EEG Signal Channel Selection using Genetic Algorithm

Author

Muhammad Sohaib Anjum

000032778

Supervisor

Dr. Ali Hassan

Department of Computer and Software Engineering

College of Electrical and Mechanical Engineering (CEME),

National University of Sciences and Technology (NUST),

Islamabad,

(August 2023)

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by NS Muhammad Sohaib Anjum Registration No. 00000327786 of College of E&ME has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the thesis.

Signature:

Name of Supervisor: Dr Ali Hassan

Date: 08-08-2023

Signature of HOD

(Dr Usman Qamar)

Date: 08-08-2023

Signature of Dean

(Brig Dr Nasir Rashid)

Date: 08 AUG 2023

I dedicate this thesis to my parents, supportive sisters, and all those who believe in the power of learning. Their unwavering support, guidance, and belief in my abilities have been crucial in achieving this milestone. I am grateful for their presence in my life and their constant encouragement. This dedication is a tribute to their love and the shared belief that knowledge is a catalyst for personal growth and transformation

Acknowledgements

Glory be to Allah (S.W.A), the Creator, the Sustainer of the Universe. Who only has the power to honor whom He please, and to abase whom He please. Verily no one can do anything without His will. From the day, I came to NUST till the day of my departure, He was the only one Who blessed me and opened ways for me and showed me the path of success. There is nothing which can payback for His bounties throughout my research period to complete it successfully.

Abstract

This thesis proposes a novel approach for selecting the optimal subset of EEG channels that can best discriminate between different types of events. The proposed method combines common spatial pattern (CSP) filtering, artifact removal, linear discriminant analysis (LDA) classification, and genetic algorithm (GA) optimization. The proposed methodology involves several steps, including preprocessing, event separation, artifact removal, initial population creation, fitness calculation, genetic algorithm iteration, and validation. The aim is to find the best set of EEG channels that can best discriminate between different types of events using a combination of machine learning and genetic algorithm techniques. The proposed approach has the potential to enhance the accuracy and reliability of EEG signal analysis, while considering the individual variability, spatial resolution, and frequency band of the signal. The approach is tested on a publicly available dataset BCI Competition IV Datasets 2a and 2b, and the results show that it outperforms existing methods in the literature with the accuracy of 87.40 % at 16 channels. Overall, the proposed methodology has important implications for advancing the state-of-the-art in EEG signal channel selection and improving the interpretability and clinical utility of EEG-based brain-computer interfaces.

Keywords: EEG channels, Optimal subset selection, , Genetic algorithm (GA), Machine learning, Event separation, Preprocessing, Spatial resolution, Frequency band, BCI Competition IV Datasets 2a and 2b, Accuracy, Reliability, Interpretability, Clinical utility, Brain-computer interfaces (BCIs),

Table of Contents

Declaration .... **iii**

Acknowledgements .... **iv**

Abstract .... **v**

List of Figures .... **viii**

List of Tables .... **ix**

Chapter 1: Introduction **1**

1.1 Problem Statement 1

1.2 Significance and Rationale 2

1.3 Research Objectives 2

1.4 Research Questions 2

1.3 Thesis Outlines 3

Chapter 2: Literature Review **4**

2.1 Literature Review 5

2.2 Summary 20

Chapter 3: Datasets **21**

3.1 BCI Competition IV Datasets 2a 21

3.3 BCI Competition IV Datasets 2b 25

3.3.1 Experimental paradigm 25

3.3.2 Data recording 26

3.3.3 Evaluation 30

3.4 Summary of Chapter 31

Chapter 4: Methodology **34**

4.1 Preprocessing 35

4.3 Event Separation ... 36
4.4 Window Extraction ... 37
4.5 Artifact Removal ... 38
4.6 Initial Population Creation ... 38
4.7 Fitness Calculation ... 39
4.8 Genetic Algorithm ... 40
4.9 Elitism ... 43
4.10 Generation Iteration ... 43
4.11 Validation ... 44
4.12 Summary of Chapter ... 45
CHAPTER 5: Results and Discussion ... 48
5.1 Result and Discussion of BCI Competition IV Datasets 2a ... 50
5.1.1 Channels selected by the genetic algorithm ... 50
5.1.2 Accuracy's of different channel selection ... 52
5.1.3 Comparison with state-of-the-art approaches ... 54
5.1.4 Discussion ... 55
5.2 Result and Discussion of BCI Competition IV Datasets 2b ... 57
5.2.1 Channels selected by the genetic algorithm ... 57
5.2.1 Accuracy of different channel selected ... 58
5.2.2 Comparison with state-of-the-art approach ... 60
5.3 Discussion ... 61
5.3 Summary of Chapter ... 63
Chapter 6: Conclusion ... 65
6.1 Future Directions ... 66

References 68

List of Figures

Figure 1:Timing Scheme of one session 26

Figure 2:Electrode montage of the three monopolar EOG channels 27

Figure 3:Timing Scheme of Paradigm. 29

Figure 4:Pre-Processing Block Diagram 35

Figure 5:Flow Chart of Genetic Algorithm 41

Figure 6:Representation of genetic algorithm rules 42

Figure 7: Flow Chart of methodology 46

Figure 8: Accuracy vs. Channel graph 54

List of Tables

Table 1:Summary of Literature Review 19

Table 2:Summary of BCI Competition IV Datasets 2a 24

Table 3:List of all files in BCI Competition IV Datasets 2b 26

Table 4:Summary of Dataset 2b 32

Table 5:Accuracy's and weight at different channels selected. 51

Table 6:Accuracy's at different channels selected. 53

Table 7:Comparison with State of the art Approaches 55

Table 8:Channel Selection by Genetic Algorithm 58

Table 9:Classification Accuracy of Different Channel Selection Strategies 59

Table 10:Comparison with State of the art Approaches 60

Chapter 1: Introduction

The Brain electrical activity can be monitored and recorded non-invasively using (EEG) electroencephalography. The electrical potentials generated by the coordinated activity of millions of neurons in the cerebral cortex are the source of EEG signal, which is recorded at the scalp. The EEG signal is a complicated and noisy signal that provides insight into normal and abnormal brain activity [1],

The choice of channels for recording the signal is a significant obstacle in EEG signal analysis. Various brain parts are responsible for a variety of activities, and the scalp is divided into distinct regions. As a result, getting the EEG recording channels right is crucial for accurate and reliable analysis [2].

Several uses, including brain-computer interfaces, clinical diagnosis, and cognitive neuroscience research, necessitate careful consideration of the channels used to record EEG signal. Through the interpretation of EEG signal, external devices can be controlled by the user, such as a robotic arm or the cursor on a computer screen, in brain-computer interfaces. The EEG signal is used to detect and diagnose a wide range of neurological disorders in the clinic, including epilepsy, dementia, and TBI. The EEG signal is used in cognitive neuroscience studies to probe the brain's inner workings in order to learn more about things like attention, memory, and decision-making [3].

Electrodes placed on the scalp are the standard method of capturing the EEG signal. Standardized placement systems, such as the International 10-20 systems, are used to determine where on the scalp to place the electrodes [4].

1.1 Problem Statement

Due to vast number of EEG channels and the need to identify the optimal set of channels that capture the most relevant information about motor imagery, selecting the appropriate channels for motor imagery-based brain-computer interfaces (BCIs) can be a difficult and challenging task. EEG channel selection has been approached from a variety of perspectives, but the majority of these approaches rely on either manual selection or heuristic methods, neither of which guarantees the discovery of the best possible combination of channels.

Therefore, there is a requirement for a method of EEG channel selection that is both automated and accurate, so that motor imagery-based brain-computer interfaces (BCIs) can perform better. It has been suggested that genetic algorithms (GA) could be used to solve this issue because they have the ability to optimize the selection of EEG channels based on the classification accuracy of motor imagery signals.

1.2 Significance and Rationale

The proper selection of EEG channels is essential to the functionality of motor imagery-based brain-computer interfaces (BCIs). An inaccurate selection of channels can result in poor performance, which can restrict the ways in which the technology can be put to use in the real world. The proposed application of GA in the motor imagery-based brain-computer interfaces (BCIs) for the purpose of EEG channel selection has the potential to enhance the BCI system accuracy by determining the best possible combination of EEG channels. The method that has been proposed has the potential to make an important contribution to the fields of EEg signal processing and motor imagery-based brain-computer interfaces.

1.3 Research Objectives

- A channel selection method is a way of choosing a subset of EEG channels that are most relevant to a particular task. This can be done for a number of reasons, such as to reduce computational complexity, improve classification accuracy, or reduce the setup time. There are many different channel selection methods available, each with its own advantages and disadvantages.
- The most relevant channels from the EEG signal are the ones that contain the most information about the task that is being performed. For example, if the task is to classify different types of brain waves, then the most relevant channels would be the ones that are associated with those brain waves.

1.4 Research Questions

- How does the performance of proposed method compare to other channel selection method?

- Can genetic algorithms be used to select the optimal set of EEG channels for motor imagery based BCIs?

1.3 Thesis Outlines

The rest of the thesis is organized as follows.

Chapter 2 present the literature review of all previous research on Channel Selection. Chapter 3 describes the data set. Chapter 4 describes the methodology used for developing the Channel Selection using Genetic Algorithm. Chapter 5 describe the Genetic algorithm. Chapter 6 presents the results obtained from the research and coding and discusses the interpretation of the results and comparison with other research. Finally, Chapter 7 summarizes the main findings of the study, discusses their implications, and suggests directions or future research.

Chapter 2: Literature Review

In the next section, we will conduct a literature review on the topic of employing genetic algorithms for EEG channel selection in the context of studies on motor imagery. The use of motor imagery as a method for researching the brain's motor system shows promise; nevertheless, identifying the most appropriate channels for signal capture is an essential step for achieving accurate measurement and categorization. In EEG-based research on motor imagery, genetic algorithms are a common method for channel selection because of their ability to scan the expansive universe of possible channel combinations in an effective manner.

To begin, we will provide a review of the relevant literature on EEG-based motor imagery research. This will include a discussion of the fundamental concepts of both motor imagery and EEG as well as an examination of the various methods that are currently used for selecting EEG channels. Following that, we will centre our attention on genetic algorithms application to the process of EEG channel selection in the context of motor imagery research. In this part, we will examine the most important research that have used genetic algorithms for channel selection and then analyse the benefits and drawbacks of using this methodology.

In addition, we will evaluate the efficacy of genetic algorithms in relation to the performance of several other techniques for EEG channel selection, such as spatial filters and many other feature selection strategies. In the final part of this presentation, we will talk about the potential uses of genetic algorithms in this subject, as well as the future directions of EEG channel selection for motor imagery research.

