface

This chapter outlines a brief description of the brain-computer interface and its applications, and also the various types of conventional BCI systems. Further, the motivation behind the proposal of the feature extraction and classification techniques in BCI system design is explained. Additionally, the architecture of the present research is described for a better understanding of the work.

1.1 Overview

Brain-Computer Interface (BCI), also known as *Brain Interface (BI), and Brain-Machine Interface (BMI)* is an artificial communication pathway between the computer and the human brain that translates neural activity as a control signal [1]. In other words, BCI acts as an alternative communication pathway for subjects to interact with the environment without the participation of peripheral nerves and muscles. BCI systems have enormous applications in restoring the movement ability of physically challenged and locked-in people [2]. The primary goal of the BCI system is to provide a quality life to such people. These systems allow patients to communicate with the environment in order to fulfil their basic needs. Moreover, it makes patients capable to control functional electrical devices (like wheelchairs, prosthetics) [3]. It has been estimated that around *1* million people worldwide are suffering from movement disability. This has motivated the researchers to explore the possibility of harnessing the alternative communication pathway between the human brain and the external world.

Conventionally, BCI systems are classified into three categories namely, dependent and independent BCI, invasive and non-invasive BCI, and synchronous and asynchronous BCI. A brief description of these state-of-art BCI systems is as follows:

a) **Dependent and independent BCI:** Dependent BCI system utilizes the brain signals which are acquired during control over motor activity such as Visually Evoked Potentials (VEP) [4], whereas an independent BCI system does not depend upon peripheral nerves and muscle activity during brain signal acquisition, e.g., a BCI system utilizing Sensorimotor Rhythms (SMRs). An independent BCI system can aid people suffering from severe neuromuscular disabilities (spinal cord injury or ALS) [5]. On the other hand, a dependent BCI system can assist a healthy person in terms of gaming and entertainment.

b) **Invasive and non-invasive BCI**: Invasive BCI systems include brain activity recording by implantation of electrodes under the scalp during neurosurgery, such as in Electrocorticography (ECoG), and Positron Emission Tomography (PET) [6]. The invasive BCI system provides brain signals of high quality with a very high Signal to Noise Ratio (SNR), but the acquisition procedure is dangerous and expensive, whereas the non-invasive BCI system includes data acquisition by electrode placement over the scalp [7]. This procedure does not require surgery for the implantation of electrodes and it is inexpensive. Data acquisition techniques such as Functional Magnetic Resonance Imaging (fMRI), Magnetoencephalography (MEG), and Electroencephalography (EEG) come under non-invasive BCI system design [8].

c) Synchronous and asynchronous BCI: Synchronous (cue-paced) BCI system processes brain signals acquired under a stipulated time-frame [7], [9]. Any electrical activity outside the pre-defined time frame is ignored. These BCI systems such as Motor Imagery (MI) BCI define certain periods for the subject to send commands. At the same time, in asynchronous (self-paced) BCI systems, complete brain activity is analysed irrespective of subject action or time limits. The asynchronous BCI systems are more complex and demand high computation [9].

BCI systems can be sub-divided within six components namely, brain activity acquisition, pre-processing, feature extraction, feature selection, classification, and applications (as shown in Figure 1.1). Human brain electrical activity can be acquired using many non-invasive procedures such as fMRI, EEG, MEG, and Near-Infrared Spectroscopy (NIRS) [8]. The fMRI and NiRS are functional neuroimaging procedures based on the criterion of infrared light and magnetic field respectively. These neuroimaging techniques are utilized to detect brain oxygen level changes in various regions of the human brain, whereas MEG based brain activity mapping criterion relies on the principle of measuring the magnetic field produced by the electrical activity of human brain. These neuroimaging techniques are less prone to artifacts but suffer from the drawback of high cost and importability. Electroencephalography (EEG) is a brain electrical activity recording procedure in which several electrodes are mounted on the human scalp as per the international standards of electrode placement (e.g., 10–20 system of electrode placement) [2]. These EEG signals contain vital neural information and can be acquired using wearable devices [1]. Therefore, these signals are used to detect dysfunctional neural dynamics of Alzheimer's, Parkinson's, Epilepsy and various other neurological disorders. EEG signals have also been utilized to detect human vigilance levels as well as to monitor mental states [10]. Considering the advantages of portability and cost-effectiveness, this research work utilizes raw brain activity acquired using EEG based data acquisition criterion.

Despite the advantages, non-linearity, non-stationarity, and small data packets in EEG signals make their processing very challenging [11]. EEG activity is often contaminated by several non-cerebral activities like Electromyogram (EMG), Electrooculogram (EOG), muscle activity, eyeball movement, eye blinks, Electrocardiogram (ECG) [12]. Noises due to instruments, external electromagnetic disturbances, and electrical circuits may affect the EEG signals adversely. The occurrence of artifactual activity weakens the quality of EEG signals and leads to errors in the process of feature extraction. Hence, it is crucial to introduce a pre-processing stage to eradicate artifacts and additive noise from EEG signals before the feature extraction stage [13].

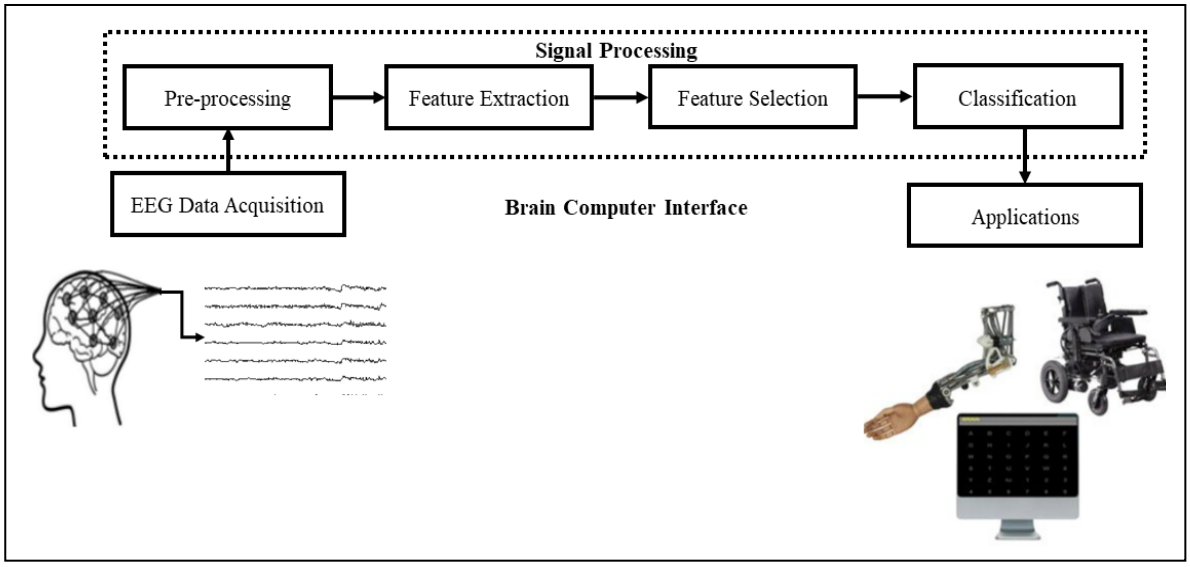


Figure 1.1 Functional model of Brain-Computer Interface

Once the artifacts are removed using the specific pre-processing technique, the EEG activity becomes noise-free, and enhanced EEG signals can be utilized for further processing. The processing pipeline of EEG activity is required to be robust and computationally efficient for competent real-time BCI system development. Also, increasing the count of EEG channels enhances the complexity of the BCI system in terms of computational time. If two or more channels contains same information, then it is required to remove redundant channels to reduce computational complexity [14]. Therefore, an optimum channel selection procedure is needed to remove channels containing redundant information. In order to extract relevant features from pre-processed EEG activity, suitable time-frequency transforms or its improved versions will be studied and implemented for high classification performance.

After extracting useful information, the classifier algorithm is trained with set of features for each class. Each class represents a kind of mental state of brain. Once training is completed, classification algorithm is ready for testing. The input EEG activity to the trained classifier algorithm is translated into a meaningful command to control assistive devices or applications.

Since biomedical signals like EEG are multi-dimensional, it is hard to develop a robust feature extraction and classification methodology for the BCI system. The present thesis aims to develop an efficient BCI system utilizing multi-class motor imagery EEG activity. The enhanced classification accuracy along with optimized intermediate steps following novel pre-processing, feature extraction, and classification methodologies is a primary concern of this work.

1.2 Motivation and Challenges

In order to develop an efficient real-time BCI system, each stage of BCI system design is required to be competent. The challenges identified after surveying the published work pertaining to BCI systems are listed below:

(a) EEG activity cleaning prior to information extraction

In order to eliminate artifacts and noise from EEG activity, researchers suggested various prevention and correction steps using thresholding criterion and Time-Frequency Representation (TFR) based approaches, etc. [14] [2]. However, these approaches lack automation and require expert interference while operating. Not all the approaches available in the literature are suitable for online/real-time EEG recording. In addition, the inappropriate parameter estimation causes loss of actual EEG activity without eliminating artifactual content. It has been observed from the literature that removing non-cerebral biological artifacts originating from Electrocardiogram (ECG), muscle and ocular activity while maintaining high quality EEG activity for real-time applications is still a challenging task and requires efforts to explore new methods for EEG artifact removal.

(b) Selection of the optimum number of EEG channels

Although plenty of methodologies for optimum EEG channel selection for specific applications have been proposed [15], [16], optimum EEG channel selection is a significant area of research. It is inferred from past studies that designing computationally efficient and highly discriminable feature extraction methodology is a challenging task in BCI system design. Therefore, developing a robust EEG channel selection criterion independent of neuro mechanism and mental tasks selection is highly essential for the better computational performance of BCI system.

(c) Efficient feature estimation and multi-class classification of EEG activity

The BCI has a wide range of applications due to which it has gained noteworthy momentum over the past few years. It is inferred from existing studies that developing a computationally efficient and highly discriminable feature extraction methodology is still a challenging task in BCI implementation [17]. In the past, researchers have suggested and implemented several EEG feature extraction techniques including time, frequency, and time-frequency domain-based features. Some of these available techniques are efficient in achieving high classification accuracy but they suffer from increased computational requirements, while the remaining techniques with low computational requirements face the challenge of achieving high classification performance. Therefore, it is important to develop a new computationally efficient feature extraction and classification methodology using advanced signal processing tools for better performance of BCI system.

The above-discussed challenges have motivated me to work on the development of an efficient methodology to classify multi-class BCI-EEG activity by achieving highly discriminating features. The methodologies that have been proposed in this work address the discussed challenges. The detail of each developed methodology is described in subsequent chapters thoroughly.

1.3 Aim of the Present Work

In the context of the research gaps discussed in section 1.2 of this chapter, the aim of this study is defined as follows:

1. Development of a methodology for detection and correction of EEG artifacts.
2. To propose an optimum channel selection criterion and evaluate the performance of the BCI system for selected EEG channels.
3. To develop and validate an effective feature generation and classification methodology for multi-class BCI system implementation.

1.4 Organization of the Thesis

This thesis is structured into six chapters delineated below:

Chapter 1 includes a brief description of the BCI system and its applications, and also the various types of conventional BCI systems. Further, the motivation that incited me to work on the proposal of feature extraction and classification techniques in BCI system design is explained.

Chapter 2 focused on the critical review of state-of-art methods in the area of BCI system design for different applications. The employed approach, extracted performance evaluation parameters, results and findings are discussed in this chapter. This chapter outlines a brief description of the datasets utilized in the present thesis. A detailed description of simulated EEG activity and experimental EEG activity utilized for artifact removal objective is provided in this chapter. Additionally, multiclass motor imagery BCI EEG activity dataset utilized for channel selection, feature extraction, and classification objectives viz. BCI Competition III 3a and BCI Competition IV 2a is also provided.

Chapter 3 describes the proposed methodology of EEG artifact removal, developed using a joint application of Fast-Power Independent Component Analysis and General Linear Chirplet Transform. In this work, the artifactual Independent Components are automatically identified from segregated Independent Components using the proposed Katz-Fractal Sparsity criterion. Further, the GLCT-based soft thresholding technique is implemented to treat the identified artifactual components and recover useful cerebral information. The Inverse Independent

Component Analysis yields artifact corrected clean EEG activity for further analysis. The effectiveness of the proposed methodology is validated with simulated and empirical EEG dataset. The experimental results establish the proposed method as a potential candidate for non-cerebral artifacts correction and noise suppression from EEG records.

Chapter 4 determines the potential of the proposed channel selection and feature optimization method on multi-channel MI-EEG datasets. Chapter-4 proposes a novel channel selection and features optimization methodology to enhance the multi-class BCI system's performance. First, the multi-channel EEG dataset is reduced to an optimum number of channels subsets using developed MDA-SOGWO based EEG channel selection criterion. Later, the discriminable feature-set is generated from the time, frequency, and FAWT based time-frequency domain of the EEG dataset. Then, an informative feature set is constructed from extracted features through CCA-RFE based feature selection criterion. Eventually, training and validation are carried out for ELM, LDA, RF, and MLP classifiers using the selected feature-set. The proposed methodology is evaluated on multi-class MI datasets of BCI Competition IV 2a and BCI Competition III 3a, yielding classification accuracy of 85.65% and 95.15%, respectively.

Chapter 5 describes the developed hybrid ensemble voting classifier based AttentionNet model for EEG-based MI-BCI activity classification. The Time-Frequency Representation (TFR) of the multi-class EEG activity is generated using Transient Extracting Transform (TET). The TFR spectrogram images are input to the designed ensemble voting classifier based AttentionNet model for classification purposes. The model is evaluated using different fusion approaches, viz. feature-level and score-level fusion of layers. The proposed model is evaluated on two MI-BCI datasets, namely, BCI competition IV 2a and BCI competition III 3a, achieving the highest classification accuracy of 88.14 % and 93.13 %, respectively. The classification results obtained on a large multi-class MI-BCI dataset confirm that the proposed hybrid ensemble voting classifier based AttentionNet model significantly outperforms the conventional algorithm and achieves a remarkable classification accuracy for the feature-level fusion of layers. The developed framework aids in identifying different tasks for multi-class MI-BCI EEG activity.

Chapter 6 summarizes the work presented in the thesis and emphasizes the future prospects of this work in the field of BCI and computer vision.

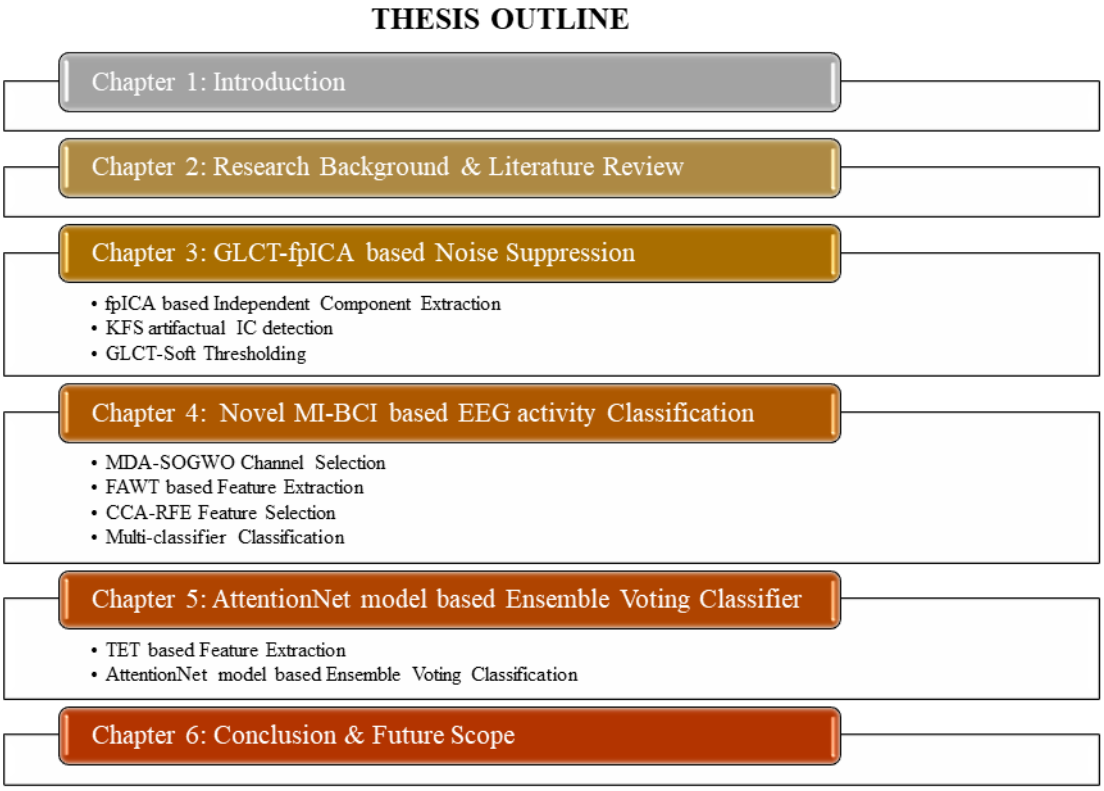


Figure 1.2 Outline of the thesis

CHAPTER 2

Research Background & Literature Review

Preface

In this chapter, the innovations made by researchers and scientists in the area of the Brain Computer Interface system to solve the problems faced during EEG activity processing stages for multi-class MI-BCI system are discussed in detail. This chapter is structures into sections dedicated to pre-processing, channel selection, feature extraction, and classification. The results achieved by state of the-art techniques in MI-BCI-based EEG activity classification have been discussed for reference and evaluation.

2.1 Introduction

A German neuropsychiatrist, Hans Berger, recorded the first human brain electrical signals (EEG) in 1924 and published the research in 1929 [18]. In 1934, Fisher and Lowenback first demonstrated Epileptiform spikes [18]. Thereafter in 1947, the American EEG society was founded [19]. Brain communication was explored by the Defence Advanced Research Projects Agency of the USA in 1970 [19]. Professor Jacques J. Vidal from the University of California, Los Angeles invented the term Brain Computer Interface in 1976 [20]. Within a few years of BCI's invention, several research groups worked on the analysis of neural electrical signals for their use in controlling machines which further led to significant growth in the BCI research domain [20]. The first high-quality brain signal generation procedure was performed in 1998 invasively. Next year, Chapin *et al.* [21] recorded the electrical activity of a monkey invasively to control the displacement of a cursor on a computer screen. The first BCI related game was developed and released to the public in 2003, followed by robotic arm control using a monkey brain in 2005 [22].

The human brain's electrical activity, EEG, plays a substantial role in the enhancement of BCI research. Electroencephalography (EEG) is the electrical records of the human brain acquired by following the international 10-20 system of electrode placement as shown in Figure 2.1. The EEG activity can be acquired corresponding to various neuromechanisms depending upon the type of BCI system development viz P300, Visually Evoked Potentials (VEPs), and Sensorimotor activity. P300 is a kind of evoked potential in which a positive peak appears after 300ms of significant auditory or visual stimuli [23]. Farwell and Donchin [25] developed the first P300 potential based BCI system to communicate through a computer [26], whereas VEP is a variation in brain activity of the occipital region due to the appearance of visual stimuli. It is also known as Steady-State Visually Evoked Potential (SSVEP) and occurs when the subject perceives periodic stimulus [24]. The Sensorimotor activity of the human brain occurs due to oscillations in the sensorimotor cortex of the human brain. These oscillations occur in terms of two major phenomena, Event-Related Synchronization (ERS)/ Event-Related Desynchronisation (ERD) which occurs due to an increase/ decrease in the spectral power respectively [25].

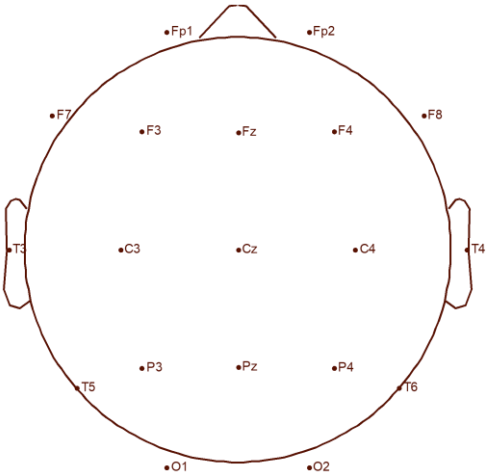


Figure 2.1 Position of 19-EEG electrodes following the International 10-20 system

2.1.1 Electroencephalography Rhythms

The information concerning the functional state of the human brain is majorly contained by five frequency bands categorized as different brain waves. Frequency bands are delta band (0.1-4 Hz), theta band (4-8 Hz), alpha band (8-13 Hz), beta band (13-25 Hz) and gamma band (26-50 Hz) [26]. The significance of each frequency rhythm is as follows:

a) Delta Band (0.1-4 Hz): The delta waves are majorly observed during a deep sleep state (3^{rd} or 4^{th} stages of sleep). The delta band are located majorly in the parietal lobes [27].
b) Theta Band (4-8 Hz): The theta waves are detected while the subject is awake and is in the deepest meditation state. It appears in the parietal and temporal lobes[28], [29].
c) Alpha Band (8-13 Hz): Alpha rhythms are observed during dreams or relaxed brain state [29]. This rhythm dominates in normal relaxed adults (>13 years of age). The alpha rhythm occurs in the occipital and parietal lobes of the human brain.
d) Beta Band (13-25 Hz): The beta waves are detected while the subject is awake and is walking attentively [30].
e) Gamma Band (26-50 Hz): The gamma rhythm is observed during the motor activity and cognitive state of the human brain. The decision-making situation triggers the gamma waves to appear [31].

2.2 Pre-processing

In BCI system design, pre-processing, an artifact removal stage preceding the EEG activity analysis, acts as an essential procedure to extract cerebral activity from artifactual EEG signals. Initially, only preventive measures

were taken during EEG recording which include removal of ocular and muscular artifacts by constraining eye movements, eye blinking and movement of other body parts [32]. Thereafter, the noisy EEG activity was directly eliminated using thresholding before processing EEG activity to reduce its effect on the study [33]. This method was time-consuming and did not assure the complete rejection of artifacts. While the thresholding was intended to remove only noisy activity, it also resulted in elimination of informative EEG activity along with that contaminated by noise [34]. Therefore, it became necessary to devise methods that remove the non-cerebral sources from EEG activity instead of eliminating EEG activity.

Various techniques were proposed to remove non-cerebral sources from EEG activity viz. filtering techniques [35], Time-Frequency based Blind Source Separation (BSS) techniques [36], and Deep-learning techniques [14]. The brief description of commonly used EEG artifact removal techniques is as follows:

(a) Filtering Techniques: Basic filtering techniques such as low-pass filter, and band-pass filter were utilized initially to remove EEG activity outside a particular frequency band [37]. As the EEG activity is non-stationary and non-linear, these techniques does not perform well on it. Therefore, Spherical Surface Laplacian (SSL) filtering [35] methods such as Current Source Density (CSD), SSL using a realistic head model (SSR), and Finite Difference Method (SSF) were utilized for EEG artifacts removal [38]. But these filtering techniques requires high-density montage, which increases preparation time and computational complexity of the system. Further, eye-blink artifact detection was performed by utilizing a Multichannel Weiner Filter (MWF) along with a subset of frontal electrodes. The estimated eye-blink artifacts using the MWF criterion were further subtracted from original EEG activity to achieve artifact-free EEG records [32]. Furthermore, a hybrid methodology of Empirical Mode Decomposition (EMD) and Canonical Correlation Analysis (CCA) for unsupervised eyeblink artifacts detection in EEG activity was developed [39]. Although these techniques efficiently removed eye-blink artifacts, computational complexity was high.

(b) **Independent Component Analysis (ICA):** ICA techniques are utilized to separate mixed signals into their source signals known as independent components. Initially, various variants of ICA were used to separate noise signals from EEG activity such as Fast-ICA [40], WASOBI ICA [41], Kurtosis-ICA [42]. Thereafter, the ICA techniques were combined with other transformation criteria to remove artifacts without EEG activity information loss. For instance, Bono *et al.* [38] 2016, explored the performance of two hybrid methodologies for the removal of EEG artifacts. The two-hybrid algorithms Wavelet Packet Transform followed by Empirical Mode Decomposition (WPTEMD) and were Wavelet Packet Transform followed by Independent Component Analysis (WPTICA). Also, Chen *et al.* [44], employed ICA based second-order and higher-order statistics extraction. Initial ICA based artifact correction methods require expertise in artifactual ICs selection, which make these methods inappropriate for real-time BCI applications. Although, thresholding criterion for artifactual ICs suppression was proposed in

successive studies, automatic artifactual ICs identification and threshold value selection is still a challenging task in real-time BCI applications [14]. In addition, entire artifactual IC rejection causes loss of useful cerebral information, which is erroneously leaked to the rejected artifactual ICs [15].

(c) Deep Learning: In the above defined techniques, an appropriate threshold value selection and sensitivity to predefined basis functions (viz. mother wavelet function) restrict the real-time performance. Therefore, various automated pre-processing techniques were required to select relevant thresholding value automatically. The most popular ones include a Generative Adversarial Network (GAN) based ocular artifact removal framework named as EEGANet [14], [45], and Deep Learning Network (DLN) to remove Ocular Artifacts (OA) [14]. Also, a 1-D convolutional neural network was developed including a hybrid Spider Monkey-Electric Fish Optimization (SM-EFO) criterion [46]. The developed criterion was tested on a benchmark dataset affected by ECG, EMG, and EOG artifacts.

2.3 Channel Selection

Researchers have developed various channel selection methods before further EEG activity processing to utilize optimum EEG channels during feature extraction and classification stages. Some utilized single-channel EEG activity [47]. However, considering a single channel in EEG activity analysis leads to scarcity of information. Therefore, other studies included C3, C4, and Cz as optimum EEG channels and utilized them for further processing [2]. In subsequent studies, the EEG channels used to acquire data varied in a large range. There are most of the studies which used 1-20 EEG channels , whereas some other studies used 21-40 EEG channels [2] There are five studies which used more than 100 EEG channels [48]. From these studies, it was observed that there is a significant increase in sensitivity and specificity when number of channels are increased till 22, but a further increase is not advantageous. A large number of EEG channels during data processing adds computational complexity to the system. Therefore, selecting appropriate channels would reduce computational complexity and optimize the system performance.

Initially, channel elimination was opted to include efficient channels by utilizing an optimization criterion such as recursive feature elimination [49], genetic algorithm [50]. For instance, Lal *et al.* [49] 2004, adapted recursive feature elimination – support-vector based EEG activity channel selection method. Yang *et al.* [50] 2012, proposed Genetic Neural Mathematic Method (GNMM) employing Genetic Algorithm (GA) for the input selection process and Multi-Layer Perceptron (MLP) classifier for fittest chromosome selection. Further, Handiru *et al.* [51] 2016, developed an Iterative Multiobjective Optimization for Channel Selection (IMOCS) criterion in which channel weights were assigned based on achieved classification accuracy using reference point based optimization method. The optimization approach considered C3 and C4 as initial candidates for left-hand and right-hand MI EEG activity classification, respectively. The proposed methodology of IMOCS and SVM classification achieved an average accuracy of 80% evaluated on 35 best performing subjects.

In recent studies, researchers have proposed various channel selection methodologies using variants of Common Spatial Patterns (CSP) viz. regularized CSP [52], Common Spatial Subspace Decomposition (CSSD) [53]. Some other channel selection techniques include covariance matrix-based channel selection [54], Bispectrum-based Channel Selection (BCS) methodology [55], Spatiotemporal Filtering based Channel Selection (STECS) method [56], Dynamic Channel Relevance (DCR) score extraction [57] and deep learning approach [58]. Besides, researchers also suggested the subject independent channel-set selection criterion. For instance, *Wu et al.* [53] suggested Spatiotemporal-filtering based channel selection (STECS) method for optimizing EEG channels prior to EEG activity analysis. Park *et al.* [59] computed the covariance matrix for channel selection and extracted Filter Bank Common Spatial Pattern Features (FBCSP) for BCI activity classification. However, such methods mostly concentrated on Subject Independent-BCI (SI-BCI) EEG channel-set selection, which reduces calibration time [2]. In a few studies, researchers considered inter-subject variability as an important factor to tune BCI explicitly for each subject and suggested opting subject-specific channel selection methodologies while proposing BCI systems [2]. Fahimi *et al.* [60] implemented a Convolutional Neural Network (CNN) based single-channel EEG activity analysis technique considering inter-subject transfer learning. While, Handiru *et al.* [51] proposed a cross-subject generalisation-based channel selection criterion for this purpose using IMOCS criterion with an average classification accuracy of 80%. The time and frequency domain features does not accurately reflect changes in power of oscillatory EEG, therefore frequency domain based features are required. Jin *et al.* [61] developed a bi-spectrum based channel selection method based on non-linear and non-gaussian information extraction from EEG activity. The extracted features provided frequency domain information of EEG activity. Thereafter, the sum of logarithmic amplitudes followed by first-order spectral moment characteristics are obtained from frequency domain information to remove redundant channels from all EEG channels [61]. The classification accuracy of 86.3% was achieved using bi-spectrum based channel selection method. The removal of redundant channels showed improvement in accuracy but still there is a need to enhance classification accuracy along with reduced computational complexity.

2.4 Feature Extraction

The next step of EEG activity processing includes the feature extraction stage. Feature extraction is a process of extracting useful information from electrical signals (EEG activity) in the form of feature vector. For good classification performance, and thus efficient BCI design, it is essential that the extracted features are appropriate and relevant. Various feature extraction methods have been stated in the literature for BCI system development broadly classified as time-domain features, frequency-domain features, and time-frequency domain features.

a) Time-domain features: The features which obtain temporal information from EEG activity are categorized as time-domain features. The time domain analysis refers to extraction of features such as mean, variance, standard deviation, Hjorth parameters, entropy, kurtosis, skewness directly from

recorded EEG time-series [4], [16]. The time-domain features have low computational complexity but these features are weak in analysing non-stationary signals like EEG. Also, signal noise adversely impact these features, therefore noise removal is necessary before considering these features.

b) Frequency-domain features: The features containing spectral information of the EEG activity are called frequency-domain features. Frequency domain features may comprise of spectral edge frequency, spectral power, and PSD, the feature computed from the transformed frequency domain of the EEG time-series [18], [19]. The features such as mean frequency, median frequency, spectral centroid, spectral flatness, and spectral crest also comes under this category [65], [66]. Additionally, researchers have extracted multivariate Auto-Regressive (AR) model [69], analytic Intrinsic Mode Function (IMF) [70] [71], and covariance matrix [72] to extract highly discriminable features for the classification task. The frequency-domain features have good spectral estimation which leads to better classification results but these features loose the temporal information from the signal. Also, frequency-domain features are mathematically more complex as compared to time-domain features.

c) Time-Frequency domain features: These features contain information from both temporal and spectral domains simultaneously. First, the EEG signal is transformed into Time-Frequency Representation (TFR) using various transformation techniques such as Short-Time Fourier Transform (STFT), Stockwell Transform (ST), and Wavelet Transform (WT) [73]. Time-Frequency Representation (TFR) based feature extraction criteria map EEG records into higher dimensional space for analysis purposes [13], [49]. To overcome the drawback of time-domain and frequency-domain-based feature extraction techniques, numerous methods have been proposed in the literature. For instance, Robinson *et al*. [75] developed a wavelet-based CSP feature extraction methodology for motor imagery task classification. So far, Siuely *et al*. [51] proposed Cross-Correlation (CC) feature extraction technique for motor imaginary signal detection. Song and Yoon [52] computed Renyi's entropy and Sample entropy features for the implementation of BCI system. Yang et al. [2], implemented Fisher's Discriminant Analysis (FDA) based f-score to concurrently optimize frequency band and time segment corresponding to multi-class classification. Some other developed TFR techniques include Linear Prediction QR decomposition (LP_QR), Sparse Group Representation Model (SGRM), and fuzzy multiclass CSP-based feature extraction [78], [79]. The time-frequency domain features preserves both temporal and spectral information and performs well for non-stationary activity but at the cost of increased computational complexity.

2.5 Feature Classification

Once the features are extracted from recorded EEG activity, the classification stage utilizes them to identify motor imagery signals. Various classification algorithms exist in the literature, e.g., Linear Discriminant Analysis

(LDA), Naïve Bayes (NB), Support Vector Machines (SVM), Artificial Neural Networks (ANN), decision trees, random forest, fuzzy logic, and K-Nearest Neighbor (KNN) [34], [44]. Support Vector Machine (SVM) classifier is the most frequently employed in biomedical applications because of their ability to consider a large number of predictors [46]. Abe et al. [46] employed SVM using Gaussian kernel function to classify spatial filtering features. Vijean et al. [48] proposed ST-based feature extraction methodology and used KNN as well as LDA classifier for mental task detection. Nguyen *et al.* [49], introduced two-stage based multi-class BCI data classification which consisted of feature extraction and fuzzy classifier stages. Zhang et al. [50], tested the effect of multi-kernel Extreme Learning Machine (ELM) for two kernel functions (i.e., Polynomial kernel and gaussian kernel).

In recent years, the deep-learning-based MI-BCI EEG activity identification gained significant attention for feature extraction and classification as it does not require human interventions and the approach can be truly automatic and subject-independent. For instance, Zhao et al. [26] proposed a 3-D CNN model for EEG-based MI-BCI task classification. Ha and Jeong [27] proposed STFT based CapsuleNet CNN network, which proved to be a good model for MI-BCI task classification. However, it is limited to analysis of binary MI tasks while considering only three EEG channels for BCI system design. Schirmeister et al. [28] developed a CNN model including deep conv and shallow conv, wherein 1-D convolution layers are employed for effective information extraction. Lawhern et al. [8] developed EEGNet based deep learning criterion for EEG task classification, which contributed as a significant improvement over other deep conv and shallow conv models. Deng et al. [15] suggested a Temporal-constrained Group Lasso-based EEGNet (TSGL-EEGNet) network, an advancement of EEGNet for multi-class MI-BCI-based EEG activity classification. These CNN techniques achieved better classification accuracy but the computational complexity is very high in these techniques, therefore significant feature extraction from noisy EEG activity via performing optimum steps for improved classification accuracy of MI-BCI tasks remains a challenge.

2.6 Inference Drawn

It has been observed from the literature survey (section 2.2(b)) that several state-of-the-art pre-processing techniques utilized thresholding criteria based on Time-Frequency Representation (TFR) for EEG activity cleaning. However, the proposed techniques require human intervention due to which automated detection and correction of noisy EEG activity is not possible. The literature on channel selection studies (section 2.3) shows that much work is still going on related to the development of a robust EEG channel-set selection criterion to balance the trade-off between computational complexity and information scarcity. The existing feature extraction and classification techniques suffer from computational inefficiency and low classification accuracy problem (section 2.5). Therefore, there is a scope in the development of highly discriminable feature extraction and computationally efficient classification methodology. Also, it can be inferred from section 2.5 that most of the studies focussed on binary class MI-BCI system and classification algorithms struggle to produce an increased classification performance for the multi-class MI-BCI system. To address these challenges, a modified and hybrid

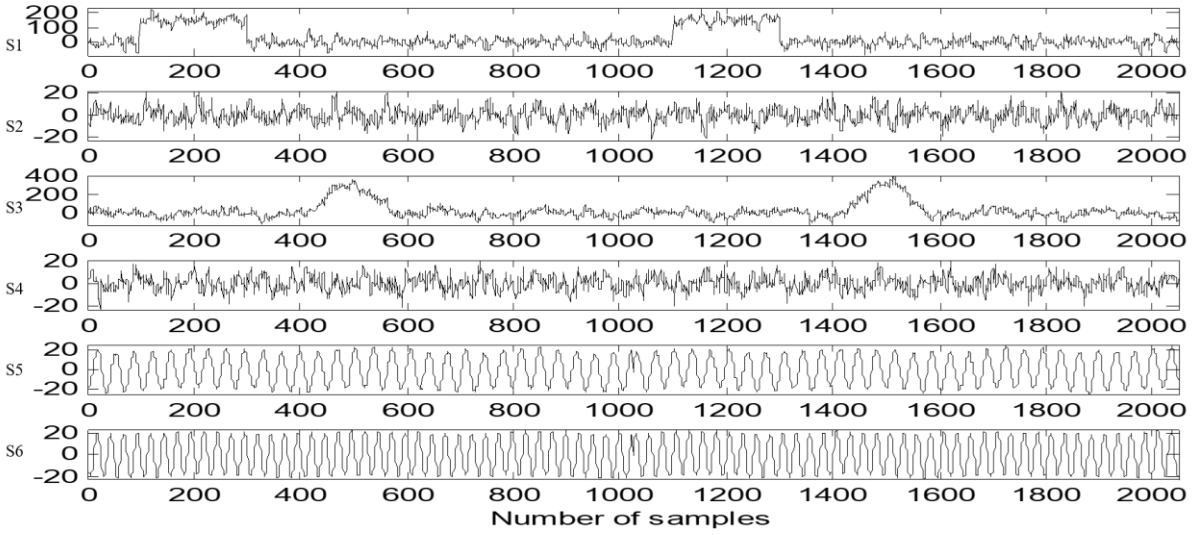
strategy is required for efficient artifact removal, feature extraction, and feature classification of multi-class MI-BCI based EEG activity. The mentioned challenges acted as a base for the evolution of this thesis and motivated us to work on the development of improved MI-BCI system design.

2.7 Artifact Removal Dataset

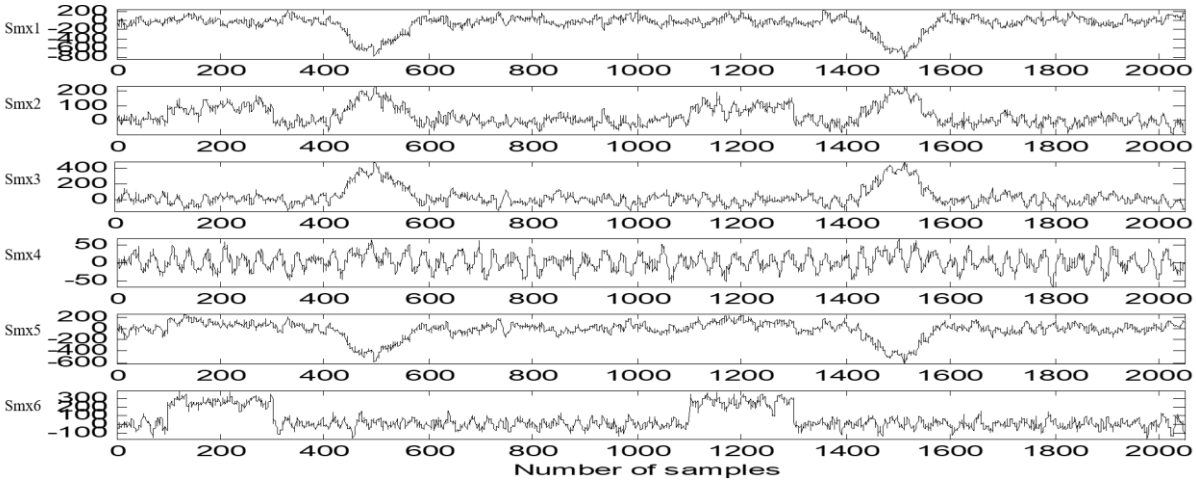
For the artifact detection and removal, simulated (*6*-channel EEG activity) and experimentally recorded (*19*-channel EEG activity) EEG activity is exploited to quantitatively evaluate the potency of proposed GLCT-PICA de-noising method in EEG artifact removal. This section briefly explains the simulated and experimentally recorded EEG datasets.

2.7.1 Simulated EEG activity

The simulated EEG-like activity helps in effectively evaluating process outcomes and hones procedural understanding of proposed GLCT-PICA de-noising method. The EEG-like sources are simulated using MATLAB toolbox developed by Yeung *et al.* [110]. In the MATLAB toolbox, function peak.m generates ocular artifact sources and noise.m generates EEG like random (cerebral) sources. In this dataset, the sources of EEG activity are simulated following work of Akhtar *et al.* [106] during which sampling rate (f_s) is fixed at *256 Hz* (f_s is consistent with experimentally available EEG dataset). In order to evaluate the efficacy of proposed GLCT-PICA EEG de-noising method, *40* sets of *6*-channel EEG activity are simulated and de-noising is performed. Figure 2.2(a) represents a typical set of realized/simulated EEG sources (S1 to S6) of ocular artifactual activity and cerebral EEG activity. In Figure 2.2(a), source S1 signifies leaked cerebral activity with eye-movement ocular artifact generated using low frequency square wave, sources S2 and S4 signify random/informative cerebral activity, source S3 signify leaked cerebral activity with ocular eye-blink artifacts, and sources S5 and S6 signify rhythmic cerebral activity of *12 Hz* and *30 Hz* frequency, respectively. In order to attain *6*-channel contaminated EEG data, the simulated source signals are multiplied with random mixing matrix $\boldsymbol{C}$ $[6 \times 6]$. Figure 2.2(b) shows *6*-channel simulated noisy scalp EEG data (Smx-1 to Smx-6) yielded after mixing of (S1 to S6) resented in Figure 2.2(a).



(a)



(b)

Figure 2.2 (a) A typical set simulated EEG sources (S1 to S6) **(b)** A typical set of 6 channel EEG activity obtained after random mixing of the sources Smx-1 to Smx-6

2.7.2 Experimentally recorded EEG activity

In this dataset, 19-channel experimentally recorded scalp EEG data is employed for evaluating the efficacy of proposed EEG artifact removal methodology. Castellanos and Markov (2006) [111] reported the performance of Wavelet enhanced ICA (wICA) based EEG de-noising technique on this dataset, hence, the dataset is suitable for comparing performance of proposed methodology with wICA based EEG de-noising technique [112]. The dataset was recorded following International *10-20* system of electrode placement from healthy subject under eyes open condition. Further, data was digitized at *256 Hz* and filtered using notch (*50 Hz*) and band-pass filters (*4-45 Hz*). A sample of filtered EEG activity $\hat{d}(t)$ processed in this work is shown in Figure2.3. It is observed from Figure

2.3 that experimentally recorded EEG activity suffers from eye blink artifact, high-frequency artifacts and other disturbances.

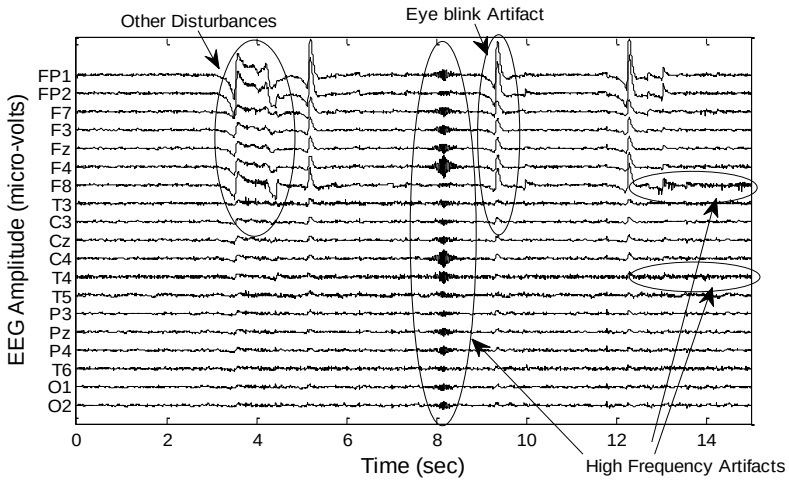


Figure 2.3 Filtered EEG (*19*-channel) activity $\hat{d}(t)$ (length of the segment = *14.5 s*)

2.7.3 Multiclass MI-BCI Dataset

In this work, publicly available BCI Competition datasets viz. BCI competition IV 2a and BCI competition III 3a datasets are examined for EEG based channel selection, feature extraction and feature selection criterion. The datasets utilized consists of four class MI-BCI tasks such as left-hand, right-hand, tongue and feet movement.

2.7.3.1 BCI Competition IV 2a Data

This dataset is composed of EEG recording of four class MI movements (i.e., left hand, right hand, tongue, and foot movement) of *9* subjects ("*A01T*", "*A02T*" …., "*A09T*") each with *25* channels. Out of *25* channels, *22* channels were Ag/AgCl electrodes which followed *10-20* system of electrode placement, and *3* channels were monopolar EOG channels. The data recording was carried out in two sessions' viz. calibration session (without feedback) and evaluation session (with feedback on different days). Each session comprises *72* trials per class resulting in *288* trials per session. The sampling frequency during EEG recording was *250* Hz. The recorded EEG data were bandpass filtered from *0.5 Hz* to *100 Hz* and a *50 Hz* notch filter was applied to eliminate line noise. The paradigm of trials is as follows: at the beginning, a fixation cross was displayed for *2 sec*. After that, a cue in the form of an arrow pointing left, right, up and down was displayed for *1.25 sec*. During which the subjects were asked to perform motor imagery task until a fixation cross appears on screen for *6 sec*. Figure 2.4(a) shows the timing paradigm of the BCI Competition IV 2a dataset. In present work, out of *6 sec* paradigms, *2 sec* data obtained after visual cue onset (from *2.5 sec* to *4.5 sec* duration) is considered for further processing as followed by BCI Competition IV winner [124][125].

2.7.3.2 BCI Competition III 3a Data

BCI competition III 3a is a *4*-class MI-BCI EEG dataset recorded from three subjects' ("*K3b*", "*K6b*" and "*L1b*") using *64* channels Neuroscan EEG system. In this dataset, the left and right mastoids were fixed as reference and ground respectively. The sampling rate of the dataset was *250 Hz* and filtered using a bandpass filter of frequency range *1 Hz* to *50 Hz*. The recording paradigm includes a blank screen for *2 sec* followed by a beep and fixation cross for *1 sec*. After that, at *3rd sec*, a visual cue of an arrow indicating left, right, up, and down was displayed for *1 sec*. Also, the subjects were asked to imagine *4*-class MI movement corresponding to visual cue until the cross disappeared. The timing paradigm corresponding to MI activity is shown in Figure 2.4(b). The recorded EEG activity from subject *K3b* includes *9* runs yielding a total of *360* trials, whereas EEG activity of subjects *K6b* and *L1b* includes *6* runs yielding a total of *240* trials. Sample EEG activity record corresponding to Left-hand, Right-hand, Tongue and Feet movement MI tasks is shown in Figure 2.5.

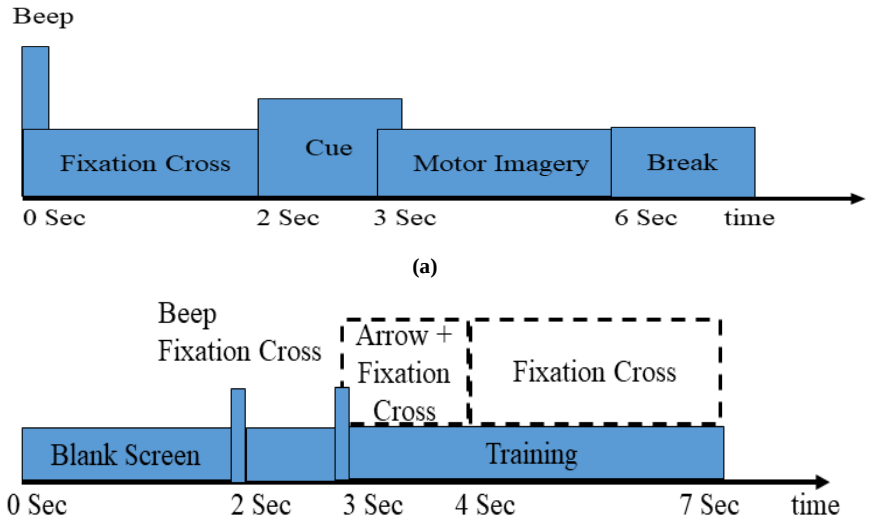


Figure 2.4 Timing Paradigm of (a) BCI Competition IV 2a dataset (b) BCI Competition III 3a dataset

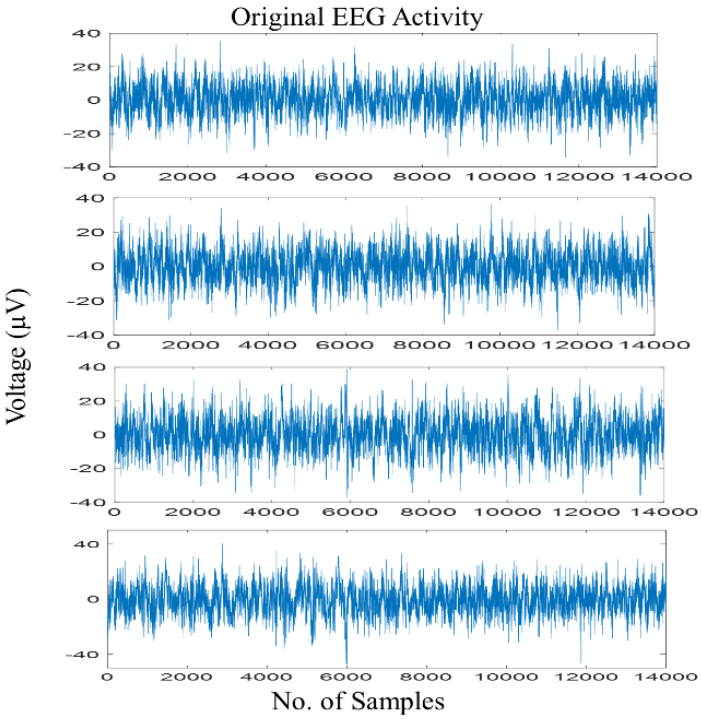


Figure 2.5 Sample EEG activity record corresponding to Left-hand, Right-hand, Tongue and Feet movement MI tasks

2.8 Summary

To achieve an enhanced classification accuracy for MI-BCI task classification, proficiency in each processing step is required. Here, the development of an efficient multi-class MI-BCI system depends on the integral stages of EEG activity processing namely, pre-processing, channel selection, feature extraction, and feature classification. The past studies focused especially on the effective development of the binary class MI-BCI system. Whereas, multi-class MI-BCI system design has not yet been explored much in the literature. The above-mentioned postulates regarding the EEG activity processing stages prompted me to for the development of a remarkable methodology for multi-class EEG activity classification in order to take a step ahead in assisting patients suffering from severe motor impairments.

CHAPTER 3

Hybrid GLCT-PICA de-noising for EEG Artifact Removal

Preface

In this chapter, a novel artifact removal methodology is developed using a joint application of fpICA and GLCT for automatic identification and rejection of artifactual origins. After segregating ongoing Electroencephalogram activity into Independent Components, the Katz-Fractal Sparsity criterion is employed to identify artifactual components. The identified artifactual components are treated by the GLCT-based EEG de-noising method to recover useful cerebral information leaked with artifactual origins. Thereafter, Inverse Independent Component Analysis yields artifact corrected clean Electroencephalogram activity for further analysis.

3.1 Introduction

The pre-processing stage is responsible for potential artifact removal and noise suppression prior to information extraction from raw EEG activity. Electroencephalogram activity is weak and non-stationary in nature, also its dynamics are prone to contaminations from non-cerebral artifactual origins and noises. Non-cerebral biological artifacts originating from Electrocardiogram (ECG), muscle and ocular activity encompass significantly high amplitude (around *10* times) when compared to cerebral activity and modify the nature of useful cerebral recordings [6]. In addition to non-cerebral sources, electrical disturbances engendered from instruments, electrical circuits, and external electromagnetic fields affect the cortical EEG activity adversely. Mixing of artifactual origins with on-going EEG recordings poses a serious problem of interpretation and a challenge of authentic statistical analysis in any EEG-based system design[7]. Therefore, it is important to eliminate artifacts and noise from the raw EEG activity before analysis and estimation of cerebral signatures

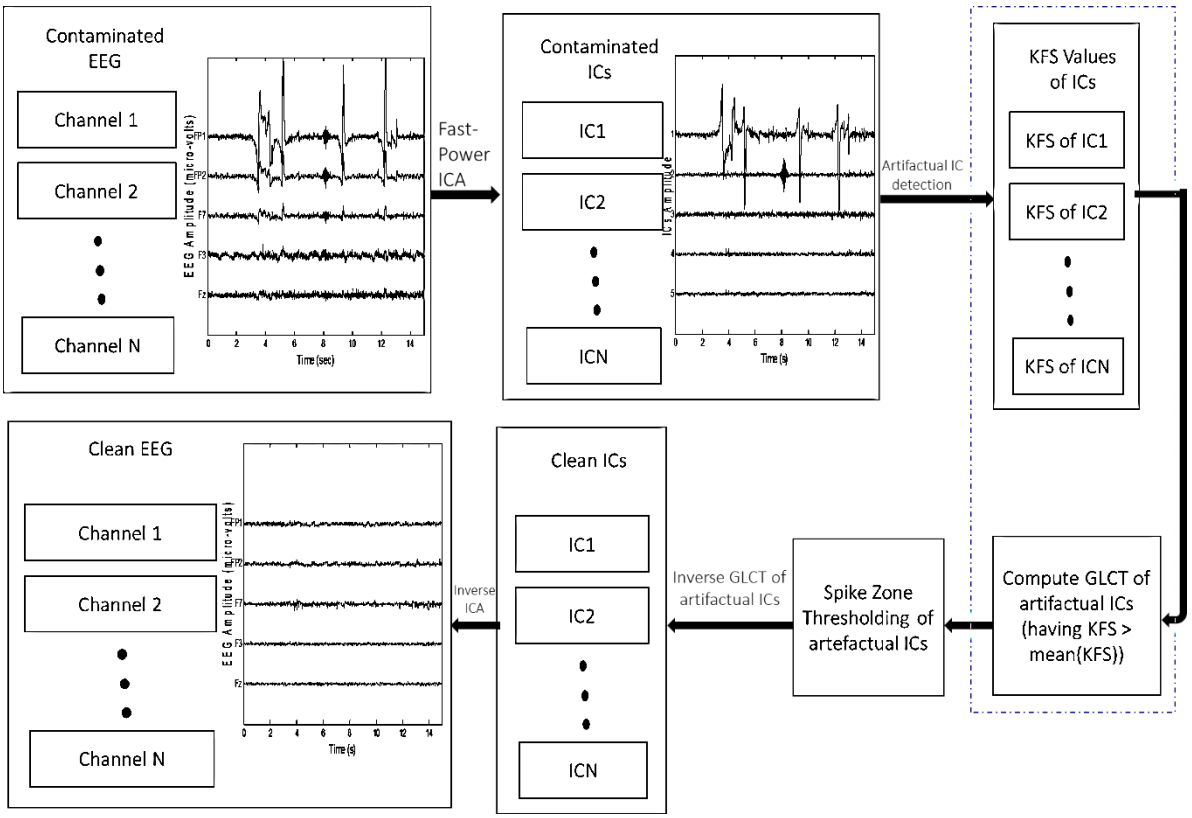


Figure 3.1 Block diagram of proposed GLCT-PICA based artifacts removal method

Considering the limitations of previously proposed EEG artifact removal methods (as described in chapter 2), this work aims to develop a novel hybrid TFR-ICA based artifact removal method combining Fast-Power Independent Component Analysis (fpICA) and General Linear Chirplet Transform (GLCT). The proposed methodology

consists of four procedural stages. In the first stage, EEG activity is decomposed into a number of Independent Components (ICs) consisting of artifactual and cerebral sources using fpICA algorithm. Fast-Power Independent Component Analysis (fpICA) algorithm ensures efficient and fast decomposition of recorded EEG activity. Thereafter, the criterion of Katz-Fractal Sparsity (KFS) is utilized to extract artifactual ICs in the second stage. The third stage comprises of Time-Frequency Representation (TFR) of identified artifactual ICs using General Linear Chirplet Transform (GLCT) and subsequently, soft thresholding of GLCT coefficients to attain artifact-free ICs. Finally, the clean EEG activity is obtained after performing inverse GLCT and inverse ICA in the fourth stage. The advantage of the proposed work lies in its capability of reducing uncertainties of the measurement process and enhancing the reliability of the system by improving the accuracy of subsequent signal analysis. The proficiency of the proposed artifact rejection and noise suppression method is verified on both simulated and real contaminated EEG data. The methodology is proven to be very effective in ocular artifact removal in the present work. The reliability and performance of the proposed work are compared with two established EEG correction methods viz. IC rejection-based EEG correction and wICA based EEG correction methods in terms of Power Spectral Densit, Root Mean Square Error, and Log Spectral Distance. Figure 3.1 represents the block diagram of GLCT-PICA (General Linear Chirplet Transform-Fast-Power Independent Component Analysis (fpICA)) based EEG de-noising method proposed in this work.

3.1.1 ICA based EEG de-noising techniques

Blind Source Separation (BSS) techniques decompose randomly mixed EEG signals into underlying source components known as Independent Components (ICs) without any physical model [4]. Independent Component Analysis (ICA) is a BSS technique to recover statistically independent components from multidimensional random vector without prior knowledge of mixing process. The postulate behind EEG signal processing techniques is that the signals are a linear mixture of the source components. The BSS model attempts to solve Equation *3.1* using ICA techniques [5], [15]:

$$u(t) = C.d(t) + z(t) \qquad (3.1)$$

where $d(t) = [d_1(t), d_2(t), d_3(t) \ldots . d_j(t)]^T$ represents *j* number of unknown independent sources (i.e., sources of EEG generation) linearly mixed to form multichannel EEG recording given as $u(t) = [u_1(t), u_2(t), u_3(t) \ldots . u_i(t)]^T$. i being mixing matrix of order $i \times j$ and $z(t) = [z_1(t), z_2(t), z_3(t) \ldots . z_i(t)]^T$is the additive noise received at the respective EEG electrodes. Considering Equation *3.1*, the ICA techniques attempt to obtain the de-mixing matrix W so as to estimate original sources as:

$$\tilde{d}(t) = W.u(t) \qquad (3.2)$$

Here, $\tilde{d}(t)$ represents recovered sources, and W is the de-mixing matrix. Once the de-mixing matrix is obtained, the mixing matrix C can be computed as $C = W^{-1}$, which helps in reconstructing clean EEG activity. The inverse of the matrix can be computed by various algorithms [102].

The recorded cerebral EEG activity $u(t)$ consists of electrical potentials corresponding to cerebral and non-cerebral artifactual sources. The ICs rejection-based EEG de-noising method works on the assumptions that the sources of cerebral and non-cerebral EEG activity are independent and are separable using ICA algorithms. Therefore, rejection of identified artifactual ICs helps in reconstructing clean EEG activity. In ICs rejection-based EEG de-noising method, the raw EEG activity is initially processed and filtered to remove baseline values, line noise and other potential low/high frequency noises. The processed EEG activity $u(t)$ is decomposed into ICs $d(t)$ representing cerebral and artifactual sources of EEG activity. Independent Components (ICs) representing artifactual origins are identified and rejected. The clean EEG activity is attained after performing inverse ICA on remaining ICs.

3.1.2 Wavelet enhanced ICA (wICA) based EEG de-noising technique

The Wavelet enhanced ICA (wICA) based EEG de-noising technique is the most frequently used technique in the recent past. Wavelet enhanced ICA (wICA) based EEG de-noising technique doesn't reject artifactual ICs completely, which reduces the risk of losing cerebral information leaked to the artifactual ICs. This technique causes less distortion and doesn't require additional EOG channel for ocular artifact rejection. Wavelet enhanced ICA (wICA) based EEG de-noising technique was originally proposed by Castellanos and Makarov [42] and its modified/advanced forms are available in [104].

In conventional wICA based EEG de-noising approach, the processed EEG activity is decomposed into ICs ($d(t)$) and artifactual ICs are marked after visual inspection of decomposed EEG activity. The identified non-cerebral ICs are transformed to Time-Frequency (TF) domain using suitable WT and mother wavelet functions. Thresholding and inverse transformation are performed on transformed wavelet coefficient to attain clean ICs. Thereafter, clean EEG activity is reconstructed from clean ICs after performing inverse ICA. The schematic diagram of wICA based EEG de-noising technique is presented in Figure 3.2.

The Wavelet enhanced ICA (wICA) based EEG de-noising technique has the capability to recover useful cerebral information present in the artifactual ICs. However, the limitation of this approach lies in the selection of optimum threshold value (τ), manual interference for artifactual ICs identification, and selection of the best base wavelet. In recent years, several threshold estimation algorithms have been proposed for the selection of optimum threshold value and an extensive survey is available in [34]. The most popularly used threshold estimation criterion

VisuShrink algorithm proposed by Donoho and Johnstone (given in Equation *3.3*) [105]. $\tau = \sigma_n\sqrt{2logN}$ *(3.3)*

Here, σ_n is the standard deviation of noise and N is the length of the data segment to be processed. Another widely accepted threshold estimation criterion was suggested in SureShrink algorithm (Stein Unbiased Risk Estimator), where the optimum threshold value is obtained by minimizing the Mean Squared Error (MSE) between the original signal and de-noised signal [1]. In recent years, various algorithms have been proposed for threshold estimation and extensive review exist [107]–[109]. Despite the fact that many studies are available for threshold (τ) selection, attaining an appropriate basis wavelet and threshold value remains still very challenging and is mostly application specific. In present work, two popularly used ICA based EEG de-noising methods (viz. ICs rejection-based EEG de-noising and wICA based EEG de-noising) have been implemented and a comparison is performed with proposed GLCT-PICA EEG de-noising method.

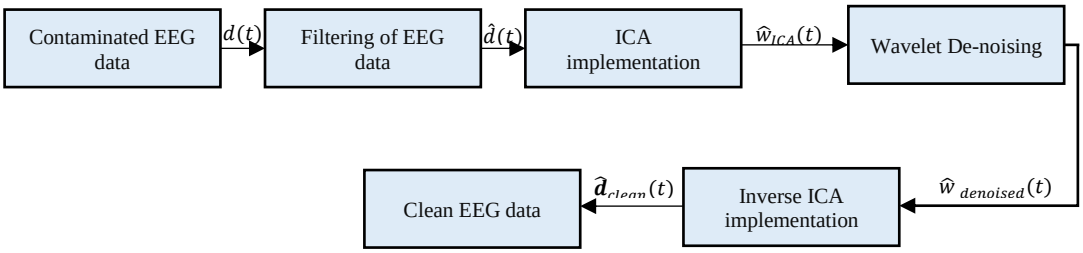


Figure 3.2 Schematic diagram of wavelet enhanced ICA (wICA) based EEG de-noising technique

In the proposed chapter, the ICs rejection based methodology and wICA based EEG denoising techniques are also implemented along with the proposed GLCT-PICA EEG de-noising method on simulated as well as experimental EEG data to analyse the difference between proposed methodology and these two available techniques efficiently.

3.2 Methods

This section provides a brief description of GLCT and fpICA methods and comprehensively illustrates GLCT-PICA based artifact correction methodology proposed in this work. Basics of GLCT and fpICA methods are explained in sub-section *3.2.1* and *3.2.2*, respectively. The proposed hybrid methodology of GLCT-PICA based EEG de-noising is presented in sub-section *3.3*.

3.2.1 General Linear Chirplet Transform (GLCT)

Signal processing techniques such as Time-Frequency Analysis (TFA) are the standard tools for characterizing time-varying nature of non-stationary EEG signals [113]. Considering the limitations of various Time-Frequency Representation (TFR) techniques, a recent TFR method viz. General Linear Chirplet Transform (GLCT) is used

to analyse transient behaviour of EEG signal in this chapter [114]. General Linear Chirplet Transform (GLCT) has ability to characterize EEG signal with high energy concentration and time-frequency resolution compared to other TFR techniques viz. Short Time Fourier Transform (STFT) [115], Synchro-squeezing Transform (SST)[116], General Parameterized Time-Frequency Analysis (GPTFA) Transform [117]. General Linear Chirplet Transform (GLCT) focuses upon distinct non-linear features of the time-varying signal and is capable of reconstructing essential component with less sensitivity to noise. Therefore, GLCT is considered a suitable tool for TFR of non-linear and non-stationary EEG signals. In GLCT method, the Instantaneous frequency (IF) of the signal is calculated from the derivative of the phase of the observed analytic signal. According to Ville's theory, an analytic signal with time varying IF given by Equation *3.4*.

$$g(t) = A(t).e^{j\int \emptyset(t)dt} \tag{3.4}$$

here, $A(t)$ is amplitude and $\emptyset(t)$ is IF of the signal. Considering short duration of time, it can be given as:

$$\emptyset(t)|t \in \tau = \varphi(\tau) = \varphi(t') + \varphi'(t')(\tau - t') \tag{3.5}$$

here, $\varphi(t')$ is IF at a specific time t' and $\varphi'(t')$ represents the first order derivative of IF at time instant t′. Given a window function $w(\tau - t')$ to truncate the signal $g(t)$ and after that Fourier Transform (FT) of $w(\tau - t').g(\tau)$ in IF $\varphi(t')$ is calculated as

$$|G(\varphi(t'))| = \left|\int_{-\infty}^{\infty} w(\tau - t').g(\tau).e^{-j\varphi(t')\tau}d\tau\right| \tag{3.6}$$

$$\left|\int_{-\infty}^{\infty} w(\tau - t').A(\tau).e^{j\varphi(t')\tau + j\varphi'(t')(\tau - t')}.e^{-j\varphi(t')\tau}d\tau\right| \tag{3.7}$$

The presence of modulated component $e^{j\varphi'(t')(\tau - t')}$ makes amplitude of the FT of $w(\tau - t').A(\tau)$ lower than its harmonic component. In order to overcome this effect of modulated component, it becomes necessary to apply demodulating operator. Since, linear modulator ratio $\varphi'(t')$ is a function of time, the demodulating operator needs to be time varying. The Short Time Fourier Transform (STFT) with time varying demodulating operator is given by Equation *3.8*.

$$G(t', \omega, c) = \int_{-\infty}^{\infty} w(\tau - t').g(\tau).e^{j\omega\tau}.e^{-jc(\tau - t')^{2}/2}d\tau \tag{3.8}$$

When the discrete demodulating operator $e^{-jc(\tau - t')^{2}/2}$ is close to modulated component of the signal for a TF point (t', ω), then TF representation across its IF will have high energy concentration and the amplitude of $|G(t', \omega, c)|$ will become maximum among all the values. After that, for each TF point, we obtain the best argument c given as

$$c' = arg\ \max_{c} |G(t', \omega, c)| \quad (3.9)$$

The TF representation of GLCT is given by Equation *3.10* and spectrum is defined by Equation *3.11*

$$GS(t', \omega) = G(t', \omega, c') \quad (3.10)$$

$$Spec(t', \omega) = |GS(t', \omega)|^2 \quad (3.11)$$

In order to determine the value of demodulating operator $e^{-jct'(\tau - t')^{2}/2}$, linear chirplet transform is used and the transform is named General Linear Chirplet Transform (GLCT). If T_s specifies sampling time and F_s specifies sampling frequency, the rotation provided by parameter α to the TF plane is given as:

$$\alpha = arctan\left(\frac{2T_s}{F_s}.c\right) \quad ; \quad \alpha \in (-\frac{\pi}{2}, \frac{\pi}{2}) \quad (3.12)$$

So that, the demodulating operator $e^{-jct'(\tau - t')^{2}/2}$ can describe all feasible modulating elements. Then equation *3.8* can be given as

$$G(t', \omega, c) = \int_{-\infty}^{\infty} w(\tau - t').g(\tau).e^{j\omega\tau}.e^{-j.tan(\alpha).\frac{F_s}{2T_s}.(\tau - t')^{2}/2} d\tau \quad (3.13)$$

If parameter α takes N values (as given in Equation *3.14*), the TF plane is divided into $N+1$ sections as given below

$$\alpha = -\frac{\pi}{2} + \frac{\pi}{N+1}, -\frac{\pi}{2} + 2.\frac{\pi}{N+1}, \ldots\ldots - \frac{\pi}{2} + N.\frac{\pi}{N+1} \quad (3.14)$$

It is to be noted here that for $N=1$, the GLCT behaves as STFT and with the increase in value of N, energy concentration of TF representation increases making it closer to IF representation. In this work, the value of N is taken as *3* as the GLCT generated TF representation at this value is suitable to illustrate true IF feature of signal.

3.2.2 EEG decomposition using fpICA

Many recently proposed artifact rejection and noise suppression methods have basis of ICA for decomposition of contaminated EEG activity into source components i.e., cerebral and artifactual. In present work, an improved Fast-ICA technique i.e., Fast-Power ICA (fpICA) comprising of power iteration algorithm is exploited for EEG signals decomposition into source components [30]. General Fast-ICA technique computes approximate solution of Newton Raphson (NR) algorithm, which is less stable for non-stationary signals such as EEG activity. In addition, Fast-ICA algorithm is often reported to have convergence problems [41]. The power iteration algorithm

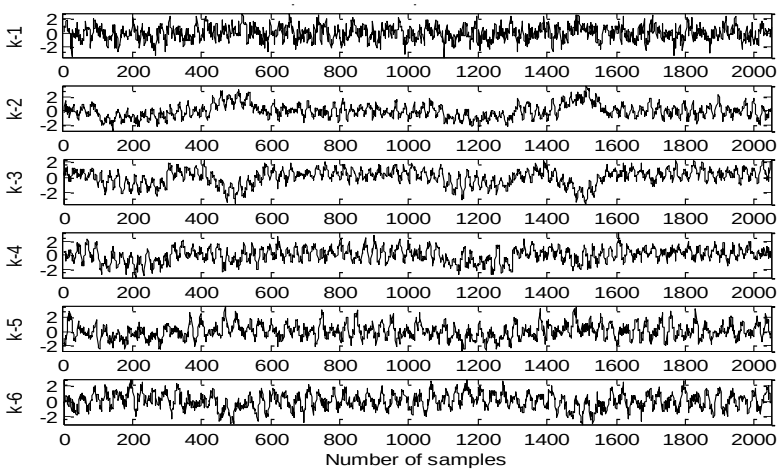
incorporated in fpICA technique overcomes the limitation of Fast-ICA algorithm by generating exact solution of *NR* algorithm. In addition, fpICA converges to the optimum solution even when the dimensionality and number of observations are of the same order. General ICA techniques maximize information transfer from input to output using nonlinear functions viz. *tanh*, *pow3* and *gauss* [118]. In this work, fpICA technique is implemented considering *pow3* non-linearity as it uses kurtosis-based approximation which has higher growth rate when compared to other non-linearities. In order to evaluate the efficacy of fpICA technique in de-mixing of EEG sources, a comparative study has been performed with three established ICA techniques viz. Fast-ICA [41], kurtosis-ICA (kICA) [43] and WASOBI [115] techniques using simulated EEG sources.

3.3 Proposed methodology

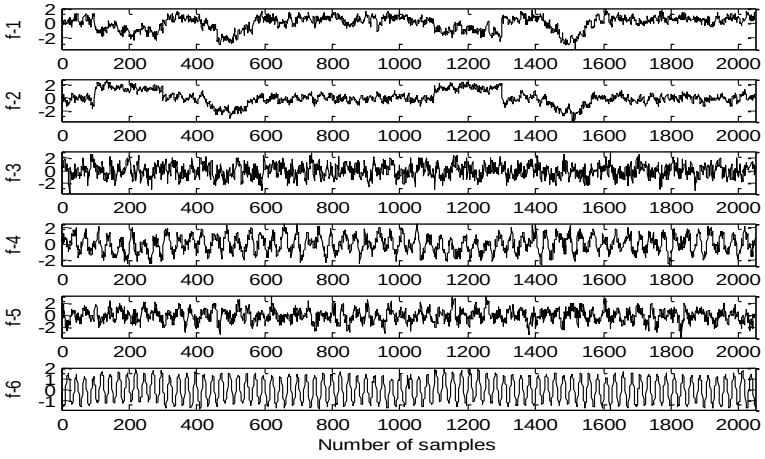
To remove artifacts and noise from EEG activity, this study proposes hybrid fpICA and GLCT based methodology (referred as GLCT-PICA EEG de-noising technique). Once the filtered EEG activity (filtered through notch and band-pass filter) is de-mixed using appropriate ICA algorithm, the extracted ICs are scrutinized to identify non-cerebral/artifactual components using proposed Katz-Fractal Sparsity (KFS) criterion. Since complete rejection of artifactual ICs causes loss of cerebral information leaked to artifactual ICs, the identified ICs are corrected by proposed GLCT de-noising technique. The complete methodology of EEG artifact removal and noise suppression using GLCT-PICA based EEG de-noising is explained in the following steps.

1. Filter (using notch and band-pass filters) *n*-channel recorded EEG activity $d(t)$ to obtain filtered EEG activity $\hat{d}(t)$:

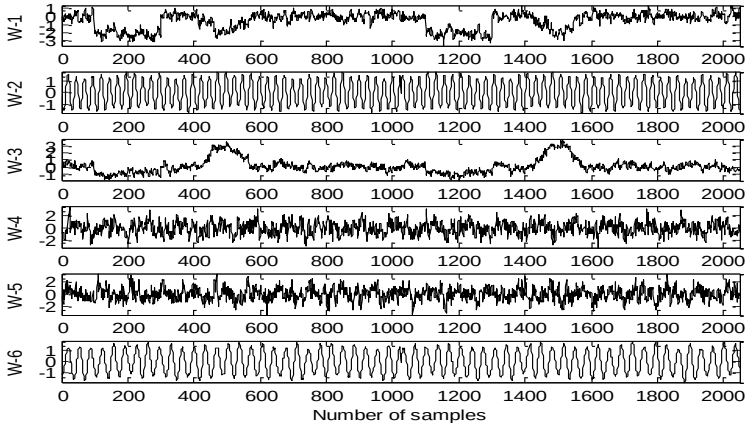
   The proposed technique of EEG de-noising is first implemented on simulated EEG data. This provides more insight to each processing step involved in GLCT-PICA based EEG de-noising. In order to illustrate the complete methodology of GLCT-PICA based EEG de-noising presented in Figure 3.1, sources of cerebral and non-cerebral/artifactual EEG activity are simulated, mixed (to attain contaminated EEG activity) and de-mixed using set of ICA algorithms. Simulated EEG data consists of *40* sets of *6*-channel contaminated EEG activity digitized at *256 Hz* and each set of *6*-channel contaminated EEG activity is the mixture of simulated cerebral and non-cerebral/artifactual sources generated using random mixing matrix.


k-1
k-2
k-3
k-4
k-5
k-6
0 200 400 600 800 1000 1200 1400 1600 1800 2000
Number of samples

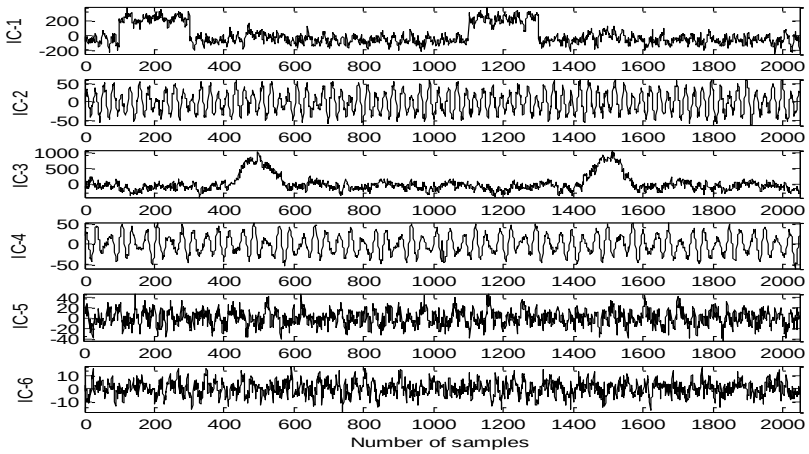
(a)


f-1
f-2
f-3
f-4
f-5
f-6
0 200 400 600 800 1000 1200 1400 1600 1800 2000
Number of samples

(b)


W-1
W-2
W-3
W-4
W-5
W-6
0 200 400 600 800 1000 1200 1400 1600 1800 2000
Number of samples

(c)



(d)

Figure 3. EEG sources (ICs) Extracted using **(a)** kICA algorithm **(b)** Fast-ICA algorithm **(c)** WASOBI algorithm **(d)** fpICA algorithm

Table 3.1 Reconstruction Error (MSE) estimated on extracted EEG sources (ICs)

MSE	**ICA algorithms**			
	Kurtosis-ICA	**Fast-ICA**	**WASOBI ICA**	**Fast-Power ICA**
Energy	9.3083	9.4883	7.8882	**0.0429**
PE	$2.6 * 10^{-6}$	$2.28 * 10^{-7}$	$2.55 * 10^{-6}$	$2.61 * 10^{-7}$
Katz FD	2.3830	2.2835	2.3752	**0.0017**

2. Apply fpICA algorithm on processed EEG activity $\hat{d}(t)$ to attain cerebral and non-cerebral/artifactual Independent Components ($\widehat{fp}_{ICA}(t)$) of brain electrical activity:

The Independent Components (ICs) are recovered from mixed EEG activity using four ICA algorithms viz. Kurtosis-ICA (kICA), Fast-ICA, WASOBI and Fast-Power ICA (fpICA). Figure 3.3(a), Figure 3.3(b), Figure 3.3(c) and Figure 3.3(d) epitomize the sets of ICs acquired after applying kICA, Fast-ICA, WASOBI and fpICA on simulated EEG activity given in Figure 3.3(b). It is observed from Figure 3.3(d) that fpICA algorithm successfully estimated eye-blink, and eye movement artifactual sources. Also, the rhythmic EEG activities are well separated using fpICA technique. On contrary, kICA, Fast-ICA and WASOBI ICA techniques failed to separate ocular EEG activity. In addition, separation of EEG activity is not appropriate using kICA and Fast-ICA algorithms. Table 3.1 represents a set of Mean Square Errors (MSE) estimated to evaluate efficacy of ICA algorithms (i.e., kICA, Fast-ICA, WASOBI and fpICA) in EEG source separation. The MSE value signifies reconstruction error present between original sources and mixed EEG sources of cerebral and non-cerebral EEG activity. The Mean Square Errors (MSE) are estimated on three

statistical measures i.e., Energy value, Permutation Entropy (PE) and Katz Fractal Dimension (FD) calculated from original and separated sources (ICs) of EEG activity. The performance analytics shown in Table 3.1 are the mean of Reconstruction Error calculated from 40 sets of simulated EEG data. It is observed from Table 3.1 that there is a significant difference between PE and katz FD values corresponding to fpICA algorithm in terms of reconstruction error as compared to other three ICA techniques. Therefore, fpICA algorithm is employed in this work for EEG source estimation.

Apply Katz-Fractal Sparsity (KFS) criterion on extracted Independent Components ($\widehat{fp}_{ICA}(t)$) for automatic marking of non-cerebral/artifactual Independent Components ($\widehat{fp}_{ICA}(t)$): The proposed methodology of GLCT-PICA based EEG de-noising aims to correct EEG activity against ocular artifacts, large EMG artifacts and other high frequency artifacts. After decomposing *N*-channel contaminated EEG activity $\hat{d}(t)$ into independent sources ($\widehat{fp}_{ICA}(t)$) using fpICA algorithm, criterion of Katz-Fractal Sparsity (KFS) is employed to automatically identify artifactual ICs. The proposed KFS criterion is an extension of sparsity criterion proposed by Zima et al. [33], which is built upon the common characteristics of artifactual components viz. large spiking activity, high kurtosis value, large variance and standard deviation, and topographic distribution of amplitude values [33]. The KFS criterion takes into consideration common characteristics of artifactual sources and marks any IC as artifactual if it shows high values of sparsity and FD value, which is calculated by KFS parameter in present work. High value of KFS parameter signifies that specific IC has high maximum absolute value and median is close to zero comparative to standard deviation, while FD value is very high. Here, FD is the measure of non-periodic and turbulent behaviour of the extracted IC. In this work, Katz's FD algorithm is employed to estimate fractal behaviour of extracted ICs. The Katz-Fractal Sparsity (KFS) estimation of an IC, $\widehat{fp}_{ICA}(t)$ is given by Equation *3.15*.

$$KFS = \frac{max\,[|\widehat{fp}_{ICA}(t)|]}{\sigma[\widehat{fp}_{ICA}(t)]} * \left(log\left\{\frac{\sigma[\widehat{fp}_{ICA}(t)]}{median[|\widehat{fp}_{ICA}(t)|]}\right\}\right) * katz\{\widehat{fp}_{ICA}(t)\} \qquad (3.15)$$

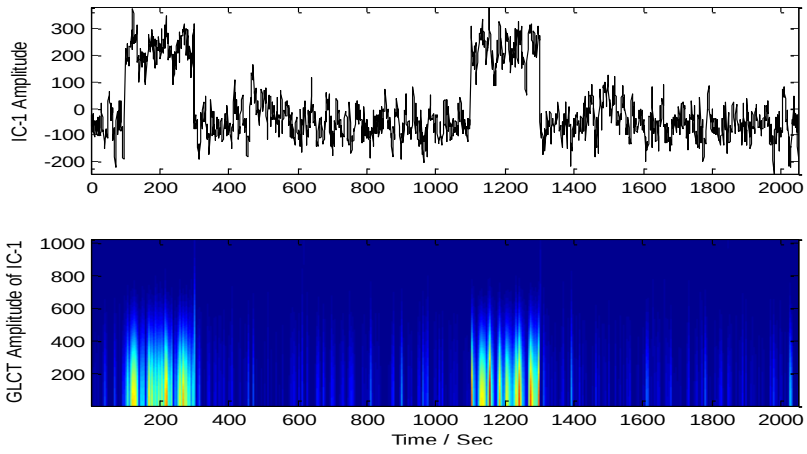
here, σ stands for standard deviation of $\widehat{fp}_{ICA}(t)$. Any IC is identified as artifactual, if its KFS parameter value is above a specified threshold. Based on the results of multiple experiments carried out on simulated EEG activity, mean of KFS value estimated from *n*-number of ICs is considered as the threshold. Table 3.2 shows the KFS parameter values estimated from ICs presented in Figure 3.3(d). It is observed from Table 3.2 that two ICs i.e. IC-1 and IC-3 are showing KFS parameter value greater than the mean (***304.69***) of *6* ICs **(i.e. KFS(ICs) > mean(KFS)).** Therefore, IC-1 and IC-3 are marked as artifactual ICs in this work.

Table 3.2 KFS parameter value computed from ICs presented in Figure 3.4(d)

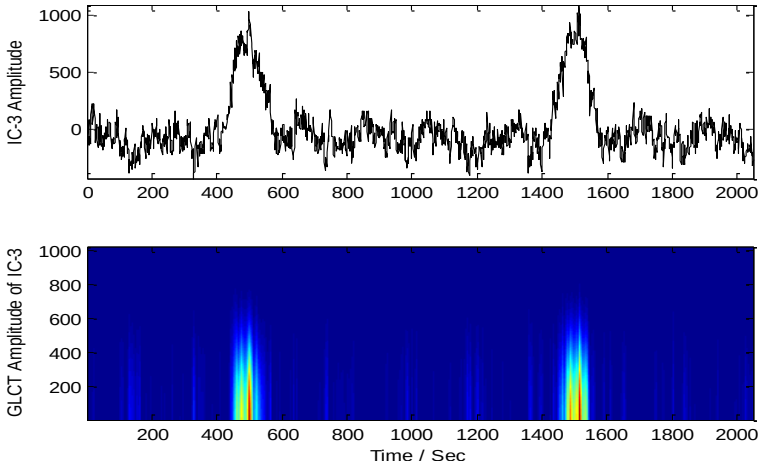
ICs	1	2	3	4	5	6
KFS value	**1414.06**	15.71	**345.11**	17.26	26.96	9.04

3. Perform GLCT on identified non-cerebral/artifactual Independent Components for obtaining Time-Frequency (TF) distribution of non-cerebral activity:

   In order to de-noise artifactual ICs, the identified ICs are transformed to TF domain using GLCT transform. Thereafter, spike zone thresholding is performed on transformed GLCT coefficients $GLCT(i,j)$, which is recently suggested by Daly et al. [34]. The identified artifactual ICs (i.e., IC-1 and IC-3) and corresponding GLCT coefficients sub-plot is shown in Figure 3.4(a-b). It is evident from Figure 3.4(a-b) that artifactual activity is accountable for higher amplitude values (spiking activity) of GLCT coefficients. Therefore, thresholding of high amplitude spiking activity $\left(GLCT(i,j)\right)$ is carried out to suppress non-cerebral artifactual origins in on-going EEG.