We hope that by the time you reach the conclusion of this chapter, you will have a complete grasp of the function that genetic algorithms play in the selection of EEG channels for motor imaging research, as well as their benefits and drawbacks in comparison to other methodologies. Our review will not only shed light on the field of current state of art, but it will also assist in directing the course of future study.

2.1 Literature Review

Systems for establishing a direct line of communication between computers and human brain are called as brain-computer interface (BCI) systems. These devices examine and interpret brain impulses to convert them into commands for computers. BCI aims to make it possible for persons with impairments to operate tools and software using their thoughts, feelings, or other brain signals. The use of many electrodes to capture brain signals has been made possible by recent developments in BCI technology. However, using a lot of electrodes can make the system more complex, take longer, and cost more to compute. For decreasing of dimensionality of data and eliminate unnecessary channels while limiting noise effects, channel selection algorithms have been created [5]The Deep GA Fitness Formation (DGAFF) technique of channel selection is a suggested genetic algorithm and sequential search strategy. Genetic algorithm is a search algorithm that takes its cues from the natural selection process, in which the most fit individuals are chosen to give rise to the following generation. With the selection of a subset of channels that are most likely to improve classification accuracy, the sequential search strategy shrinks the search space. Using a fitness creation mechanism, the DGAFF approach quickens the genetic algorithm's convergence. Depending on how well each channel contributes to classification accuracy, the fitness generation mechanism modifies the selection probability for each channel. This decreases system time and computational expenses while increasing search process efficiency. It is easy to use Deep neural network that starts the entire model training process serves as the foundation for the evaluation of the suggested approach. Based on the provided job, the neural network takes the chosen channels as input that generates the relevant output. Based on the classification task's accuracy as well as the time and computational expense necessary to complete it, the system's performance is assessed [6]

A device or software known as brain-computer interface (BCI) that uses brain impulses instead of the neuromuscular pathways that ordinarily carry commands from the brain to the body to control other devices or software. BCI's major objective is to restore or replace funtional function for persons with neuromuscular illness such as stroke, spinal cord damage, amyotrophic lateral sclerosis and cerebral palsy. BCI technology has developed from straightforward single neuron-based and electroencephalography-based spelling device control to more sophisticated systems that use a variety of brain signals to operate wheelchairs, robotic arms, prosthetic limbs, cursors, and other devices. BCIs may potentially be useful for improving the performance of medical professionals and for stroke rehabilitation. BCI technology is an exciting area of study and development that is attracting the attention of scientists, engineers, physicians, and the general public [7]Yet, in order for BCI technology to reach its full potential, three important issues must be resolved. Hardware for signal-acquisition is the first area. BCI systems need hardware that is practical, transportable, secure, and adaptable to any situation. Currently available BCI systems collect brain data from a variety of sensors, including electroencephalography, intracortical, and electrocorticographic signals. The evolution of BCI technology depends on the creation of novel, trustworthy, and useful signal-acquisition techniques. The second aspect is the validation of BCI systems by extensive research on their application to the daily lives of people with severe disabilities. It is crucial to comprehend how BCI systems can be used in the daily lives of individuals with disabilities, as well as how efficient and practical they are. For this, BCI devices must undergo comprehensive testing and evaluation in both clinical trials and real-world situations. The third area is the BCI system's reliability. It is necessary to increase the daily and moment-to-moment reliability of BCI performance in order for it to be as reliable as naturally occurring muscle-based function. To do this, sophisticated signal processing algorithms, machine learning strategies, and adaptive control approaches must be created. These methods must be able to modify in response to changes in the user's brain signals over time [8]

Brain-computer interface (BCI) helps to enable patients with paralysis or severe neuromuscular problems to communicate and manage themselves by converting ideas into actions without the use of voluntary muscle movement. BCI technology is a new and quickly emerging discipline that has ability to greatly improve the life quality for people with extreme disabilities. While various BCI prototypes have been produced, the majority of them only work in laboratory settings. Many hurdles must be overcome before it can be used at home. One critical step is to develop protocols that make it simple to set up and utilize BCI systems in a realistic setting. This would entail picking features such as electrode placements and frequency components for specific motor imaging tasks automatically [9]

Another difficulty is to employ as few recording electrodes as feasible, aiming for an optimal single EEG channel. This would lower the system's complexity, time, and cost while decreasing noise impacts. Furthermore, training time must be reduced, possibly through game-like feedback and automatic detection of artifacts such uncontrolled muscular activity. Solving these issues will open the way for usable BCI systems for a diverse range of users and applications. We may anticipate to see practical and dependable devices that allow to communicate and control for those with extreme disabilities, dramatically increasing their quality of life, as BCI technology advances [10].

Brain function and its other electrophysiological measure and EEG have been theorised to be a potential non-musculer conduit for sending information and commands to the outside world, or a "brain-computer interface" (BCI). During the course of the previous 15 years, fruitful BCI research initiatives have arise. New insights into brain fucntion, the introduction of low price powerful computer equipment and growing rate of needs and potentials of people with disabilities are all motivating factors in the development of these programmes, which aim to create control technology and new augmentative communication for people with extreme neuromuscular disorders such as brainstem stroke, spinal cord injury and amyotrophic lateral sclerosis. The short-term objective is to give these users, who may be thoroughly paralysed or "locked in," with basic communication abilities so that they can make requests of carers and use technology like word processors and neuroprostheses. Today's BCIs can interpret a wide variety of electrophysiological signals to infer the user's goals. These signals include cortical neuronal activity as well as scalp recordings of slow potentials cortical, mu and bets rhythms and P300 potentials. Instantaneously, these are translated into actions that can be taken on a computer screen or other device. These signals are used by the user to encode commands, which are then decoded by the BCI [11] As a result, consistent functioning necessitates initial and ongoing adaptations from both the user and the BCI system. The top data rates of contemporary BCIs are around 25 bits per minute. Those with severe impairments that avoid them from using standard augmentative communication methods may benefit from this constrained capacity. Many possible BCI applications, such as neuroprosthesis control, may, however, require faster data transmission rates. The following variables will determine the course of future events: development of training methods for assisting users in gaining and maintaining control; it shows that BCI development and research is and interdisciplinary issue including psychology, mathematics, engineering, neurobiology and computer science; identification of signals,such as spontaneous rhythms, evoked potentials, or neuronal firing rates, that users are easily control independently of activity in conventional motor output pathways; development of BCIs. A larger emphasison peer-reviewed scientific articles and a lack of dramatic and often inaccurate attention of media, which helps to encourage unrealistic public thinkings and scepticism among other academics, would be beneficial to BCI technology development. BCI systems have the potential to be a game-changing new communication and control alternative for people with motor disabilities, and even for people with no disabilities who may find it beneficial in certain contexts, if these concerns are properly recognized and addressed [12].

This article tackles a prevalent issue with EEG based Brain-computer interfaces (BCI): the requirement for a channels of large numbers to obtain excellent classification performance, limiting its practical application [13]. The authors suggest a new method termed (STECS) Spatiotemporal-filtering-based channel selection, which uses EEG data's spatiotemporal information to ideintify automatically disciminative small numbers of channels. The problem of channel selectionis formulated by STECS as a spatiotemporal filter optimization problems with a group sparsity constraint to ensure that only a limited subset of channels is chosen. To overcome this optimization problem, the authors devised a computationally efficient technique. The effectiveness of STECS was assessed using 3 motor imagery EEG channels. When compare to state-of-the-art spatiotemporal filtering algorithms used full EEg channels, the results demonstrated that STECS achieved comparable classification performance using only the half channels. Additionally, STECS beat conventional channel selection algorithms significantly. Overall, the findings indicate that STECS has the potential to simplify BCI setups and improve their practical value by lowering the number of needed channels while maintaining good classification performance [14]

The focus of this research is on improving brain-computer interfaces through the useof (MI) motor imagery categorization using common spatial pattern (CSP)-based spatial filtering for extracting EEG features (BCIs). Although several methods have been developed to increase the selection of the time window is often done heuristically, which can result in inefficient feature extraction. This is because the time window and band frequency of EEG segments affect the performance of CSP. To improve the EEG classification accuracy, the authors propose a new method they call (TSGSP) temporally constrained sparse group spatial pattern, which simultaneously optimises time frames and filter bands within CSP. The first step in the method is to generate spectrally-specific signals by bandpass filtering raw EEG data across a series of overlapping filter bands. Sliding window analysis further divides each of these signals into numerous subseries. The next step, to extract strong CSP features, is to employ a multitask learning framework to carry out a joint sparse optimisation time windows and filter bands, the latter of which is constrained by a need for temporal smoothness. Finally, On refined EEG data, a linear support vector machine classifier is trained to accurately detect MI tasks [15]. Using three publicly available EEG data-sets (BCI Competition IV datasets IIa, BCI Competition III datasets IIIa, and BCI Competition IV dataset IIb), the effectiveness of TSGSP is evaluated and compare to many othr competing methodologies.

The result of experiment shows that TSGSP achieves superior classification performance than these other methods, with average accuracies of 88.5%, 83.3%, and 84.3% across the three datasets. By adjusting temporal windows and filter bands within CSP to extract robust EEG properties, the proposed TSGSP approach has potential to enhance the performance of MI-based BCIs [16]

This study provides an in-depth examination of recent advances in electroencephalography (EEG) channel selection algorithms and their applications. With so many EEG channels accessible, effective channel selection methods are required to extract the most important information while minimizing computing complexity, overfitting, and setup time in various EEG applications. In most approaches, signal processing technologies that includes power spectral estimation, time domain analysis and wavelet transform have been employed for the channel selection and feature extraction [17]For the evaluation of the selected subset of channels, many evaluation mechanisms including as hybrid, embedding, wrapper, filtering and human based techniques have been widely used. The study discusses the many approaches and strategies utilized in EEG channel selection, as well as their benefits and drawbacks. The approaches are also classified in the study based on the evaluation approach. The study gives an overview of current advances in EEG channel selection methods, as well as their applications in motor imagery classification, mental task classification, emotion classification, seizure detection/prediction, sleep state classification and medication effects diagnosis. The poll also analyzes the field's difficulties and potential research goals [18]

Recent developments in brain imaging technology, storag capacity and measurement approaches have resulted in a previously unheard-of supply of high temporal resolution neural data. These findings offer a rare chance to develop a mechanistic knowledge of not only structure of circuit, but also dynamics of circuit and their roles in diseae and cognition. To make sense of this massive amount of data, a variety of modeling models that take into account the brain's temporally developing interconnected structure have been developed. These models seek to capture the brain's dynamic nature and summarize it in a dynamic graph. The authors review recent breakthroughs in dynamic graph modeling methodologies and their applications in this study. Latest research to represent dynamic patterns of connection, activity of dynamic patterns, and activity of dynamic patterns atop connectivity are described by the authors. They also go over fundamental aspects of statistical testing, such as parametric and non-parametric techniques [19]. Lastly, the authors discuss how to understand dynamic graph architecture carefully and accurately, as well as describe crucial future areas for method improvement. Further integration of numerous imaging modalities, refining of model selection criteria, and development of models that can capture both temporal and spatial dynamics of brain were among the future prospects. Overall, this study emphasizes the significance of dynamic graph modeling in understanding the brain's complex and dynamic nature, and it provides a useful overview of recent achievements in this subject [20].