(a)



(b)

Figure 3.4(a) Artifactual IC (IC-1) and corresponding GLCT coefficient
(b) Artifactual IC (IC-3) and corresponding GLCT coefficient

4. Perform spike zone thresholding of the transformed GLCT coefficients and then the inverse transform of thresholded GLCT coefficients to obtain de-noised ICs ($\widehat{fp}_{ICA\_denoised}(t)$):

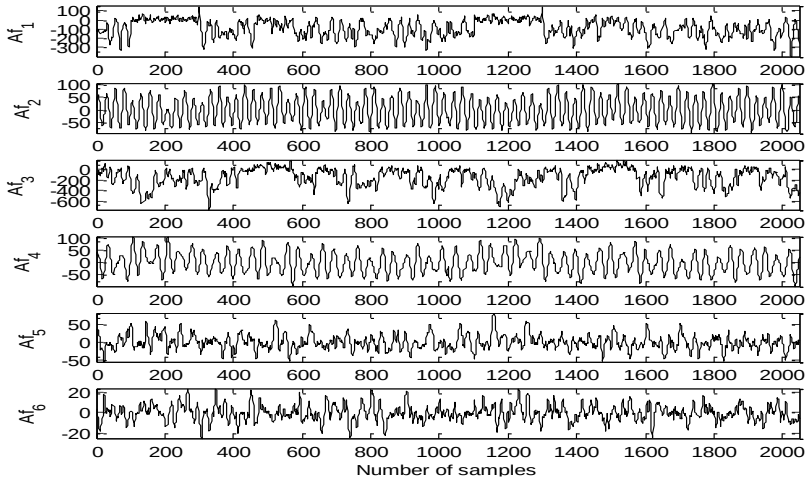
   Following the spike zone thresholding criterion, any GLCT coefficient having amplitude greater than $AG * (GLCT_m + GLCT_{std})$ is treated as coefficient of spiking activity; where $GLCT_m$ signifies mean of GLCT coefficients, $\mathrm{GLCT_{std}}$ signifies standard deviation of GLCT coefficients and *AG* is the adjustable gain responsible for controlling overall threshold value. In order to de-noise artifactual ICs, spike zone thresholding criterion is applied on GLCT coefficients $\mathrm{GLCT}(i,j)$:

$$GLCT(i,j)=\begin{cases} W \times CT(i,j) & if\ CT(i,j) > AG * (Coeff_m + Coeff_{sd}) \\ CT(i,j) & otherwise \end{cases} \quad (3.16)$$

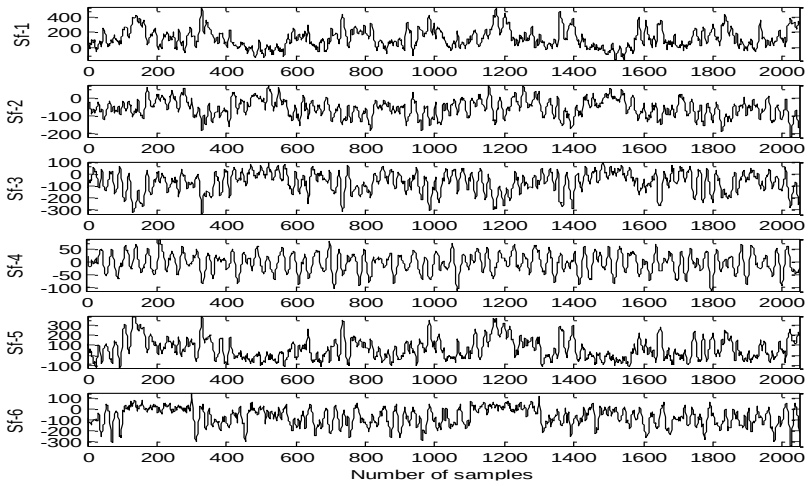
   here, W denotes the weight given to GLCT coefficient $GLCT(i,j)$ for suppressing artifactual activity. Following the results of multiple experiments carried out on simulated data, the weight W and adjustable gain AG are set to $W = 0.01$ and $AG = 1$, respectively.

5. In final step, perform inverse ICA on de-noised ($\widehat{fp}_{ICA\_denoised}(t)$) and cerebral ICs to obtain artifact corrected clean EEG data ($\hat{d}_{clean}(t)$):

   Inverse GLCT transform is performed on thresholded GLCT coefficients to reconstruct artifact-free/clean ICs.



(a)



(b)

Figure 3.5(a) De-noised/artifact-free ICs $\widehat{fp}_{ICA\_denoised}(t)$ obtained from contaminated EEG activity shown in Fig. 2.2(b)

(b) De-noised/artifact-free EEG (*6*-channel) activity $\hat{d}_{clean}(t)$ obtained from contaminated EEG shown in Fig. 2.2(b)

3.4 Results & Discussion

The artifact removal and noise suppression methodology proposed in this study is the hybrid approach of GLCT and fpICA algorithms in EEG de-noising application. The performance of proposed EEG de-noising methodology is primarily verified on *6*-channel simulated EEG activity and later applied on experimentally recorded *19*-channel EEG activity. Since, fpICA algorithms out-performed among other ICA algorithms in present study (as illustrated in section 3.2.2), fpICA algorithm is applied for Blind Source Separation (BSS) of simulated and experimentally recorded EEG activity in this work. The results of hybrid GLCT-PICA based EEG de-noising methodology are following subsections.

3.4.1 Artifact correction of simulated EEG activity

In order to illustrate effectiveness of proposed EEG de-noising methodology, results attained posterior to each methodological step are presented for visual examination. The evaluation metrics such as root mean square error, log spectral distance, signal to noise ratio and normalised root mean square difference are also computed. The de-noising results obtained on *6*-channel contaminated EEG activity present in Figure 2.2(b) are presented here for this purpose. The separated ICs i.e., $\widehat{fp}_{ICA}(t)$ obtained after application of fpICA BSS algorithm are presented in Figure 3.3(d). The estimated KFS parameter values of separated ICs are given in Table 3.2. Following the KFS criterion of artifactual ICs identification, IC-1 and IC-3 are marked as artifactual ICs and processed through GLCT based de-noising stage. The marked artifactual ICs and corresponding transformed GLCT coefficients are given in Figure 3.4(a-b). The remaining ICs (cerebral ICs) are by-passed through GLCT de-noising stage and combined with de-noised/clean ICs i.e., $\widehat{fp}_{ICA\_denoised}(t)$ for reconstructing clean EEG activity $\hat{d}_{clean}(t)$ by performing inverse ICA operation. Figure 3.5(a) shows a set of de-noised/artifact-free ICs $\widehat{fp}_{ICA\_denoised}(t)$ and Figure 3.5(b)

shows a set of *6*-channel denoised/artifact-free EEG activity $\hat{d}_{clean}(t)$ obtained after de-noising of contaminated EEG activity shown in Figure 2.2(b).

3.4.2 Artifact correction of experimental EEG activity

Experimentally recorded EEG data provided by Castellanos and Makarov (2006) [119] is used to show efficacy of proposed artifacts correction and noise suppression method in this work. A sample of filtered EEG activity $\hat{d}(t)$ processed in this work is shown in Figure 2.3. It is observed from Figure 3.6 that experimentally recorded EEG activity suffers from eye blink artifact, high-frequency artifacts and other disturbances. Thereafter, filtered EEG activity is separated in independent sources (i.e., $\widehat{fp}_{ICA}(t)$) using fpICA algorithm and extracted ICs are presented in Figure 3.6. After visual examination of separated ICs presented in Figure 3.6, it is determined that IC-1, IC-2 and IC-3 are the sources of artifactual activity in recorded EEG data.

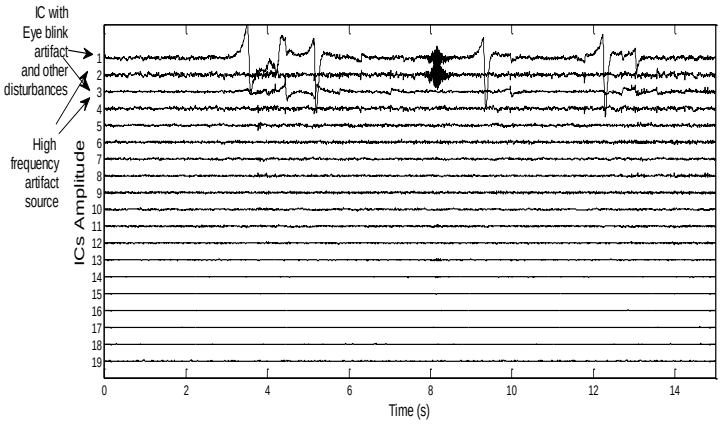


Figure 3.6 ICs $\widehat{fp}_{ICA}(t)$ extracted from filtered EEG activity using fpICA algorithm

In order to identify artifactual/non-cerebral ICs, criterion of KFS is applied on $\widehat{fp}_{ICA}(t)$ and computed values of KFS parameter are shown in Figure 3.7. The estimated mean of KFS parameter for extracted ICs is 82.13 and all ICs having KFS > $mean(KFS)$ are marked as artifactual. It is shown in Figure 3.7 that IC-1, IC-2 and IC-3 are having KFS $>$ 82.13 and these ICs are marked as artifactual for given set of EEG activity. Thereafter, IC-1, IC-2 and IC-3 are processed through proposed EEG de-noising methodology to remove strong artifacts and noise from $\widehat{fp}_{ICA}(t)$. The General Linear Chirplet Transform (GLCT) is computed on marked ICs i.e., IC-1, IC-2 and IC-3 respectively. Identified artifactual ICs (IC-1, IC-2 and IC-3) and respective GLCT coefficients $GLCT(i,j)$ amplitude plots are shown in Figure 3.8(a-c). It is inferred from Figure 3.8(a-c) that artifactual/non-cerebral components contribute to high amplitude of GLCT coefficients $GLCT(i,j)$ and therefore, thresholding of $GLCT(i,j)$ needed to be performed for de-noising purpose. Spike zone thresholding is performed on transformed on GLCT coefficients $GLCT(i,j)$ and de-noised ICs $\widehat{fp}_{ICA\_denoised}(t)$ are obtained after inversion of respective GLCT coefficients. The comparison of identified artifactual ICs and corrected/denoised ICs $\widehat{fp}_{ICA\_denoised}(t)$

using GLCT de-noising method is presented in Figure 3.9(a-c). It is inferred from Figure 3.9(a-c) that GLCT coefficient thresholding technique effectively suppressed eye blink and other high frequency artifacts, while preserved the useful EEG information. Thereafter, inverse ICA is performed on clean ICs (i.e., de-noised $\widehat{fp}_{ICA\_denoised}(t)$ and cerebral ICs), which helps in recovering clean EEG activity $\hat{d}_{clean}(t)$ from contaminated EEG data. Figure 3.10 shows set of clean ICs obtained after de-noising of ICs presented in Figure 3.6. Figure 3.11 shows *19*-channel clean EEG activity $\hat{d}_{clean}(t)$ reconstructed after applying inverse ICA on clean ICs presented in Figure 3.10. On comparing clean EEG activity $\hat{d}_{clean}(t)$ with contaminated EEG record shown in Figure 2.3(a), it can be deduced that proposed GLCT-PICA based EEG de-noising method effectively removed renounce artifactual component from contaminated EEG activity and successfully recover useful cerebral information.

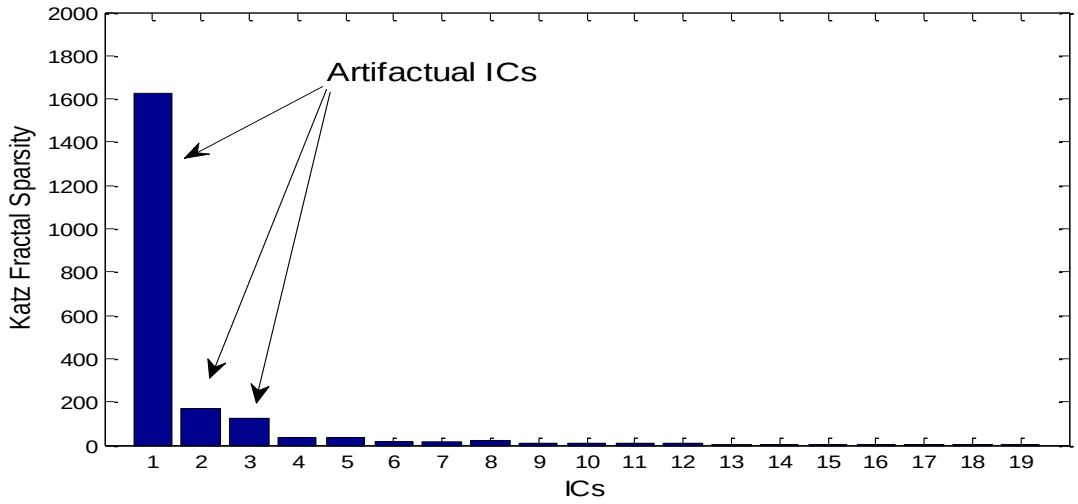
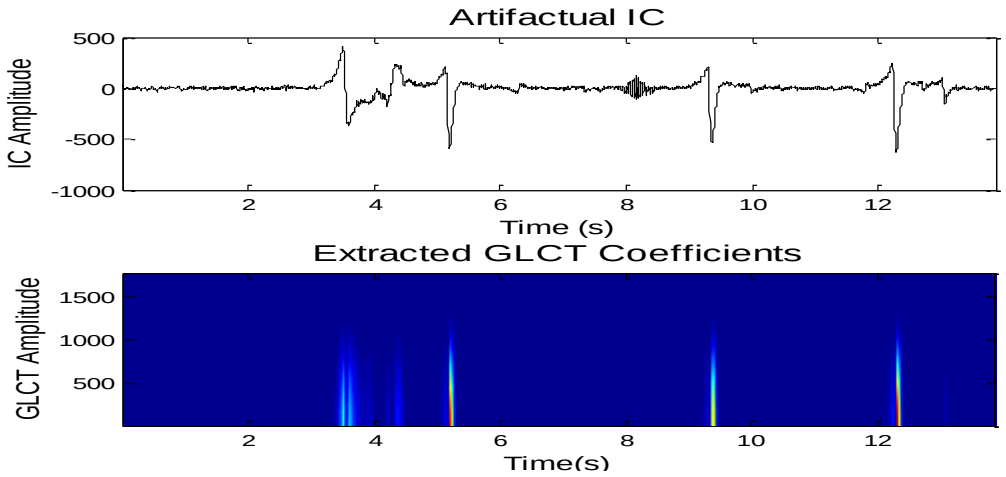
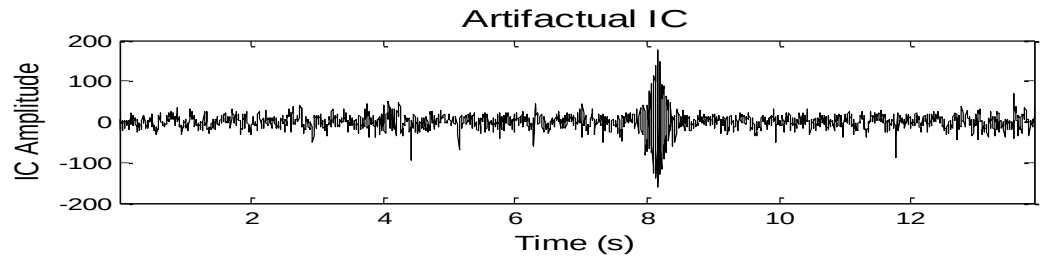
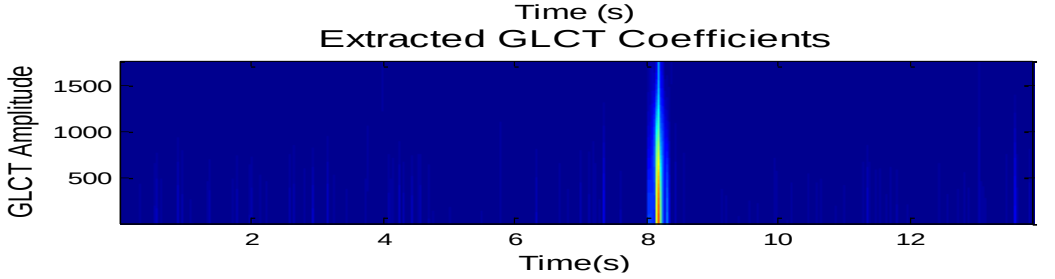


Figure 3.7 KFS parameter value computed from ICs presented in Figure 3.7(b)

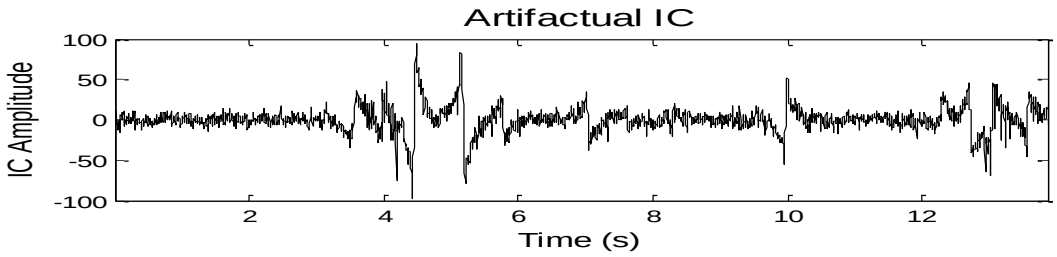


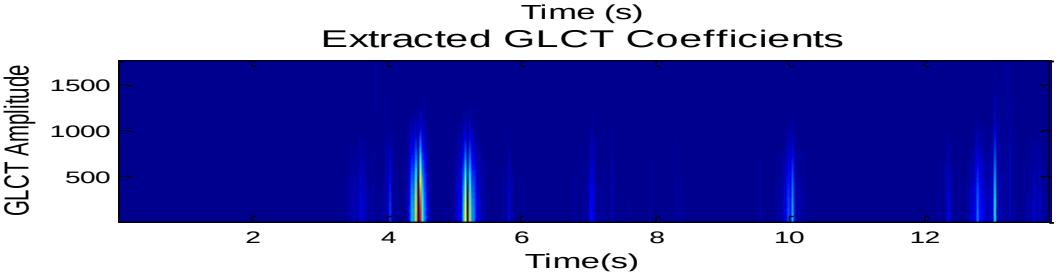
(a)





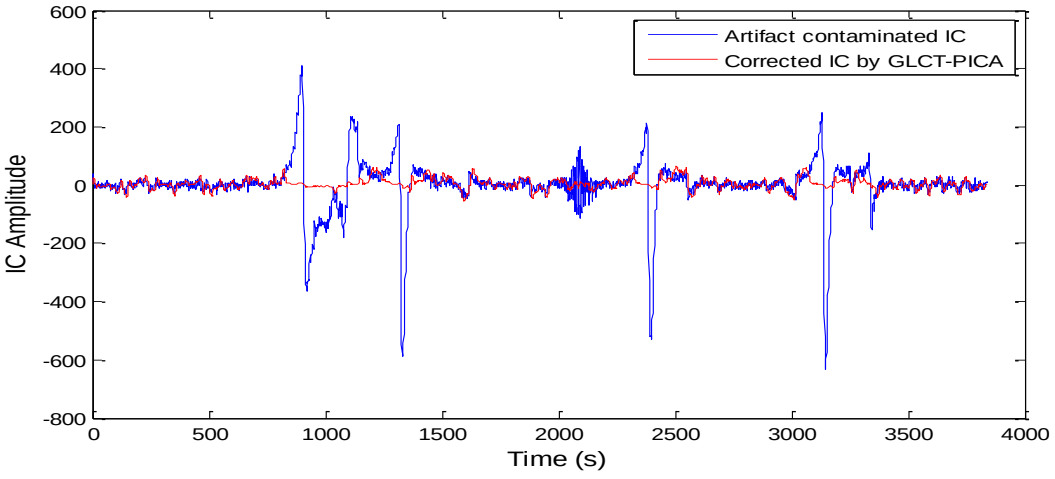
(b)



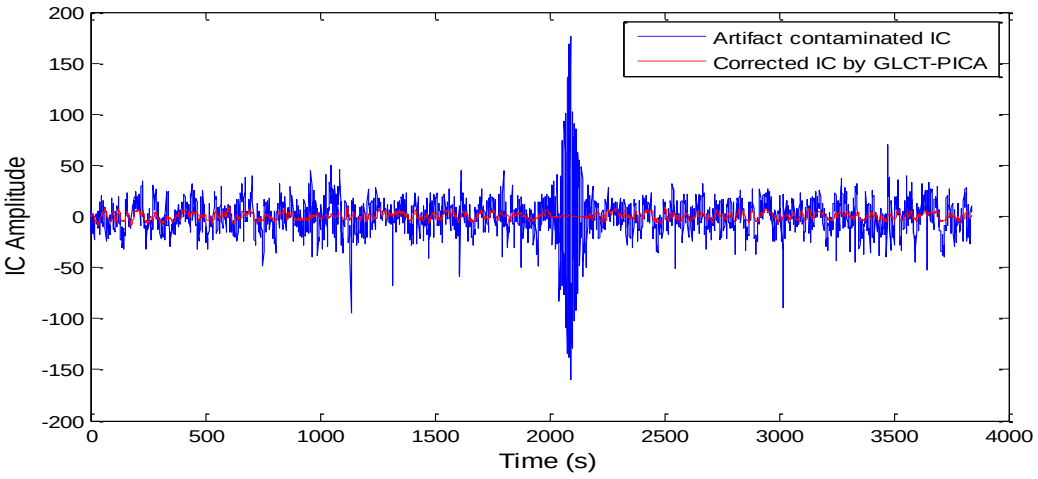


(c)

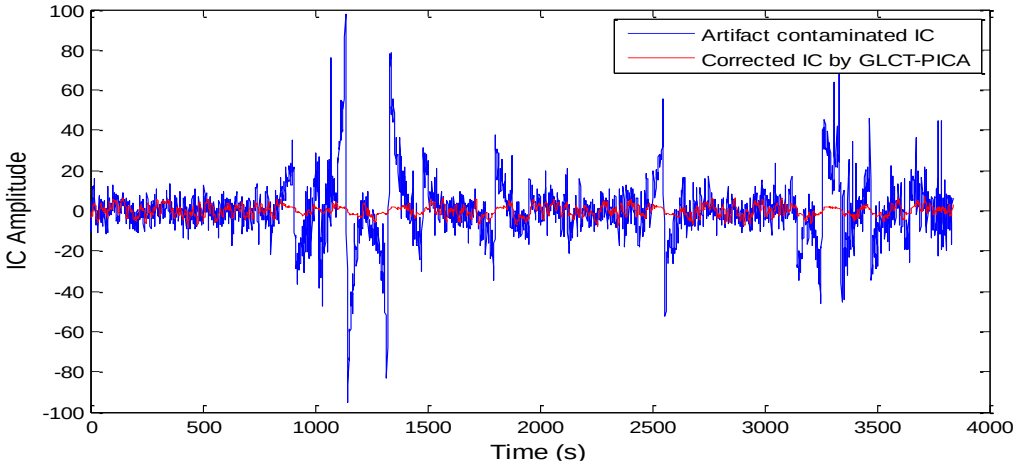
Figure 3.8 Marked artifactual ICs and respective GLCT coefficient amplitude plots for **(a)** IC-1 **(b)** IC-2 **(c)** IC-3



(a)



(b)



(c)

Figure 3.9 Comparison of identified artifactual ICs and corrected/denoised ICs using GLCT de-noising method **(a)** IC-1 **(b)** IC-2 **(c)** IC-3

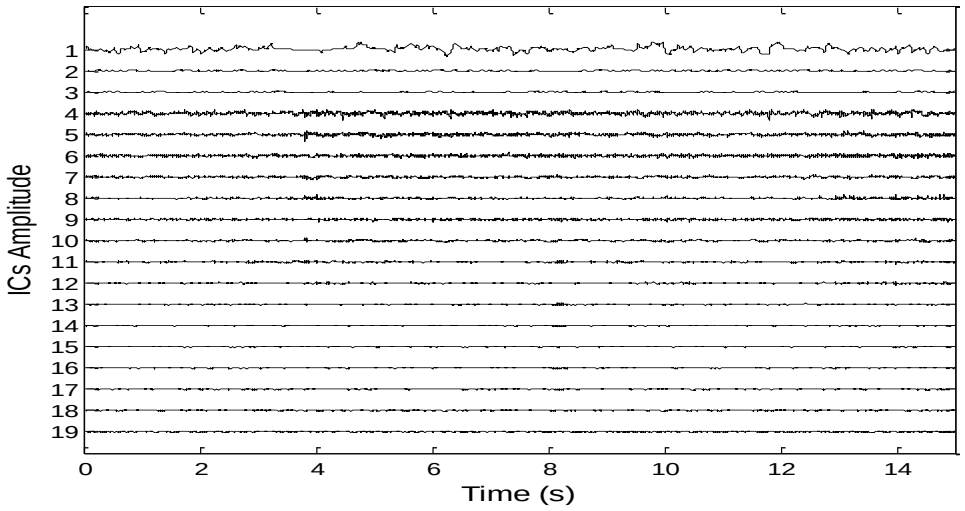


Figure 3.10 Clean ICs obtained after de-noising of ICs presented in Fig. 3.6

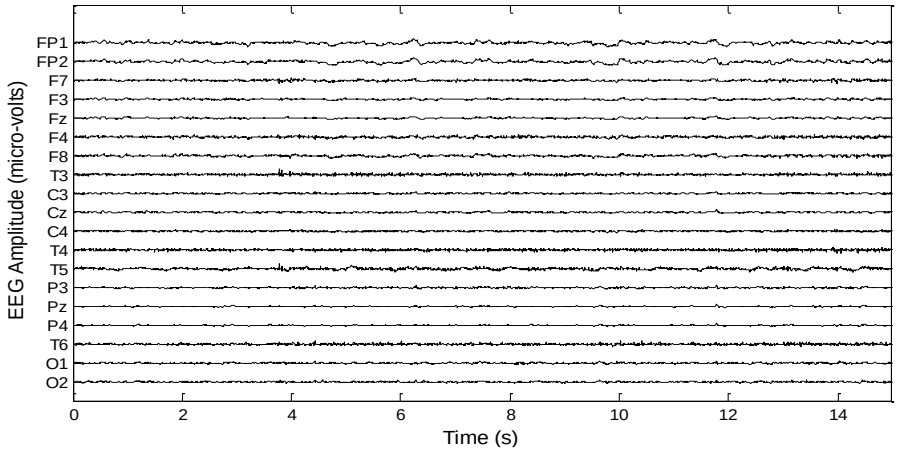
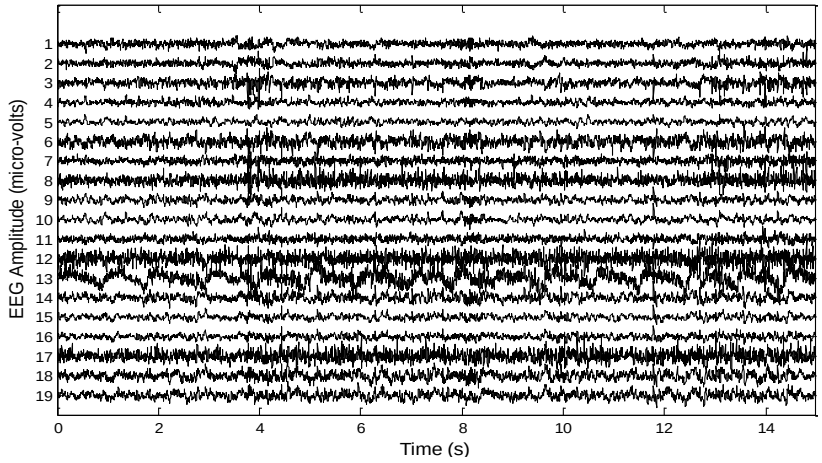
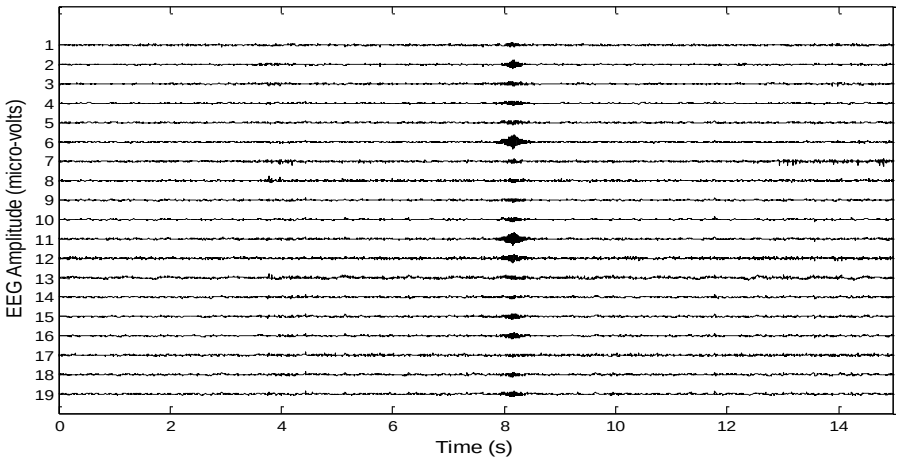


Figure 3.11 Clean EEG activity $\hat{\boldsymbol{d}}_{clean}(t)$ obtained after GLCT-PICA based de-noising of contaminated EEG data presented in Figure 2.3



(a)

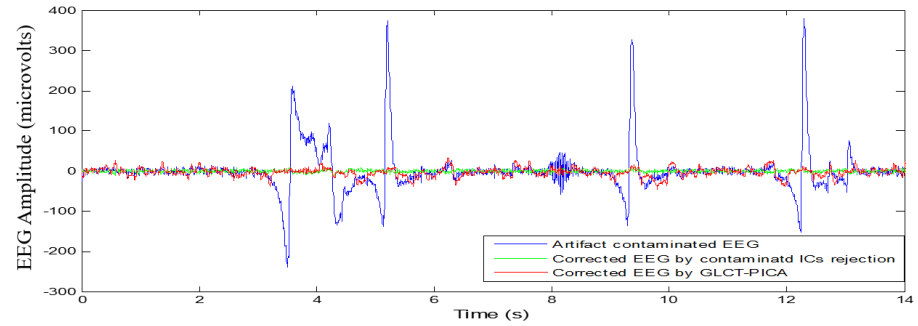


(b)

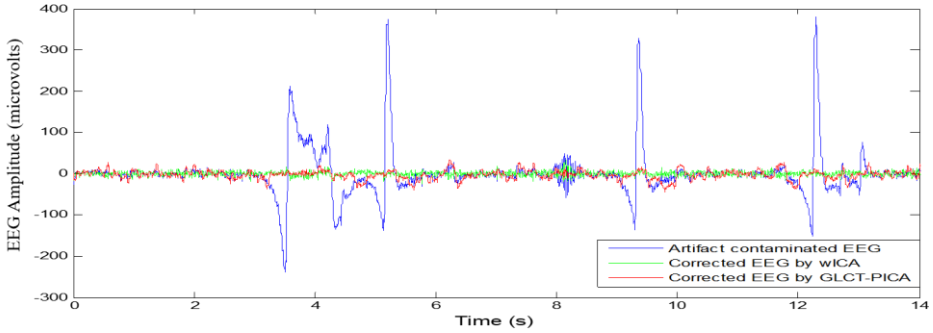
Figure 3.12 *19*-channel clean/artifact corrected EEG activity obtained after processing through (a) IC rejection technique (b) wICA based EEG de-noising technique

The performance of proposed GLCT-PICA EEG de-noising methodology is compared with two conventional ICA based EEG de-noising techniques viz. IC rejection-based EEG de-noising technique and wICA based EEG de-noising technique. The contaminated EEG activity presented in Figure 2.3 is processed through both conventional techniques and results are presented for visual examination in both time and frequency domains. The clean/artifact corrected EEG activity obtained after processing through IC rejection and wICA based EEG de-noising techniques is shown in Figure 3.12(a) and Figure 3.12(b), respectively.

Figure (3.13-3.15) show comparison plot of clean EEG activity (for channel-1, channel-2 and channel-3) obtained after processing through GLCT-PICA de-noising technique, IC rejection based de-noising technique and wICA based de-noising technique. It is observed from visual inspection of Figure (3.13-3.15) that the proposed technique of GLCT-PICA based EEG de-noising performed better than conventional ICA based EEG de-noising methods. Clean EEG activity obtained through proposed GLCT-PICA de-noising method strictly followed original EEG pattern until arrival of contaminated section. Subsequently, EEG activity significantly changed its shape to suppress artifactual content and maintained small variations to preserve useful cerebral information. Further, it is inferred from Figure (3.13-3.15) that ICA rejection and wICA based EEG de-nosing techniques suppressed high amplitude artifactual component significantly, but both these methods failed to preserve useful cerebral information present alongside contaminated EEG activity.

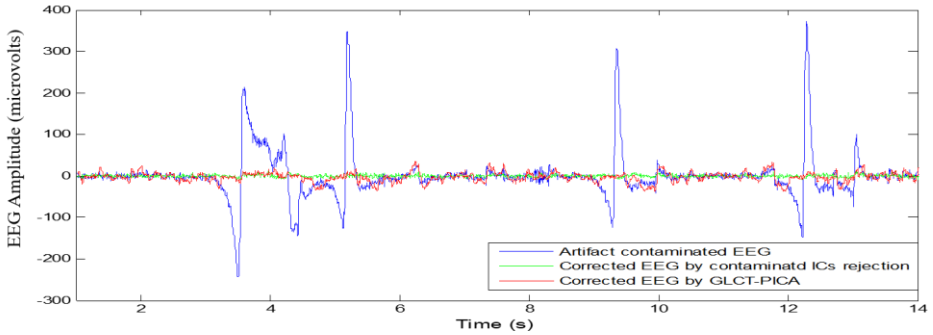


(a)

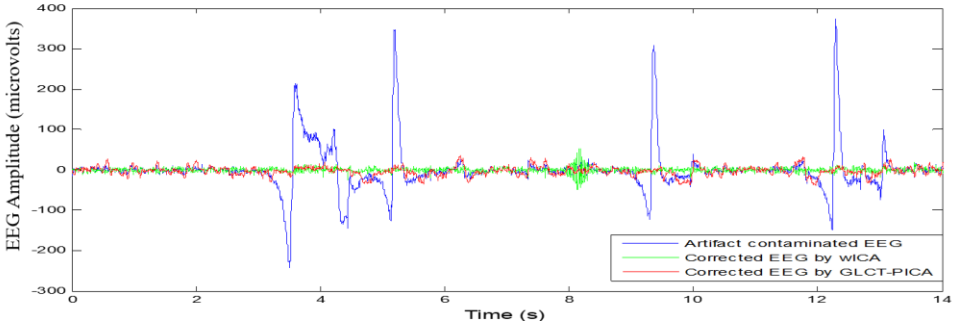


(b)

Figure 3.13 Comparison of clean EEG activity (channel-1) **(a)** Corrected by IC rejection and GLCT-PICA de-noising techniques **(b)** Corrected by wICA and GLCT-PICA de-noising techniques

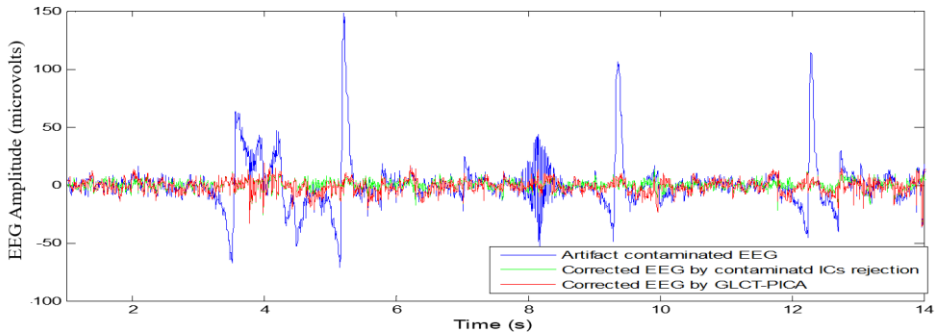


(a)

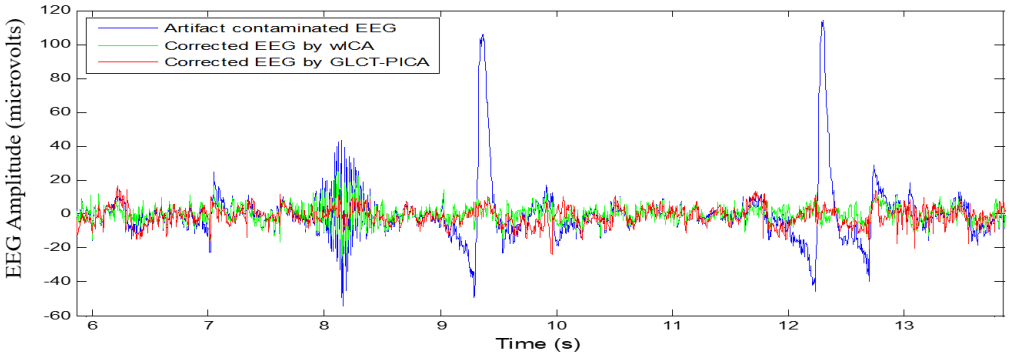


(b)

Figure 3.14 Comparison of clean EEG activity (channel-2) **(a)** Corrected by IC rejection and GLCT-PICA de-noising techniques **(b)** Corrected by wICA and GLCT-PICA de-noising techniques



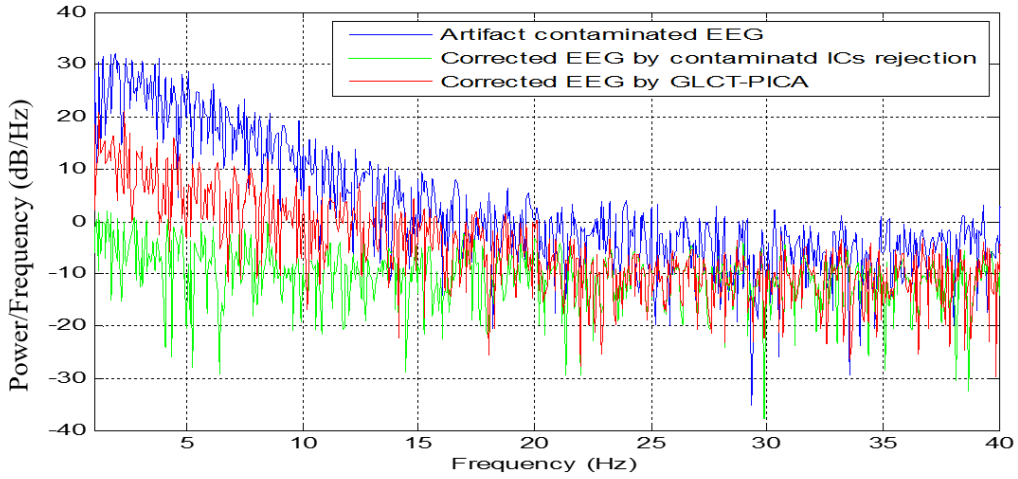
(a)



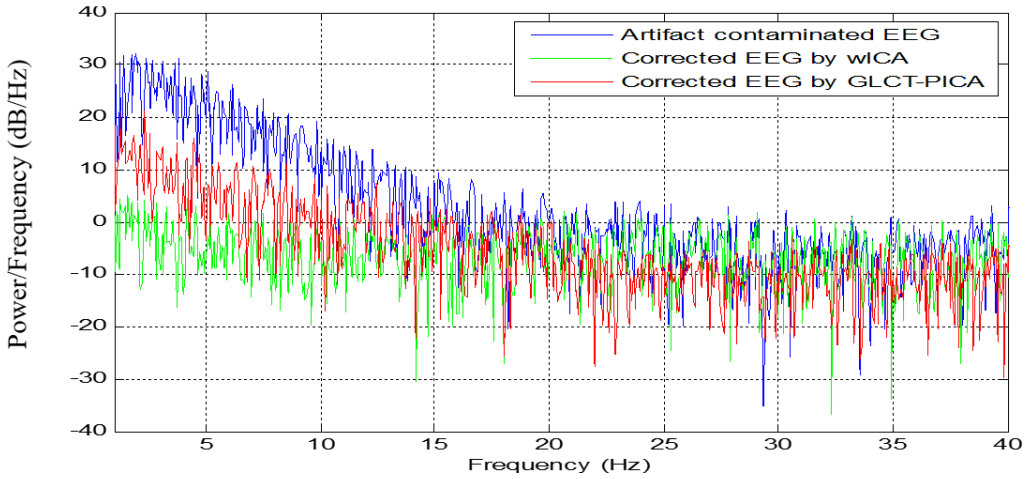
(b)

Figure 3.15 Comparison of clean EEG activity (channel-3) **(a)** Corrected by IC rejection and GLCT-PICA de-noising techniques **(b)** Corrected by wICA and GLCT-PICA de-noising techniques

In order to evaluate the effect of GLCT-PICA de-noising algorithm on spectral properties of contaminated EEG signals, Power Spectral Density (PSD) of contaminated and clean EEG data is plotted and comparison is performed. Figure (3.16-3.18) show comparison plot of PSD of clean EEG activity (for channel-1, channel-2 and channel-3) obtained after processing through GLCT-PICA de-noising technique, IC rejection based de-noising technique and wICA based de-noising technique. It is observed from visual inspection of Figure (3.16-3.18) that compared to other two conventional de-noising techniques; GLCT-PICA method slightly modified the power spectrum of contaminated EEG in cleaning process. The response of GLCT-PICA de-noising method is significantly better than conventional methods (IC rejection method and wICA method) for lower frequencies of EEG spectra.