Motorimagery brain-computer interfaces (MI-BCI) system development necessitates precise, quick, and dependable multiclass categorization of electroencephalography (EEG) information, which is a difficult issue. To improve execution time and classification accuracy during testing and training, the authors suggest upgrades to two well-known feature extractors, Riemannian covariance approaches and common spatial pattern (CSP), as well as classifier support vector machine (SVM). These two feature extractors are extensively extended by the authors to multiscale temporal and spectral scenarios. The multiscale CSP features outperform the state-of-the-art approach with classification accuracy of 73.7015.90% (mean standard deviation across 9 patients) for the 4-class BCI competition IV-2a dataset. [21] The Riemannian covariance features outperform the CSP with an accuracy of 74.2715.5% with execution times that are 4x faster in testing and 9x faster in training. Utilizing additional temporal windows for Riemannian features yields 75.4712.8% accuracy while testing at 1.6x the speed of CSP [22].

To extract brain signals from electroencephalography (EEG) data, blind source separation (BSS) methods are used. Unfortunately, quantifying BSS performance is challenging since there is no criterion for distinguishing between neuronal impulses and noise in EEG readings. This paper presents a method for measuring BSS effectiveness by comparing EEG signals with electrocorticography (ECoG) signals obtained simultaneously. The ECoG electrodes covers lateral cortical surface mostly and catch the majority of neuronal original sources in EEG signals. Using real ECoG and EEG data, the authors creates for evaluating BSS performance an algorithm [23] To begin, BSS algorithm divides the EEG data into EEG components. Second, the EEG components are ordered using the regression ECoG correlation coefficients, based on their ranks components were sorted into subsets. Third, canonical correleation analysis determine how much information is shared between ECoG and EEG components subsets. The algorithm is used to compare the five BSS algorithms performances (AMUSE, SOBI, PCA, fastICA AND JADE) on anesthetized nonhuman monkeys' EEG and ECoG data. The findings shows that proposed algorithm effectively evaluates BSS performance and distinguishes between BSS algorithm performance [24]

Complex signal processing techniques are required for electroencephalographic (EEG) source imaging due to the difficulties associated with source separation, localization and data cleaning. Problems like these are often described in stages using different methods, limited using of EEg images in variety of settings. In this study, we describe a single empirical Bayes algorithm that can handle all of these different types of situations. By imposing spatial sparsity constraints, we are able to adaptively partition brain inputs into a large number of anatomically distinct components with minimal overlapping artifactual activity. The system generates an offline/online-capable recursive inverse spatiotemporal filter. Recursive Sparse Bayesian Learning describes this filter (RSBL). We theoretically demonstrate how Infomax Independent Component Analysis is related to Rapid Scan Body Lidar [21]. Using simulations, we show how RSBL can extract spatio-temporally overlapping cortical and artefact components from noisy data. We use RSBL to explore in cingulate gyrus origins of single-trial error-related potentials using data from the actual world. We also ran ours algorithm on two different EEG datasets and found that it 1) outperforms Informax for source separation on short time scales and 2) can reduce artefacts without degrading clean epochs, whereas the popular Artifact Subspace Elimination approach cannot do either of these things. In conclusion, we replicate the most important findings from past experimental work by using body imaging/mobile brain data to identify the brain dynamics that support heading computation while full-body rotations. [25]

A brain-computer interface (BCI) allows to communicate b/w the computer and human brain or devices that do not require physical contact. The technology making use of EEG signals generating by brain activity. The selection of EEG signal processing algorithms at each level is critical for developing a robust BCI system. This study discusses the numerous BCI approaches utilized at each level, such as feature extraction, pre-processing and classification [26]. At each stage, the authors address the benefits, drawbacks, and current trends in BCI. In addition, unlike earlier survey publications, they disclose preliminary experimental results at each BCI level. Several techniques including as filtering, artifact removal, and spatial filtering are described in the pre-processing step. Each technique's benefits and drawbacks are discussed. The authors address the most often utilized techniques, such as time frequency analysis, frequency domain, time domain during feature extraction stage [27] They also discuss the current state-of-the-art in deep-learning based feature extraction approach. The authors describe the most often used classifiers in the classification stage, includes k-nearest neighbour (k-NN), support vector machine (SVM) and artificial neural network (ANN) (ANN). They also discuss the current state-of-the-art in deep-learning based classification approach. Lastly, the authors give preliminary experimental data at each level of the BCI process. Using a motor imagery based BCI task, they illustrate the efficiency of various processing approaches for each level. overall, this paper gives a briefly overview of the numerous techniques employed in various stages of BCI, as well as highlighting current developments and future directions in the field. The preliminary experimental results given in this research shed light on the efficacy of various processing strategies in the robust BCI system developments [28]

Individuals' brain impulses are used to operate computer devices in Brain-computer interface (BCI) systems. The introduction of deep learning techniques has substantially enhanced the BCI systems performance in recent years. The authors analyze the various forms of brain signals utilized for BCI and give a detailed assessment of deep learning techniques applicable to BCI in this article. They provide a detailed examination of each technique and summarize over 230 contibutions publishing in the last 5 years. The authors discussed deep-learning techniques such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), autoencoders and generative adversarial networks (GANs) applied to different types of brain signals such as magnetoencephalography (MEG), electroencephalography (EEG) and funtional magnetic resonance imaging (fMRI) [29]. The authors also go over some of the applications of deep learning-based BCI, such as assistive technologies, gaming, and virtual reality. They emphasize the initial hurdles, such as the need for improved feature extraction and preprocessing approaches, as well as the difficulty of designing individualized BCI systems. Finally, the authors discuss the future directions of deep learning-based BCI, such as the need for more reliable and accurate brain signal acquisition systems, the development of hybrid models that combine multiple deep learning techniques, and the integration of BCI with other emerging technologies like (IoT) internet of things and augmented reality [30].

In this study, the authors present a new channel selection technique to boost the efficiency of (CSP) common spatial pattern associated attributes for (MI) motor imagery classification. Unlike traditional approaches to channel selection, the suggested method prioritises channels based on the relative strength of their signals when used for classification. The proposed method, on the other hand, selects channels based on the strengths of their respective correlations. For binary MI tasks, a channels's uniqueness is measured by the all number of channels with which it generates a large correlation coefficient value dispersion. The suggested method builds a cluster of highly correlated channels for each individual channel and then applies the (FBCSP) filter-bank CSP only to the cluster [31]The Fisher's score is calculated from the feature output, and the channels in the group with the highest score are picked. The proposed method on average selects fewer channels, making it superior than previously used methods. With two openly available BCI datasets—BCI competition IV dataset I and BCI competition III dataset IVa- the effectiveness of proposed technique is proved. For MI classification in general, this work provides a novel channel selection technique to enhance CSP-related features. The proposed method picks fewer channels overall, but it is more effective than earlier methods. Findings indicate the proposed approach may enhance the functionality of MI-based BCI devices [32]

The authors present (GAs) three multi-objective genetic algorithms in this study to number of channels optimization chosen and system accuracy in Brain-computer interface (BCI) systems. The goal of this study is to determine the best tradeoff between a BCI system's classification accuracy and all the numbers of channels used. This tradeoff is significant because various BCI application have varied objectives; for example, small number of channels are prefered by some implementations while others prioritize classification accuracy. The authors also look into the efficacy of genetic algorithms (GAs) as a channel selection method for BCI systems based on different forms of brain activities [33]The suggested GAs are tested on three BCI Competition datasets to see how well they select the optimal numbers of channels while retaining good classification accuracy. The authors perform a non-parametric Friedman test, which shows that a large reduction in the number of channels have no effect on the categorization accuracy of the assessment data. This validates the using of genetic algorithms (GAs) for channel selection in both motor imaging data and P300. Overall, the suggested multi-objective GAs show promise for optimizing the tradeoff between channel number and system accuracy in BCI systems. The authors show how GAs can be used to pick channels for BCI system based on various forms of brain activity, which can aid in the creation of hybrid BCI systems that mix multiple brain activity patterns for better functionality [32]

The authors evaluate three signal processing techniques for decomposition of Electroencephalography (EEG) data (BCI) Brain-computer interface systems for a classification job in this study: Discrete Wavelet Transform, Empirical Mode Decomposition and Wavelet Packet Decomposition. They test the performance of these approaches using BCI Competition III dataset IVa available publicly, a multichannel 2-class motor imagery dataset. The authors use Multiscale Principal Component Analysis (MSPCA) to removes noise and create various sets of features to investigates the effect of different groups of features on the classification job. They also go into great detail on the parameter selection procedure for the signal decomposition methods [34]. According to their findings, higher order statistical features of the combination of MSPCA de-noising recovered from wavelet packet-decomposition sub-bands yields the highest average classification accuracy of 92.8%. The authors explains the importance of high frequency ranges in better EEG signal categorization in BCI systems. The suggested model has the ability to acquire an accurate categorization of motor imagery (MI) EEG signals and can be employed as a actual system for operating a wheelchair or complementing present rehabilitation therapies by providing appropriate feedback after the execution of correct movement by individual [35].

A common pre-processing method for extracting motor-based Electroencephalography (EEG) signals is (CSP) common spatial pattern algorithm. In order to classify EEG data into their respective categories, CSP extracts spatial filters. However, a regularized CSP method can further improve these filters' performance. The regularized CSP algorithm modifies the CSP algorithm's objective function by incorporating regularization terms based on previously acquired data knowledge. In this research, we employed the time-correlation between the different classes of EEg signals as a source of initial information for regularization. By elucidating the degree of similarity between various classes of signals, this background knowledge aids in regularizing the CSP method. The proposed goal function can be generalized to issues with more than two classes [36]. Three industry-standard datasets were utilized to measure the effectiveness of the suggested technique. When compared to other regularized CSP algorithm like PCA and FDA to original CSP algorithm, the correlation-based CSP (CCSP) fared better in two class and multi-class scenarios. The present strategy improved mean classification accuracy by 6.9% for the two-class problem and by 2.23% for the multi-class problem, according to the simulation results. When used to motor-based EEG signals in BCI applications, the suggested CCSP algorithm with previous information about time correlation can be an effective method for enhancing classification accuracy [37].

This article explains a technique for channel reduction in BCIs that use motor imagery (BCIs). (EEG) Electroencephalography signals can be used to detect and categorize motor imagery, in which a person imagine doing a particular movement, such as moving their hands or feet. Multi-channel EEGs can be used to spot spatial patterns, although the process can be laborious and time-consuming. The authors propose uses the (CSP) common spatial pattern technique to locate crucial channels for detecting mental hand and foot motions. The CSP technique examines the spatial patterns of EEG signals and draws out the channels that are most useful for distinguishing between the various motor imagery tasks [38]. After deciding on which channels to analyze, the authors employ linear discriminant analysis to categorize ERD and RP. This demonstrates the potential for channel reduction in MI-based BCIs, as the results show that excellent classification accuracy's may be achieved with only four ideal channels using the suggested strategy. Overall, this study offers a promising strategy for enhancing the effectiveness and usability of MI-based BCIs, which may have significant consequences for the creation of more widely available and user-friendly BCI technology [39].

In brain-computer interfaces, selecting the appropriate EEG channels is presented here (BCIs). Multichannel EEG is widely utilized in BCIs, however choosing the best channels for the task at hand is crucial to the success of the BCI and the user experience. Sparse common spatial pattern (SCSP) is present method that treats selecting the optimal number of channels for EEG analysis as a constrained optimization issue. As a result, the method can be tweaked to enhance classification accuracy by keeping only the most useful channels or to lessen the number of channels without sacrificing accuracy. Two motor imaging datasets, one with large number of channels and one with moderate number of channels, were used to test the proposed SCSP technique [40]. Classification accuracy was improved over previous channel selection approaches using the mutual information, fisher criterion, support vector machines, common spatial patterns and regularized common spatial patterns, all of which were shown to be superseded by the present SCSP algorithm. On average, the suggested algorithm SCSP improved classification accuracy by 10% compared to when only three channels were used (C3, C4 and Cz)). The overall performance of BCIs and the user experience could benefit from the implementation of the suggested algorithm SCSP for EEG channel selection [41]

The Motor Imagery of Two Types (CSP) the common spatial pattern algorithm is frequently uses for the classifying of EEG data. The usefulness of this technique is depends on the frequency range relevant to the study subject. In order to finds the sweet spot in terms of frequency range for CSP, this research presents the FBCSP method. Dataset 2a containing 22 channel EEG data from 9 participants organised into 4 classes, while Dataset 2b has 9-subject EEG data organised into 2 groups using bipolar channels. Dataset 2a, which contains 4-class EEG data, is used to improve the FBCSP algorithm by employing Divide-and-Conquer, One-Versus-Rest and Pair-Wise [42]. Using Dataset 2b, author employ MIBIF and MIRSR to pick informative CSP features for further analysis. Classification accuracy is demonstrated session-to-session transfer on the evaluation data and 10-fold cross-validations on training data from both datasets. The FBCSP approach has the highest mean kappa in both Datasets 2a and 2b, across all participants, as the test data labels are exposed (0.569 and 0.600, respectively). Subject-specific frequency spectra for CSP in BCI applications are optimally calculated using the FBCSP algorithms [30].

This research suggests a strategy for enhancing the discriminatory power of the Rayleigh coefficient feature in EEG signals based on motor imagery. While maximizing of the Rayleigh coefficient is an useful characteristic, redundant electrode channels can reduce its efficacy, leading to poor classification accuracy, an increase in computing load, and a lack of insight into the important brain activity [43]. To solve these problems, the research introduces a genetic algorithm enhancement to pick the best possible subset of channels for the Rayleigh coefficient feature. Experiments on two various motor imagery EEG datasets validate the effectiveness of the suggested technique. In sum, the findings prove the method's efficacy in selecting channels for classification motor imagery EEG signals [44]

Noise, outliers, non-stationarity of EEG data, lengthy calibration durations, and inconsistent results are only some of the problems that plague brain-computer interfaces (BCIs). Riemannian methods, which rely on covariance matrices, show promise as instruments for overcoming such restrictions. Following a brief introduction to Riemannian geometry and BCI-relevant manifolds, this article discusses the application of these methods to EEG-based BCI. Calibration time reduction, classifier construction, and feature representation and learning are discussed in this article [5] The Riemannian method takes the geometry of covariance matrices as input and processes them on Riemannian manifolds. This opens the door for the application of geodesics and tangent spaces, two powerful techniques from Riemannian geometry, to the analysis of BCI signals. EEG-based BCIs that employ Riemannian methods for feature extraction and classification have shown considerable promise. Calibration time is a key drawback of current BCIs, however Riemannian techniques can help alleviate this problem. Classifiers can be learned using relatively little quantities of calibration data by taking advantage of the geometry of covariance matrices, thus cutting down on calibration times [45]. Many difficulties and potential avenues for future study of EEG signal classification in BCIs are highlighted in this article. There is a need for adaptive and online learning algorithms that can deals with the non-stationarity of EEG data, as well as multi-task learning and feature tracking [46]

(BCI) Brain-computer interface technology permits brain-machine contact without nerves or muscles. Electroencephalography (EEG) is the most frequent way BCI systems acquire brain data.

As they picture moving, EEG-based BCI users generate motor imagery (MI) brain activity. EEG electrodes on the scalp can detect and interpret these patterns into device commands. In clinical settings with limited preparation time, Standard EEG-based BCI systems require a vast number of scalp electrodes to classify MI accurately. This research presents a channel selection strategy for EEG-based BCI systems to improve classification accuracy and reduce scalp electrodes [47]. The suggested method iteratively optimizes the number of relevant channels using MI task information such the frequency range and spatial distribution of relevant brain activity. The suggested method lowers signal acquisition electrodes by selecting the most informative channels, making the BCI system more practical and efficient. Two public datasets— BCI Competition IV dataset 2a and BCI Competition III dataset IVa- with hand and foot Mi tasks are uses to assess the present approach. The proposed method outperformed state of the art technologies while requiring fewer channels for signal capture [48]

BCIs allow brain signals to control external equipment. Motor-Imagery (MI)-based BCI detects movement-imaging patterns in EEG signals. Conventional MI-based BCIs require many EEG channels to achieve excellent classification accuracy. However, employing more channels can result in redundant data and lengthier preparation times, which may be impractical.

This study suggests Effect-sized-based CSP channel selection to address this issue (E-CSP). E-CSP removes superfluous channels that do not distinguish two MI tasks to improve classification performance. Non-iterative Cohen's d effect-size computation determines redundant channels [49]. The E-CSP technique calculates Cohen's d effect-sized for each channel after removing noise trials. This effect-size computation ranks channels by their importance in distinguishing the two MI tasks. The algorithm then chooses a classification-informative subset of channels. E-CSP is compared to CSP and SCSP on two publicly available BCI competition datasets. E-CSP exceeds other approaches in classification accuracy while requiring fewer channels. E-CSP improves performance non-iteratively, making it more feasible for real-world applications [50].

The research presents an EEG-based BCI channel selection strategy to improve classification accuracy and reduce computational complexity. Dynamic Channel Relevance (DCR) selects relevant channels from raw EEG signals using a minimum redundancy maximum relevance paradigm. Using three available EEG datsets (BCI Competition IV- dataset I, BCI Competition IV-2008-IIA, BCI Competition III-dataset IVa) the method outperforms state-of-the-art methods with less channels. The procedure has numerous steps. The Savitzky-Golay filter preprocesses the EEG signals, and a sliding window technique decomposes them into fixed-length chunks. Second, three known channels are used to find new linked channels. Each channel's DCR score indicates its importance to MI tasks. Analyze the channels with the highest DCR values [51]. MEMD is use to extracts spatial- temporal information from the specified channels. Support Vector Machines classify 4 MI tasks—right hand, left hand, tongue and foot-using the retrieved features (SVM). The suggested technique is validates on 3 public EEG datasets (BCI Competition IV- dataset I, BCI Competition IV-2008-IIA, BCI Competition III-dataset IVa) and exhibits greater classification accuracy with methods of state of the art than less number of channels. The proposed strategy saves calculation time without compromises classification accuracy [51] Finally, Topographical mapping is employed to determine the link between cognitive regions and selected channels that perform MI tasks during physical activity. During physical exercise, the frontal, central and parietal lobes performs MI task [50]

Reference	**Approach**	**Year**	**Dataset**	**Channels**	**Accuracy**
[6]	DGAFF	2023	BCI IV 2a	22	82.03
[6]	HICS	2023	BCI IV 2a	22	78.43
[6]	DGAFF and HICS	2023	BCI 1V 2a	22	84.41
[16]	Sparse representation	2019	BCI IV 2a	22	77.5
[13]	EEGNet	2018	BCI IV 2a	22	80.07
[52]	Filter bank CSP	2023	BCI IV 2a	22	82.03
[39]	CSP	2023	BCI IV 2a	22	82.03

Table 1:Summary of Literature Review

In summary, as shown in Table 1 selecting relevant channels is crucial for reducing data dimension, eliminating noise and irrelevant data, and improving BCI system performance. Channel selection techniques include filtering, wrapping, and hybrid approaches; each has advantages and restrictions of its own. Using these techniques with signal cleaning and preprocessing procedures can improve the accuracy of MI-BCI systems.

2.2 Summary

In conclusion, selecting the appropriate channels for the electroencephalogram (EEG) is an necessary step in the study of MI motor imagery because it has the potential to significantly impact both the EEG signal's quality and the precision of the motor imagery classification. Because genetic algorithms are able to scan the vast universe of all possible channel combinations in an efficient manner, they have recently emerged as a usable tool for channel selection in EEG-based motor imagery (MI) research.

In this study, we have focused on the most important research that have utilised genetic algorithms for channel selection, and we have examined the benefits of as well as the drawbacks to using this methodology. We have also investigated the efficacy of genetic algorithms in comparison to the performance of other techniques for EEG channel selection, such as spatial filters and several other feature selection strategies.

According to the findings of our review, genetic algorithms appear to be a valid and successful strategy for selecting EEG channels in motor imagery research. However, it is essential to keep in mind that there is no one strategy that is ideal in all circumstance. The selection of the channel selection method should be determined by the particular research question as well as the dataset.

In upcoming studies, it will be interesting to investigate the potential of hybrid methods that combine genetic algorithms with other methods for EEG channel selection, as well as the application of these methods to other kinds of EEG data. Furthermore, It will be fascinating to search potential of hybrid methods that combine genetic algorithms with other techniques for EEG channel selection. Our analysis lays a solid groundwork for further investigation in this fascinating area of study.