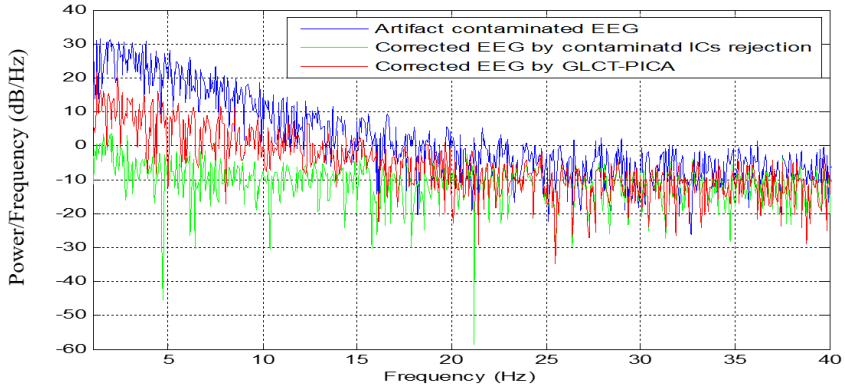


(a)

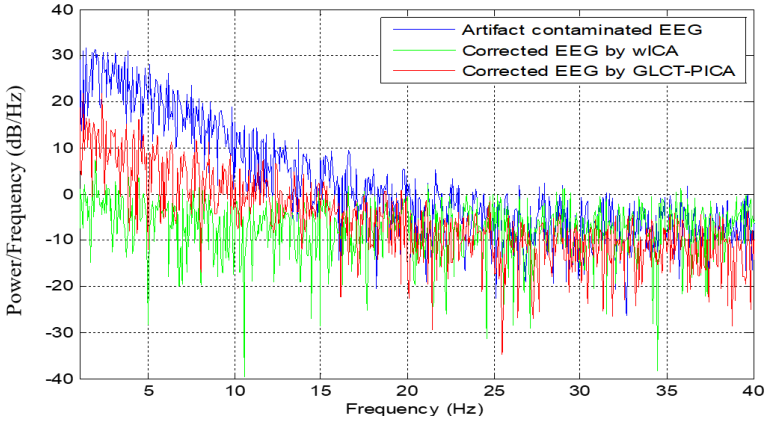


(b)

Figure 3.16 Comparison of PSD of clean EEG activity (channel-1) **(a)** Corrected by IC rejection and GLCT-PICA de-noising techniques **(b)** Corrected by wICA and GLCT-PICA de-noising techniques denoising

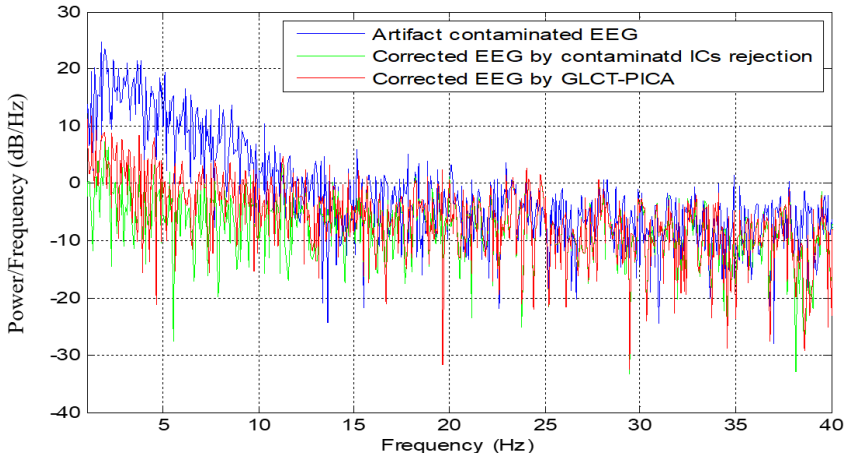


(a)

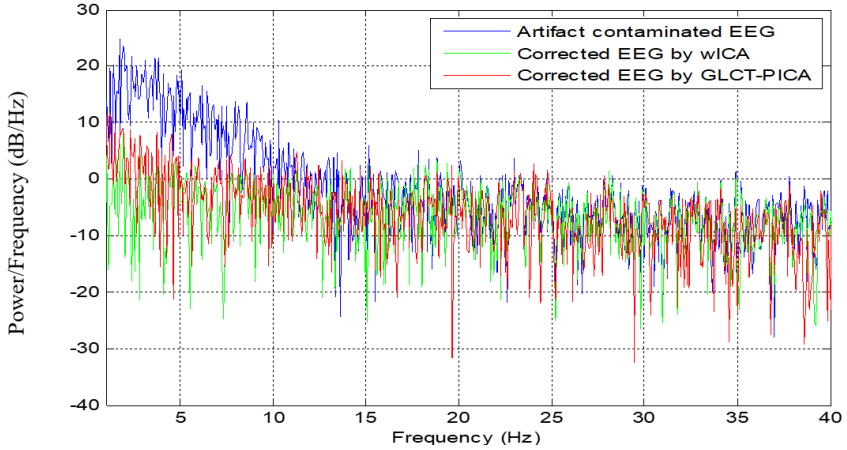


(b)

Figure 3.17 Comparison of PSD of clean EEG activity (channel-2) **(a)** Corrected by IC rejection and GLCT-PICA de-noising techniques **(b)** Corrected by wICA and GLCT-PICA de-noising techniques denoising



(a)



(b)

Figure 3.18 Comparison of PSD of clean EEG activity (channel-3) **(a)** Corrected by IC rejection and GLCT-PICA de-noising techniques **(b)** Corrected by wICA and GLCT-PICA de-noising techniques

In addition to ICA rejection and wICA based EEG artifact removal methods, the performance of the proposed methodology is also compared with recently proposed non-ICA based EEG artifact removal methods viz. VMD [35], mMSE_Regression [36] and EEMD_Renyi's Entropy [37] methods. These methods are implemented on the similar 19-channel experimental EEG dataset used for qualitative analysis of the proposed methods with ICA rejection and wICA based EEG artifact removal methods. The quantitative analysis of the EEG de-noising methods is carried out by estimating different objective quality assessment metrics viz. Signal to Noise Ratio (SNR), Root Mean Square Error (RMSE), Normalized Maximum Absolute Error (NMAX) and Normalized Root Mean Square Difference (NRD) [35]. The performance metrics computed from the original EEG signal $xo(i)$ (i.e.d(t)) and reconstructed EEG activity $xr(i)$ (i.e. $\boldsymbol{d}clean\ (t)$) are as follows:

$$SNR = 10log_{10}[\frac{\sum_{i=1}^{N}(x_0(i)-u_0)^2}{\sum_{i=1}^{N}(x_0(i)-x_r(i))^2}] \quad (3.17)$$

$$RMSE = \sqrt{\sum_{i=1}^{N}(x_0(i)-x_r(i))^2 / N} \quad (3.18)$$

$$NMAX = \frac{max_{i=1}^{N}|x_0(i)-x_{r0}(i)|}{\max(x_0(i))-\min(x_r(i))} \quad (3.19)$$

$$NRD = \frac{\sqrt{\sum_{i=1}^{N}(x_0(i)-x_r(i))^2 / N}}{\max(x_0(i))-\min(x_r(i))} \quad (3.20)$$

here, μo refers to the mean of original EEG activity. The comparative mean ± std results of proposed method with existing methods for selected objective metrics is presented in Table 3.3. It is evident from Table 3.3 that proposed GLCT-PICA based EEG artifact removal method outperforms existing artifact removal methods for given performance metrics. The RMSE value plot estimated from individual EEG channels for proposed and existing artifact correction methods is presented in Figure 3.19. It is observed from Figure 3.19 that proposed GLCT-PICA method performs better than existing methods for most channels, except for few channels where mMSE_Regression method performed better than proposed method.

Table 3 Comparative *mean* ± *std* performance metrics for real EEG activity using proposed and existing methods

Performance Metrics	GLCT-PICA Mean± Std	VMD [35] Mean± Std	mMSE_Regression [36] Mean± Std	EEMD_renyi [37] Mean± Std	wICA [21] Mean± Std	IC_rejection [12] Mean± Std
SNR (dB)	**3.4329± 3.7652**	-17.7329 ± 3.8559	0.5831± 1.7507	-3.7594 ± 5.3108	-16.2524± 3.4066	0.6296± 5.1188
RMSE	**0.2264± 0.1781**	0.3709 ± 0.1945	0.2398 ± 0.1459	0.3807 ± 0.2379	0.3925 ± 0.2126	0.3376 ± 0.2286

NMAX	**2.2857 ± 1.8931**	15.7554 ± 16.1642	2.9224 ± 2.9021	6.0905 ± 9.6933	17.4289 ± 26.4941	3.6943 ± 4.2612
NRD	**0.0176 ± 0.0062**	0.1366 ± 0.0900	0.0258 ± 0.0148	0.0378± 0.0480	0.1342± 0.1381	0.0312 ± 0.0220

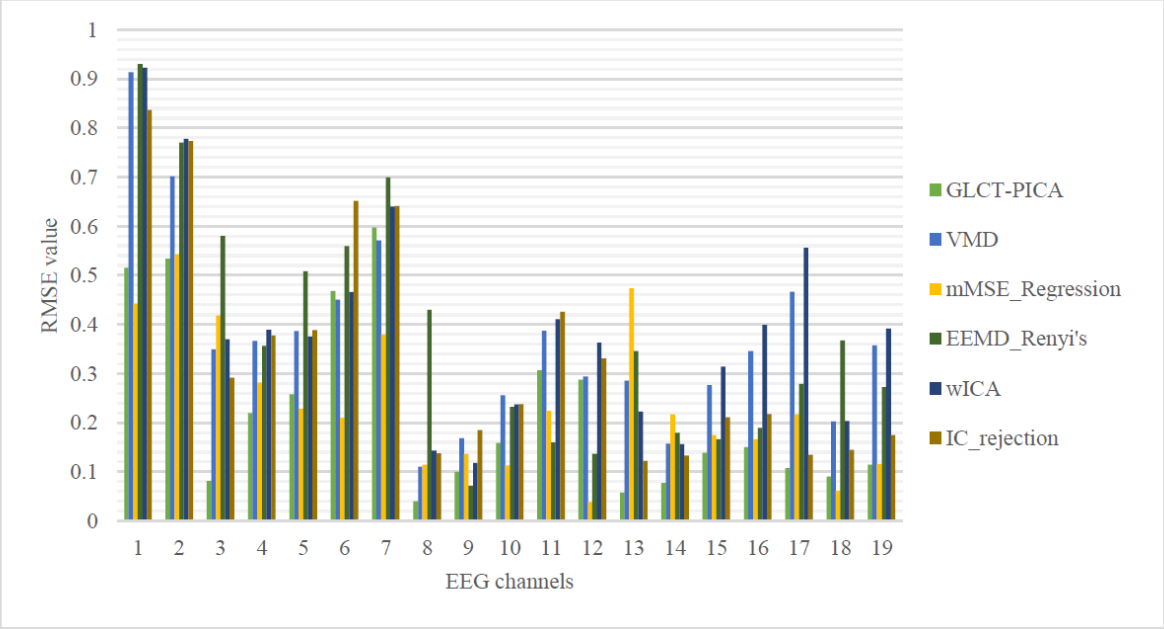


Figure 3.19 Channel-wise RMSE plot for proposed and existing methods of EEG de-noising

In addition to objective quality assessment metrics, the Log Spectral Distance (LSD) parameter is estimated to measure signal distortion elicited in each EEG channel by proposed and existing methods. Signal distortion is attained by estimating Power Spectral Density (PSD) of an artifact-free EEG data segment and another equal- length reconstructed EEG segment received after artifact processing. If, $Po(w)$ is the PSD of the original artifact- free EEG segment and $Pr(w)$ is the PSD of reconstructed EEG segment, the LSD parameter estimation is given by Equation 3.21. Log Spectral Distance (LSD) values estimated for 19-channel EEG activity using proposed and existing methods are presented in Table 3.4. It is observed from Table 4 that proposed GLCT-PICA EEG de-noising method yielded least LSD value for eleven out of nineteen EEG channels. After, GLCT_PICA method, the VMD method was found to perform better than other methods, which yielded least LST values for five out of nineteen EEG channels.

$$LSD = \sqrt{\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[10log_{10}\frac{P_0(w)}{P_r(w)}\right]^2}\,dw \qquad (3.21)$$

Table 3.4 Log Spectral Distance (LSD) value for 19-channel EEG activity using proposed and existing methods

Channel	GLCT-PICA (Proposed method)	VMD [35]	mMSE_ Regression [36]	EEMD_ Renyi's [37]	wICA [21]	IC-rejection [12]
Ch_1	1.928208	**1.762976**	2.2713	1.971934	2.18937	2.138418
Ch_2	1.678612	**0.920957**	2.433305	2.152548	2.218652	2.087933
Ch_3	1.359534	1.534922	**1.288607**	1.576305	1.607515	1.413608
Ch_4	1.261466	1.493881	1.672261	1.634322	**0.866642**	1.361536
Ch_5	**1.203692**	1.572605	1.47679	1.667242	1.786788	1.244171
Ch_6	**1.536941**	1.605163	1.648129	1.624695	1.551272	1.617848
Ch_7	1.761306	**0.933722**	1.958158	1.997607	1.802048	2.033965
Ch_8	1.02777	**0.929957**	1.236107	1.081511	1.203373	1.059559
Ch_9	**0.978309**	1.22239	1.029921	1.252818	1.023896	1.011638
Ch_10	**0.953661**	1.228696	1.178374	1.280154	1.580139	0.977779
Ch_11	1.273238	1.057535	**1.003046**	1.175937	1.353146	1.376665
Ch_12	1.009962	**0.871412**	1.2121	0.970349	1.221743	1.031418
Ch_13	**0.636978**	0.851833	1.293891	1.057838	1.024942	0.695379
Ch_14	**0.802547**	0.99321	1.364632	1.083953	1.271065	0.868706
Ch_15	**0.849908**	1.03081	1.133108	0.967032	1.268802	1.024385
Ch_16	**0.879073**	0.933047	1.140529	0.969822	1.249607	0.951567
Ch_17	**0.697141**	0.85446	1.377332	0.979507	1.318534	0.728598
Ch_18	**0.810638**	0.712359	1.365262	0.830673	1.121381	0.855444
Ch_19	**0.780957**	0.762149	1.28234	0.923966	1.240992	0.842345

3.5 Summary

In this work, GLCT-PICA based EEG de-noising methodology is proposed for automatic identification and rejection of non-cerebral EEG artifacts from contaminated recordings. Proposed methodology is successfully implemented on simulated and experimentally recorded EEG activity. The effectiveness of methodology is validated by accessing time-domain and frequency-domain features of contaminated and cleaned/processed EEG activity. In this work, fpICA BSS technique is found highly suitable for EEG source separation process. Fast-Power ICA algorithm successfully separated mixed source signals of simulated EEG activity. The criterion of Katz Fractal Sparsity (KFS) established that ICs with KFS value greater than mean of all KFS values should be considered as artifactual/non-cerebral and processed through GLCT based de-noising stage. Spike zone thresholding criterion helped in suppressing spiking areas of artifactual ICs, while keeping cerebral information intact during processing. The time-domain and frequency-domain analysis of contaminated and clean EEG activity established the fact that proposed GLCT-PICA EEG de-noising method performed better than conventional ICA based EEG de-noising methods and preserved necessary cerebral information in the processed EEG signals. In addition, proposed methodology has capability to be implemented in real time environment for improving performance of EEG based BCI/disease diagnosis system.

CHAPTER 4

EEG Channel Selection for Multi-Class Motor Imagery-based BCI System

Preface

In this chapter, to enhance the multi-class BCI system's performance, a novel channel selection, and features optimization methodology have been proposed. First, the multi-channel EEG dataset is reduced to an optimum no. of channels subset using developed MDA-SOGWO based EEG channel selection criterion. After that, the discriminable feature-set is generated from the time, frequency, and FAWT based time-frequency domain of the EEG dataset. Thereafter, an informative feature set is constructed from extracted features through CCA-RFE based feature selection criterion. Finally, training and validation of selected feature-set are carried out using ELM, LDA, RF, and MLP classifiers.

4.1 Introduction

Motor Imagery (MI) BCI activity refers to performing limb, tongue, or leg movements mentally without performing the task in reality [120]. The real or MI activity leads to alter dynamics of EEG activity significantly over the Sensorimotor Region (SMR) of the brain [121]. The capability to detect MI activity through EEG signals enables a BCI system to control assistive applications to support disabled people or patients suffering from motor impairments such as multiple sclerosis, amyotrophic lateral sclerosis, etc. [122]. Although BCI research is still in its early stages, it has been established as a useful technology for a wide range of applications and has achieved significant impetus in the past few years. It is inferred from past studies that designing computationally efficient and highly discriminable feature extraction methodology is a challenging task in BCI system design. Therefore, designing an optimum EEG channel selection criterion is highly essential for the better computational performance of BCI system.

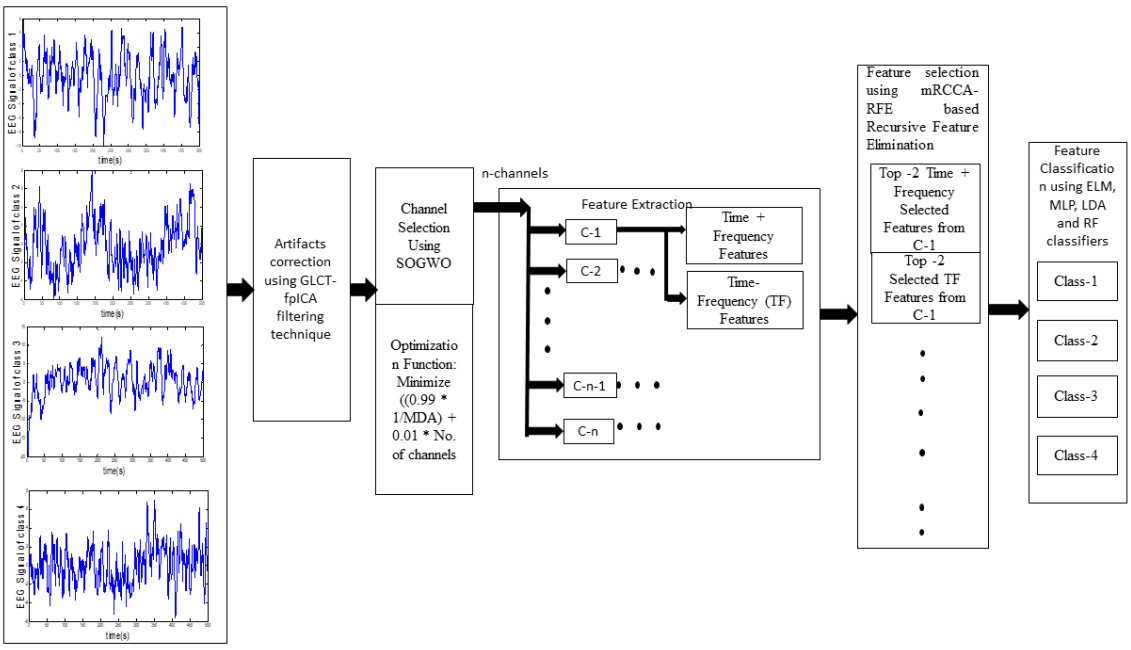


Figure 4.1. Schematic diagram of the developed EEG channel selection and optimum feature-set generation methodology

In this chapter, a channel selection and feature optimization-based EEG processing and classification pipeline is developed while focusing on multi-class MI-BCI tasks. As the first stage of the developed methodology, raw EEG activity is treated to reject severe EEG artifacts following our previously proposed hybrid GLCT-PICA based EEG de-noising pipeline [123]. Thereafter, in order to estimate the most informative EEG electrodes "Multi-class Discriminant Analysis (MDA) based Selective Opposition Grey Wolf Optimization (SOGWO)" criterion is applied on the clean EEG activity. A detailed feature set is estimated from the selected EEG channel that includes

time-domain, frequency–domain, and Time-Frequency (TF) domain-based features. More often, such large feature sets lead to redundancy and results in reduced classification accuracy. Therefore, CCA-RFE based feature selection criterion is proposed in the present work for vital features selection. At the final stage, the proposed methodology of EEG channel selection and optimized feature generation is validated by a set of machine learning methods viz. Extreme Learning Machine (ELM), Multi-Layer Perceptron (MLP), Linear Discriminant Analysis (LDA), and Random Forest (RF) tree classifiers. The schematic diagram of the proposed methodology is presented in Figure 4.1.

4.2 Methods

4.2.1 Data Pre-Processing

Since, EEG activity is contaminated by non-cerebral artifact origins, cleaning the recorded EEG activity is the first step towards EEG signal analysis. Therefore, before implementing any feature extraction and classification methods, a suitable pre-processing technique is needed [126]. In this work, Fast power ICA based artifact detection and correction method is utilized for cleaning of recorded EEG activity [127]. The pre-processing pipeline consists of three steps. In the first step, Fast-power ICA (fpICA) based Blind Source Separation (BSS) is performed to decompose artifactual EEG activity into independent sources (i.e., ICs) of artifactual/non-cerebral and cerebral activities [127]. After separation of EEG activity into ICs, automatic identification of artifactual ICs is carried out using the Katz-Fractal-Sparsity (KFS) criterion, in the second step. In the third step, the identified ICs are corrected using General Linear Chirplet Transform (GLCT) based Time-Frequency coefficient suppression criterion and thereafter, inverse ICA is performed on the corrected ICs to attain clean EEG activity [128][121]. Figure 4.2 shows the schematic diagram of GLCT-PICA based EEG artifact correction pipeline.

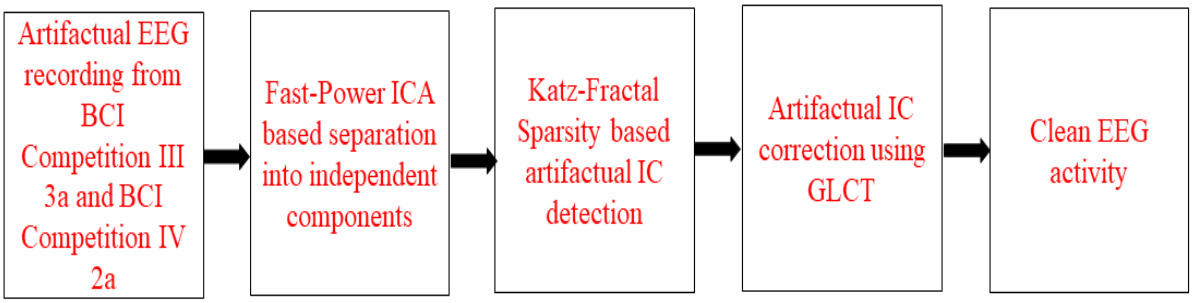


Figure 4.2 Schematic diagram of GLCT-PICA based artifact correction pipeline

4.2.2 Channel Selection using MDA-SOGWO

Multi-channel EEG is extensively used for motor-imagery BCI. However, multi-channel EEG activity leads to redundant information extraction and would lessen the distinction ability among various MI tasks [28]. Therefore,

it is needed to limit number of EEG channels for enhanced usability of modern BCI systems without compromising on classification accuracy. An optimum channel selection criterion reduces processing burden and increases overall efficiency of the system. In this study, MDA-SOGWO based EEG channel selection method is developed and successfully implemented on the multi-channel EEG dataset. In proposed method, Multi-class Discriminant Analysis (MDA) is considered as a part of the objective function. Following section briefly explains Selective Opposition Grey Wolf Optimization (SOGWO) and MDA based EEG channel selection method proposed in the present work:

4.2.2.1 Selective Opposition GWO (SOGWO)

The Selective Opposition Grey Wolf Optimization (SOGWO) algorithm combines Opposition Based Learning (OBL) with GWO and enhances its exploratory nature while providing a fast convergence rate [36]. In GWO algorithm, every iteration includes the best solution generation using α, β, δ and ω wolves; where, ω are the least fit candidates and α are the best fit candidates [130]. The α candidates are closer to optima, followed by β and δ as shown in Figure 4.3. Therefore, considering the possibility that the optimum solution is in opposite position to the ω wolves' position i.e., least fit candidates' position, best fit solution corresponding to α candidates is searched.

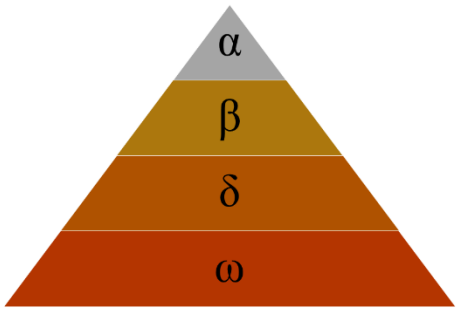


Figure 4.3 Social hierarchy of grey wolves

Opposition Based Learning (OBL) considers the best solution search in opposite/reverse direction of the worst solution using Equation *4.1*.

$$x_{opi} = ub_i + lb_i - x_i \qquad (4.1)$$

here, x_i is the worst solution vector between $[lb_i, ub_i]$ and x_{opi} is opposite direction vector. Figure 4.5 shows SOGWO based optimum solution search strategy for x being the real number in the interval $[a1, a2]$. The opposite number of x viz. x_{op} is given in Equation *4.2*. Similarly, this search strategy can be extended to more than *2* dimensions. In this chapter, the EEG channels corresponding to each subject are optimized by implementing SOGWO search strategy where, ω represents the least fit channel and α denotes best fit channel.

$$x_{op} = a1 + a2 - x \qquad (4.2)$$

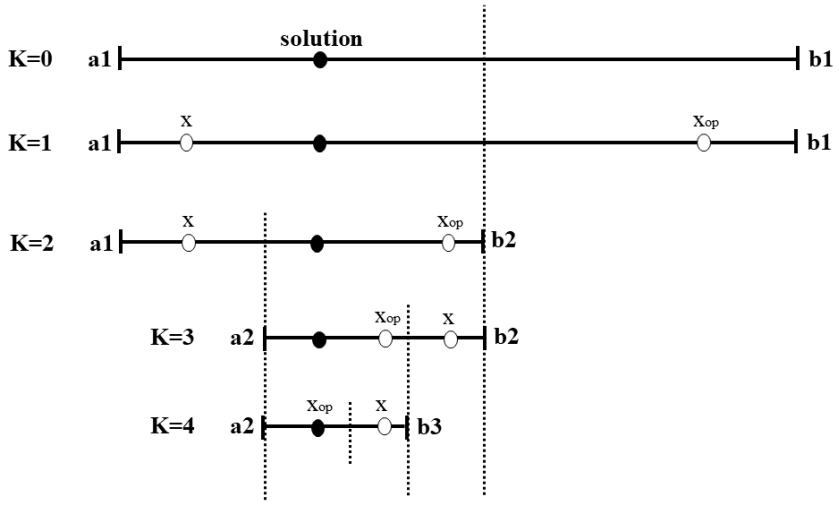


Figure 4.4 Optimum solution search strategy of one-dimensional equation for each iteration (i.e., *K=0* to *K=4*) with respect to current estimate *x* and its opposite *xop* using SOGWO.

Also, SOGWO includes Spearsmen's Correlation Coefficient (SCC), which assists in selecting the wolves corresponding to which dimensions are required to be opposed. After each iteration, the ranked correlation coefficient of ω is calculated with respect to α and if this is less than zero i.e., α and ω have negative correlation, the position of ω is opposed. This provides a direction to the search and tends towards fast convergence without affecting the probability of best solution search.

In order to select the dimensions which are required to be opposed, the threshold variable '*th*' initialized as *d* is linearly decreased with each iteration. For each dimension (k) of the given vector x and α vector (x_α), the difference $df(k)$ is calculated by Equation *4.3* as:

$$df(k) = |x(k) - x_\alpha(k)| \qquad (4.3)$$

If the difference $df(k)$ between ω and k^{th} dimension of α is greater than threshold '*th*', the no. of dimensions g with value greater (g) is incremented by one i.e., $g = g + 1$. however, if $n - g < g$; where *n* is total number of dimensions, the dimensions with $df(k)$ greater than '*th*' are opposed using OBL as given by Equation *4.4*.

$$x_{opi} = ub_i + lb_i - x_i; \qquad \text{where } i \in \{k : df(k) \leq th\} \qquad (4.4)$$

If selected opposite vectors add up in optimization only then, these vectors are selected otherwise vectors are kept as they were before. The advantage of SOGWO technique as compared to other optimization techniques is that it

does not perform opposition in all dimension of ω wolf but only in those which are with difference more than threshold '*th*'. Figure 4.5 shows the schematic diagram of the SOGWO methodology.

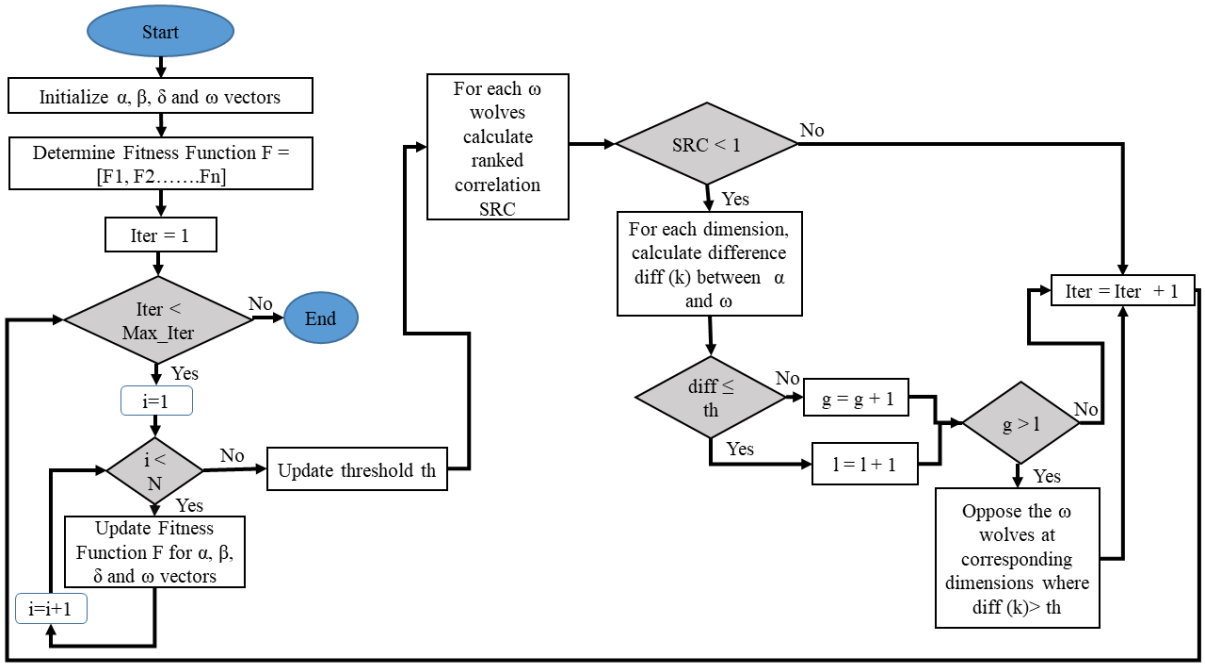


Figure 4.5 Flow diagram of the SOGWO methodology

4.2.2.2 Multi-class Fisher's LDA (MDA)

Binary Fisher's LDA is one of the popular classification algorithms in the BCI domain due to its excellent discrimination capability [131]. It can effectively project high-dimensional data towards one direction and then perform linear classification [132]. In binary Fisher's LDA, maximum separation is searched between two classes for optimum projection [133]. In this study, a new four-class Fisher's Linear Discriminant Analysis (LDA) criterion is proposed as an objective function and referred as multi-class Fisher's LDA (MDA). The developed MDA acts as the objective function in SOGWO based optimization to remove irrelevant and redundant channels from the available dataset. In binary Fisher's LDA, the separation between two classes is expressed by Equation *4.5*.

$$FDA = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \qquad (4.5)$$

here, μ_1, μ_2 and σ_1^2, σ_2^2 denotes mean and variance values of the feature over all channels corresponding to class-1 and class-2 respectively. The higher value of Fisher's LDA corresponding to each channel denotes the high discrimination ability of that channel between the classes. In this work, the log-power of each time segment is considered as a feature for which FDA is computed [72]. Also, the concept of binary class FDA is extended to

multi-class FDA by computing FDA between each of two classes' viz. left hand-right hand, left hand- tongue, right hand-tongue, right hand-feet, tongue-feet, and feet-left hand and denoted as FDA_1, FDA_2, FDA_3, FDA_4, FDA_5, and FDA_6, respectively. After that, the mean of all the computed LDAs corresponding to each channel is generated using Equation *4.6*. Also, the multi-class FDA is computed using one vs rest strategy along with above mentioned one vs one strategy. In one vs rest strategy, the FDA of each class is computed with rest of the classes and mean of all the computed LDAs with respect to each channel is generated using Equation *4.7*.

$$MDA_{ovo} = \sum_{n=1}^{6} FDA_{n1} / 6 \quad (4.6)$$

$$MDA_{ovr} = \sum_{n=1}^{4} FDA_{n2} / 4 \quad (4.7)$$

In this work, SOGWO criterion is implemented using MDA (as a part of the objective function) as given by Equation *4.8*. The proposed objective function in Equation *4.8* is iterated over *500* times for optimum EEG channel selection in this study.

$$Objective\ function = k_a * mean(MDA) + k_b * \frac{c}{N};\ \ k_a = 0.05\ \&\ k_b = 0.95 \quad (4.8)$$

here, k_a and k_b parameters denote predefined normalized weight functions for MDA and the selected number of channels respectively, and $k_a + k_b = 1$ [41] . Different weights were allotted to the features depending upon their significance. Multi-class Discriminant Analysis (MDA) represents discriminant value of channel for different classes. The mean value corresponding to all the selected channels is utilized in objective function and $\frac{c}{N}$ denotes the ratio of number of selected channels using SOGWO to the total number of channels. The impartial function tries to maximize MDA with minimum number of selected EEG channels. This objective function finds the best solution for the α wolves. The objective function will have values in between *[0, 1]*, which leads to the global best α wolves' position.

4.2.3 Feature Extraction

This section provides detailed description of extracted features and feature-set construction from selected EEG channels. The EEG activity is non-linear, non-stationary time-series data and more often the time-domain based feature extraction methods limit the amount of information from on-going EEG activity [74]. In previous studies, various non-linear feature extraction methods have been proposed to transfer non-linear dynamical information of EEG signals [135]. In this work, with intentions to carry forward the concept of non-linearity and dynamical information estimation, time-domain, frequency-domain and Time-Frequency (TF) domain-based EEG features

have been extracted. Also, much commonly used feature estimators viz. Power Spectral Density (PSD), entropy, hjorth parameters such as mobility, complexity and activity have been included in the generated feature set.

4.2.3.1 Time-domain and Frequency-domain Features

The various time-domain and frequency-domain based features estimated from the selected EEG channels are presented in Table 4.1.

4.2.3.2 Time-Frequency Features

Wavelet Transform (WT) has been extensively used for Time-Frequency (TF)-based feature estimation of non-stationary signals in the past studies [136]. However, WT and its extended methods have various limitations, for instance, low-resolution at high frequencies and shift-variance, to count a few. The various limitations of classical WT are overcome by its latest extension viz. Flexible Analytical Wavelet Transform (FAWT) [45].

Flexible Analytical Wavelet Transform (FAWT)

Flexible Analytical Wavelet Transform (FAWT) employs iterative filter banks that include one low pass (f_l) and two high *pass* (f_h) filters along with flexibly controlling the filter parameters viz. redundancy (r), dilation factor, Quality Factor (Q_F). High pass filter (f_h) splits the signal into positive and negative frequencies [138]. The up-sampling and down-sampling parameters corresponding to f_l and f_h are denoted by s_1, s_2 and s_3, s_4, respectively [139]. The frequency response of f_l and f_h are expressed in Equation *4.9* and Equation *4.10* respectively.

$$H(w) = \begin{cases} \sqrt{s_1 s_2} & |w| < w_p \\ \sqrt{s_1 s_2}\,\theta\left[\frac{w-w_p}{w_s-w_p}\right] & w_p \le w \le w_s \\ \sqrt{s_1 s_2}\,\theta\left[\frac{\Pi-w+w_p}{w_s-w_p}\right] & -w_s \le w \le -w_p \\ 0 & |w| > w_s \end{cases} \tag{4.9}$$

$$G(w) = \begin{cases} \sqrt{s_3 s_4}\,\theta\left[\frac{\Pi-w+w_0}{w_1-w_0}\right] & w_0 \le w \le w_1 \\ \sqrt{s_3 s_4} & w_1 < w < w_2 \\ \sqrt{s_3 s_4}\,\theta\left[\frac{w+w_0}{w_1-w_0}\right] & w_2 \le w \le w_3 \\ 0 & w \epsilon [-\Pi, w_0\,] \cup [w_3\,, \Pi] \end{cases} \tag{4.10}$$

where, $w_p = \frac{1-\beta}{s_1}\pi + \frac{\epsilon}{s_1}$, $w_s = \frac{\pi}{s_2}$, $w_0 = \frac{1-\beta}{s_3}\pi + \frac{\epsilon}{s_3}$,

$$w_1 = \frac{s_1}{s_2 s_3}\,\Pi, \quad w_2 = \frac{\Pi}{s_3} - \frac{\epsilon}{s_3}\,, \quad w_3 = \frac{\Pi}{s_3} + \frac{\epsilon}{s_3}, \epsilon = \frac{1}{32}\left(\frac{s_1 - s_2 + \beta s_2}{s_1 + s_2}\right)\pi$$

here, β and ϵ denote non-negative constants related to perfect reconstruction, here $\beta < 1$. For optimum FAWT bases selection i.e., s_1, s_2, s_3, s_4 and β, values of these parameters are iterated in the range of $s_1(1 \leq s_1 \leq 10)$, $s_2(1 \leq s_2 \leq 10)$, $s_3(1 \leq s_3 \leq 10)$, $s_4(1 \leq s_4 \leq 10)$, while satisfying the Equation *4.11* as follows:

$$\beta = m.\frac{s_3}{s_4}; \qquad 0 < m \leq 1 \tag{4.11}$$

Flexible Analytical Wavelet Transform (FAWT) is unable to separate lower rhythmic activities when the number of decomposition levels *j* is kept below *5*. Whereas keeping *j* above *5*, tends to add more complexity to the methodology instead of improving the results significantly. Therefore, value of *j* is kept as *5* in this study. After FAWT decomposition, features are estimated from decomposed EEG activity and a feature set is generated. Table *4.1* shows the TF and TF based non-linear dynamical features estimated in this work. The combined feature set provides enhanced capability to extract hidden information in the ongoing EEG activity.

Table 4.1 List of estimated Time-domain/ Frequency-domain and Time-Frequency domain/ Non–Linear Dynamical Features

Time/Frequency domain Features [71] [66]	
Time-domain features	1. Kurtosis 2. Skewness 3. Mobility 4. Complexity 5. Activity 6. Maximum 7. Minimum 8. Mean 9. Standard Deviation
Frequency-domain features	1. Mean frequency 2. Median Frequency 3. Spectral Centroid 4. Spectral Spread 5. Spectral flatness 6. Spectral Crest
Time-Frequency/ Non–Linear Dynamical Features [140] [74]	
Time-Frequency Domain Features	1. Mean Absolute Value (MAV) 2. Sample Sign Integral (SSI) 3. Root Mean Square (RMS) 4. Power Spectral Density (PSD) 5. Peak to Peak Mean (P2P) 6. Peak to RMS Mean (P2R).
Non–Linear Dynamical Features	1. Renyi's Entropy (Renyi_En) 2. Tsallis Entropy (Tsalli_En) 3. Approximate Entropy (App_En) 4. Distribution Entropy (Dstrbn_En) 5. Sample Entropy (Samp_En) 6. Fuzzy Entropy (Fuzzy_En) 7. Higuchi's Fractal Dimension (Higu_FD) 8. Katz Fractal Dimension (Katz_FD) 9. Fuzzy Entropy (Fuzzy_En)

4.2.4 Feature Selection using CCA-RFE

Though, feature extraction stage helps in extracting hidden information from the ongoing EEG activity, it may yield redundant and irrelevant features as well. Taking the redundant features to the classification stage leads to higher complexity and reduced classification performance. Therefore, feature selection step is introduced in the processing pipeline to optimize the estimated features. In literature, various feature selection methods have been proposed for instance Locally Robust Feature Selection (LRFS) [141], CSP based feature selection [142], Universum Support Vector Machine-Recursive Feature Elimination (USVM-RFE) [143], and maximum relevance CCA [71]. In this study, an amalgamation of two feature selection criteria viz. Canonical Correlation Analysis (CCA) and Recursive Feature Elimination (RFE) is proposed and implemented on the BCI Competition

dataset. The base technique viz. canonical correlation between the features is computed while reducing the features using RFE.

4.2.4.1 Canonical Correlation Analysis- Recursive Feature Elimination (CCA-RFE)

This methodology is developed to eliminate redundant EEG features from estimated feature set. The developed algorithm of CCA-RFE based EEG feature selection is given in *Algorithm-1* section. Canonical Correlation Analysis-Recursive Feature Elimination (CCA-RFE) methodology is an iterative procedure that focuses on maximizing the correlation between selected feature vectors and feature sets via recursively eliminating redundant features. Canonical Correlation Analysis (CCA) computes the canonical correlation between the selected feature vector and already selected feature subset using Equation *4.12* as follows:

$$CCA = \rho_{CCA}(x_j, S_{k-1}) \qquad (4.12)$$

Algorithm-1: CCA- RFE

F= Selected feature set corresponding to one channel
S= [1, 2, …. N] *[N is no. of features]*
R= [] *[Feature ranked list]*
for *S*=1: N
    W[*s*] = CCA(*F*[*k-1*], *F*[*S*]) *[here, k-1 indicate all features except S]*
end for
$C_i = (W_i)^2$ for *i* =1: N
L=argmin(*W*) [*Features with smallest rank]*
R =(*S*[*L*], *R*) *[Update feature ranked list R]*
S= [1: L-1, L+1: length(S)] [Eliminate features with smallest rank]
Output: ***R*** *[feature ranked list]*

here, x_j is the candidate feature and S_{k-1} denotes already selected feature subset with *k-1* features. The RFE algorithm [54] is defined as a procedure that recursively assigns a weight to each feature vector based on the various criteria such as efficiency, variance, and correlation [143]. In this work, depending upon the CCA value between feature set and each feature vector, weights are assigned to each feature vector. After computing weights for each feature vector, the assigned weights are used to rank each feature vector using RFE algorithm. The feature vector with lowest rank (i.e., highest rank number) viz. rank-15 and rank-14 are selected for optimum feature-set, the feature with highest rank number represent least correlation and have high discrimination ability. Whereas, the features with lowest rank viz. rank-1 tend to have maximum correlation between feature vectors of various classes and acts as redundant features.

4.2.5 Classification Methods

The classification stage helps to validate the developed methodology of EEG channel selection, feature extraction and optimum features selection. The classifiers used in the present work include Multi-Layer Perceptron (MLP), Linear Discriminant Analysis (LDA), Random Forest (RF), and Extreme Learning Machine (ELM). Multi-Layer Perceptron (MLP) classifier is a feed-forward neural network-based classifier with one input layer, more than two hidden layers and one output layer with specific functions [144]. The input layer is responsible for receiving features i.e., one per neuron and each hidden layer is responsible for the certain illustration of the problem [145]. The output layer gets stimulus from hidden layers and generates predicted class linearly[146]. The number of hidden layers is selected after computing classification accuracy for complete data trials received from all subjects. In this work, *200* hidden layers with *'lbfgs'* optimizer are used for the classification task. On contrary, the Linear Discriminant Analysis (LDA) classifier draws the decision region between four classes while optimizing the ratio of inter-class data to intra-class scatter [65]. The data separation is carried out using a linear solution. The parameters of a discriminant function are obtained using training feature set. In Random Forest (RF) classifier, the class is predicted depending upon the random vectors [147]. In this study, the best performance in terms of accuracyis obtained corresponding to *100* forest trees. Conversely, Extreme Learning Machine (ELM) is an extension of a feed-forward neural network with a single hidden layer. In ELM, the classification problem is considered as a linear system and Moore-Penrose based inverse operations are employed to predict output weights [138]. It has a single hidden layer with the fastest learning capability. In ELM, parameters used to obtain optimum results are number of hidden layers = *50*, utilization function = '*multiquadric*', alpha = *2* and *rbf_width =3*.

4.3 Proposed Framework

The proposed methodology is implemented on two multi-class MI-BCI datasets viz. BCI Competition IV 2a and BCI Competition III 3a. First, the EEG activity is pre-processed using earlier developed GLCT-PICA EEG artifact removal methodology [148]. Thereafter, EEG channel optimization is performed using developed MDA-SOGWO method, which helps to remove redundant EEG channels and improve the overall efficiency and performance of the BCI system. From the subject specific, optimum EEG channels, a set of time, frequency, and Time-Frequency (TF) features are estimated. The estimated feature sets are analysed independently to compare the efficacy of features. Thus, the estimated feature set consists of *15 features/EEG channel* for time and frequency and same is the case for TF features as well. It is evident that the large feature set will lead to redundancy of information at the stage of classification.

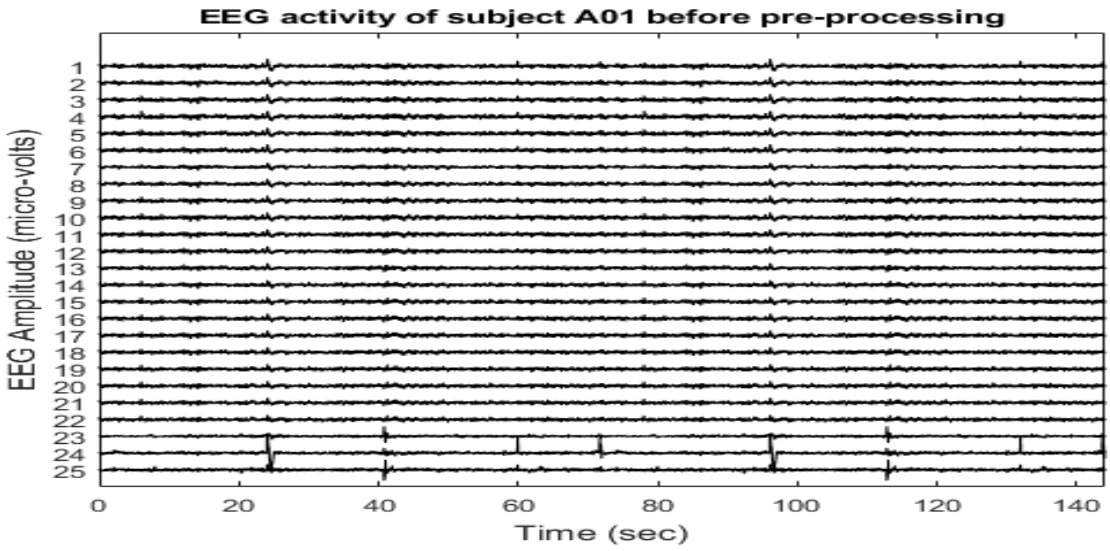
Algorithm-2 Proposed Approach
for *i* = 1: Number of channels **do**
Clean EEG activity using GLCT-PICA methodology

```
    Compute alpha_pos using MDA-SOGWO
    if alpha_pos(i) == 1
      for k = 1: Number of Trials do
          Compute wavelet coefficients M(k, j)
          for j=1: J+1 do      [here, J is no. of decomposition levels of FAWT]
              Compute time and frequency and time-frequency features of M(k, j)
          end for
      end for
    end if
    for l = 1: number of features: length (M (k, j))
          Apply CCA-RFE and select features with highest rank number viz. rank-15 and rank-14
      end for
end for
Compute efficiency using MLP, RF, LDA and ELM Classifiers
```

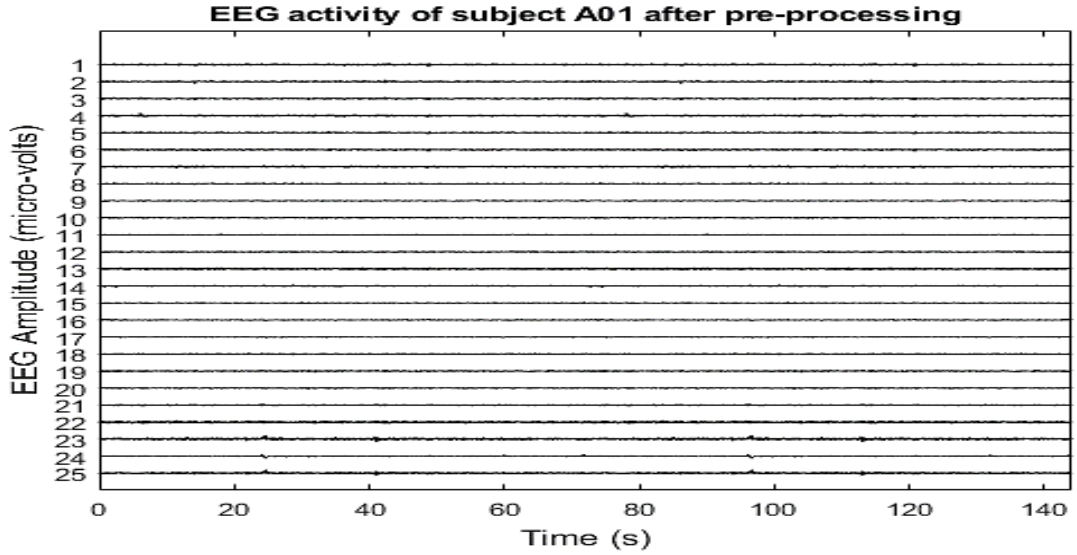
Therefore, only vital features are selected from each feature set by following proposed feature selection criterion i.e., CCA-RFE. The developed methodology recursively eliminates the features depending upon their canonical correlation with the remaining feature set and allots a rank to each feature vector. The two highest rank number features viz. rank-15 and rank-14 are considered to correspond to each channel for further processing. The selected features are further passed to different classifier algorithms for presenting a comparative study at the classification stage. The steps involved in the pipeline of complete methodology are summarized in *Algorithm-2* section.

4.4 Results and Discussion

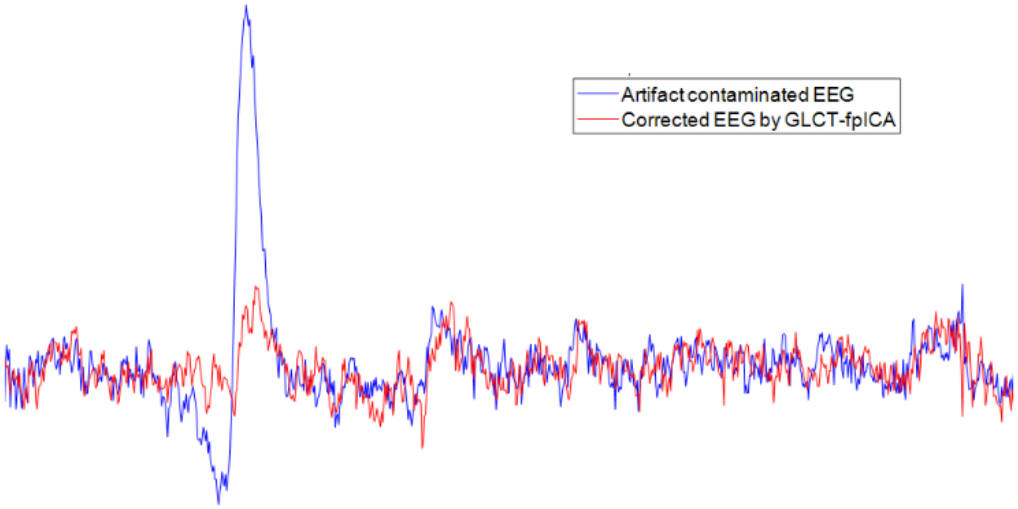
The generalizability of MDA-SOGWO based channel optimization and CCA-RFE based feature selection methodology is investigated on two multi-channel datasets. First, both raw EEG datasets are pre-processed using GLCT-PICA processing pipeline. Two snippets showing raw EEG activity and corresponding processed EEG activity using GLCT-PICA method are given in Figure 4.6(a) and 4.6(b). Figure 4.6(c) shows the zoomed-in representation of original EEG activity and artifact-free EEG activity. The processed and clean EEG activity is further passed through channel selection criterion i.e., MDA-SOGWO and the EEG channels maximally contributing for discriminative information are selected for each subject independently. Figure 4.7 shows the selected EEG channels for subject "*A01*" of BCI Competition IV 2a dataset.


EEG activity of subject A01 before pre-processing
EEG Amplitude (micro-volts)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
0
20
40
60
80
100
120
140
Time (sec)

(a)


EEG activity of subject A01 after pre-processing
EEG Amplitude (micro-volts)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
0
20
40
60
80
100
120
140
Time (s)

(b)



(c)

Figure 4.6 EEG activity record of subject "*A01*" (a) Raw EEG activity (b) Processed EEG activity (c) Zoomed In view of original EEG w.r.t artifact corrected EEG

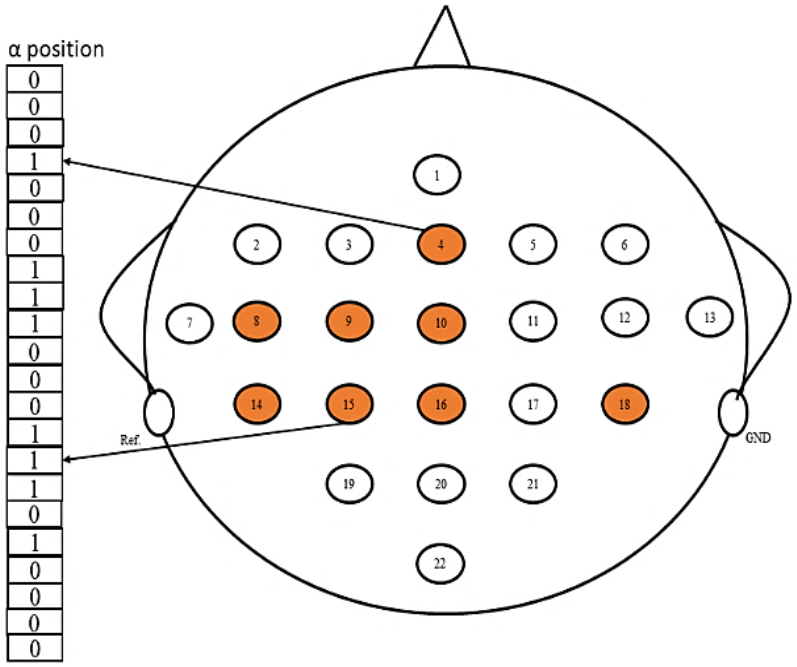


Figure 4.7 Selected EEG channels for subject "*A01*" based on alpha position using MDA-SOGWO channel selection methodology

After pre-processing and EEG channel selection, the feature set is generated using the time, frequency and TF based feature extraction methodology and thereafter, the vital feature set is selected out of generated features using CCA-RFE criterion. Table 4.2 shows selected EEG features for subject "*A01*" estimated from selected EEG channels. Similarly, the other feature sets are generated corresponding to selected EEG channels for all subjects and employed to train the classifier algorithms separately. The generated feature set is divided into training and testing sets in the *70:30* ratio. The proposed work is carried out on a desktop machine having *Intel(R) Core™ i5-*

5200U CPU @2.20GHz with 8 GB RAM. The overall developed methodology of EEG signal processing and analysis is implemented in *MATLAB 2020a*.

Table 4.2 Selected features for subject "A01" using proposed feature selection method

	Subject-A01 rank-wise selected feature							
Rank	**Channel-4**	**Channel-8**	**Channel-9**	**Channel-10**	**Channel-14**	**Channel-15**	**Channel-16**	**Channel-18**
1	**PSD**	**Katz_FD**	**App_En**	**MAV**	**App_En**	**SSI**	**Renyi_En**	**Dstrbtn_En**
2	**SSI**	**Fuzzy_En**	**P2P**	**Renyi_En**	**Tsalli_En**	**Dstrbtn_En**	**Tsalli_En**	**Samp_En**
3	P2P	SSI	P2R	P2P	Fuzzy_En	PSD	App_En	Higu_FD
4	Fuzzy_En	P2P	Tsalli_En	Katz_FD	P2R	Katz_FD	Katz_FD	Katz_FD
5	MAV	RMS	SSI	SSI	P2P	Higu_FD	Higu_FD	Tsalli_En
6	RMS	P2R	MAV	Fuzzy_En	Renyi_En	Tsalli_En	Samp_En	App_En
7	P2R	Log_En	RMS	Dstrbtn_En	SSI	Renyi_En	Dstrbtn_En	Renyi_En
8	Log_En	PSD	Higu_FD	Log_En	RMS	Samp_En	PSD	Log_En
9	Samp_En	Renyi_En	Renyi_En	PSD	Katz_FD	P2R	Log_En	PSD
10	Dstrbtn_En	Dstrbtn_En	PSD	Samp_En	Dstrbtn_En	App_En	P2R	Fuzzy_En
11	Katz_FD	Higu_FD	Dstrbtn_En	Higu_FD	MAV	Fuzzy_En	MAV	MAV
12	App_En	Katz_FD	Katz_FD	Tsalli_En	Samp_En	Log_En	P2P	SSI
13	Tsalli_En	Tsalli_En	Fuzzy_En	App_En	PSD	RMS	Fuzzy_En	P2R
14	Higu_FD	MAV	Samp_En	P2R	Higu_FD	MAV	RMS	RMS
15	Renyi_En	Samp_En	Log_En	RMS	Log_En	P2P	SSI	P2P

***P2P**: Peak to Peak ratio; **Fuzzy_En**: Fuzzy Entropy; **MAV**: Mean Absolute Value; **RMS**: Root Mean Square; **P2R**: Peak to RMS ratio; **Log_En**: Log Entropy; **Samp_En**: Sample Entropy; **Dstrbtn_En**: Distribution Entropy; **Katz_FD**: Katz Fractal Dimension; **App_En**: Approximate Entropy; **Tsalli_En**: Tsalli's Entropy; **Higu_FD**: Higuchi's Fractal Dimension; **Renyi_En**: Renyi's Entropy*

In the classification stage, a wide set of classifier algorithms is used to validate the proposed EEG channel selection, feature extraction and optimum features selection methodology. Four classifiers are implemented in the present study including MLP, LDA, RF, and ELM. The classification performance obtained for each subject of BCI Competition IV 2a and BCI Competition III 3a datasets is shown in Figure 4.8(a), (b) and (c), (d), respectively. In Figure 4.8(a) and Figure 4.8(b), the classification results of BCI Competition IV 2a dataset corresponding to the two strategies viz. one vs one and one vs rest respectively are presented. Similarly, in Figure 4.8(c) and Figure 4.8(d), the classification results of BCI Competition III 3a dataset corresponding to the two strategies viz. one vs one and one vs rest respectively are shown. It can be deduced from the classification results that most of the subjects outperformed in multi-class classification of MI-BCI tasks. The classification performance corresponding to all the four classifiers is comparable. Also, Table 4.3 shows comparison of one vs one strategy with one vs rest for BCI Competition IV 2a dataset and BCI Competition III 3a dataset utilizing time + frequency features and Time-Frequency (TF) features. For a detailed analysis of the classification performance, other performance measures such as Specificity (*Spec*), Sensitivity (*Sens*), G-measure (*G-meas*), Precision (*Prec*), and F1-Score (*F1-sc*) have been computed. Here, *G-meas* represents the geometric mean of *Sens* and *Spec* values. The expressions of estimated performance measures are given as follows:

$$Accuracy\ (Accy) = \frac{TN+TP}{TN+TP+FN+FP} \quad (4.13)$$

$$Sensitivity\ (Sens) = \frac{TP}{TP+FN} \quad (4.14)$$

$$Specificity\ (Spec) = \frac{TN}{TN+FP} \quad (4.15)$$

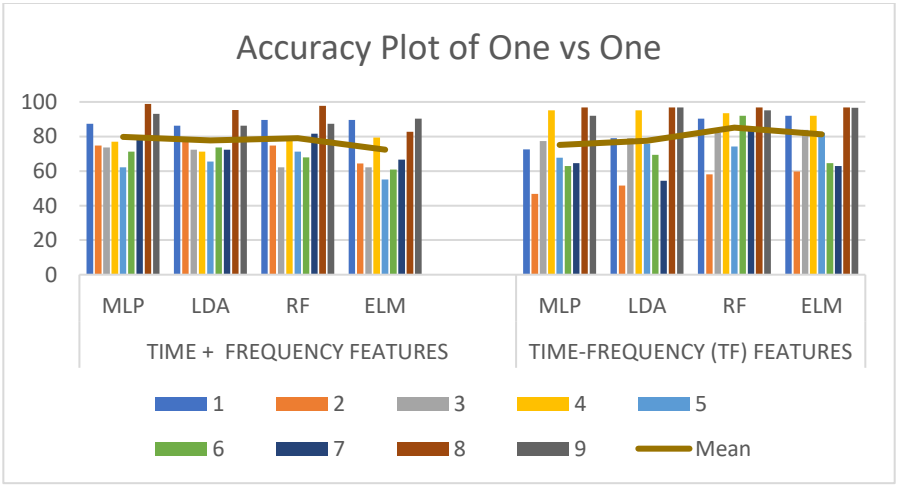
$$G-measure\ (G-meas) = \sqrt{Sensitivity * Specificity} \quad (4.16)$$

$$Precision\ (Prec) = \frac{TP}{TP+FP} \quad (4.17)$$

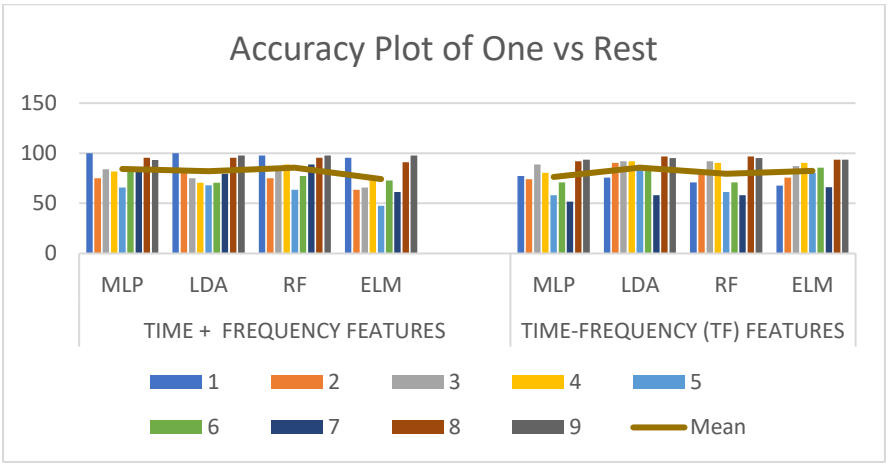
$$recall = \frac{TP}{TP+FN} \quad (4.18)$$

$$F1-Score\ (F1-sc) = 2 * \frac{Precision*recall}{Precision+recall} \quad (4.19)$$

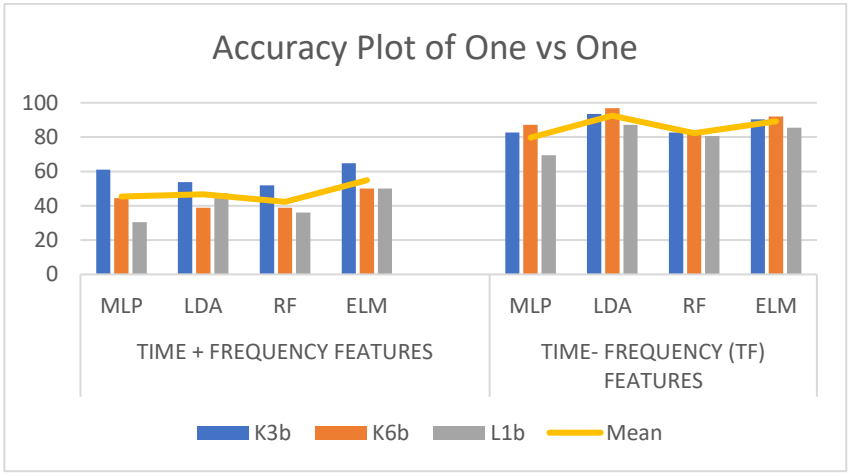
here, *TP* shows a True-Positive class, TN is True-Negative, FP is False-Positive, and *FN* shows a False-Negative class. In addition, the t-distributed Stochastic Neighbor Embedding (SNE) plot is utilized to describe the discrimination between extracted features from different classes of the dataset. The t-SNE plots obtained from subjects "*A04*", "*A05*", "*A08*" establish clear discrimination among feature set of left-hand and right-hand movement classes as shown in Figure 4.9. However, a clear hyperprint cannot be drawn between feature set of tongue and feet movement classes.



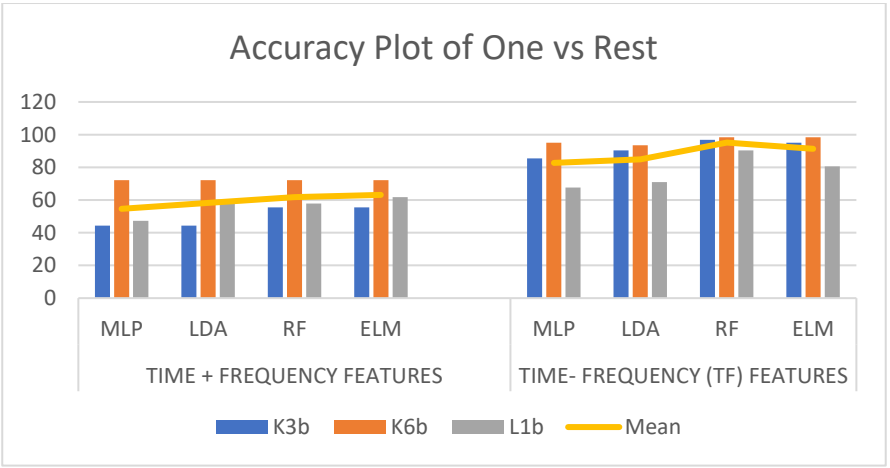
(a)



(b)



(c)



(d)

Figure 4.8 Classification accuracy obtained for all subjects of (a) BCI Competition IV 2a dataset for one vs one strategy (b) BCI Competition IV 2a dataset for one vs rest strategy (c) BCI Competition III 3a dataset for one vs one strategy (d) BCI Competition III 3a dataset for one vs rest strategy

Table 4.3 Mean Classification accuracy comparison of both datasets for one vs one and one vs rest corresponding to all classifiers

Dataset	Time + Frequency Features				Time-Frequency (TF) Features			
	One vs One							
	MLP	**LDA**	**RF**	**ELM**	**MLP**	**LDA**	**RF**	**ELM**
BCI Competition IV 2a	79.81	77.77	78.92	72.35	75.08	77.54	**85.11**	81.15
BCI Competition III 3a	45.35	46.6	42.28	54.93	79.67	**92.46**	82.33	89.24
	One vs Rest							
BCI Competition IV 2a	84.33	82.06	85.60	74.23	76.33	**85.65**	79.41	82.25
BCI Competition III 3a	54.63	58.17	61.87	63.17	82.79	84.94	**95.15**	91.39

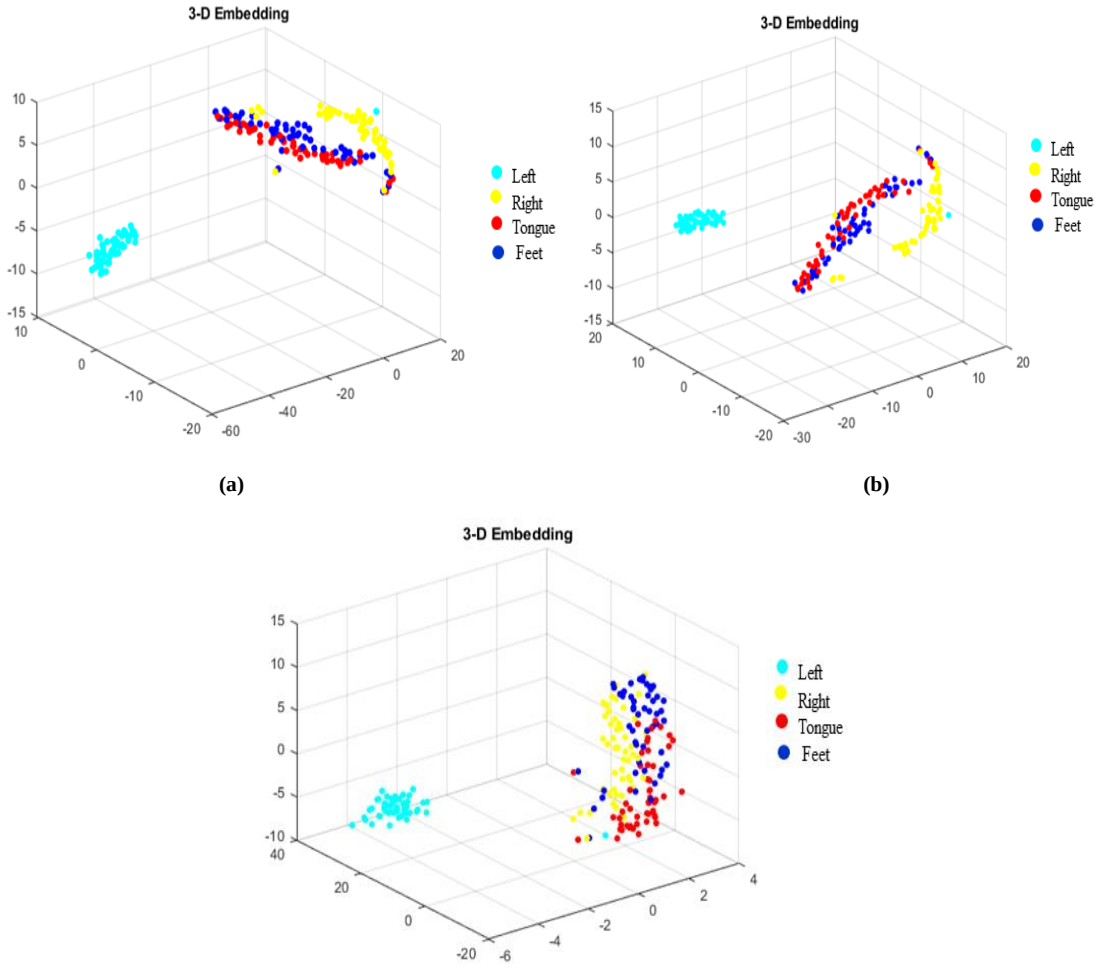


Figure 4.9 t-SNE plot estimated for subjects (a) *A04* (b) *A05* (c) *A08* of BCI Competition IV 2a dataset

Table 4.4 Performance measures obtained for subjects before applying proposed methodology of EEG channel and feature selection

BCI Data	Subject No.	*Accy*	*Spec*	*Sens*	*G-meas*	*Prec*	*F1-sc*
	A01	60.91	86.62	62.61	73.64	62.07	61.18
	A02	50.57	83.45	51.73	65.71	51.74	50.51

	A03	42.52	80.86	41.72	58.08	43.70	42.96
	A04	78.16	92.79	79.70	86.00	81.90	77.43
	A05	67.81	89.31	69.55	78.81	69.77	67.70
	A06	52.87	84.52	54.02	67.57	55.46	52.66
BCI Competition IV 2a	A07	65.51	88.45	65.31	76.01	66.26	64.96
	A08	94.25	98.11	94.70	96.39	94.25	94.26
	A09	90.80	96.91	90.40	93.60	90.77	90.72
BCI Competition III 3a	K3b	26.85	75.91	27.96	46.07	28.52	26.59
	K6b	34.72	78.12	35.37	52.57	35.36	34.67
	L1b	33.33	77.67	33.33	50.88	35.08	33.74

Accy: Accuracy; ***Spec***: Specificity; ***Sens***: Sensitivity; ***G-meas***: G-measure; ***Prec***: Precision; ***F1-sc***: F1-Score.

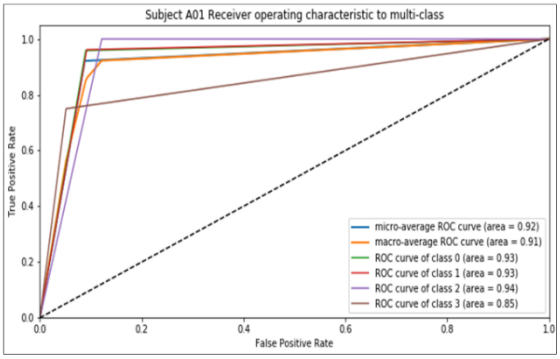
Table 4.5 Performance Measure of various subjects after applying the proposed approach of EEG channel and feature selection

BCI Data	**Subject**	***Accy***	***Spec***	***Sens***	***G-meas***	***Prec***	***F1-sc***
	A01	91.93	97.31	91.77	94.50	92.24	91.85
	A02	59.67	86.56	59.58	71.81	59.19	59.28
	A03	83.87	94.61	83.75	89.01	83.94	83.78
	A04	91.93	97.32	91.87	94.56	92.13	91.98
	A05	82.25	94.07	82.29	87.98	81.89	82.03
	A06	64.51	88.15	64.16	75.21	66.01	64.89
BCI Competition IV 2a	A07	62.90	87.65	62.50	74.01	62.43	62.31
	A08	96.77	98.92	96.66	97.78	96.96	96.70
	A09	93.54	97.87	93.33	95.55	94.10	93.19
BCI Competition III 3a	K3b	90.32	96.76	90.41	93.53	90.83	90.34
	K6b	91.35	97.24	91.87	94.54	92.47	92.01
	L1b	85.48	95.16	85.52	90.21	85.84	85.52

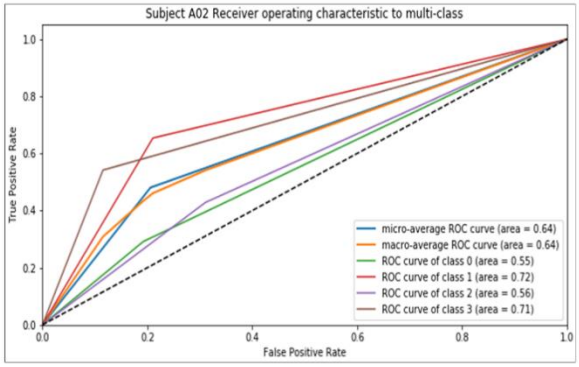
Accy: Accuracy; ***Spec***: Specificity; ***Sens***: Sensitivity; ***G-meas***: G-measure; ***Prec***: Precision; ***F1-sc***: F1-Score.

In order to evaluate the efficacy of developed EEG channel selection (MDA-SOGWO criterion) and optimum features selection (CCA-RFE criterion) methodologies, the classification performance measures are estimated for both cases i.e., applying and without applying the two methodologies before training of the classifier. The Table 4.4 shows the classification performance measures obtained for training the classifier without applying the complete methodology. The classification performance measures computed for the complete methodology of EEG channel selection, feature extraction and optimum features selection are presented in Table 4.5. On closely observing the Table 4.4 and Table 4.5, it is deduced that the proposed methodology of EEG channel and feature selection plays an important role in attaining better classification performance.

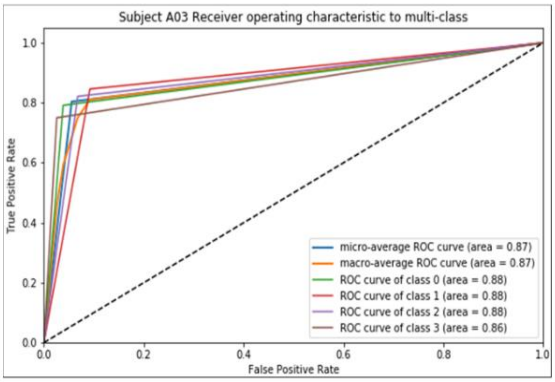
The classification performance achieved for each subject is analysed independently. The ROC curve generated for various subjects (nine subjects in BCI Competition IV 2a dataset and three subjects in BCI Competition III 3a dataset are shown in Figure 4.10. It can be inferred from Figure 4.10(a-i) that most of the subjects are efficiently supporting in discrimination of four class MI-BCI EEG activity. However, in Figure 4.10(j-l), there is an increase in discrimination of multi-class MI-BCI activity for most of the subjects.



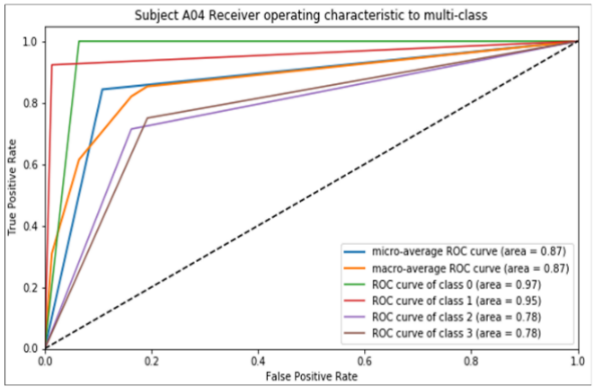
(a)



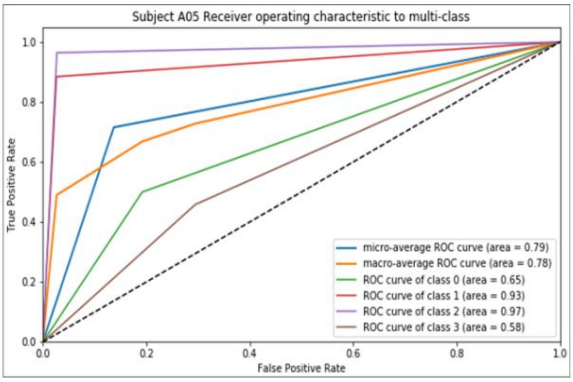
(b)



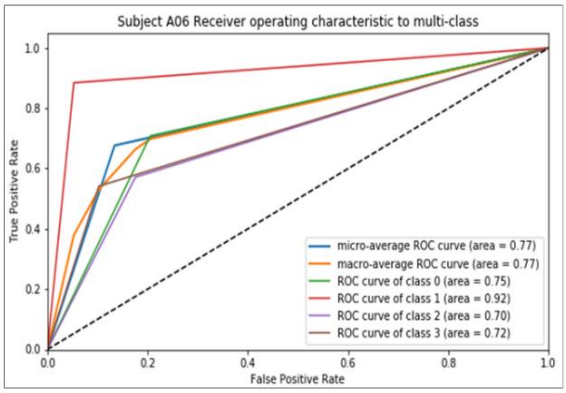
(c)



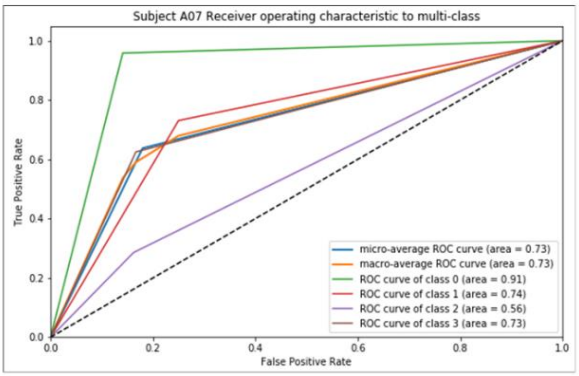
(d)



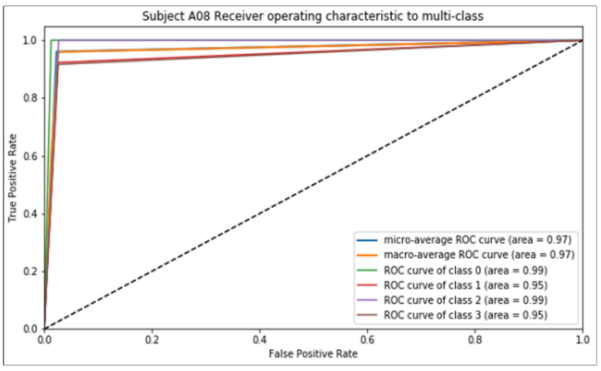
(e)



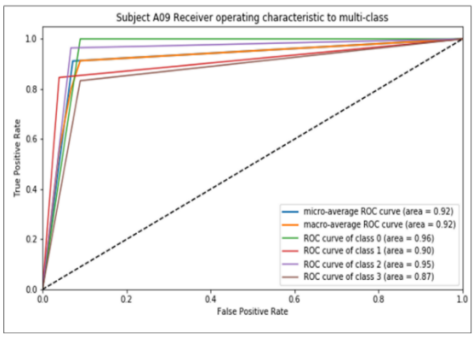
(f)



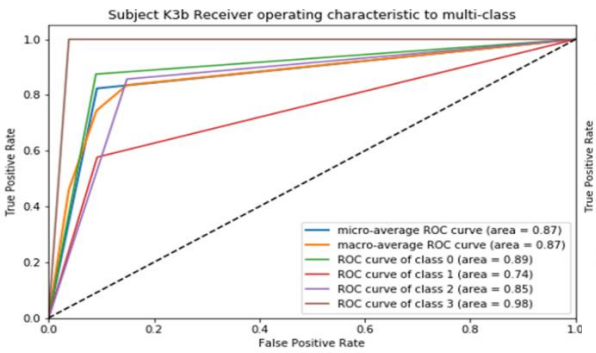
(g)



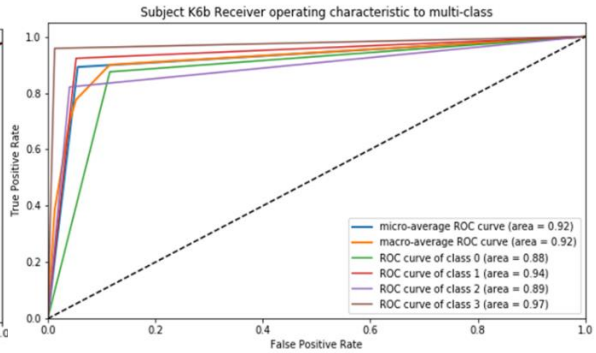
(h)



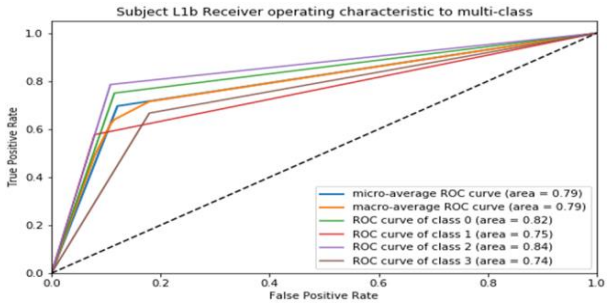
(i)



(j)



(k)



(l)

Figure 4.10 (a)-(i) ROC plots for all Subjects of BCI Competition IV 2a dataset **(j)-(l)** BCI Competition III 3a dataset

A comparison of developed methodology with recently proposed techniques in literature on same dataset is provided in Table 4.6. In [149], an amalgamation of Neural Structured Learning (NSL) and EEGNet architecture was proposed to design an MI-BCI system. Electroencephalography Topographical Representation (ETR) based CNN technique was suggested by Xu [150] for a multi-class MI-BCI classification task. Olias *et al.* [151] suggested Gaussian shrinkage LDA and nCSP based criterion to differentiate between different MI-BCI tasks. In [64] and [81], K-means clustering based criterion and Phase State Reconstruction (PSR) criterion were proposed respectively and the dynamical behaviour of brain activities was analysed. It can be deduced from Table 4.6 that foe the 6 out of 9 subjects, accuracy obtained by developed methodology outperforms state-of-art. In present work, the individual accuracy obtained is also highest for most of the subjects and is found to be effective for multi-class MI-BCI classification task.

In order to evaluate the efficacy of MDA-SOGWO channel selection methodology, a comparison is made with state-of-art EEG channel selection techniques. To maintain the uniformity in the comparison results, same feature set is prepared from the EEG channels obtained after implementing the channel selection techniques. It is inferred from the comparison results presented in Table 4.7 that the proposed MDA-SOGWO based channel selection methodology is efficient for selecting optimum number of EEG channels and provides higher classification performance for the estimated features. This study provides an efficient way to develop a composite methodology for non-redundant channel selection and optimum feature set selection for efficient multi-class MI-BCI task classification. Developed methodology outperformed over other competing channel selection techniques. The limitation of proposed methodology is the size of dataset. In the follow-up studies, the developed methodology can be tested on more available datasets. Another advantage that the present methodology posed includes development of separate modules for channel selection and feature extraction, which can be applied or skipped as per the application requirement. In addition, the proposed channel selection criterion can be applied for other applications as well viz., neurological disease diagnosis using EEG signals.

Table 4.6 Accuracy comparison with literature for BCI Competition IV 2a and III 3a

BCI Competition IV 2a				
Subject	**NSL-EEGNet** [149]	**ETRCNN**[150]	**nCSP + Gaussian Shrinkage LDA** [151]	**Proposed Methodology**
A01	82.29	88.02	89.23	**90.32**
A02	51.39	73.21	**76.15**	58.06
A03	85.07	**90.68**	90.60	82.25
A04	67.01	81.71	71.38	**93.5**
A05	58.33	65.63	59.82	**74.19**
A06	56.25	91.72	63.26	**91.93**
A07	83.33	85.47	**91.70**	83.87
A08	73.96	91.67	89.18	**96.77**
A09	78.47	93.0	85.26	**95.16**
Mean	70.68	84.57	79.62	**85.11**
BCI Competition III 3a				

Subject	K-means Clustering [64]	PSR [81]	nCSP + Gaussian Shrinkage LDA [151]	Proposed Methodology
K3b	74.9	75	**96.31**	93.54
K6b	70.6	**98.20**	79.69	96.77
L1b	61.1	**95**	90.29	87.096
Mean	69.8	89.20	88.76	**92.46**

Table 4.7 Accuracy comparison with state-of-art channel selection techniques for BCI Competition IV 2a and III 3a

BCI Competition IV 2a				
Subject	**Covariance_matrix** [152]	**Modified_GWO** [16]	**EEG-BCI-correlation** [153]	**Proposed Methodology**
A01	**100**	62.06	74.71	90.32
A02	57.87	48.27	48.27	**58.06**
A03	**82.75**	42.52	34.48	82.25
A04	85.05	64.36	67.81	**93.5**
A05	56.32	57.47	45.97	**74.19**
A06	32.18	54.02	25.28	**91.93**
A07	40.22	52.87	**86.20**	83.87
A08	78.16	81.60	80.45	**96.77**
A09	50.57	81.60	71.26	**95.16**
Mean	64.79	60.53	59.38	**85.11**
BCI Competition III 3a				
K3b	60	79.69	65.18	**93.54**
K6b	41.66	58.33	68.88	**96.77**
L1b	65	56.11	57.77	**87.09**
Mean	55.55	64.71	63.94	**92.46**

4.5 Summary

In this chapter, a complete processing pipeline of MDA-SOGWO based EEG channel selection, feature extraction (i.e., time, frequency and TF feature estimation), CCA-RFE based feature selection and feature classification is developed and analysed on two different BCI competition datasets. The results of the study illustrate that the developed methodology enables the extraction of significant information from multi-class MI-BCI EEG activity and reduction of redundancy. The MDA-SOGWO criterion is able to select contributing EEG channels from multi-channel EEG activity for each subject independently. The features extracted from optimized channel locations are selected using CCA-RFE based feature selection criterion. The developed feature selection criterion helps in ranking features according to their discrimination ability and further reducing the dimensionality of feature vector. The comparative analysis reveals that the proposed channel and feature selection methodology is a good candidate for enhanced efficiency and classification accuracy in BCI task. Classification results obtained using all four classifiers provide a clear indication of the significant improvement in classification performance as compared to state-of-art methods. Out of all four classifiers used in this study, the highest average classification accuracy of *85.65*% and *95.15*% is yielded in one vs rest strategy for BCI Competition IV 2a and BCI Competition III 3a datasets, respectively. In future, it may be interesting to introduce an optimized deep learning model at the classification stage for more than four classes of MI-BCI based EEG activity.