Chapter 3: Datasets

In this chapter, we will discuss the dataset that was utilised in our research on the application of genetic algorithm to the selection of EEG Motor Imaginary signal Channels. The dataset is an essential component of any research study due to the fact that the quality of the data as well as its characteristics has a outstanding impact on the validity and generalizability of the findings.

Our research makes use of a dataset that is accessible to the public and goes by the name BCI IV 2a and 2b. This dataset was compiled by BCI. The dataset is made up of electroencephalogram (EEG) recordings made by 13 and 9 healthy volunteers.

For the purpose of our research, selecting a dataset that is typical of the population of interest and that contains a suitable amount of both variability and complexity is essential in order to judge the usefulness of the channel selection method that we have developed. It has been demonstrated that the BCI IV 2a and 2b dataset are reliable and valid measure of brain activity that is connected to motor imagery. This dataset has been utilised extensively in prior research on motor imagery.

In this chapter, we will present a complete description of the dataset, which will include information on the demographics of the members, (MI) the motor imagery task, and the EEG recordings. In addition, we will talk about the preprocessing stages that are used to assemble the data for analysis. These steps include excluding artefacts, filtering the data, and segmenting it.

In general, this chapter serves to assure the validity and dependability of our findings by providing a foundation that is crucial for our research and providing it here. The aim of this chapter is to give the reader with a comprehensive grasp of the attributes and quality of the dataset that was utilised in our research, as well as the actions that were take to ready the data for analysis, by the time they reach the end of this chapter.

3.1 BCI Competition IV Datasets 2a

Within the domain of (BCI) Brain-computer interface research, one of the most well-known datasets is referred to as the BCI Competiton IV Dataset 2a. In 2008, the dataset was made available as a part of the fourth (BCI) Brain-computer interface competition, which were held with the intention of advancing the progress of EEG-based BCI systems. The dataset has been extensively utilized in a wide variety of studies and has made a sizeable contribution to the progress of BCI systems [53]

The dataset is comprised of EEG recordings taken from 9 subjects while they carried out a total of 14 distinct types of (MI) motor imagery tasks. In the motor imagery tasks, participants were request to imaginary moving either their right or left hand, their tongue, or their foot. The EEG signals were recorded using the conventional 10-20 system from 22 electrodes that were positioned on the subject's scalp. The EEG signal was examined at rate of 250 Hz, and each trial lasted for a total of 6 seconds. Additionally, there was a resting period of 2 seconds at the beginning and the end of each trial. Additionally, the dataset contains information regarding the class labels of each trial, which indicate the type of (MI) motor imagery task that were performed [54].

Reason of the high degree of inter-subject variability and noise that is present in the BCI Competition IV Dataset 2a, it is a difficult dataset to work with when conducting BCI research. Inter-subject variability is caused by differences in brain activity and individual characteristics of each subject, both of which have the potential to influence how well the BCI system functions. This variability can be reduced by using multiple subjects. The noise in the dataset may have been caused by a number of different factors, including muscle artifacts, eye movements, and environmental noise, amongst others [53]

The dataset has been subjected to a number of different pre-processing and feature extraction strategies in order to address these challenges. The objective of the pre-processing techniques is to get rid of noise and artifacts in the EEG signals. For example, one technique uses notch filters to get rid of interference from powerlines, while another technique uses spatial filters to get rid of artifacts caused by muscle activity and eye movements. The aim of the feature extraction methods is to glean useful data from the EEG signals so that it can be applied to the classification of (MI) motor imagery tasks. Some of these methods include the use of features based on the time domain or the frequency domain.

Several studies have make use of BCI Competition IV Dataset 2a in order to progress and evaluate machine learning algorithms for BCI classification. These studies can be found here. These studies have demonstrated that the dataset can be utilized to achieve high classification accuracy, which shows the capacity of EEG based BCI systems for use in applications in the real world. The dataset has also been used to investigate the effect of various factors, such as electrode placement or frequency band, on the progress of the CI systems. This has been done by comparing the effectiveness of various feature extraction and classification algorithms [54] }.

Investigating the effect that various classifiers such as linear discriminant analysis (LDA) and support vector machines (SVMs), have on classification performance is one of the specific studies that have have make usedof the BCI Competition Dataset 2a. In other research, the effect of utilizing various feature sets, such as frequency domain or time domain features, or the effect of utilizing various subsets of electrodes was investigated [53] .

Research into the application of transfer learning to BCI classification has also been conducted using the dataset. Transfer learning is when you take the skills and knowledge you've acquired from one activity or dataset and apply them to another activity or dataset that is distinct but related. In the context of BCI, this could mean applying the knowledge gained from one subject to better classification progress on a different subject. Alternatively, it could mean applying the knowledge gained from 1-type of MI motor imagery task to improve performance on others type of MI motor imagery tasks.

In general, the BCI Competition Dataset 2a has proven to be an invaluable resource for the design, testing, and evaluation of BCI systems, and it has made a significant contribution to the progression of the field as a whole. Because of its widespread application and popularity, a standard has been established that can be used to judge the efficacy of newly developed BCI algorithms and methods.

- The dataset was recorded with an amplifier system from g.tec Medical Engineering GmbH known as g.USBamp. This system is a 16-bit bioamplifier and has a bandwidth that ranges from 0.1 to 100 Hz.
- All nine of the subjects who were included in the dataset were adults who were healthy and showed no signs of any neurological or psychiatric conditions. During the trials, the members were request to perform the (MI) motor imagery task while seated at a restful chair. Additionally, they were instructed to maintain stillness and to refrain from making any movements that were deemed to be unnecessary.

- Randomization was used to determine the order in which each participant would complete the total of 140 trials (14 tasks times 10 repetitions). A single trial was obtained for each of the 10 tasks by averaging the previous trials; it lead to the total number of trials for each one subject to 126.
- The class labels for each trial were binary, indicating whether the task required movement of the rightor left hand (for tasks 1–4 and 9–12), the foot (for tasks 5–8), or the tongue (for tasks 1–4). (for tasks 13-14).

The summary of dataset is shown in Table 2 below

Dataset Name	BCI Competition IV 2a
Acquisition System	g.USBamp
Recording EEG Channels	22
EEG Sampling Rate	250 Hz
Subjects	9 subjects
Tasks	Motor imagery of right and left hand, 2 classes
Trials	72 trials per class per subject
Trial Length	3-4 seconds
Preprocessing	Common Spatial Pattern (CSP)
Availability	Publicly available

Table 2:Summary of BCI Competition IV Datasets 2a

- The dataset underwent several preprocessing procedures, such as filtering based on the common average reference (CAR), notch filtering to get rid of powerline interference, and artifact rejection based on amplitude and variance thresholds.
- In order to withdraw information from the EEg signals, a number of feature extraction techniques were utilized. These feature extraction techniques included frequency domain

charatcers such as Spectral coherence and power spectral density (PSD) as well as time domain characters such as slope, variance and mean.

- The dataset has been utilized in a great number of studies, including a few benchmark studies that have compared the effectiveness of various BCI systems and algorithms. The linear (SVM) support vector machine classifier, in conjunction with time-frequency features and spatial filters, were used to achieve the highest reported classification accuracy on the dataset, which was approximately 90%.

3.3 BCI Competition IV Datasets 2b

It is usual practise to made use of the BCI IV 2b dataset, which is made available to the public, while conducting research on Brain-Computer Interfaces (BCI) that involves motor imaging tasks. The EEG recordings of healthy subjects engaged in various motor imagining tasks make up the dataset, which was compiled by the Graz University of Technology and contains these recordings.

3.3.1 Experimental paradigm

The EEG data come from 9 different people who took part in the study and includes in the data set. Everyone who took part in the study had their right hand dominant, had corrected to normal vision and normal vision, and was financially compensated for their time. Over the course of the research, each participant are guides to remain sitting at the chair that provided adequate comfort while they observed a screen that was approximately one metre away and positioned at eye level. The data set contains five sessions for each individual, the first two of which were used for training without receiving any response, and the last 3 of which was recorded with response [55]

As can be seen in Figure 1, each session included in the dataset is composed of several separate runs. An initial recording of around 5 minutes were made at the beginning of each period in order to determine the influence that eye movements have on the EEG data. This recording was broken up into three sections: the first two minutes consisted of the subjects being guides to keep their eyes opened and concentrate on a cross that was displayed on the screen; the next two minutes consisted of the subjects either closing their eyes or moving their gaze. After that, the artefact block was cut into four portions, each of which had 15 seconds of eye movement followed by 5 seconds of rest. It was informed to the participants that they were to conduct rolling, eye blinking, up-down, or left-right motions while a high or low warning tone was provided at the beginning and finish of each exercise, respectively. Nevertheless, due to technical issues that arose during sessions B0102T and B0504E, the EOG block could not be accessed during those sessions. The participants in the study are outlined in Table 8, which can be found here.

ID	Training	Evaluation
1	B0101T, B0102T, B0103T	B0104E, B0105E
2	B0201T, B0202T, B0203T	B0204E, B0205E
3	B0301T, B0302T, B0303T	B0304E, B0305E
4	B0401T, B0402T, B0403T	B0404E, B0405E
5	B0501T, B0502T, B0503T	B0504E, B0505E
6	B0601T, B0602T, B0603T	B0604E, B0605E
7	B0701T, B0702T, B0703T	B0704E, B0705E
8	B0801T, B0802T, B0803T	B0804E, B0805E
9	B0901T, B0902T, B0903T	B0904E, B0905E

Table 3:List of all files in BCI Competition IV Datasets 2b

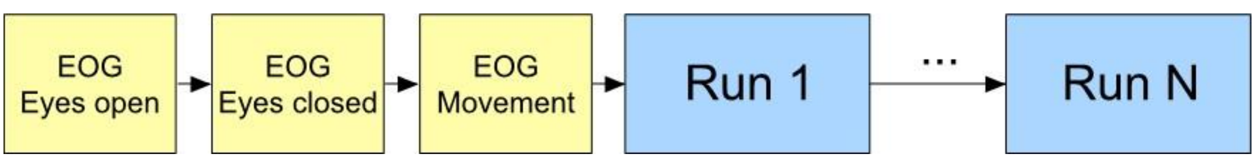

Figure 1:Timing Scheme of one session

3.3.2 Data recording

Frequency of 250 Hz wereuses as sampling throughout the recording process for all (C3, Cz and C4) all 3 bipolar recordings. During the screening, the recordings have a dynamic range of 100 V, and during the feedback sessions, the range was 50 V. They were given a bandpass filter with a frequency range of 0.5-100 Hz, and a notch filter set at 50 Hz were enabled. For each participant, the 3 bipolar recordings were positioned somewhat differently (at larger or smaller distances, in a more anterior or posterior orientation, respectively; for further information, see [55]). Ground electrode were situated at Fz location of the electrode array.

furthermore to the EEG channels, an electrooculogram (EOG) was noted but with a dynamic range of 1 mV as shown in Figure 2, left mastoid serving as refrence) using the same amplifier settings. The EOG is noted with 3 monopolar electrodes as shown in Figure 5, left mastoid acting as refrence). It is imperative that the EOG channels not be used for categorization purposes since they have been supplied for the eventual application of artefact processing algorithms [56].

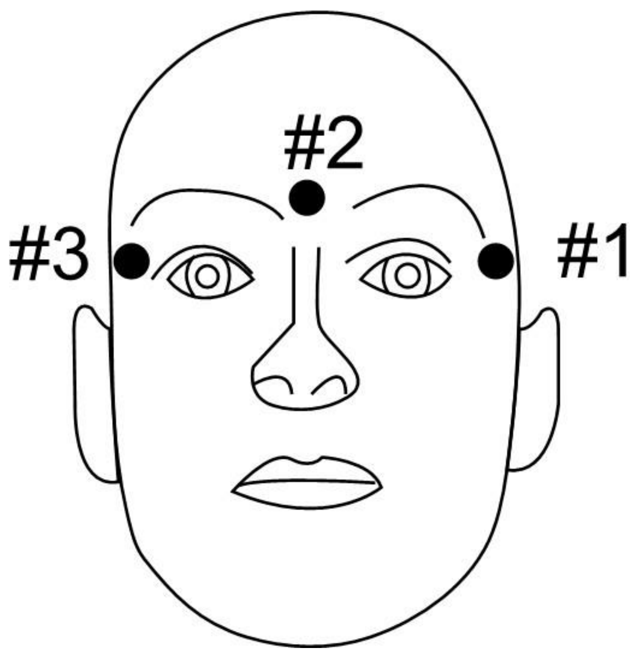

Figure 2:Electrode montage of the three monopolar EOG channels

The klaxon diagnostic approach as shown in Figure 3 included two classes: the first class focused on the left hand's motor imagery (MI), while the second class focused on the right hand's motor imagery (MI). Each member take part in two screening periods on separate days within a span of two weeks, however none of the sessions included feedback recording. Each session included two classes of imagery as well as six runs that were each comprised of 10 trials. This resulted in 20 trials being conducted during each run and 120 being conducted throughout each session. Data from a total of 120 repetitions of each MI lesson were accessible for every individual participant. Before beginning the first session of moor imagery training, the subject imagined and performed a variety of movements corresponding to each part of body, and then chose the one that they could think up the most easily (for example, crushing a ball or pushing on a brake).

At the beginning of each test, brief acoustic warning tone, a fixation cross and an additional, were presented (1 kHz, 70 ms). When a few seconds had passed, a visible indication in the shape of an arrow that pointed towards in right and left, depending on the class that was being requested, was displayed for 1.25 seconds. After that, the participants were tasked with visualising the equivalent hand movement for a duration of four seconds. After every round of testing, there was a brief rest lasting at least 1.5 seconds. In order to prevent people from adapting to the break, a random amount of time, up to one second, was added to it.

During the online 3 response sessions, a total of 4 runs with smiling response were noted as shown in Figure 3. Each run contains of twenty individual trials for each of the different types of motor imagery. The feedback, which was represented on the screen by a grey smiley, was centred on to screen at the each trial beginning (second 0). At the second 2 mark, a brief warning tone of 1 kHz and 70 milliseconds was played. From second three to second seven and a half, the cue was provided. The members were guided to imagined either moving their right or left hand the in order to moves the smiley face to the right or left side of screen, depending on the signal that was given to them. During the feedback phase, the smiley would turn from red to green if it wasn't pushed in the appropriate direction, but it would remain red otherwise. The distance of the smiley from its starting point was determined based on the output of the integrated categorization during the course of the previous two seconds (more details see [55]). In addition, the output of classifier according to curvature of the smiley mouth was mapped, which determined whether smiley should be sad (with the corners of mouth turns downwards) or happy (corners of mouth turns upwards). The screen turned black at the 7.5 second mark, and randomly interval remains for 1.0 and 2.0 seconds were add on to the test at that point . The participant was given the directive to keep the smiley on the appropriate side for the longest amount of time possible, and consequently to continue the MI for the longest amount of time feasible.

Participants are required to deliver a continuing classification output in form of classlabels (1, 2) for each sample. This output should include labelled trials as well as trials that have been tagged as artefacts. After that, an uncertainty matrix will be constructed using all of the artifacts free trials for each time point. The kappa coefficient and the time trend of accuracy will be determined [57]from these confusion matrices. Also acquired will be the kappa coefficient. The computational method that was applied to this assessment will be detailed in BioSig. The winning algorithm is the one that has a kappa value that is greater than X.KAP00.

In light of the facts that the judgement sets of data will not be made available until after the conclusion of the competition, the program-mes that are entered into the contest must be able to take EEG data as input (the formation of this data must be identical as that which is uses in all training sets3) and build the aforementioned class label vector.

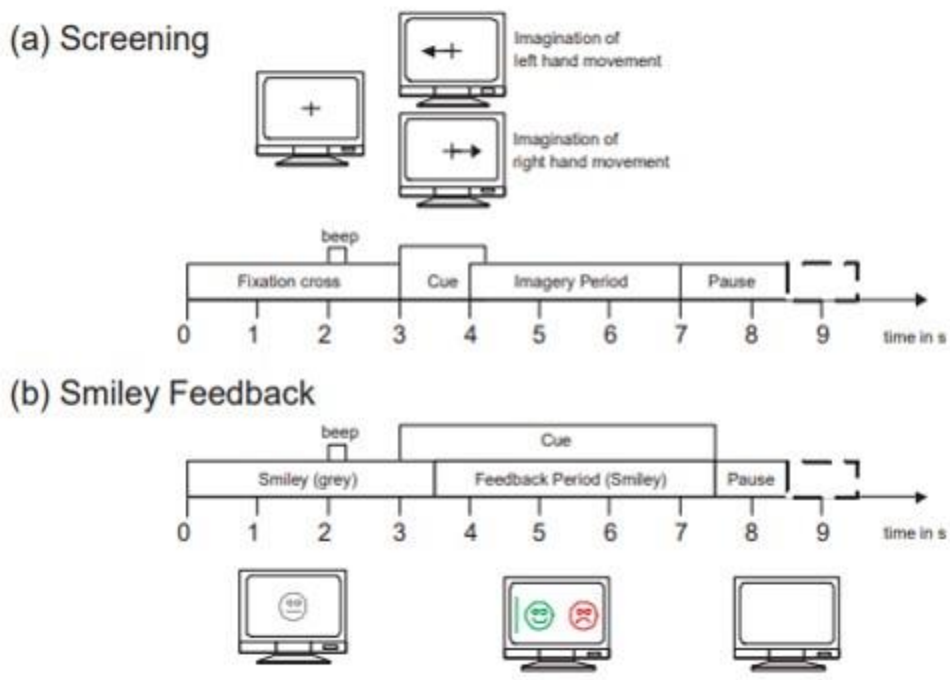

Figure 3:Timing Scheme of Paradigm. (a) The first two sessions (01T, 02T) contain training data without feedback, and (b) the last three sessions (03T, 04E, 05E) with smiley feedback

3.3.3 Evaluation

Due to the fact that there are three EOG channels available, it is necessary to clean the data of any EOG artefacts using artefact elimination techniques such as linear regression or high-pass filtering before continuing with the further data processing [58]In order to make it possible for various correction methods to be used, we have chosen an approach that provides the highest amount of transparency and have made the EOG channels available. At same time, we urge that artefacts have no impact on the classification result.

The following information pertains to the BCI IV 2b dataset:

- This dataset comprises information from nine healthy subjects ranges from 20-35 years old, three of which were female and six of whom were male.
- The participants were given two different types of motor imaging tasks to complete. The first task required them to imagine right or left hand movement, while the second task required them to imagined right to left foot movement. Following a period of rest for two seconds, each activity was carried out for a duration of four seconds.
- The EEG recordings was obtained by using a g.Tec device that has 118 channels and a sample rate of 250 Hz. Following the guidelines of the worldwide 10-10 systems, the electrodes were situated.
- The dataset can be downloaded in GDF format, which is a common format for electroencephalogram (EEG) data. The information gathered from each participant is organised into five separate sessions, and each of those sessions contains nine individual iterations of motor imagery tasks.
- The dataset has undergone pre-processing, during which artefacts such eye movements, muscle activity, and electrooculogram (EOG) artefacts have been eliminated. The data are subjected to filtering, segmentation, and down-sampling as part of the pre-processing step.
- Annotation files are included in the dataset, and they include information regarding the duration and nature of each motor imagery exercise that the participants completed. Labels for each activity, such as "left hand," "right hand," "left foot," or "right foot," are included in the annotation files, which are then used for categorization.

- The BCI IV 2b dataset is designed to include variability in the EEG signals due to factors such as changes in electrode impedance, fatigue, and individual differences in brain activity. The dataset also includes trials with varying levels of difficulty, where the members were guided to perform the (MI) motor imagery tasks with varying levels of intensity.
- The dataset includes pre-defined training and testing sets, with each training set consisting of 72 trials (9 runs x 8 tasks) and each testing set consisting of 36 trials (4 runs x 9 tasks). The performance of a classification algorithm on the testing set is estimated by using different metrics such as kappa, F1-score ad accuracy.
- The BCI IV 2b dataset was created for the purpose of evaluating the progress of (MI) motor imagery based BCI using EEG signals. The dataset has been used to test various methods and algorithms for feature extraction, classification, and channel selection in EEG-based motor imagery research. The dataset is publicly available and has been widely used in the research community.

Overall, the BCI IV 2b dataset is a valuable resource for EEG-based motor imagery research and motor imagery-based BCI development. The dataset includes annotation files, performance metrics, variability, and was designed for a specific purpose. These features make it suitable for evaluating the performance of different algorithms and methods, as well as for improving our understanding of motor imagery-related brain activity.

Figure 3summarizes the key attributes of the BCI IV 2b dataset, including the number and demographics of participants, the motor imagery tasks performed, the EEG recording details, the data format, preprocessing steps, annotation, performance metrics, variability, and purpose of the dataset.

3.4 Summary of Chapter

In conclusion, the dataset that was used in our study on EEG Motor Imaginary signal Channel selection using genetic Algorithm was a publicly available dataset called BC1 IV 2a and 2b. This dataset has been widely used in several previous researches on (MI) motor imagery. Our study focused on the selection of EEG Motor Imaginary signal Channels. The dataset comprises of EEG recordings taken from nine healthy volunteers while they performed motor imagining tasks under the following conditions and instructions: [insert specifics on the settings and directions for the task.