CHAPTER 5

Hybrid Ensemble Voting based Residual Attention Model

Preface

In this chapter, a hybrid AttentionNet ensemble voting classifier model is developed for EEG-based MI-BCI task classification. The Time-Frequency Representation (TFR) of the multi-class EEG activity is generated using Transient Extracting Transform (TET). The TFR spectrogram images are input to the designed AttentionNet ensemble voting classifier model for training and classification purposes. The model is evaluated using different fusion strategies viz. feature-level and score-level fusion of layers.

5.1 Introduction

A Motor Imagery (MI) BCI (MI-BCI) task involves extracting motion-related EEG activity from the human brain and rendering it to external commands for controlling BCI applications [154]. Such BCI systems have enormous applications in restoring the movement ability of physically challenged and locked-in people [155]. Though EEG records represent brain activity variation conveniently, EEG signal processing is challenging due to its non-linearity, non-stationarity, and small data packets. In addition, EEG activity is easily affected by external noise that further leads to a low signal-to-noise ratio (SNR) [155]. Efficient feature extraction for high classification accuracy in EEG-derived MI-BCI tasks remains a challenge. The conventional algorithms sometimes suffer from low classification accuracy and do not support adaptive hyperparameter learning [155]. In this context, a time-frequency domain-based novel CNN model is developed in this chapter, that is capable of learning from high-level features while achieving high classification accuracy for multi-class EEG-based MI-BCI tasks.

In this chapter, a hybrid AttentionNet ensemble voting classifier model with bi-level fusion is developed to identify EEG- based multi-class MI-BCI activity. The significant contribution of the proposed chapter is as follows:

1. In this study, a framework is developed for EEG-based multi-class MI-BCI task classification via Time-Frequency Representation (TFR) spectrogram image generation using Transient Extracting Transform (TET).
2. A novel deep CNN architecture called AttentionNet is designed to extract task-specific features from TFR spectrogram images. The proposed model learns the frequency of task occurrence during an EEG epoch with its deep convolutional layers and thereafter, integrates it with non-linear adaptive features using residual_attention module.
3. Two fusion strategies are examined while integrating the layers of the AttentionNet model: (1) Feature-level fusion by concatenating feature vectors of convolution layer (2) Score-level fusion by concatenating activation group of fully-connected layers.
4. The proposed model is extended to classify feature-set, extracted from the AttentionNet model employing ensemble voting classifier-based hard voting criterion corresponding to each MI-BCI task.
5. Experimental results obtained from developed study shows the efficacy for EEG-based multi-class MI-BCI tasks classification.

5.2 Methods

The proposed AttentionNet model-based ensemble voting classifier methodology is developed in three stages. At first, the EEG activity is denoised using General Linear Chirplet Transform-fast power Independent Component Analysis (GLCT-PICA) based denoising methodology [123]. In the second stage, the multi-channel EEG activity corresponding to four-class MI-BCI tasks is input to TET for TFR spectrogram image generation corresponding to each channel. After that, in the third stage, the extracted TFR spectrogram images are input to the developed AttentionNet ensemble voting classifier model in different training and validation ratios. Further, the developed AttentionNet model is extended to two different fusions viz. feature-level and score-level fusion for feature-set generation. The obtained feature-set is fed into an ensemble voting classifier for MI-BCI task classification. Figure 5.1 shows the structure of the proposed hybrid AttentionNet ensemble voting classifier model to classify multi-class EEG-based MI-BCI tasks. The proposed methodology is evaluated on two multi-class MI-BCI datasets viz. BCI Competition IV 2a and BCI Competition III 3a.

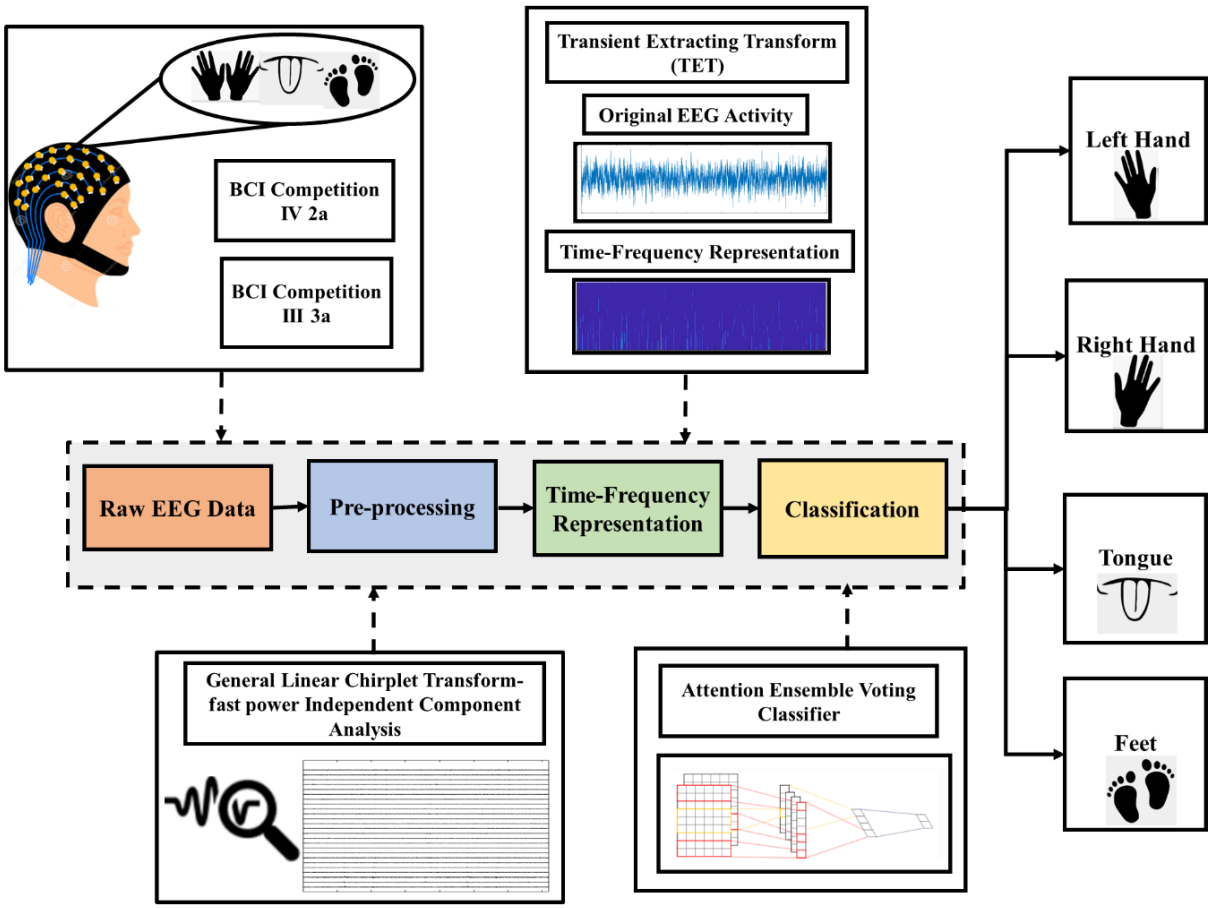


Figure 5.1 Block diagram of TET-based AttentionNet Ensemble Voting Classifier Model

5.2.1 Data Pre-processing

Raw EEG activity is likely to be contaminated by non-cerebral artifact origin such as Electrocardiogram (ECG) activity, Electrooculogram, Electromyogram (EMG) activity and other electrical disturbances. The non-cerebral artifact origins interfere and alter the original cerebral activity making it difficult to analyse and estimate hidden cerebral signatures. Therefore, cleaning the raw EEG activity is crucial prior to feature estimation and classification.

In this study, the contaminated EEG activity is corrected using General Linear Chirplet Transform-fast power Independent Component Analysis (GLCT-PICA) pre-processing method [123]. The block diagram of the GLCT-PICA method is shown in Figure 5.2. Raw multi-channel artifactual EEG activity is first separated into Independent Components (ICs) i.e., artifactual and cerebral components using the fpICA method. Further, artifactual ICs are identified using automatic Katz-Fractal Sparsity (KFS) criterion. In order to retain the information leaked to the artifactual ICs, the identified ICs are treated using General Linear Chirplet Transform (GLCT) criterion [127], as the complete elimination of such ICs may lead to loss of useful cerebral information. The GLCT based Time-Frequency Representation (TFR) can efficiently enhance artifactual components, and soft-thresholding helps suppressing them in Time-Frequency domain. Finally, the thresholded TFR activity is fed to the Inverse-GLCT for the generation of clean ICs, which is further passed to inverse fpICA to produce artifact-free clean EEG activity [114]. Thereafter, the clean EEG activity is processed and analysed for MI-BCI task classification.

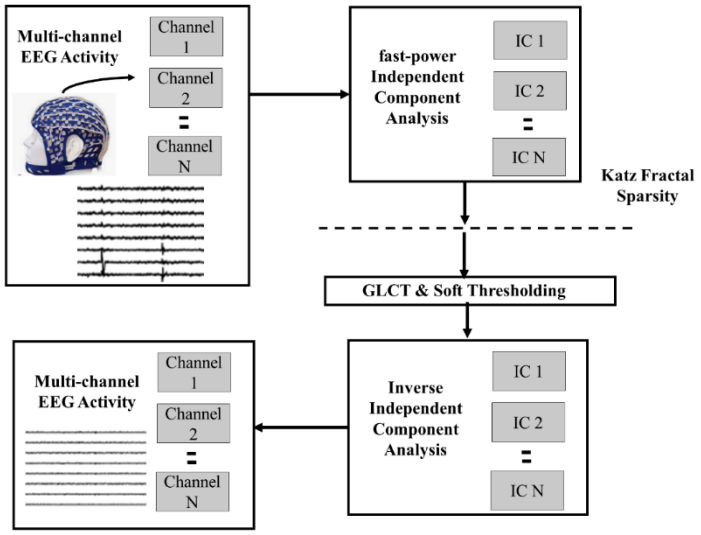


Figure 5.2 Block Diagram of GLCT-PICA based artifact rejection criterion

5.2.2 Time Frequency Representation

This section delineates the extraction of Transient Extracting Transform (TET) based Time-Frequency Representation (TFR) from EEG activity of four class MI-BCI tasks [156]. The conventional TFR methods i.e., Short-Time Fourier Transform (STFT), the Wavelet Transform (WT), and the S-transform, change *1*-D time-domain data to *2*-D time-frequency plane utilizing the concept of signal decomposition [66]. However, as per the Heisenberg uncertainty principle, the generated TFR of the decomposed activity is blurr in nature and sometimes lead to imprecise time-frequency information extraction from non-stationary signals. The limitations of conventional TFR methods such as STFT, WT, S-transform, *etc.* is overcome by Transient Extracting Transform (TET), which locates transient components precisely while providing better energy concentration over the conventional methods [156]. The TET is derived from a raw Dirac function following the steps explained here. The raw Dirac signal is defined as given in Equation *5.1*.

$$g_\delta(t) = A.\delta(t - t_0) \tag{5.1}$$

The expression for STFT computation is given as:

$$S(t,\omega) = \int_{-\infty}^{\infty} s(u-t).A.g(u)e^{-i\omega u}du \tag{5.2}$$

here, $g(u)$ is the time-varying signal and $s(u-t)$ is the moving time window.

$$S(t,\omega) = \int_{-\infty}^{\infty} s(u-t).A.\delta(u-t_0).e^{-i\omega u}du \tag{5.3}$$

$$= A.s(t_0 - t).e^{-iwt_0} \tag{5.4}$$

The STFT of a Dirac function consists of various Dirac functions with equal Group Delay (GD) t_0. The precise GD computation guides us to first calculate the derivative of $S(t,\omega)$ with respect to frequency variable which is given as Equation *5.5*. The final expression for first derivative of $S(t,\omega)$ is given in Equation *5.7*.

$$\partial_w S(t,w) = \partial_w(A.s(t_0 - t).e^{-iwt_0}) \tag{5.5}$$

$$= -it_0.A.s(t_0 - t).e^{-iwt_0} \tag{5.6}$$

$$= -it_0.S(t,w) \tag{5.7}$$

For the given raw dirac signal in Equation *5.1*, ideal Time-Frequency Representation (TFR) should appear only at one time instant without being spread in a large region [156]. To achieve energy concentrated TFR, Transient Extracting Operator (TEO) (given in Equation *5.8*) is multiplied with the extracted STFT output.

$$TEO(t,\tau) = S(t - t_0(t,\tau)) \tag{5.8}$$

$$here, t_0(t,w) = \begin{cases} t_0, & t \in [t_0 - \triangle, t_0 + \triangle], w \in R^+ \\ 0, & otherwise \end{cases} \tag{5.9}$$

For the dirac signal $S(t - t_0(t,\tau))$ is given as,

$$\delta\big(t - t_0(t,\tau)\big) = \delta(t - t_0) \quad (5.10)$$

Therefore, after multiplication of TEO with extracted STFT output, an energy concentrated TFR extraction transform viz. Transient Extracting Transform (TET) is obtained as:

$$T_e(t,\tau) = S(t,\tau).TEO(t,\tau) \quad (5.11)$$

$$|T_e(t,\omega)| = |S(t,\omega).TEO(t,\omega)| \quad (5.12)$$

$$= \left|A.s(t_0 - t).\ e^{-i\omega t_0}.\delta(t - t_0(t,\omega))\right| \quad (5.13)$$

$$= \left|A.s(t_0 - t).\ e^{-i\omega t_0}.\delta(t - t_0)\right| \quad (5.14)$$

$$= \left|A.s(0).\ e^{-i\omega t_0}.\delta(t - t_0)\right| \quad (5.15)$$

$$= |A.s(0).\delta(t - t_0)| \quad (5.16)$$

The expression for STFT and TEO corresponding to dirac signal is input to Equation *5.11* and the TET result for dirac function is obtained in Equation *5.16*. The results of STFT and TET as given in Equation *5.2* and Equation *5.16* are compared and it is observed that the energy is highly concentrated at time t_0 for TET.

5.2.3 Architecture of AttentionNet based ensemble voting classifier

This section presents the architecture of the developed AttentionNet ensemble voting classifier model. Initially, the extracted TFR spectrogram corresponding to four class EEG activity is fed to the developed "*AttentionNet*" model for hidden feature-set extraction. Further, the extracted feature-set is input to the ensemble voting classifier for classification purposes.

The developed AttentionNet model consists of convolution layer of various filter sizes, max-pooling layer, residual, and attention layers. The frequency of task occurrence during an epoch determines which task the present EEG signal belongs, therefore, the convolution layer of various filter sizes viz. *1*, *3*, and *5* is utilized for meaningful feature extraction from the convolution layer. The output of the convolutional layer is input to each max-pooling layer of filter size *2* to extract temporal features. The non-linear layers are required for the efficient mapping of EEG activity into performed tasks. Therefore, an amalgamation of residual and attention layers named as residual_attention module is utilized after each max-pooling layer. The architecture and parameter details corresponding to each layer are given in Table 5.1.

Table 5.1 Architecture and parameter details for AttentionNet model

Convolution Layer	
# of horizontal convolution layers	
Filter size	1, 3, 5
# of filters	32
Stride size	1
Padding	Same
Activation function	ReLU
Max-pooling Layer	
Pool size	[2 x 2]
Stride	[2 x 2]
padding	Same
Residual _Attention module	
Filter size	[32 x 64 x 128]
Dense Layer	
# of Units	256, 128, 64

5.2.3.1 Residual_Attention Module

The residual_attention module is a composition of residual and attention layers. It is a popular approach for image classification problems [157]. The residual_attention module is sub-divided into two fragments named as attention module and residual connection [158]. The basic structure of the residual_attention layer is shown in Figure 5.3. To learn compact features, a downsampling layer is added to the *3*-D CNN layer of the attention module to limit the features extracted from preceding layer into the corresponding dimensions. After that, an upsampling layer is added to reconstruct the original dimensions of the feature maps. This downsampling-upsampling procedure assures that the succeeding calibration process of feature maps can accurately designate the weight coefficient to the corresponding position.

Then, the sigmoid function is used to limit the optimized feature attention maps to *[0, 1]* for each position's weight coefficient allocation [159]. The coefficients leading towards *0* shows that the amount of learning at this position is small, and *1* indicates that the position comprises more critical information. The weight allocation procedure is utilized to obtain adaptively recalibrating features and weaken the unimportant features. Therefore, the output obtained is multiplied with the feature maps of *3*-D CNN to rescale the final output of the attention module. At last, the residual connection and extra convolution layer are added to attention module to carry forward the information of original features to the developed AttentionNet model.

Furthermore, for the integration of the AttentionNet model, two different fusion strategies viz. feature-level fusion and score-level fusion are investigated as shown in Figure 5.4. The architecture of feature-level and score-level fusion criteria are described in detail here.

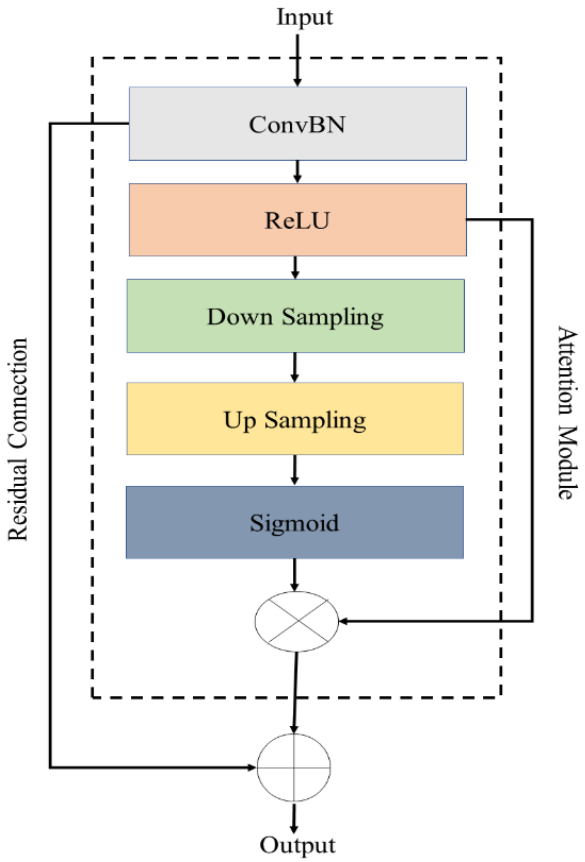


Figure 5.3 Residual_Attention module layer-wise description

A. **Feature-level fusion:** The extracted feature maps from the residual_attention module corresponding to various filter sizes of convolutional layers are multiplied to obtain only contributing features and the output is fed to the max-pooling layer. Further, the output of max-pooling layer is flattened and concatenated to generate a single feature vector. The extracted feature maps from the concatenation layer are fed to the dense layers with varying unit parameters to approximate a non-linear feature mapping function. The extracted feature map corresponding to each class are further utilized for classification purpose. The developed feature-fusion strategy based AttentionNet model is shown in Figure 5.4.

B. **Score-level fusion:** In the score-level fusion, the three extracted residual_attention modules corresponding to various filter sizes of convolutional layers are processed independently using flatten and dense layers. The three independent outputs of the dense layer are then fed into multiplication and concatenation layers to generate the non-linearly mapped single feature vector. The extracted single

domain feature vector corresponding to each class is utilized for classification. The developed score level fusion strategy based AttentionNet model is shown in Figure 5.5.

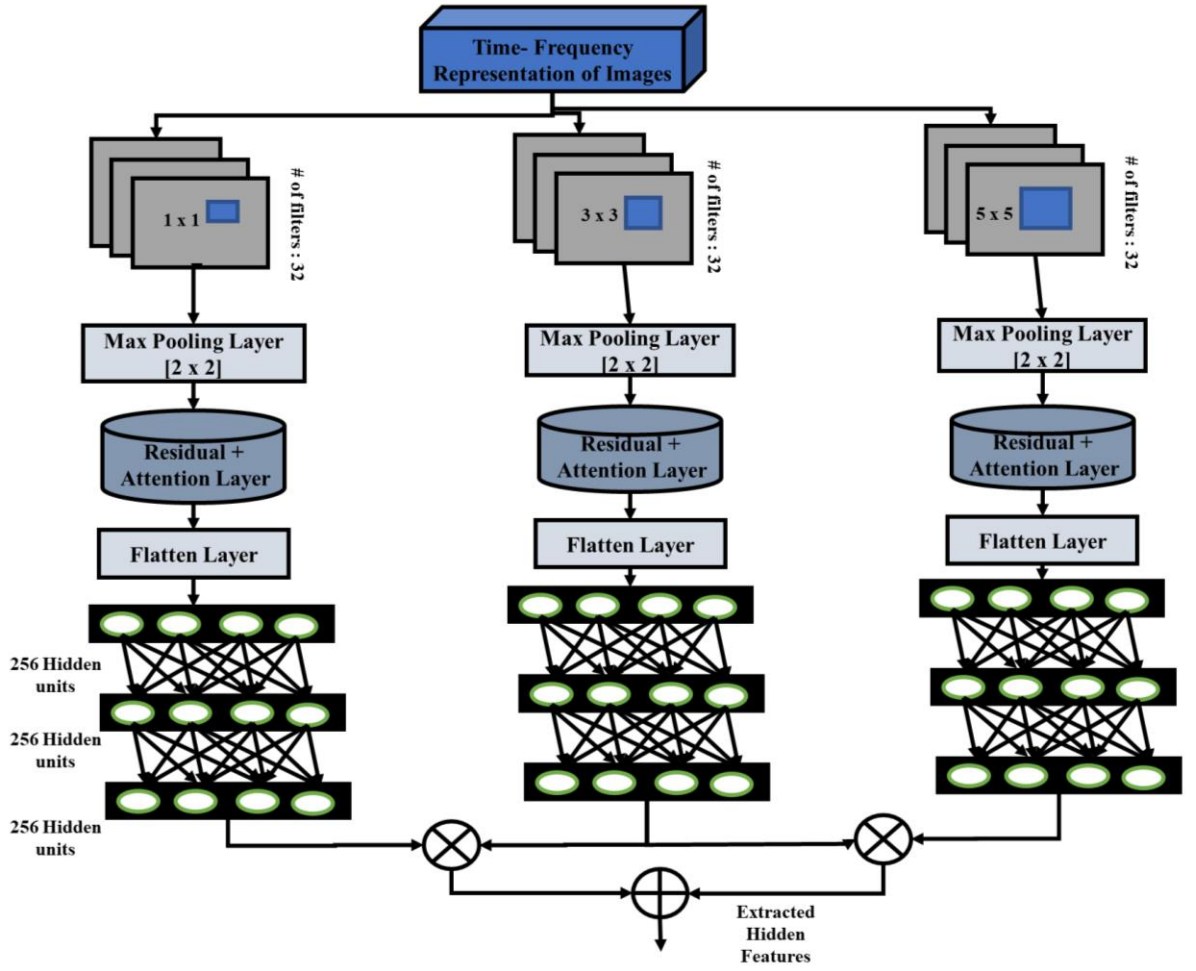


Figure 5.4 Architecture of AttentionNet model for feature-level fusion criterion

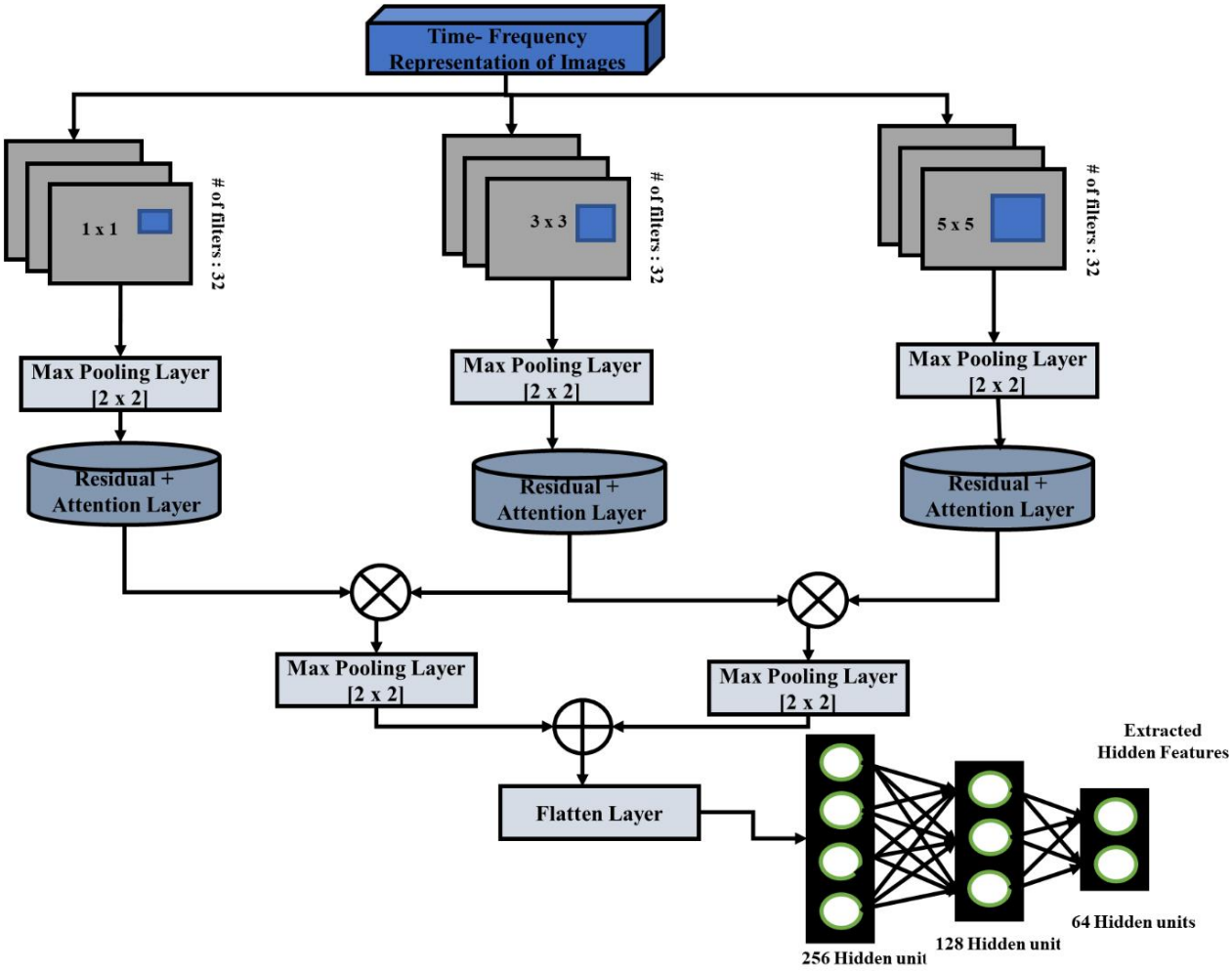


Figure 5.5 Architecture of AttentionNet model for score-level fusion criterion

5.2.3.2 Ensemble Voting Classifier:

The purpose of the developed framework is to identify different tasks corresponding to multi-class MI-BCI EEG activity. The classifier is built based on different training and validation ratios corresponding to both the MI-BCI datasets viz. BCI Competition IV 2a and BCI Competition III 3a. It is evident that independent machine learning classifiers generate unknown target vectors and, in certain scenarios, leads to unreliability during classification. Therefore, in this study, an ensemble voting approach is applied which aims to aggregate the classification results while utilizing the highly voted target vector from all the classifiers in multi-class MI-BCI classification task. The various classifiers utilized this work for target vector generation include Extreme Learning Machine (ELM) [138], Random Forest (RF) [66],

Logistic Regression (LR) [160], *K*-Nearest Neighbour (KNN) [161], and Multi-Layer Perceptron (MLP) [144].

Extreme Learning Machine (ELM) extends the feed-forward neural network with a single hidden layer [63]. In ELM, the classification problem is considered as a linear system, and Moore-Penrose-based inverse operations are employed to predict output weights. It has a single hidden layer with the fastest learning capability. In ELM, parameters used to obtain optimum results are number of hidden layers = *10*, utilization function = '*multiquadric*', alpha = *2* and *rbf_width =3*. Random Forest (RF) is a decision tree estimator which combines outcomes of various decision trees implemented on different sub-samples of the original training set [162]. It utilizes the averaging principle to improve prediction accuracy along with controlled over-fitting. The size of each sub-sample is the same as that of the input dataset drawn with replacement. Logistic Regression classifier works on the principle of maximum-likelihood estimation to extract coefficients from training data [160]. It is a linear model that establishes a linear relationship between input and output variables in order to create a more accurate classification model. The probabilities of possible outcomes are described using a logistic function for a single experiment. *K*-Nearest Neighbour is a non-parametric method of classification that utilizes training data to compute data points by following the principle of distance measurement [163]. First, the number of nearest neighbors is computed, and based on that the data samples are extracted using distance measurement criteria such as Euclidean, Manhattan, and Minkowski. The testing data points are calculated using the distance method from training data, and the class with minimum distance is predicted. Multi-Layer Perceptron (MLP) classifier is a feed-forward neural network-based classifier with one input layer, more than two hidden layers, and one output layer with specific functions [144]. The input layer is responsible for receiving features i.e., one per neuron and each hidden layer is responsible for a certain illustration of the problem. The output layer gets stimulus from hidden layers and generates predicted class linearly. The number of hidden layers is selected in present work after computing classification accuracy for data trials received from all subjects. The various parameters of different classifiers utilized in this work during classification are provided in Table 5.2. These machine learning classifiers are further employed for the development of ensemble voting classifier.

The ensemble voting classifier is built using a group G comprising N different classifiers set is given by Equation 5.17 on which a concluding decision is made [160].

$$G = [C_1, \dots, C_k]; k = 1 \dots . . N \qquad (5.17)$$

Let the decision of k^{th} classifier to choose j^{th} class is given as,

$$d_{k,j} = \{0, 1\}; k = 1 \ldots . N; j = 1 \ldots . M \quad (5.18)$$

here, N is the number of classifiers and M is the number of classes. If k^{th} classifier selects class j, then $d_{k,j}$ = 1, else $d_{k,j}$ = 0. The ensemble voting classifier developed in this study is utilizing the hard voting criterion. Hard voting criterion is based on labels only i.e., $d_{k,j}$ *=1* or *0* depending upon whether the classifier k chooses j or not, respectively [164]. Then, ensemble selects class j that receives largest number of votes. Ensemble voting classifier predict the final class label that has been predicted most frequently by the individual classifiers. The concluding decision is made on the basis of D_G as given in Equation *5.19*. Figure 5.6 shows the schematic of ensemble voting classifier-based classification approach utilized in this work.

$$D_G = \underset{j = 1, \ldots . M}{argmax} \sum_{k=1}^{N} d_{k,j} \quad (5.19)$$

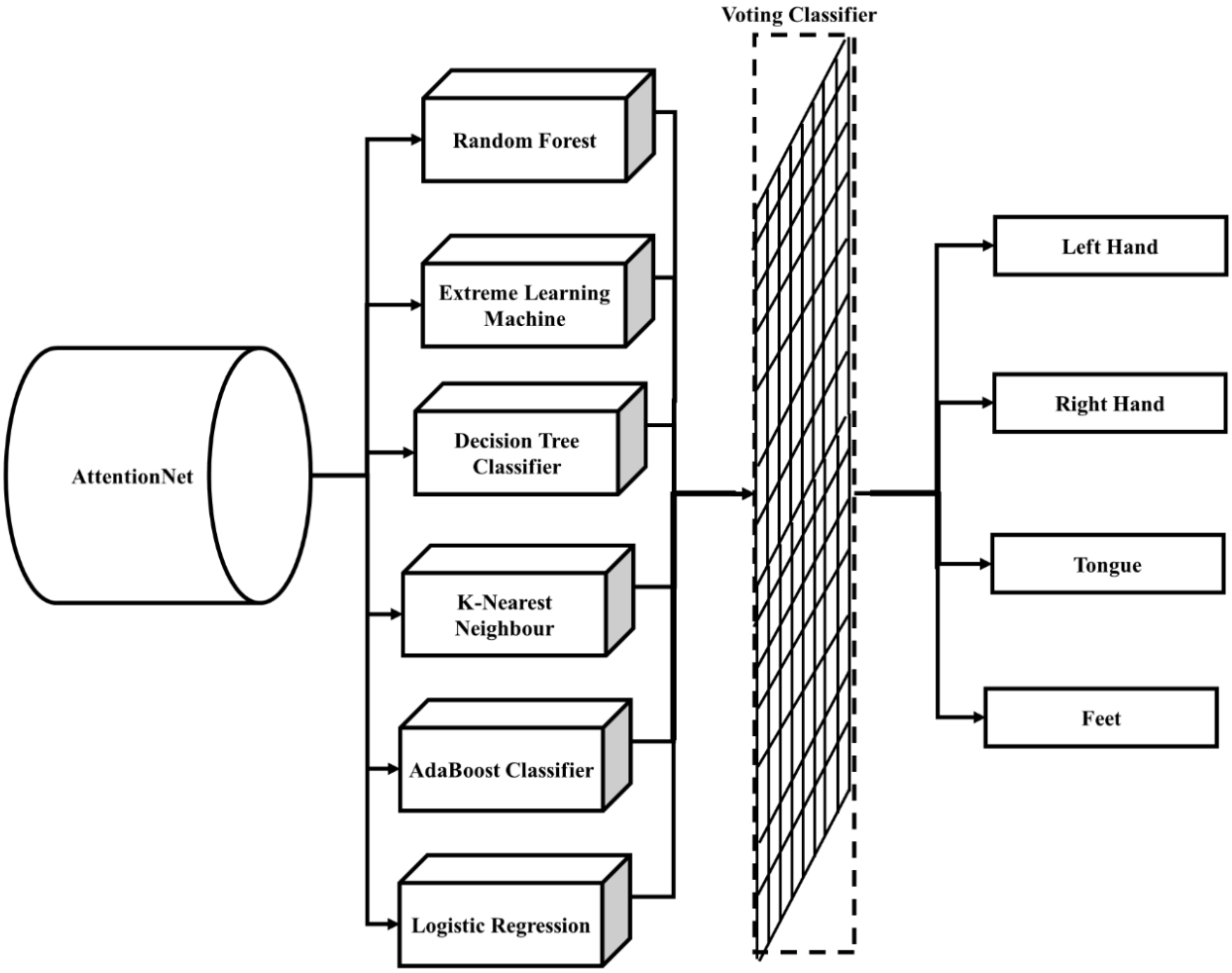


Figure 5.6 Ensemble voting classifier-based classification approach

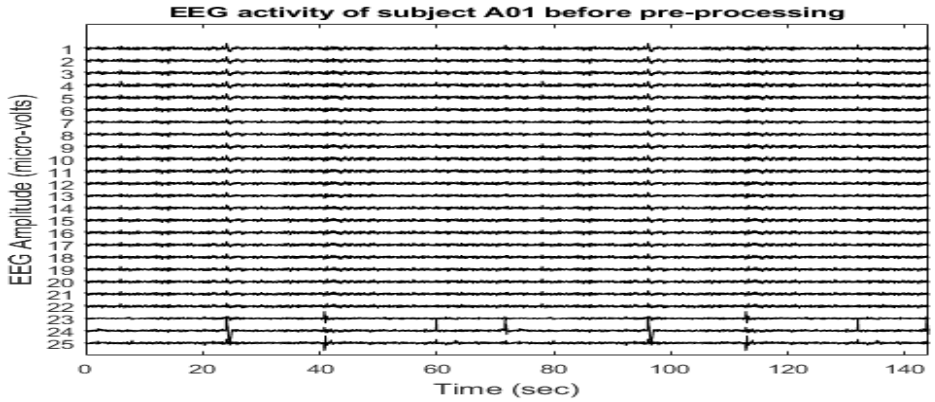
Table 5.2 Parameter description of various classifiers utilized in this chapter

Classifier	**Parameters**
Extreme Learning Machine	*hidden_layer*=10
K-Nearest Neighbour	*metric*='minkowski' *n_neighbors* = 6
	metric='manhattan' *n_neighbors* = 6
	metric='euclidean' *n_neighbors* = 6
Multi-Layer Perceptron	*solver* = 'lbfgs'

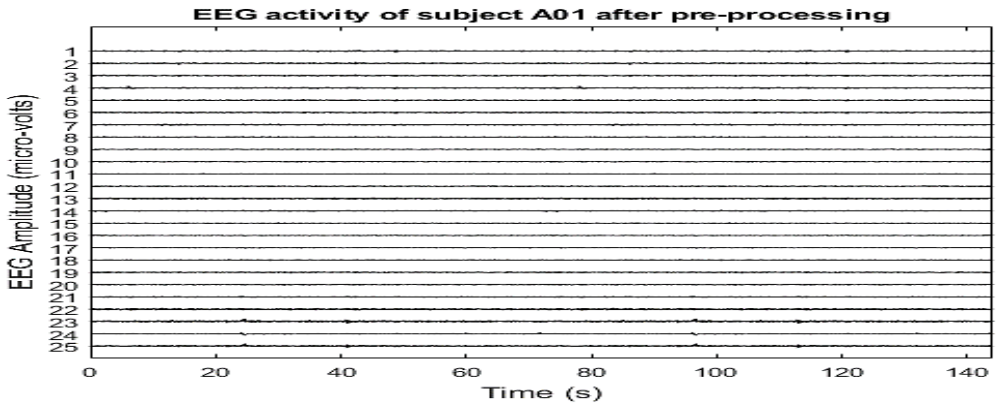
	alpha = 10
	hidden_layer = 200

5.3 Results and Discussion

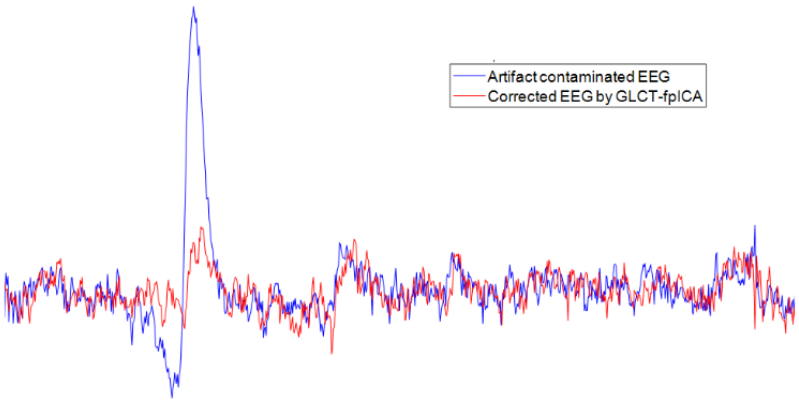
No filtering is carried out prior to pre-processing of EEG activity. Figure 5.7(a-b) represents original EEG activity and artifact-corrected/clean EEG activity using GLCT-PICA technique respectively. Figure 5.7(c) shows the zoomed-in representation of artifact-free/clean single channel EEG activity overlayed on original EEG records of same channel. Further, the artifact-free/clean EEG activity is processed through TET and decomposed into TFR spectrograms. The raw EEG activity and its corresponding TFR spectrogram are shown in Figure 5.8(a-d) for four classes i.e., left-hand, right-hand, tongue, and feet movement, respectively. It can be depicted from Figure 5.8 that, the variation in the time-domain, and frequency-domain of EEG activity are enhanced in TFR spectrogram corresponding to each MI-BCI task.



(a)

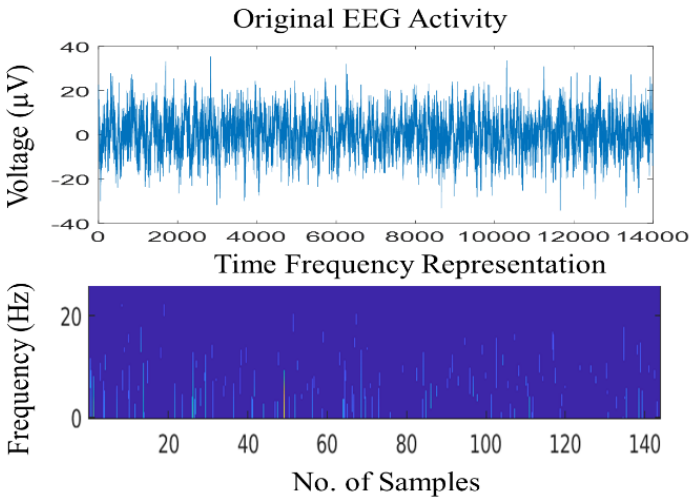


(b)

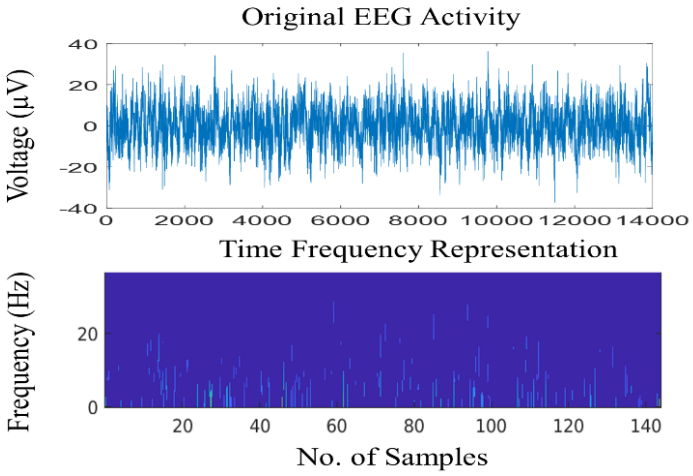


(c)

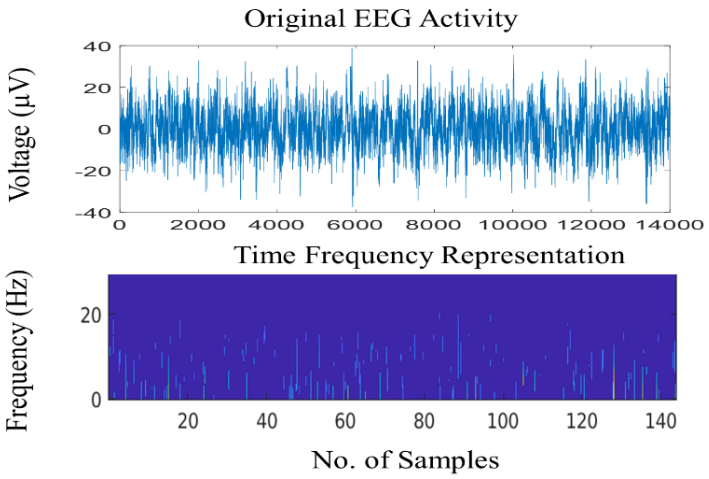
Figure 5.7 EEG activity record of subject "A01" (a) Raw EEG activity (b) Processed EEG activity (c) Zoomed In view of original EEG w.r.t artifact corrected EEG



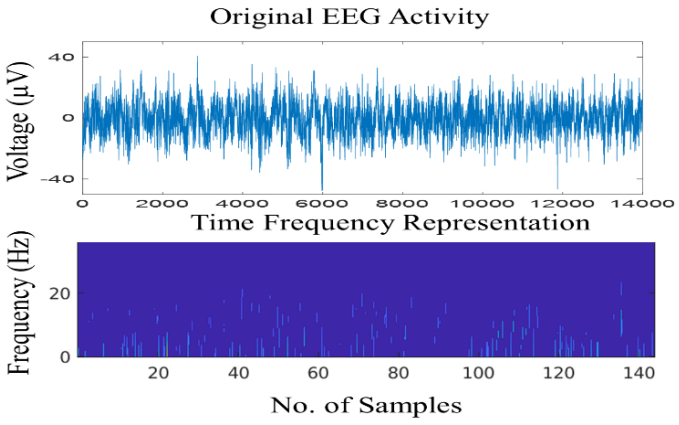
(a)



(b)



(c)



(d)

Figure 5.8 Raw EEG activity and its corresponding TFR spectrogram computed using TET for (a) Left-hand movement (b) Right-hand movement (c) Tongue movement (d) Feet movement

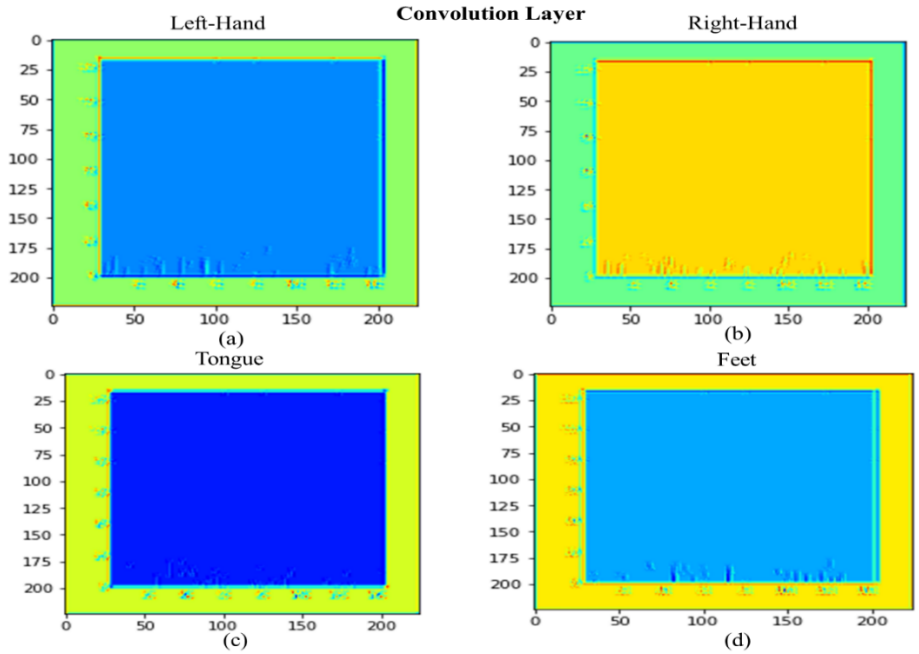
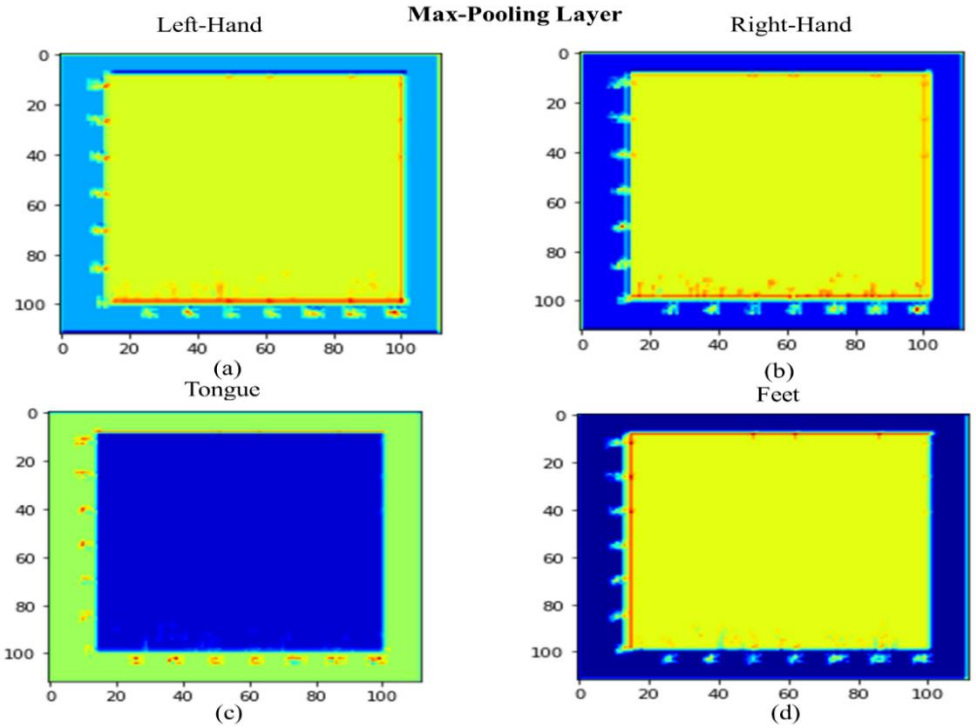


Figure 5.9 Visualization of 1 feature map out of *32* for first convolutional layer in AttentionNet generated from EEG activity based TFR for (a) Left-Hand movement (b) Right-Hand movement (c) Tongue movement (d) Feet movement



9

Figure 5.10 Visualization of *1* feature map of Max Pooling Layer in AttentionNet generated from EEG activity derived TFR for (a) Left-Hand movement (b) Right-Hand movement (c) Tongue movement (d) Feet movement

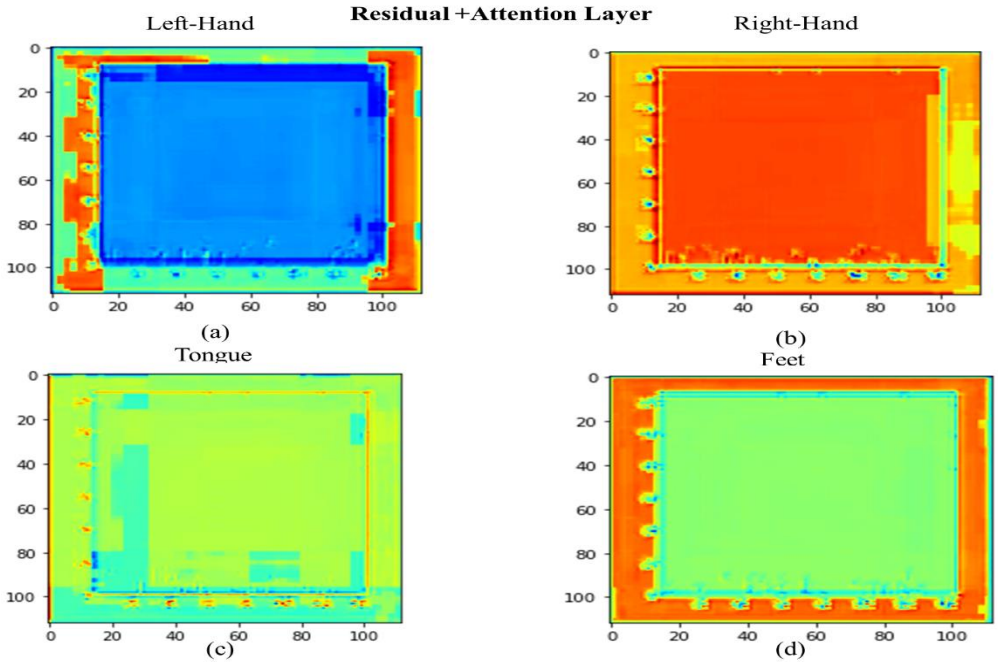


Figure 5.11 Visualization of output feature map of Residual +Attention Layer in AttentionNet generated from EEG activity derived TFR for (a) Left-Hand movement (b) Right-Hand movement (c) Tongue movement (d) Feet movement

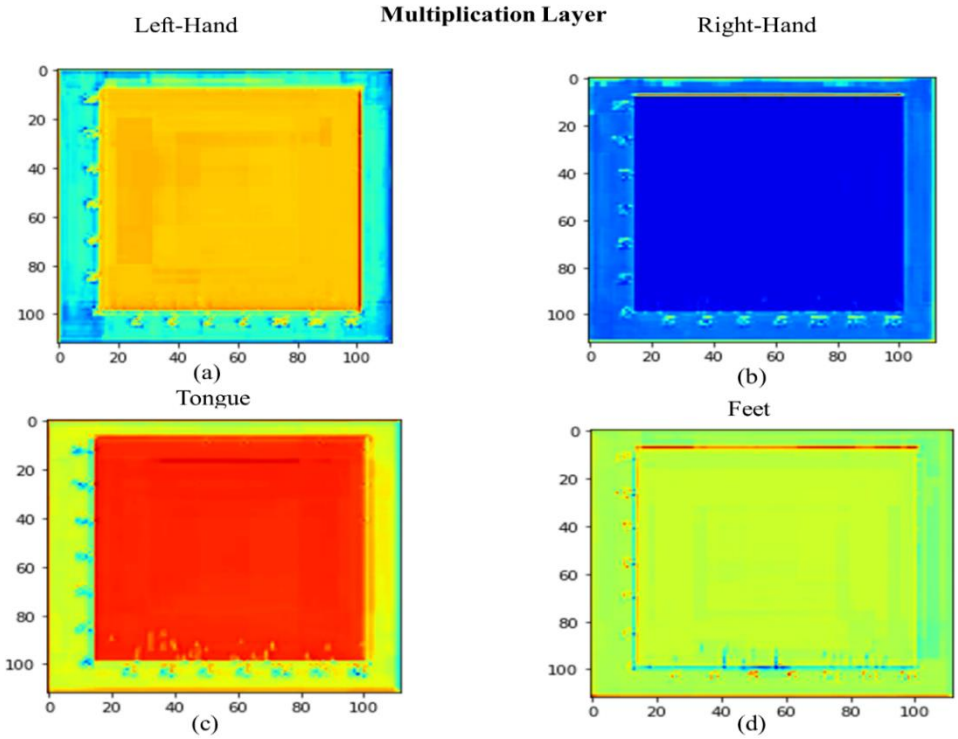


Figure 5.12 Visualization of Multiplication Layer in AttentionNet generated from EEG activity derived TFR for (a) Left-Hand movement (b) Right-Hand movement (c) Tongue movement (d) Feet movement

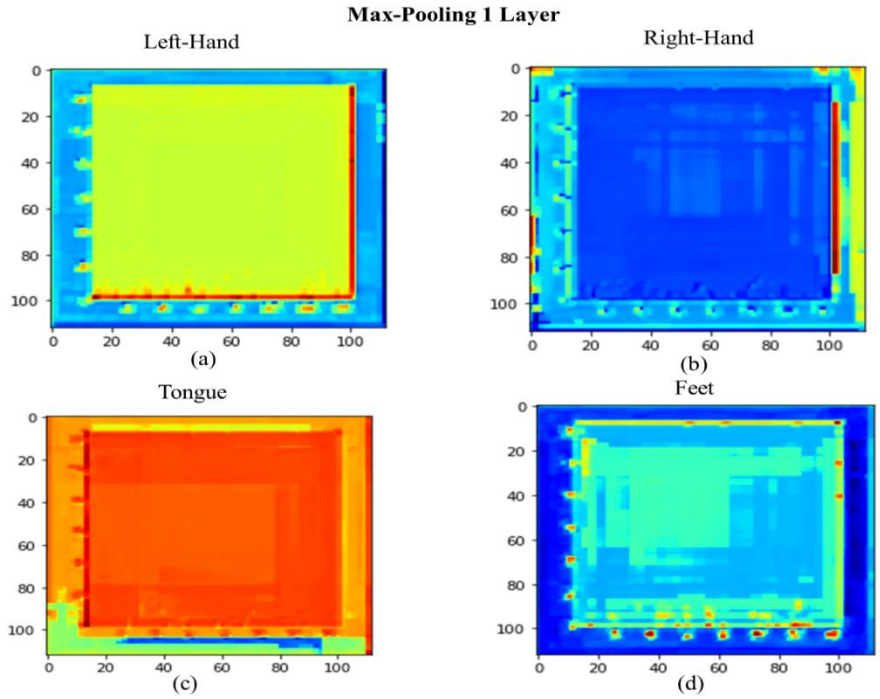


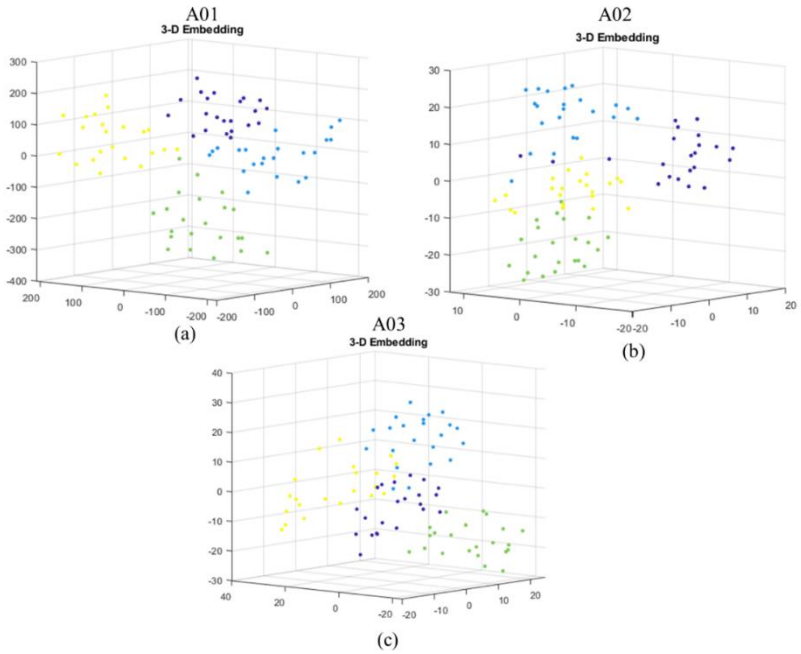
Figure 5.13 Visualization of Max Pooling 1 Layer in AttentionNet generated from EEG activity derived TFR for (a) Left-Hand movement (b) Right-Hand movement (c) Tongue movement (d) Feet movement

At last, the subject independent TFR spectrogram images of each class are fed into the developed AttentionNet model for feature-set generation relative to each MI-BCI task. Figures (5.9-5.13) provide the visualization for output feature map corresponding to each layer of AttentionNet model for left-hand, right-hand, tongue, and feet movement classes of EEG activity. It can be deduced from Figure (5.9-5.13) that apparent difference is stimulated by the AttentionNet model within feature patterns of four classes while progressing from the first layer to the last layer. The significant difference between the feature vector of all four classes is perceptible from Figure 5.13, and it infers the ability of the developed model to extract intrinsic features for class discrimination and classification efficiency improvement.

Also, the discrimination ability of the AttentionNet model for all four classes is validated using the Kruskal-Wallis statistical test. The Kruskal-Wallis statistical test is a non-parametric variance analysis test based on the assumption that data samples are not distributed normally. The probability-value (p-value) of the Kruskal Wallis test corresponding to each subject of BCI competition IV 2a, and BCI competition III 3a datasets are provided in Table 5.3. The feature matrix to be fed to the Kruskal-Wallis test is generated using the mean value of *64* features obtained from the output layer of the AttentionNet model. The lesser p-values lead to high discrimination between features and vice-versa [165]. It can be observed from Table 5.3 that the p-value of all the subjects is close to zero. Therefore, the developed features can effectively discriminate between the left-hand, right-hand, tongue, and feet movement classes of both BCI datasets

Table 5.3 p-value computed from the Kruskal Wallis Test

BCI Competition IV 2a	
Subject	**p-value**
A01	0.002
A02	0
A03	0.0023
A04	0.0563
A05	0
A06	0.119
A07	0.0002
A08	0
A09	0.0003
BCI Competition III 3a	
K3b	0.0011
K6b	0
L1b	0.0069



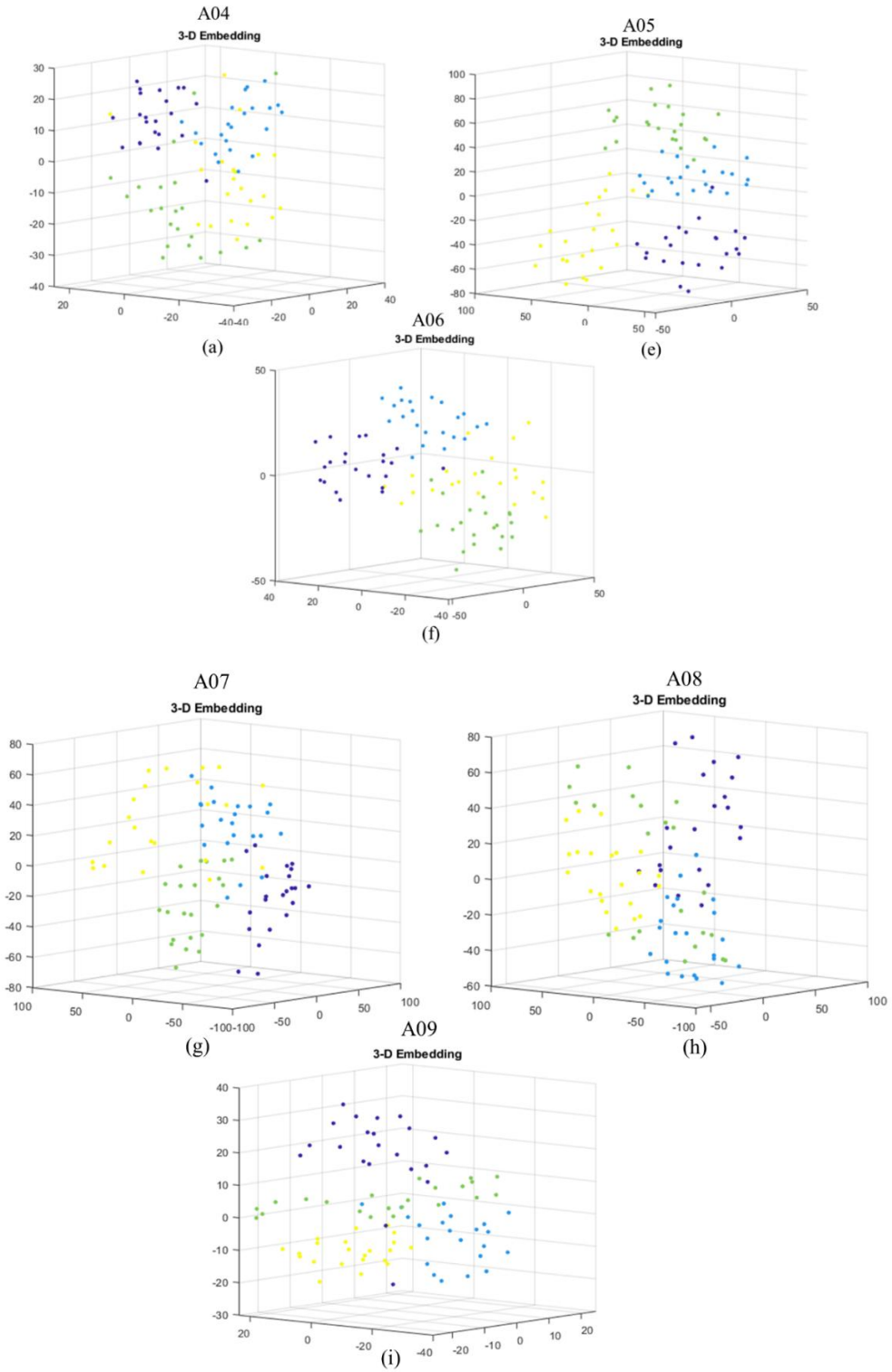


Figure 5.14 t-SNE plot estimated from extracted features before inputting into ensemble voting classifier corresponding to (a) *A01* (b) *A02* (c) *A03* (d) *A04*(e) *A05* (f) *A06* (g) *A07* (h) *A08* (i) *A09*

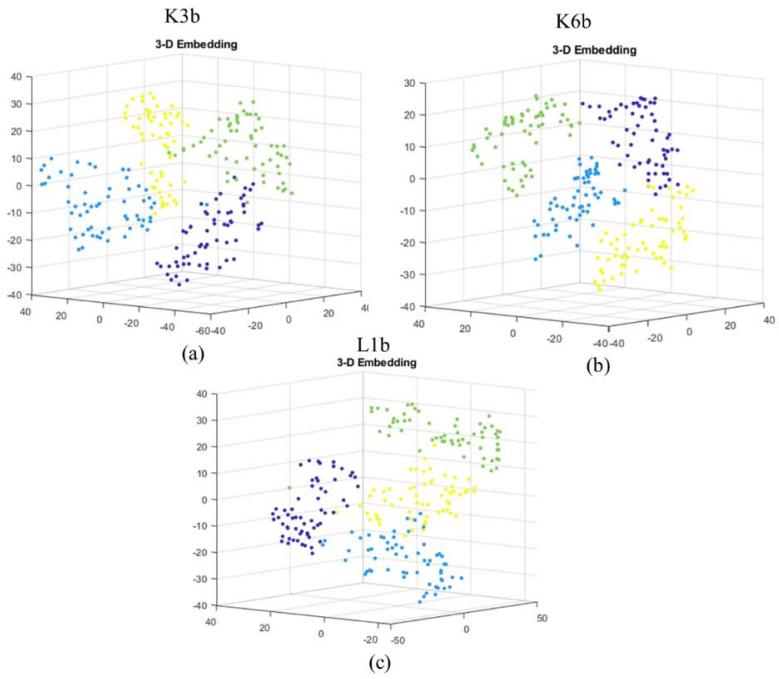


Figure 5.15 t-SNE plot estimated from extracted features before inputting into ensemble voting classifier corresponding to (a) *K3b* (b) *K6b* (c) *L1b*

Table 5.4 Kullback-Leibler divergence loss estimated while computing t-SNE plot

BCI Competition IV 2a	
Subject	**Kullback-Leibler divergence Loss**
A01	0.2189
A02	0.6151
A03	0.162
A04	0.256
A05	0.1575
A06	0.2496
A07	0.1733
A08	0.2111
A09	0.2356
BCI Competition III 3a	
K3b	0.3915
K6b	0.3651
L1b	0.3931

In addition, the t-distributed Stochastic Neighbor Embedding (SNE) plot is utilized to illustrate the discrimination among extracted features from the AttentionNet model. The t-SNE approach reduces high dimensionality data to lower dimensions by translating Euclidean distance between high dimensional data points into conditional probabilities that signifies similarities [166]. If the data points are closer, it leads to high similarity between classes. The t-SNE plots generated from all subjects of BCI competition IV 2a and BCI competition III 3a datasets establish clear bias among four class movement features, as shown in Figure 5.14 and Figure 5.15. Further, the

deviation among the extracted features is evaluated while computing t-SNE using the Kullback-Leibler divergence loss measure. Table 5.4 shows the Kullback-Leibler divergence loss estimated for both datasets, i.e., BCI Competition IV 2a, and BCI Competition III 3a.

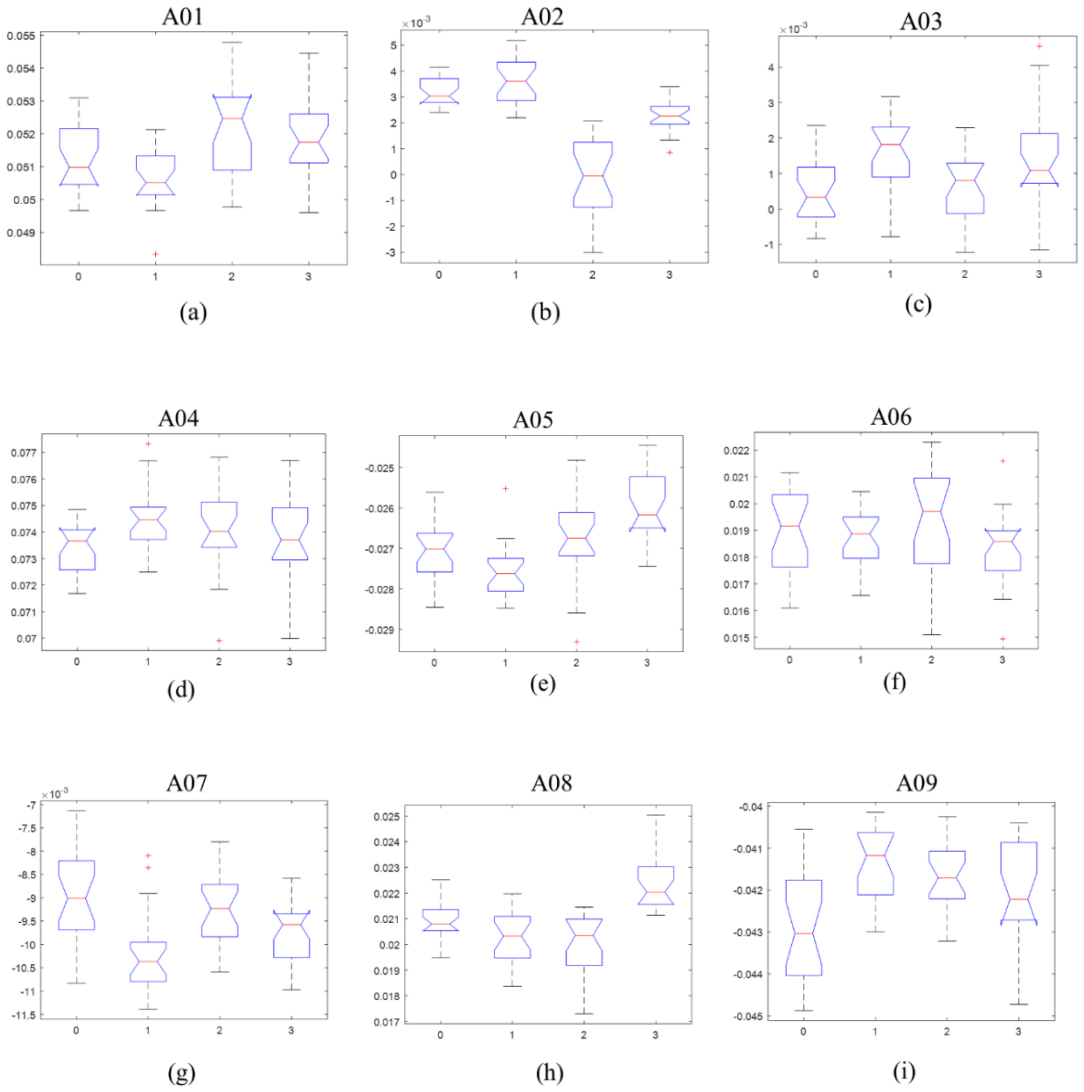


Figure 5.16 Boxplot generated from extracted features before inputting into ensemble voting classifier corresponding to (a) *A01* (b) *A02* (c) *A03* (d) *A04*(e) *A05* (f) *A06* (g) *A07* (h) *A08* (i) *A09*

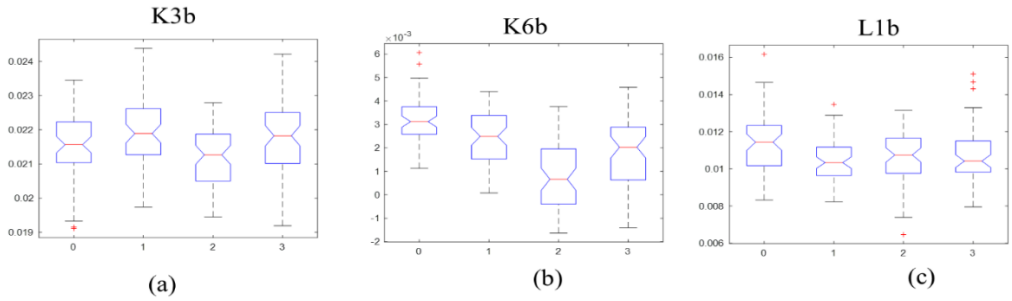


Figure 5.17 Boxplot generated from extracted features before inputting into ensemble voting classifier corresponding to (a) *K3b* (b) *K6b* (c) *L1b*

The box-plot of feature-set generated from all subjects of both datasets is shown in Figure (5.16-5.17). It can be deduced from the subjects' box plots that there is noteworthy difference between four classes of MI-BCI tasks. The extracted feature-set is then passed through an ensemble voting classifier to classify multi-class MI-BCI EEG activity. The different fusion levels of the developed AttentionNet model, i.e., feature-level and score-level fusion, are also evaluated using *10*-fold cross-validation accuracy of all the subjects along with various performance metrics viz. precision, recall and F1-score. The classification results obtained using various fusion strategies is shown in Table 5.5 and Table 5.6. The classification results of the feature-level fusion of the AttentionNet model are more efficient as compared to score-level fusion. One possible reason behind the better performance of feature-level fusion criterion is that it leverages richer information of features as compared to score-level fusion. For further validation, feature-level fusion strategy is utilized considering its good performance than score-level fusion. Also, the feature set is divided into various ratios of training and testing subsets, viz. *90*:*10*, *70*:*30*, *60*:*40*, and *50*:*50* for validating the developed methodology. The various train/test split ratios are utilized to evaluate the performance of developed AttentionNet ensemble voting classifier model independent of split ratio. The classification results for all subjects of both datasets for different separation ratios are shown in Figure 5.17 and Figure 5.18. The classification results show that the proposed methodology can efficiently discriminate between all the four classes of MI-BCI data.

The proposed work is performed using a desktop computer with a configuration of *Intel(R) CoreTM i5-5200U* CPU running at *2.20GHz* and *8 GB* of RAM. The TFR based spectrogram generation is made utilizing *MATLAB 2020a*, and the AttentionNet ensemble voting classifier model is developed and validated using *Python 3.7*.

Table 5.5 Performance measure of various subjects for both datasets after classification using ensemble voting classifier with feature-fusion AttentionNet model

BCI Data	Subject No.	Accuracy Mean(std)	Precision	Recall	F1-Score
	A01	85.27(0.09)	85.27	86.38	85.82

	A02	96(0.1)	95.55	94.44	94.99
	A03	93(0.05)	95.55	94.44	94.99
BCI Competition IV 2a	A04	81(0.12)	80.69	80.69	80.69
	A05	90(0.14)	89.86	88.75	89.3
	A06	82(0.14)	78.05	78.05	78.05
	A07	92(0.09)	91.11	91.11	91.11
	A08	87(0.1)	88.61	87.5	88.05
	A09	86(0.11)	86.25	87.5	86.87
BCI Competition III 3a	K3b	95.83(0.05)	95.8	95.8	95.8
	K6b	91.66(0.07)	91.66	91.7	91.7
	L1b	95.83(0.06)	95.83	95.8	95.8

Table 5.6 Performance measure of various subjects for both datasets after classification using ensemble voting classifier with score-fusion AttentionNet model

BCI Data	**Subject No.**	**Accuracy Mean(std)**	**Precision**	**Recall**	**F1-Score**
	A01	79.3(0.091)	70.37	70.4	70.4
	A02	88.6(0.07)	81.48	81.5	81.5
	A03	79.7(0.107)	66.66	66.7	66.7
BCI Competition IV 2a	A04	71.5(0.115)	81.4	81.5	81.5
	A05	88.3(0.12)	81.48	81.5	81.5
	A06	71.5(0.104)	77.7	77.8	77.8
	A07	80.6(0.079)	66.66	66.7	66.7
	A08	87.4(0.108)	88.88	88.9	88.9
	A09	51.1(0.106)	51.8	51.8	51.9
BCI Competition III 3a	K3b	78.7(0.066)	76.38	76.4	76.4
	K6b	77.9(0.062)	80.55	80.6	80.6
	L1b	90(0.057)	90.27	90.3	90.3

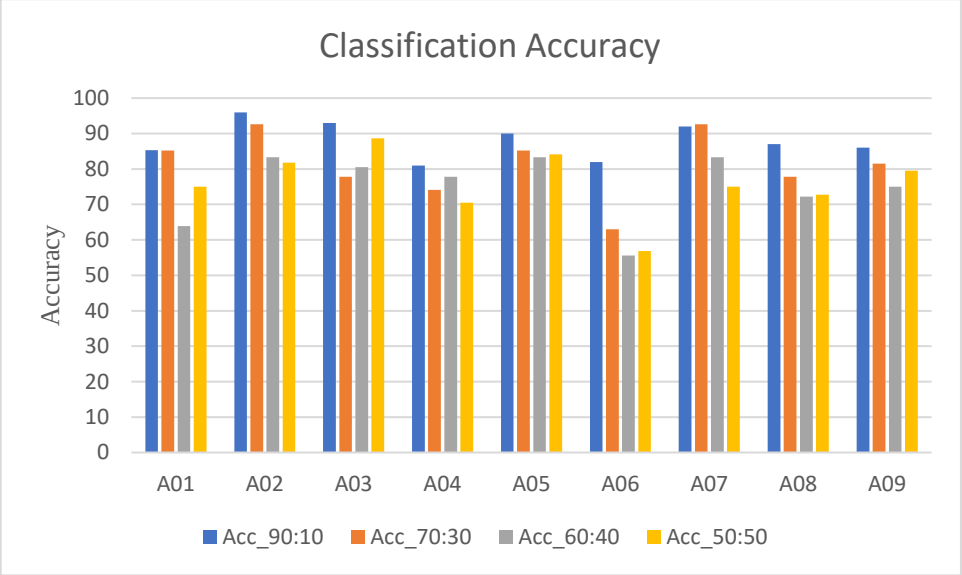


Figure 5.21 Classification accuracy obtained for varying training and testing size of BCI Competition IV 2a

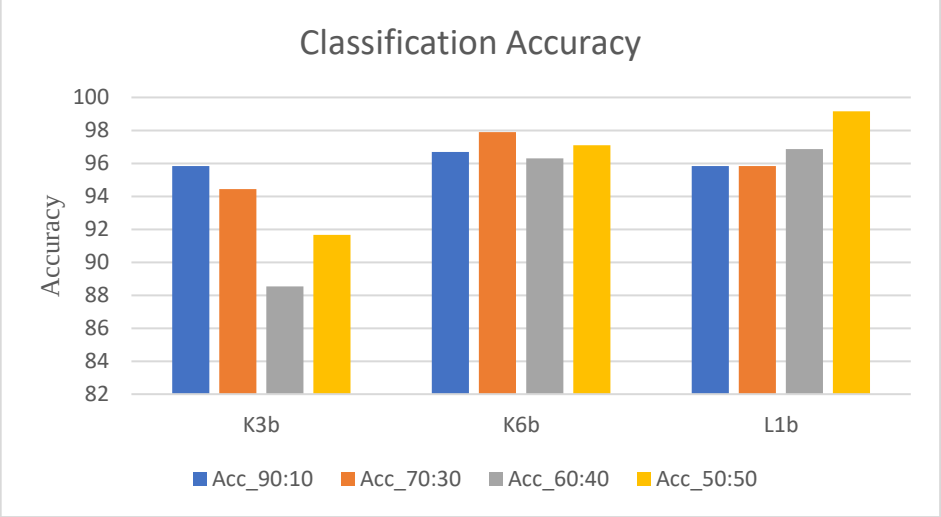


Figure 5.19 Classification accuracy obtained for varying training and testing size of BCI Competition III 3a

Table 5.7. Accuracy Comparison with literature for BCI Competition IV 2a and III 3a

	BCI Competition IV 2a			
Subject	**NSL-EEGNet** [149]	**ETRCNN**[167]	**nCSP + Gaussian Shrinkage LDA** [151]	**Proposed Methodology**
A01	82.29	88.02	**89.23**	85.27
A02	51.39	73.21	76.15	**96**
A03	85.07	90.68	90.60	**93**

A04	67.01	**81.71**	71.38	81
A05	58.33	65.63	59.82	9**0**
A06	56.25	**91.72**	63.26	82
A07	83.33	85.47	91.70	**92**
A08	73.96	**91.67**	89.18	87
A09	78.47	**93.0**	85.26	86
Mean	70.68	84.57	79.62	**88.14**

BCI Competition III 3a

Subject	**K-means Clustering** [168]	**PSR** [81]	**nCSP + Gaussian Shrinkage LDA** [151]	**Proposed Methodology**
K3b	74.9	75	**96.31**	95.83
K6b	70.6	98.20	79.69	**91.66**
L1b	61.1	95	90.29	**95.83**
Mean	69.8	89.20	88.76	**93.31**

A comparison of the developed methodology with recently proposed techniques from literature is provided in Table 5.7. In [149], an amalgamation of Neural Structured Learning (NSL) and EEGNet architecture was proposed to design an MI-BCI system. Electroencephalography Topographical Representation (ETR) based CNN technique was suggested by Xu [167] for a multi-class MI-BCI classification task. Olias et al. [151] proposed gaussian shrinkage LDA and nCSP based criterion to differentiate between different MI-BCI tasks. In [168] and [81], K-means clustering-based criterion and Phase State Reconstruction (PSR) criteria were proposed, respectively. Also, the dynamical behaviour of brain activities using EEG records is analysed. It can be deduced from Table 5.7 that the mean accuracy obtained by the proposed methodology outperforms state-of-art. In the present work, the individual accuracy obtained is also highest for most subjects and is effective for multi-class MI-BCI classification tasks.

5.4 Summary

The study results illustrate that the developed methodology enables the extraction of significant information from multi-class MI-BCI EEG activity along with redundancy reduction. The comparative analysis reveals that the proposed AttentionNet ensemble voting classifier model is suitable for enhanced performance and classification accuracy in identifying BCI tasks. Classification results obtained using all four classifiers indicate a significant improvement in classification performance compared to the state-of-art methods. The highest accuracy yielded by

ensemble voting classifier for BCI Competition IV 2a and BCI Competition III 3a datasets is *88.14%* and *93.31%*, respectively.

CHAPTER 6

Concluding Remarks & Future Scope

Preface

This chapter outlines the salient features of developed pre-processing, feature extraction, and feature classification methodologies. The concluding remarks corresponding to each developed methodology of EEG-based multi-class MI-BCI system are also postulated. Furthermore, the chapter emphasizes the future scope of the proposed work.

6.1 Conclusion

The prime objective of this work was to develop an efficient BCI system utilizing the multi-class motor imagery EEG activity. The classification accuracy enhancement was a foremost concern of this work while specifically targeting the intermediate steps optimization throughout the experimentation. The major signal processing techniques utilized time-domain, frequency-domain, and time-frequency domain analyses. In this research work, we had focused on employing all three types of analyses along with posing a comparison between time, frequency, and time-frequency domains. The multi-class MI-EEG activity classification was performed using machine learning as well as deep learning techniques. Also, to analyse the performance of classification approaches developed for MI-BCI system design, various performance metrics had been employed.

Electroencephalography (EEG) recordings containing non-cerebral EEG artifacts were cleaned by developing a GLCT-PICA based EEG de-noising methodology. The developed methodology was tested and validated on simulated as well as experimentally recorded EEG activity. First, the EEG source separation process was performed by implementing the Fast Power Independent Component Analysis (fpICA) BSS technique. After that, the artifactual ICs were automatically segregated from separated ICs using the developed Katz Fractal Sparsity (KFS) criterion. Consequently, the artifactual ICs were processed using General Linear Chirplet Transform (GLCT) and spike zone thresholding was performed to suppress spiking areas of artifactual ICs. This procedure led to keeping cerebral information intact while cleaning the contaminated EEG records. The developed artifact removal criterion i.e., fpICA-GLCT EEG denoising methodology is proficient to be employed in a real-time environment for improved performance of EEG-based BCI system.

Further, BCI system development was performed by developing a complete processing pipeline of MDA-SOGWO based EEG channel selection, feature extraction, CCA-RFE based feature selection, and feature classification. The proposed methodology helps in optimizing various stages involved in EEG activity processing during BCI system design. The results of the study illustrate that the present methodology enables the extraction of significant information from multi-class MI-BCI EEG activity and reduction in the redundancy. The optimum channel-set selection is performed using the MDA-SOGWO technique. Thereafter, time, frequency, and time-frequency domain-based feature set extraction is performed. The most contributing features from the extracted features are selected using the developed CCA-RFE based feature selection criterion. The obtained classification accuracy reveals the efficacy of the developed pipeline in multi-class MI-BCI task classification. The highest average classification accuracy of 85.65% and 95.15% was achieved for BCI Competition IV 2a and BCI Competition III 3a datasets, respectively.

Afterward, a complete processing pipeline of TET-based TFR generation and AttentionNet ensemble voting classifier model for EEG-based MI-BCI activity classification was developed and analysed on two types of BCI competition datasets. The developed methodology aids in efficient feature-set extraction from multi-class MI-BCI EEG activity. The TFR extracted from TET was utilized as spectrogram images and fed into the developed AttentionNet model. The proposed AttentionNet ensemble voting classifier model was employed to extract

efficient features and identify the various MI-BCI tasks corresponding to their respective EEG activity. The results achieved from the developed AttentionNet ensemble voting classifier model revealed its efficacy as compared to state-of-art techniques. The classification accuracy achieved by employing the AttentionNet ensemble voting classifier model for BCI Competition IV 2a and BCI Competition III 3a datasets was 88.14% and 93.31%, respectively.