In this chapter, we have included a full description of the dataset. This description includes information on the demographics of the candidates, the (MI) motor imagery tasks, as well as the EEG recordings. We have also gone over the pre-processing stages that are used to ready the data for analysis. These steps include the rejection of artefacts, the filtering of the data, and the segmentation of the data.

Attribute	Description
Dataset Name	BCI Competition IV 2b
Participants	9 healthy participants (3 females, 6 males)
Age range	20-35 years
Tasks	Imagining left or right hand and foot movement
EEG Recordings	118 channels, 250 Hz sampling rate
Data Format	GDF format
Preprocessing	Artifact rejection, filtering
Annotation	Timing of each motor imagery task
Performance Metrics	F1-score
Variability	Includes trials with varying, multiple runs
Purpose	Evaluation of motor imagery signals

Table 4:Summary of Dataset 2b

In general, the datasets BC1 IV 2a and 2b are an appropriate and representative sample of the population that we are interested in researching for this particular study. The preprocessing methods that we implemented helped to ensure the quality and authenticity of the data, and they provided a solid platform for the subsequent analyses that we carried out.

In the next chapters, we will make use of this dataset in order to estimate of our genetic algorithm-based channel selection strategy, as well as study the neurological correlates of brain activity connected to motor imagery.

Chapter 4: Methodology

In the following chapter, we will discuss the methodology that was applied in our research on the EEG Motor Imaginary signal Channel selection utilising genetic Algorithm. The methodology is an essential part of every research study since it defines the particular actions and processes that were carried out in order to answer the research query and accomplish the aims of the study.

EEG signals will be used in this research project of ours to determine whether or not genetic algorithm-based channel selection is an effective technique for studying the (MI) motor imagery. In this chapter, we will provide a comprehensive explanation of the approach that was taken in order to accomplish the aforementioned goal. First, we provide an overview of the research issue and aims, and then we detail the design of the study, which includes the selection of the dataset as well as the particular motor imagery tasks that were utilised.

Next, we will provide a comprehensive explanation of the pre-processing stages that were carried out on the EEG data. These steps include the elimination of artefacts, the application of filters, and the division of the data into segments. In addition to this, we discuss the feature extraction techniques that were applied, frequency-domain features, time domain features and time frequency features.

Next, we will describe the genetic algorithm-based channel selection method that was utilised in our research. This will include a discussion of the particular parameters and operators that were applied, as well as the fitness function that was utilised to check the effectiveness of the channel subsets.

Finally, we discussed the classification algorithm that was applied in order to estimate the progress of choosen channel subsets. This discussion includes the particular machine learning algorithm that was applied, as well as the performance metrics that were applied in order to estimate the classification's robustness and accuracy.

At the conclusion of this chapter, the audience will have a thorough comprehension of the particular actions and procedures that were carried out throughout our research, as well as an appreciation for how the research question and objectives were handled by utilising our methodology. The approach lays a sturdy groundwork for the succeeding chapters of the thesis, which discuss the outcomes of our investigation and the conclusions that may be drawn from it.

4.1 Preprocessing

The (CSP) Common Spatial Pattern filter bank is a broadly used method in EEG signal processing that aids in the preprocessing of raw EEG data by reducing noise power and boosting useful signals. The CSP filter bank extracts pertinent information from the EEG signals using a spatial filter as shown in Figure 4.

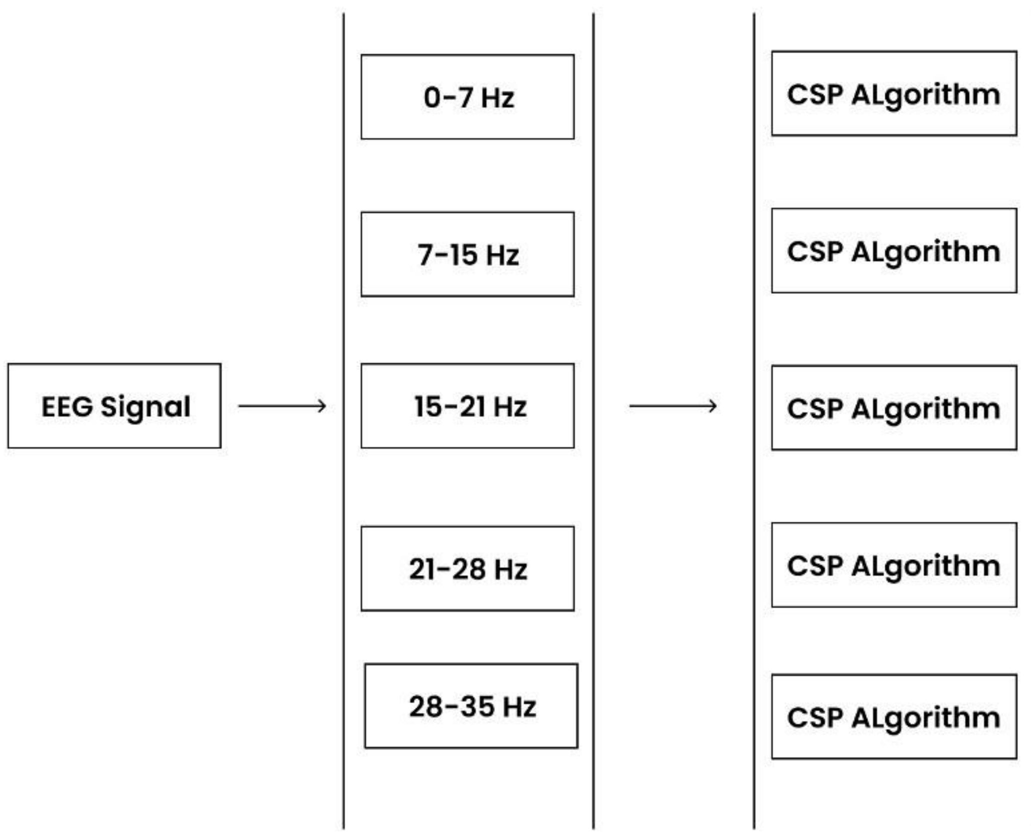

Figure 4:Pre-Processing Block Diagram

In this instance, the CSP filter bank has a frequency range of 7-35 Hz with a 7 Hz step size. This indicates that the filter bank is designed to extract information from EEG signals within this frequency range, with each filter covering a distinct frequency band. There will be five filters in the filter bank, covering the frequencies 7-14 Hz, 14-21 Hz, 21-28 Hz, 28-35 Hz, and a final filter covering frequencies above 35 Hz.

The CSP filter bank is utilized to increase the signal-to-noise ratio of the raw EEG data. Raw EEG signals can contain a significant amount of noise, which can interfere with the useful signal and make it challenging to extract meaningful information. The CSP filter bank aids in the reduction of this noise by employing a spatial filter that eliminates unwanted signals and boosts useful ones.

Finding a linear transformation of the EEG data that enhances the separation between two classes of EEG signals is the fundamental principle of CSP. In other words, the CSP filter bank is designed to extract the characteristics that are most pertinent to the investigated research question. Using a filter bank with a frequency range of 7-35 Hz and an interval of 7 Hz, this technique is able to selectively withdraw the suitable information from the EEG signals, thereby improving the signal to noise ratio and increasing the accuracy of the results.

Overall, the application of the CSP filter bank is a crucial step in the initialization of raw EEG data, as it helps to reduce noise and boost useful signals. By employing this technique, researchers can withdraw the relevant characters from the EEG signals and improve the accuracy of their analysis, resulting in deeper insights and a greater comprehension of the underlying neural processes.

4.3 Event Separation

After applying techniques such as filtering and noise reduction to the raw EEG data, the next step in EEG signal processing is to categorize the data. In this instance, the preprocessed data is categorized as left hand, right hand, foot and tongue.

The objective of segmentation is to identify the specific events or conditions associated with the analyzed EEG signals. By categorizing the data, researchers can identify the EEG signals associated with particular actions or stimuli, such as the motion of the left hand, right hand, foot and tongue.

After segmenting the data into these categories, the next step is to determine the event-related potential (ERP) components associated with each type of occurrence. ERP components are neural responses that occur in the brain in feedback to particular stimuli or events, and they can be measured using EEG signals. Researchers can gain insight into the underlying neural processes underlying these actions or stimuli by identifying the ERP components associated with each type of occurrence.

For instance, by analyzing the ERP components associated with right hand movement, researchers may be able to identify specific neural activity patterns associated with this action. Similarly, by analyzing the ERP components associated with left-hand, tongue, and foot movements, researchers can gain insight into the neural processes underlying these actions.

The segmentation of preprocessed EEG data into categories is a crucial step in EEG signal processing, as it enables researchers to isolate and analyze the EEG signals associated with particular events or stimuli. By identifying the ERP components associated with each type of occurrence, researchers can gain insight into the neural processes underlying these actions or stimuli, thereby enhancing their understanding of the brain's functions.

4.4 Window Extraction

(EEG) electroencephalography is a neurophysiological technique that estimates electrical activity in the brain using scalp electrodes. In cognitive and clinical neuroscience research, EEG signals are frequently employed to examine brain function and activity.

(ERPs) Event related potentials were type of EEG signal created in feedback to specific events, such as stimulus presentation. ERP analysis entails separating the EEG signal into distinct event types, such as target and non-target stimuli in a task, and then examining the electrical activity patterns that occur in response to each event type.

A time frame window of 1 to 4 seconds is frequently extracted from the EEG signal for each event type in order to analyze ERPs. This time window is chosen because it is commonly used for ERP analysis in EEG signal processing.

This time window is deemed long enough to capture the most significant aspects of the ERP waveform, such as the peaks and troughs that correspond to different stages of cognitive processing, but short enough to avoid contamination from unrelated brain activity or artifacts.

Depending on the nature of the experiment, the stimuli employed, and the research question being investigated, the specific time frame window chosen may vary. In EEG signal processing, the 1 to 4 second time window is widely used and well-established for ERP analysis.

4.5 Artifact Removal

Eye blinks, muscle movements, and electrical interference from the environment are examples of artifacts that frequently contaminate EEG signals. These artifacts can obscure or distort the underlying brain activity of interest, which can be problematic for EEG data analysis.

To ensure that the EEG signal is clean and reliable for further analysis, it is standard practice to remove these artifacts from the data using methods such as (ICA) Independent Component Analysis or Automatic Artifact Removal (AAR).

ICA is a signal processing technique that separates a complex signal into independent, non-Gaussian components. In the context of EEG data analysis, ICA can be utilized to identify and remove signal components corresponding to artifacts, such as eye blinks or muscle movements. By isolating and removing these components, ICA can ensure that the remaining EEG signal is clean and trustworthy for further analysis.