6.2 Future Scope

The classification methodologies employed in the present work have indicated a significant improvement in terms of performance in comparison to state-of-the-art methods. Though this study shows great potential in developing a BCI system utilizing the multi-class motor imagery EEG activity, there is still scope for further improvement and extension. It would be interesting to enhance the classification accuracy by introducing an optimized deep learning model by tuning various hyper-parameters for the classification of MI-BCI-based EEG activity. Moreover, the potential use of data from multiple modalities can also be explored as one of the future aspects. As this is an ever-evolving field with new ideas leading the way to more complex and efficient models which will certainly come up to make use of these opportunities. Also, the techniques proposed in present thesis work can be utilized to explore various other biomedical signals such as Electrocardiogram signals, Electromyogram signals, Respiratory Signals, Galvanic Skin Response etc.

BIBLIOGRAPHY

[1] P. Gaur *et al.*, "A sliding window common spatial pattern for Enhancing Motor Imagery Classification in EEG-BCI," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, 2021.

[2] Y. Yang, S. Chevallier, J. Wiart, and I. Bloch, "Subject-specific time-frequency selection for multi-class motor imagery-based BCIs using few Laplacian EEG channels," *Biomed Signal Process Control*, vol. 38, pp. 302–311, 2017, doi: 10.1016/j.bspc.2017.06.016.

[3] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," 1988.

[4] B. Z. Allison, D. J. McFarland, G. Schalk, S. D. Zheng, M. M. Jackson, and J. R. Wolpaw, "Towards an independent brain-computer interface using steady state visual evoked potentials," *Clinical Neurophysiology*, vol. 119, no. 2, pp. 399–408, Feb. 2008, doi: 10.1016/j.clinph.2007.09.121.

[5] E. C. Lalor *et al.*, "Steady-State VEP-Based Brain-Computer Interface Control in an Immersive 3D Gaming Environment," 2005.

[6] L. R. Hochberg and J. P. Donoghue, "Sensors for brain-computer interfaces: Options for turning thought into action," *IEEE Engineering in Medicine and Biology Magazine*, vol. 25, no. 5. Institute of Electrical and Electronics Engineers Inc., pp. 32–38, 2006. doi: 10.1109/MEMB.2006.1705745.

[7] J. Kalcher, I. D. Flotzinger, C. Neuper, S. G611y, and G. Pfurtscheller, "Graz brain-computer interface I1: towards communication between humans and computers based on online classification of three different EEG patterns," 1996.

[8] J. Sui *et al.*, "Combination of FMRI-SMRI-EEG Data Improves Discrimination of Schizophrenia Patients by Ensemble Feature Selection," no. 2, pp. 3889–3892, 2014.

[9] S. Mason *et al.*, "Evaluating the Performance of Self-Paced Brain-Computer Interface Technology Revision: 1.0 (draft)," 2006.

[10] B. Güntekin, T. Aktürk, E. Yıldırım, N. Mantar, F. Cadırcı, and L. Hanoglu, "P07-F Event related EEG brain oscillations differentiate cognitive decline in patients with Parkinson's disease," *Clinical Neurophysiology*, vol. 130, no. 7, p. e68, 2019, doi: 10.1016/j.clinph.2019.04.460.

[11] A. Bria, C. Marrocco, and F. Tortorella, "Sinc-based convolutional neural networks for EEG-BCI-based motor imagery classification," no. Mi, pp. 1–10.

[12] M. Rakibul Mowla, S. C. Ng, M. S. A. Zilany, and R. Paramesran, "Artifacts-matched blind source separation and wavelet transform for multichannel EEG denoising," *Biomed Signal Process Control*, vol. 22, pp. 111–118, Jul. 2015, doi: 10.1016/j.bspc.2015.06.009.

[13] J. Hou *et al.*, “An improved artifacts removal method for high dimensional EEG,” *J Neurosci Methods*, vol. 268, pp. 31–42, Aug. 2016, doi: 10.1016/j.jneumeth.2016.05.003.

[14] B. Yang, K. Duan, C. Fan, C. Hu, and J. Wang, “Automatic ocular artifacts removal in EEG using deep learning,” *Biomed Signal Process Control*, vol. 43, pp. 148–158, May 2018, doi: 10.1016/j.bspc.2018.02.021.

[15] H. R. Hussien, E.-S. M. El-Kenawy, and A. I. El-Desouky, “EEG Channel Selection Using A Modified Grey Wolf Optimizer,” *European Journal of Electrical Engineering and Computer Science*, vol. 5, no. 1, pp. 17–24, 2021, doi: 10.24018/ejece.2021.5.1.265.

[16] R. Upadhyay, A. Manglick, D. K. Reddy, P. K. Padhy, and P. K. Kankar, “Channel optimization and nonlinear feature extraction for Electroencephalogram signals classification,” *Computers and Electrical Engineering*, vol. 45, pp. 222–234, Jul. 2015, doi: 10.1016/j.compeleceng.2015.03.015.

[17] A. Khorshidtalab, M. J. E. Salami, and R. Akmeliawati, “Motor imagery task classification using transformation based features,” *Biomed Signal Process Control*, vol. 33, pp. 213–219, 2017, doi: 10.1016/j.bspc.2016.12.006.

[18] V. Professor Hans Berger, “Nach JBechterew: Die Energie des lebenden,” 1875.

[19] P. Andersen and A. Lundberg, “(1951) Doubt and Certainty in Science,” Oxford University Press 8 Young, 1997.

[20] J. J. Vidal, “TOW ARD DIRECT BRAIN-COMPUTER COMMUNICATION,” 1973. [Online]. Available: www.annualreviews.org

[21] J. K. Chapin, K. A. Moxon, R. S. Markowitz, and M. A. L. Nicolelis, “Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex,” 1999. [Online]. Available: http://neurosci.nature.com

[22] B. Z. Allison, E. W. Wolpaw, and J. R. Wolpaw, “Brain-computer interface systems: Progress and prospects,” *Expert Review of Medical Devices*, vol. 4, no. 4. pp. 463–474, Jul. 2007. doi: 10.1586/17434440.4.4.463.

[23] R. Mertens and J. Polich, “P300 from a single-stimulus paradigm: passive versus active tasks and stimulus modality,” 1997.

[24] Y. Zhu, Y. Li, J. Lu, and P. Li, “EEGNet with Ensemble Learning to Improve the Cross-Session Classification of SSVEP Based BCI from Ear-EEG,” *IEEE Access*, vol. 9, pp. 15295–15303, 2021, doi: 10.1109/ACCESS.2021.3052656.

[25] G. Pfurtscheller, “Functional brain imaging based on ERD/ERS,” 2001. [Online]. Available: www.elsevier.com

[26] B. Z. Allison, E. W. Wolpaw, and J. R. Wolpaw, “Brain-computer interface systems: Progress and prospects,” *Expert Review of Medical Devices*, vol. 4, no. 4. pp. 463–474, Jul. 2007. doi: 10.1586/17434440.4.4.463.

[27] R. Acharya U., O. Faust, N. Kannathal, T. Chua, and S. Laxminarayan, “Non-linear analysis of EEG signals at various sleep stages,” *Comput Methods Programs Biomed*, vol. 80, no. 1, pp. 37–45, Oct. 2005, doi: 10.1016/j.cmpb.2005.06.011.

[28] S. Taran, P. C. Sharma, and V. Bajaj, “Automatic sleep stages classification using optimize flexible analytic wavelet transform,” *Knowl Based Syst*, vol. 192, p. 105367, 2020, doi: 10.1016/j.knosys.2019.105367.

[29] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, “Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection,” *IEEE Trans Biomed Eng*, vol. 54, no. 9, pp. 1545–1551, Sep. 2007, doi: 10.1109/TBME.2007.891945.

[30] B. J. Roach and D. H. Mathalon, “Event-related EEG time-frequency analysis: An overview of measures and an analysis of early gamma band phase locking in schizophrenia,” *Schizophr Bull*, vol. 34, no. 5, pp. 907–926, 2008, doi: 10.1093/schbul/sbn093.

[31] F. A. Alturki, K. Alsharabi, M. Aljalal, and A. M. Abdurraqeeb, “A DWT-Band power-SVM Based Architecture for Neurological Brain Disorders Diagnosis Using EEG Signals,” *2nd International Conference on Computer Applications and Information Security, ICCAIS 2019*, vol. 8, pp. 1–4, 2019, doi: 10.1109/CAIS.2019.8769492.

[32] A. Borowicz, “Using a multichannel Wiener filter to remove eye-blink artifacts from EEG data,” *Biomed Signal Process Control*, vol. 45, pp. 246–255, Aug. 2018, doi: 10.1016/j.bspc.2018.05.012.

[33] S. K. Noorbasha and G. F. Sudha, “Removal of EOG artifacts and separation of different cerebral activity components from single channel EEG—An efficient approach combining SSA–ICA with wavelet thresholding for BCI applications,” *Biomed Signal Process Control*, vol. 63, Jan. 2021, doi: 10.1016/j.bspc.2020.102168.

[34] Y. Zou, V. Nathan, and R. Jafari, “Automatic identification of artifact-Related independent components for artifact removal in EEG recordings,” *IEEE J Biomed Health Inform*, vol. 20, no. 1, pp. 73–81, Jan. 2016, doi: 10.1109/JBHI.2014.2370646.

[35] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, “Optimal spatial filtering of single trial EEG during imagined hand movement,” *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000, doi: 10.1109/86.895946.

[36] M. Rakibul Mowla, S. C. Ng, M. S. A. Zilany, and R. Paramesran, “Artifacts-matched blind source separation and wavelet transform for multichannel EEG denoising,” *Biomed Signal Process Control*, vol. 22, pp. 111–118, Jul. 2015, doi: 10.1016/j.bspc.2015.06.009.

[37] A. Jafarifarmand and M. A. Badamchizadeh, “Artifacts removal in EEG signal using a new neural network enhanced adaptive filter,” *Neurocomputing*, vol. 103, pp. 222–231, Mar. 2013, doi: 10.1016/j.neucom.2012.09.024.

[38] A. Borowicz, “Biomedical Signal Processing and Control Using a multichannel Wiener filter to remove eye-blink artifacts from EEG data,” vol. 45, pp. 246–255, 2018.

[39] H. Kaya, F. Eyben, A. A. Salah, and B. Schuller, “CCA BASED FEATURE SELECTION WITH APPLICATION TO CONTINUOUS DEPRESSION RECOGNITION FROM ACOUSTIC SPEECH FEATURES Bogazici University , Department of Computer Engineering , Istanbul / Turkey Technische Universit ¨ at M ¨ unchen , Institute for Human-Machine,” *Proceedings of the International Conference on Acoustic, Speech and Signal Processing*, pp. 3757–3761, 2014, doi: 10.1109/TIP.2008.924980.

[40] H. Nolan, R. Whelan, and R. B. Reilly, “FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection,” *J Neurosci Methods*, vol. 192, no. 1, pp. 152–162, Sep. 2010, doi: 10.1016/j.jneumeth.2010.07.015.

[41] R. Upadhyay, P. K. Padhy, and P. K. Kankar, “EEG artifact removal and noise suppression by Discrete Orthonormal S-Transform denoising,” *Computers and Electrical Engineering*, vol. 53, pp. 125–142, Jul. 2016, doi: 10.1016/j.compeleceng.2016.05.015.

[42] R. Mahajan and B. I. Morshed, “Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, kurtosis, and wavelet-ICA,” *IEEE J Biomed Health Inform*, vol. 19, no. 1, pp. 158–165, Jan. 2015, doi: 10.1109/JBHI.2014.2333010.

[43] V. Bono, S. Das, W. Jamal, and K. Maharatna, “Hybrid wavelet and EMD/ICA approach for artifact suppression in pervasive EEG,” *J Neurosci Methods*, vol. 267, pp. 89–107, Jul. 2016, doi: 10.1016/j.jneumeth.2016.04.006.

[44] X. Chen, A. Liu, Q. Chen, Y. Liu, L. Zou, and M. J. McKeown, “Simultaneous ocular and muscle artifact removal from EEG data by exploiting diverse statistics,” *Comput Biol Med*, vol. 88, pp. 1–10, Sep. 2017, doi: 10.1016/j.compbiomed.2017.06.013.

[45] P. Sawangjai *et al.*, “EEGANet: Removal of Ocular Artifact from the EEG Signal Using Generative Adversarial Networks,” *IEEE J Biomed Health Inform*, 2021, doi: 10.1109/JBHI.2021.3131104.

[46] M. Mathe, M. Padmaja, and B. Tirumala Krishna, “Intelligent approach for artifacts removal from EEG signal using heuristic-based convolutional neural network,” *Biomed Signal Process Control*, vol. 70, Sep. 2021, doi: 10.1016/j.bspc.2021.102935.

[47] S. Khatun, R. Mahajan, and B. I. Morshed, “Comparative Study of Wavelet-Based Unsupervised Ocular Artifact Removal Techniques for Single-Channel EEG Data,” *IEEE J Transl Eng Health Med*, vol. 4, 2016, doi: 10.1109/JTEHM.2016.2544298.

[48] B. Darkhovsky, A. Piryatinska, and A. Kaplan, “Binary classification of multi-channel EEG records based on the ϱ -complexity of continuous vector functions,” pp. 1–14, 2018.

[49] T. N. Lal *et al.*, “Support vector channel selection in BCI,” *IEEE Trans Biomed Eng*, vol. 51, no. 6, pp. 1003–1010, Jun. 2004, doi: 10.1109/TBME.2004.827827.

[50] J. Yang *et al.*, “Channel selection and classification of electroencephalogram signals: An artificial neural network and genetic algorithm-based approach,” *Artif Intell Med*, vol. 55, no. 2, pp. 117–126, Jun. 2012, doi: 10.1016/j.artmed.2012.02.001.

[51] V. S. Handiru and V. A. Prasad, “Optimized Bi-Objective EEG Channel Selection and Cross-Subject Generalization with Brain-Computer Interfaces,” *IEEE Trans Hum Mach Syst*, vol. 46, no. 6, pp. 777–786, Dec. 2016, doi: 10.1109/THMS.2016.2573827.

[52] A. Jafarifarmand, M. Badamchizadeh, and S. Khanmohammadi, “Biomedical Signal Processing and Control Real-time ocular artifacts removal of EEG data using a hybrid ICA-ANC approach,” vol. 31, pp. 199–210, 2017.

[53] V. S. Handiru and V. A. Prasad, “Optimized Bi-Objective EEG Channel Selection and Cross-Subject Generalization with Brain-Computer Interfaces,” *IEEE Trans Hum Mach Syst*, vol. 46, no. 6, pp. 777–786, 2016, doi: 10.1109/THMS.2016.2573827.

[54] Y. Park and W. Chung, “Optimal channel selection using covariance matrix and cross-combining region in EEG-based BCI; Optimal channel selection using covariance matrix and cross-combining region in EEG-based BCI,” 2019.

[55] J. Jin *et al.*, “Bispectrum-Based Channel Selection for Motor Imagery Based Brain-Computer Interfacing,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 10, pp. 2153–2163, Oct. 2020, doi: 10.1109/TNSRE.2020.3020975.

[56] F. Qi *et al.*, “Spatiotemporal-Filtering-Based Channel Selection for Single-Trial EEG Classification,” *IEEE Trans Cybern*, pp. 1–10, 2020, doi: 10.1109/tcyb.2019.2963709.

[57] A. Tiwari and A. Chaturvedi, “A Novel Channel Selection Method for BCI Classification Using Dynamic Channel Relevance,” *IEEE Access*, vol. 9, pp. 126698–126716, 2021, doi: 10.1109/ACCESS.2021.3110882.

[58] T. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, and I. Ahmad, “A review of channel selection algorithms for EEG signal processing,” *EURASIP J Adv Signal Process*, vol. 2015, no. 1, Dec. 2015, doi: 10.1186/s13634-015-0251-9.

[59] Y. Park and W. Chung, “Optimal channel selection using covariance matrix and cross-combining region in EEG-based BCI,” *7th International Winter Conference on Brain-Computer Interface, BCI 2019*, pp. 11–14, 2019, doi: 10.1109/IWW-BCI.2019.8737257.

[60] F. Fahimi, Z. Zhang, W. B. Goh, T. S. Lee, K. K. Ang, and C. Guan, “Inter-subject transfer learning with an end-to-end deep convolutional neural network for EEG-based BCI,” *J Neural Eng*, vol. 16, no. 2, 2019, doi: 10.1088/1741-2552/aaf3f6.

[61] J. Jin *et al.*, "Bispectrum-Based Channel Selection for Motor Imagery Based Brain-Computer Interfacing," vol. 28, no. 10, pp. 2153–2163, 2020.

[62] V. Bajaj, S. Taran, S. K. Khare, and A. Sengur, "Feature extraction method for classification of alertness and drowsiness states EEG signals," *Applied Acoustics*, vol. 163, p. 107224, 2020, doi: 10.1016/j.apacoust.2020.107224.

[63] S. Taran, V. Bajaj, D. Sharma, S. Siuly, and A. Sengur, "Features based on analytic IMF for classifying motor imagery EEG signals in BCI applications," *Measurement (Lond)*, vol. 116, no. October 2017, pp. 68–76, 2018, doi: 10.1016/j.measurement.2017.10.067.

[64] A. R. Hassan and M. I. H. Bhuiyan, "A decision support system for automatic sleep staging from EEG signals using tunable Q-factor wavelet transform and spectral features," *J Neurosci Methods*, vol. 271, pp. 107–118, 2016, doi: 10.1016/j.jneumeth.2016.07.012.

[65] R. Sharma, P. Sircar, and R. B. Pachori, "Automated focal EEG signal detection based on third order cumulant function," *Biomed Signal Process Control*, vol. 58, p. 101856, 2020, doi: 10.1016/j.bspc.2019.101742.

[66] C. Sravani, V. Bajaj, S. Taran, and A. Sengur, "Flexible Analytic Wavelet Transform Based Features for Physical Action Identification Using sEMG Signals," *IRBM*, vol. 41, no. 1, pp. 18–22, Feb. 2020, doi: 10.1016/j.irbm.2019.07.002.

[67] Z. Chen, G. Lu, Z. Xie, and W. Shang, "A Unified Framework and Method for EEG-Based Early Epileptic Seizure Detection and Epilepsy Diagnosis," *IEEE Access*, vol. 8, pp. 20080–20092, 2020, doi: 10.1109/access.2020.2969055.

[68] A. Gautam, N. Simões-capela, G. Schiavone, A. Acharyya, W. De Raedt, and C. Van Hoof, "A Data Driven Empirical Iterative Algorithm for GSR Signal Pre-Processing," no. November, 2018, doi: 10.23919/EUSIPCO.2018.8553191.

[69] R. Sharma, P. Sircar, and R. B. Pachori, "Automated focal EEG signal detection based on third order cumulant function," *Biomed Signal Process Control*, vol. 58, p. 101856, 2020, doi: 10.1016/j.bspc.2020.101856.

[70] P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, "A multi-class EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry," *Expert Syst Appl*, vol. 95, pp. 201–211, 2018, doi: 10.1016/j.eswa.2017.11.007.

[71] V. Vijean, M. Hariharan, A. Saidatul, and S. Yaacob, "Mental tasks classifications using S-transform for BCI applications," in *2011 IEEE Conference on Sustainable Utilization Development in Engineering and Technology, STUDENT 2011*, 2011, pp. 69–73. doi: 10.1109/STUDENT.2011.6089327.

[72] B. R. Nayana and P. Geethanjali, "Analysis of Statistical Time-Domain Features Effectiveness in Identification of Bearing Faults from Vibration Signal," *IEEE Sens J*, vol. 17, no. 17, pp. 5618–5625, 2017, doi: 10.1109/JSEN.2017.2727638.

[73] N. Robinson, A. P. Vinod, C. Guan, K. K. Ang, and T. K. Peng, "A Wavelet-CSP method to classify hand movement directions in EEG based BCI system," *ICICS 2011 - 8th International Conference on Information, Communications and Signal Processing*, pp. 11–15, 2011, doi: 10.1109/ICICS.2011.6174210.

[74] S. Siuly, S. K. Khare, V. Bajaj, H. Wang, and Y. Zhang, "A Computerized Method for Automatic Detection of Schizophrenia Using EEG Signals," vol. 4320, no. c, pp. 1–11, 2020, doi: 10.1109/TNSRE.2020.3022715.

[75] X. Song and S. C. Yoon, "Improving brain-computer interface classification using adaptive common spatial patterns," *Comput Biol Med*, vol. 61, pp. 150–160, Jun. 2015, doi: 10.1016/j.compbiomed.2015.03.023.

[76] Y. Jiao *et al.*, "Sparse Group Representation Model for Motor Imagery EEG Classification," *IEEE J Biomed Health Inform*, vol. 23, no. 2, pp. 631–641, Mar. 2019, doi: 10.1109/JBHI.2018.2832538.

[77] T. K. Reddy, V. Arora, L. Behera, Y. K. Wang, and C. T. Lin, "Multiclass Fuzzy Time-Delay Common Spatio-Spectral Patterns with Fuzzy Information Theoretic Optimization for EEG-Based Regression Problems in Brain-Computer Interface (BCI)," *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 10, pp. 1943–1951, 2019, doi: 10.1109/TFUZZ.2019.2892921.

[78] T. Zikov, S. Bibian, G. A. Dumont, M. Huzmezan, and C. R. Ries, "A wavelet based de-noising technique for ocular artifact correction of the electroencephalogram," in *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] [Engineering in Medicine and Biology*, IEEE, pp. 98–105. doi: 10.1109/IEMBS.2002.1134407.

[79] N. Bagh and M. R. Reddy, "Improving the Performance of Motor Imagery Based Brain-Computer Interface Using Phase Space Reconstruction," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, no. Mi, pp. 3075–3078, 2019, doi: 10.1109/EMBC.2019.8857066.

[80] V. Vijean, M. Hariharan, A. Saidatul, and S. Yaacob, "Mental tasks classifications using S-transform for BCI applications," in *2011 IEEE Conference on Sustainable Utilization Development in Engineering and Technology, STUDENT 2011*, 2011, pp. 69–73. doi: 10.1109/STUDENT.2011.6089327.

[81] T. Nguyen, I. Hettiarachchi, A. Khatami, L. Gordon-Brown, C. P. Lim, and S. Nahavandi, "Classification of Multi-Class BCI Data by Common Spatial Pattern and Fuzzy System," *IEEE Access*, vol. 6, pp. 27873–27884, May 2018, doi: 10.1109/ACCESS.2018.2841051.

[82] Y. Zhang *et al.*, "Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces," *Expert Syst Appl*, vol. 96, pp. 302–310, Apr. 2018, doi: 10.1016/j.eswa.2017.12.015.

[83] R. Zhang, Q. Zong, and X. Zhao, “A new convolutional neural network for motor imagery classification,” *Chinese Control Conference, CCC*, vol. 2019-July, no. 10, pp. 8428–8432, 2019, doi: 10.23919/ChiCC.2019.8865152.

[84] K. W. Ha and J. W. Jeong, “Motor imagery EEG classification using capsule networks,” *Sensors (Switzerland)*, vol. 19, no. 13, 2019, doi: 10.3390/s19132854.

[85] R. T. Schirrmeister *et al.*, “Deep learning with convolutional neural networks for EEG decoding and visualization,” *Hum Brain Mapp*, vol. 38, no. 11, pp. 5391–5420, 2017, doi: 10.1002/hbm.23730.

[86] D. verma Atul, “EEGNet: A Compact Convolutional NN for EEG-based BCI,” *On certain distance and degree based topological indices of Zeolite LTA frameworks*, no. December 2016, pp. 11–14, 2018.

[87] X. Deng, B. Zhang, N. Yu, K. Liu, and K. Sun, “Advanced TSGL-EEGNet for Motor Imagery EEG-Based Brain-Computer Interfaces,” *IEEE Access*, vol. 9, pp. 25118–25130, 2021, doi: 10.1109/ACCESS.2021.3056088.

[88] N. Yeung, R. Bogacz, C. B. Holroyd, and J. D. Cohen, “Detection of synchronized oscillations in the electroencephalogram: An evaluation of methods,” *Psychophysiology*, vol. 41, no. 6, pp. 822–832, 2004, doi: 10.1111/j.1469-8986.2004.00239.x.

[89] M. T. Akhtar, W. Mitsuhashi, and C. J. James, “Employing spatially constrained ICA and wavelet denoising, for automatic removal of artifacts from multichannel EEG data,” *Signal Processing*, vol. 92, no. 2, pp. 401–416, Feb. 2012, doi: 10.1016/j.sigpro.2011.08.005.

[90] N. P. Castellanos and V. A. Makarov, “Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis,” *J Neurosci Methods*, vol. 158, no. 2, pp. 300–312, Dec. 2006, doi: 10.1016/j.jneumeth.2006.05.033.

[91] S. Khatun, R. Mahajan, and B. I. Morshed, “Comparative Study of Wavelet-Based Unsupervised Ocular Artifact Removal Techniques for Single-Channel EEG Data,” *IEEE J Transl Eng Health Med*, vol. 4, 2016, doi: 10.1109/JTEHM.2016.2544298.

[92] M. Miao, A. Wang, and F. Liu, “Application of artificial bee colony algorithm in feature optimization for motor imagery EEG classification,” *Neural Comput Appl*, vol. 30, no. 12, pp. 3677–3691, 2018, doi: 10.1007/s00521-017-2950-7.

[93] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, “Filter Bank Common Spatial Pattern (FBCSP) in brain-computer interface,” *Proceedings of the International Joint Conference on Neural Networks*, pp. 2390–2397, 2008, doi: 10.1109/IJCNN.2008.4634130.

[94] R. Romo Vázquez, H. Vélez-Pérez, R. Ranta, V. Louis Dorr, D. Maquin, and L. Maillard, “Blind source separation, wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling,” *Biomed Signal Process Control*, vol. 7, no. 4, pp. 389–400, 2012, doi: 10.1016/j.bspc.2011.06.005.

[95] X. Chen, A. Liu, Q. Chen, Y. Liu, L. Zou, and M. J. McKeown, "Simultaneous ocular and muscle artifact removal from EEG data by exploiting diverse statistics," *Comput Biol Med*, vol. 88, pp. 1–10, Sep. 2017, doi: 10.1016/j.compbiomed.2017.06.013.

[96] M. Rakibul Mowla, S. C. Ng, M. S. A. Zilany, and R. Paramesran, "Artifacts-matched blind source separation and wavelet transform for multichannel EEG denoising," *Biomed Signal Process Control*, vol. 22, pp. 111–118, Jul. 2015, doi: 10.1016/j.bspc.2015.06.009.

[97] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *J Neurosci Methods*, vol. 158, no. 2, pp. 300–312, Dec. 2006, doi: 10.1016/j.jneumeth.2006.05.033.

[98] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *J Neurosci Methods*, vol. 158, no. 2, pp. 300–312, Dec. 2006, doi: 10.1016/j.jneumeth.2006.05.033.

[99] S. Khatun, R. Mahajan, and B. I. Morshed, "Comparative Study of Wavelet-Based Unsupervised Ocular Artifact Removal Techniques for Single-Channel EEG Data," *IEEE J Transl Eng Health Med*, vol. 4, 2016, doi: 10.1109/JTEHM.2016.2544298.

[100] F. Xiao and Y. Zhang, "A comparative study on thresholding methods in wavelet-based image denoising," in *Procedia Engineering*, 2011, pp. 3998–4003. doi: 10.1016/j.proeng.2011.08.749.

[101] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J Am Stat Assoc*, vol. 90, no. 432, pp. 1200–1224, 1995, doi: 10.1080/01621459.1995.10476626.

[102] A. Antoniadis, "Wavelet methods in statistics: Some recent developments and their applications," *Statistics Surveys*, vol. 2. pp. 16–55, 2008. doi: 10.1214/07-SS014.

[103] D. L. Don0h0 and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," 1994. [Online]. Available: http://biomet.oxfordjournals.org/

[104] R. Upadhyay, P. K. Padhy, and P. K. Kankar, "A comparative study of feature ranking techniques for epileptic seizure detection using wavelet transform," *Computers and Electrical Engineering*, vol. 53, pp. 163–176, Jul. 2016, doi: 10.1016/j.compeleceng.2016.05.016.

[105] G. Yu and Y. Zhou, "General linear chirplet transform," *Mech Syst Signal Process*, vol. 70–71, pp. 958–973, 2016, doi: 10.1016/j.ymssp.2015.09.004.

[106] R. Upadhyay, P. K. Padhy, and P. K. Kankar, "EEG artifact removal and noise suppression by Discrete Orthonormal S-Transform denoising," *Computers and Electrical Engineering*, vol. 53, pp. 125–142, Jul. 2016, doi: 10.1016/j.compeleceng.2016.05.015.

[107] H. T. Wu, Y. H. Chan, Y. T. Lin, and Y. H. Yeh, "Using synchrosqueezing transform to discover breathing dynamics from ECG signals," *Appl Comput Harmon Anal*, vol. 36, no. 2, pp. 354–359, Mar. 2014, doi: 10.1016/j.acha.2013.07.003.

[108] Y. Yang, Z. K. Peng, X. J. Dong, W. M. Zhang, and G. Meng, "General parameterized time-frequency transform," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 2751–2764, Jun. 2014, doi: 10.1109/TSP.2014.2314061.

[109] S. Basiri, E. Ollila, and V. Koivunen, "Alternative Derivation of FastICA with Novel Power Iteration Algorithm," *IEEE Signal Process Lett*, vol. 24, no. 9, pp. 1378–1382, Sep. 2017, doi: 10.1109/LSP.2017.2732342.

[110] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *J Neurosci Methods*, vol. 158, no. 2, pp. 300–312, Dec. 2006, doi: 10.1016/j.jneumeth.2006.05.033.

[111] H. Raza, D. Rathee, S. M. Zhou, H. Cecotti, and G. Prasad, "Covariate shift estimation based adaptive ensemble learning for handling non-stationarity in motor imagery related EEG-based brain-computer interface," *Neurocomputing*, vol. 343, pp. 154–166, 2019, doi: 10.1016/j.neucom.2018.04.087.

[112] M. Jochumsen *et al.*, "Quantification of movement-related EEG correlates associated with motor training: A study on movement-related cortical potentials and sensorimotor rhythms," *Front Hum Neurosci*, vol. 11, no. December, pp. 1–12, 2017, doi: 10.3389/fnhum.2017.00604.

[113] Y. Yu, Y. Liu, E. Yin, J. Jiang, Z. Zhou, and D. Hu, "An Asynchronous Hybrid Spelling Approach Based on EEG – EOG Signals for," vol. 27, no. 6, pp. 1292–1302, 2019.

[114] K. Jindal, R. Upadhyay, and H. S. Singh, "Application of hybrid GLCT-PICA de-noising method in automated EEG artifact removal," *Biomed Signal Process Control*, vol. 60, p. 101977, 2020, doi: 10.1016/j.bspc.2020.101977.

[115] K. Jindal, R. Upadhyay, and H. S. Singh, "EEG Artifact Removal and Noise Suppression Using Hybrid Glct -Ica Technique," *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, vol. 2018-Novem, pp. 2–6, 2019, doi: 10.1109/ICUMT.2018.8631219.

[116] K. Jindal, R. Upadhyay, and H. S. Singh, "Application of hybrid GLCT-PICA de-noising method in automated EEG artifact removal," *Biomed Signal Process Control*, vol. 60, p. 101977, 2020, doi: 10.1016/j.bspc.2020.101977.

[117] G. Yu and Y. Zhou, "General linear chirplet transform," *Mech Syst Signal Process*, vol. 70–71, pp. 958–973, 2016, doi: 10.1016/j.ymssp.2015.09.004.

[118] S. Dhargupta, M. Ghosh, S. Mirjalili, and R. Sarkar, "Selective Opposition based Grey Wolf Optimization," *Expert Syst Appl*, vol. 151, p. 113389, 2020, doi: 10.1016/j.eswa.2020.113389.

[119] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[120] V. Vakharia, V. K. Gupta, and P. K. Kankar, “A comparison of feature ranking techniques for fault diagnosis of ball bearing,” *Soft comput*, vol. 20, no. 4, pp. 1601–1619, 2016, doi: 10.1007/s00500-015-1608-6.

[121] E. B. Sadeghian and M. H. Moradi, “Continuous detection of motor imagery in a four-class asynchronous BCI,” *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pp. 3241–3244, 2007, doi: 10.1109/IEMBS.2007.4353020.

[122] U. Talukdar, S. M. Hazarika, and J. Q. Gan, “Adaptation of Common Spatial Patterns based on mental fatigue for motor-imagery BCI,” *Biomed Signal Process Control*, vol. 58, p. 101829, 2020, doi: 10.1016/j.bspc.2019.101829.

[123] A. Bablani, D. R. Edla, and S. Member, “Novel Approach for EEG Channel Selection and SVM Parameter Optimization,” vol. 14, no. 11, pp. 3057–3068, 2019.

[124] R. Yao *et al.*, “Dynamic Changes of Brain Networks during Working Memory Tasks in Schizophrenia,” *Neuroscience*, vol. 453, pp. 187–205, 2021, doi: 10.1016/j.neuroscience.2020.11.007.

[125] Y. Zhang, C. S. Nam, G. Zhou, J. Jin, X. Wang, and A. Cichocki, “Temporally constrained sparse group spatial patterns for motor imagery BCI,” *IEEE Trans Cybern*, vol. 49, no. 9, pp. 3322–3332, 2019, doi: 10.1109/TCYB.2018.2841847.

[126] K. Jindal and R. Upadhyay, “Epileptic seizure detection from EEG signal using Flexible Analytical Wavelet Transform,” *2017 International Conference on Computer, Communications and Electronics, COMPTELIX 2017*, pp. 67–72, 2017, doi: 10.1109/COMPTELIX.2017.8003940.

[127] C. Sravani, V. Bajaj, S. Taran, and A. Sengur, “Flexible Analytic Wavelet Transform Based Features for Physical Action Identification Using sEMG Signals,” *Irbm*, vol. 41, no. 1, pp. 18–22, 2020, doi: 10.1016/j.irbm.2019.07.002.

[128] K. Jindal and R. Upadhyay, “Epileptic seizure detection from EEG signal using Flexible Analytical Wavelet Transform,” *2017 International Conference on Computer, Communications and Electronics, COMPTELIX 2017*, no. July 2017, pp. 67–72, 2017, doi: 10.1109/COMPTELIX.2017.8003940.

[129] J. Sun *et al.*, “Complexity analysis of EEG, MEG, and fMRI in mild cognitive impairment and Alzheimer’s disease: A review,” *Entropy*, vol. 22, no. 2. 2020. doi: 10.3390/e22020239.

[130] Z. Yin, L. Liu, J. Chen, B. Zhao, and Y. Wang, “Locally robust EEG feature selection for individual-independent emotion recognition,” *Expert Syst Appl*, vol. 162, p. 113768, 2020, doi: 10.1016/j.eswa.2020.113768.

[131] J. T. Lindgren, A. Merlini, A. Lecuyer, and F. P. Andriulli, “SimBCI-A framework for studying BCI methods by simulated EEG,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 11, pp. 2006–2105, 2018, doi: 10.1109/TNSRE.2018.2873061.

[132] B. Richhariya, M. Tanveer, and A. H. Rashid, “Diagnosis of Alzheimer’s disease using universum support vector machine based recursive feature elimination (USVM-RFE),” *Biomed Signal Process Control*, vol. 59, 2020, doi: 10.1016/j.bspc.2020.101903.

[133] H. Göksu, “BCI oriented EEG analysis using log energy entropy of wavelet packets,” *Biomed Signal Process Control*, vol. 44, pp. 101–109, 2018, doi: 10.1016/j.bspc.2018.04.002.

[134] X. Gao, X. Yan, P. Gao, X. Gao, and S. Zhang, “Automatic detection of epileptic seizure based on approximate entropy, recurrence quantification analysis and convolutional neural networks,” *Artif Intell Med*, vol. 102, no. June 2019, p. 101711, 2020, doi: 10.1016/j.artmed.2019.101711.

[135] C. Barros, C. A. Silva, and A. P. Pinheiro, “Advanced EEG-based learning approaches to predict schizophrenia: Promises and pitfalls,” *Artif Intell Med*, vol. 114, no. December 2020, p. 102039, 2021, doi: 10.1016/j.artmed.2021.102039.

[136] C. Saitis and K. Kalimeri, “Multimodal Classification of Stressful Environments in Visually Impaired Mobility Using EEG and Peripheral Biosignals,” *IEEE Trans Affect Comput*, vol. XX, no. X, pp. 1–12, 2018, doi: 10.1109/TAFFC.2018.2866865.

[137] K. Jindal, R. Upadhyay, and H. S. Singh, “EEG Artifact Removal and Noise Suppression Using Hybrid Glct -Ica Technique,” *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, vol. 2018-Novem, 2019, doi: 10.1109/ICUMT.2018.8631219.

[138] H. Raza, A. Chowdhury, S. Bhattacharyya, and S. Samothrakis, “Single-Trial EEG Classification with EEGNet and Neural Structured Learning for Improving BCI Performance,” 2020.

[139] L. Li, G. Xu, J. Xie, and M. Li, “Classification of single-trial motor imagery EEG by complexity regularization,” *Neural Comput Appl*, vol. 31, no. 6, pp. 1959–1965, 2019, doi: 10.1007/s00521-017-3174-6.

[140] J. Olias, R. Martin-Clemente, M. A. Sarmiento-Vega, and S. Cruces, “EEG signal processing in mi-bci applications with improved covariance matrix estimators,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 5, pp. 895–904, 2019, doi: 10.1109/TNSRE.2019.2905894.

[141] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, “Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection,” *IEEE Trans Biomed Eng*, vol. 54, no. 9, pp. 1545–1551, 2007, doi: 10.1109/TBME.2007.891945.

[142] F. Lotte and C. Guan, “Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms,” *IEEE Trans Biomed Eng*, vol. 58, no. 2, pp. 355–362, 2011, doi: 10.1109/TBME.2010.2082539.

[143] J. Jin, Y. Miao, I. Daly, C. Zuo, D. Hu, and A. Cichocki, “Correlation-based channel selection and regularized feature optimization for MI-based BCI,” *Neural Networks*, vol. 118, pp. 262–270, 2019, doi: 10.1016/j.neunet.2019.07.008.

[144] S. Khare and V. Bajaj, “Adaptive Tunable Q Wavelet Transform based Emotion Identification,” *IEEE Trans Instrum Meas*, vol. 9456, no. c, pp. 1–1, 2020, doi: 10.1109/TIM.2020.3006611.

[145] L. He, B. Liu, D. Hu, Y. Wen, M. Wan, and J. Long, “Motor imagery EEG signals analysis based on Bayesian network with Gaussian distribution,” *Neurocomputing*, vol. 188, pp. 217–224, 2016, doi: 10.1016/j.neucom.2015.05.133.

[146] G. Yu, “A Concentrated Time-Frequency Analysis Tool for Bearing Fault Diagnosis,” *IEEE Trans Instrum Meas*, vol. 69, no. 2, pp. 371–381, 2020, doi: 10.1109/TIM.2019.2901514.

[147] F. Wang *et al.*, “Residual Attention Network for Image Classification.”

[148] R. Yao *et al.*, “Residual attention graph convolutional network for web services classification,” *Neurocomputing*, vol. 440, no. 1, pp. 45–57, 2021, doi: 10.1038/s41598-021-83350-6.

[149] W. Qu *et al.*, “A Residual Based Attention Model for EEG Based Sleep Staging,” *IEEE J Biomed Health Inform*, vol. 24, no. 10, pp. 2833–2843, Oct. 2020, doi: 10.1109/JBHI.2020.2978004.

[150] M. Saqlain, B. Jargalsaikhan, and J. Y. Lee, “A voting ensemble classifier for wafer map defect patterns identification in semiconductor manufacturing,” *IEEE Transactions on Semiconductor Manufacturing*, vol. 32, no. 2, pp. 171–182, 2019, doi: 10.1109/TSM.2019.2904306.

[151] Y. You, W. Chen, M. Li, T. Zhang, Y. Jiang, and X. Zheng, “Automatic focal and non-focal EEG detection using entropy-based features from flexible analytic wavelet transform,” *Biomed Signal Process Control*, vol. 57, p. 101761, 2020, doi: 10.1016/j.bspc.2019.101761.

[152] G. Żabiński, J. Gramacki, A. Gramacki, E. Miśta-Jakubowska, T. Birch, and A. Disser, “Multi-classifier majority voting analyses in provenance studies on iron artefacts,” *J Archaeol Sci*, vol. 113, 2020, doi: 10.1016/j.jas.2019.105055.

[153] D. Rahmawati, N. R. U. Chasanah, and R. Sarno, “Classify epilepsy and normal Electroencephalogram (EEG) signal using wavelet transform and K-nearest neighbor,” *Proceeding - 2017 3rd International Conference on Science in Information Technology: Theory and Application of IT for Education, Industry and Society in Big Data Era, ICSITech 2017*, vol. 2018-Janua, pp. 110–114, 2017, doi: 10.1109/ICSITech.2017.8257094.

[154] U. Budak, V. Bajaj, Y. Akbulut, O. Atila, and A. Sengur, “An effective hybrid model for EEG-based drowsiness detection,” *IEEE Sens J*, vol. 19, no. 17, pp. 7624–7631, 2019, doi: 10.1109/JSEN.2019.2917850.

[155] D. Das, T. Bhattacharjee, S. Datta, A. D. Choudhury, P. Das, and A. Pal, “Classification and quantitative estimation of cognitive stress from in-game keystroke analysis using EEG and GSR,” *2017 IEEE Life Sciences Conference, LSC 2017*, vol. 2018-Janua, pp. 286–291, 2018, doi: 10.1109/LSC.2017.8268199.

[156] M. C. Cieslak, A. M. Castelfranco, V. Roncalli, P. H. Lenz, and D. K. Hartline, “t-Distributed Stochastic Neighbor Embedding (t-SNE): A tool for eco-physiological transcriptomic analysis,” *Mar Genomics*, vol. 51, no. November, p. 100723, 2020, doi: 10.1016/j.margen.2019.100723.

[157] M. Xu *et al.*, “Learning EEG topographical representation for classification via convolutional neural network,” *Pattern Recognit*, no. xxxx, p. 107390, 2020, doi: 10.1016/j.patcog.2020.107390.

[158] A. Salazar-Ramirez, J. I. Martin, R. Martinez, A. Arruti, J. Muguerza, and B. Sierra, “A hierarchical architecture for recognising intentionality in mental tasks on a brain-computer interface,” *PLoS One*, vol. 14, no. 6, pp. 1–18, 2019, doi: 10.1371/journal.pone.0218181.