Automatic Artifact Removal (AAR) is another technique commonly used in EEG data analysis to remove artifacts. AAR employs an algorithm to automatically detect and remove artifacts from the EEG signal based on predefined criteria, such as amplitude or frequency thresholds. AAR can be quicker and more effective than manual artifact removal, and is particularly useful when working with large datasets.

In EEG data analysis, both ICA and AAR are well-established techniques for artifact removal. By removing artifacts from the EEG signal, these techniques ensure that the signal is clean and reliable for further analysis, and can help to reveal important patterns of brain activity that artifacts may have obscured.

4.6 Initial Population Creation

The EEG signal is frequently represented as a matrix of electrode channels by time points following preprocessing and artifact removal. This matrix may be large and contain redundant or irrelevant data, making efficient analysis challenging. Selecting a subset of the most informative channels for analysis is one strategy for dealing with this issue.

To generate this subset, a population of 20 candidates were created at random. Each candidate is estimated based on his or her ability to differentiate between the four types of events (or conditions) of interest. Typically, discrimination relies on statistical measures, such as the amplitude or latency of particular ERP components.

In a process known as evolutionary optimization or genetic algorithms, each candidate's fitness score is calculated based on their discrimination performance, and this score is used to guide the selection of subsequent generations of individuals.

Evolutionary optimization entails repeatedly generating new populations of individuals, selecting the most fit individuals from each generation, and recombining their characteristics to create new individuals for the following generation. This procedure is repeated until a convergence criterion is satisfied, such as a predetermined number of generations or a plateau in the fitness score.

In addition to lessen the dimensionality of the data and enhances the efficiency of subsequent analysis, this method also selects the most informative EEG channels for the particular task or research question at hand. However, it is essential to note that the selection of initial population size, fitness function, and optimization parameters can have a significant effect on the outcomes and should be carefully considered and validated.

4.7 Fitness Calculation

This step describes a method for evaluating an individual's ability to distinguish between four types of events in an electroencephalography (EEG) dataset.

The first step is to calculate each individual's fitness by averaging their performance in distinguishing between the four types of events. The weight assigned to each event is proportional to the number of occurrences of that event type in the dataset. This indicates that events with more trials will have a greater impact on the overall fitness score of an individual.

Next, each individual's performance is assessed using a (LDA) linear discriminant analysis classifier. In EEG signal processing, LDA is a commonly used supervised machine learning algorithm. It operates by discovering a linear combination of characters that maximally distinguishes between classes (in this case, the four types of events). A subset of the data is uses to train the classifier, which is then uses to anticipate the class labels for the remaining data. The accuracy of the classifier in predicting the correct class labels is used to assess an individual's ability to distinguish between the four types of events.

Fitness Calculation describes, in summary, a method for evaluating the performance of individuals in discriminating between four types of events in an EEG dataset. Calculating the fitness of each individual using a weighted average of their performance and evaluating their performance using an LDA classifier are the steps involved in this method.

4.8 Genetic Algorithm

Using a genetic algorithm, which is a examining optimization technique inspired by natural selection and genetic inheritance, the section describes the process of creating new populations of individuals.

This process begins by generating the initial population of individuals. The method for generating the initial population involves selecting EEG channels at random from the dataset. EEG channels were uses to measure the electrical activity of the brain, and selecting different channels can lead to the extraction of distinct data features.

After the initial population has been generated, the genetic algorithm is used to generate new populations of individuals. The genetic algorithm mimics natural selection and genetic inheritance in its operation. Each generation produces a new population of individuals by selecting the most fit individuals from the previous generation and using them to generate new individuals via crossover and mutation.

In this particular instance, the population size is set to 20. This implies that each new generation will consist of 20 people. The process of generating new populations of individuals will continue until a stopping standard is met, such as a predetermined number of generations or a certain level of fitness.

The processes of natural selection and genetics, which are responsible for the diversity of life on Earth, serve as inspiration for the GA. Similar to how genetic variation arises through random mutation, the GA begins by randomly generating a population of potential solutions. Similar to how natural selection favors organisms with traits that are better adapted to their environment, the fitness of each solution is evaluated, and the fitter individuals are more likely to be chosen as parents for the subsequent generation [41].

The processes of crossover, mutation, and selection as shown in Figure 5

used to produce new individuals in the GA. Crossover is the process of combining the genetic material of two parents in order to produce offspring with characteristics from both parents. This process is analogous to sexual reproduction, in which offspring inherit characteristics from both parents [59].

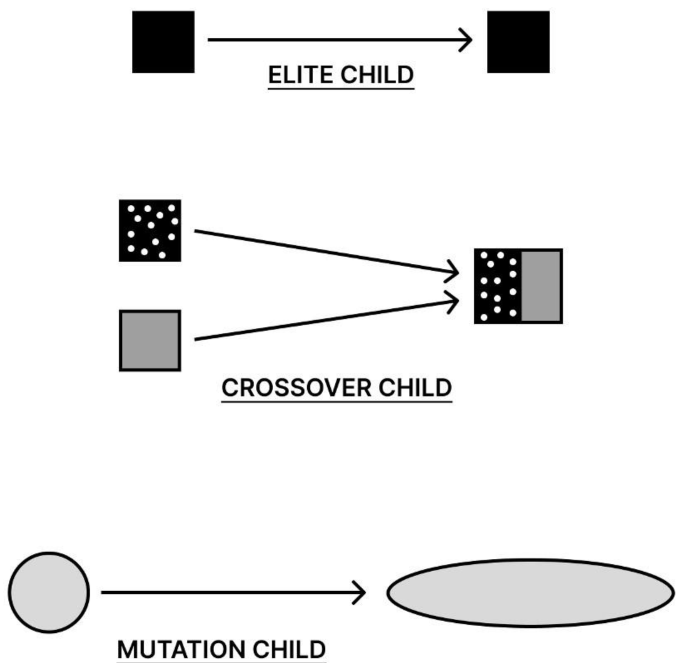

Figure 5:Flow Chart of Genetic Algorithm

There are three primary rules as shown in Figure 6 that are adhered to throughout each iteration of the GA in order to generate the next generation from the existing population.

- The first rule is to determine each individual's fitness within the population based on how well they solve the problem. This is the basis for the assessment. Typically, the evaluation of a person's fitness will be based on some kind of objective function that will measure the performance of that person.

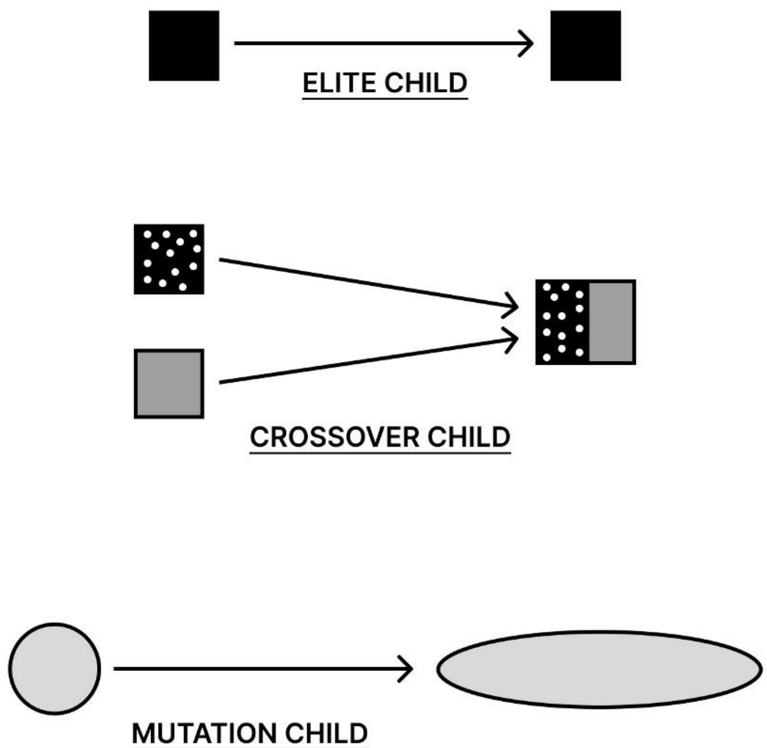

Figure 6:Representation of genetic algorithm rules

- The second guideline stipulates that individuals from the existing population must be chosen to play the role of parents to future generations. The process of selection can be based on a variety of criteria, such as each individual's level of fitness, the group's diversity, or a combination of both of these factors. The selection procedure's goal is to raise the likelihood that the healthiest and most mature members of the population will be chosen to raise the following generation.
- The third and final guideline is to generate new individuals for the subsequent generation by making use of genetic operators like crossing over and mutation. In the process of crossover, the genetic material of two parents is combined in order to produce offspring that inherit characteristics from both of their parents. A person's genes can be randomly

altered through the process of mutation, which then results in the introduction of new variations into the population. Observation of solution space and the production of new individuals with the potential to be superior to their parents are the objectives of the genetic operators.

In conclusion, a genetic algorithm is uses to create new populations of candidates. The initial population is generated by selecting EEG channels at random, and the population size is set to 20. The genetic algorithm simulates natural selection and genetic inheritance in order to generate new populations of individuals.

4.9 Elitism

This section describes a method for selecting the most successful members of a population to ensure their survival and continued success in future generations.

The next step, following the creation of a new population of individuals using a genetic algorithm, is to extract the elites from this population. The elites are those with the highest fitness scores, indicating that they performed the best at differentiating between the four types of events in the EEG dataset.

The elites are then used as parents for the subsequent generation. This indicates that the genetic algorithm will generate new individuals by performing crossover and mutation on the genes of the elites, as opposed to randomly selecting a new set of EEG channels.

Elitism ensures that the highest-achieving individuals are preserved and passed down to future generations. This can result in a faster merging to maximum solutions, as the genetic algorithm can build on the success of the highest-performing individuals in each generation.

In summary, an elitist process in which the highest-achieving members of a population are chosen as parents for the next generation. This procedure ensures that the success of these individuals is preserved and built upon, resulting in a quicker convergence on the optimal solution.

4.10 Generation Iteration

The section describes a process in which the steps of selecting elites as parents and creating new populations are repeated iteratively until the prefered number of EEG channels is achieved.

The genetic algorithm initiates by randomly selecting EEG channels from the dataset to generate an initial population of individuals. The fitness function is then applied to each generation of the genetic algorithm to determine its fitness. The fitness function computes the weighted mean of the individual's performance in discriminating between the four types of EEG dataset events.

After evaluating the fitness of every member of a generation, the top performers, or elites, are chosen as the parents of the following generation. This process of elitism ensures that the highest-performing individuals are preserved and passed down to future generations, thereby increasing the probability of discovering a superior solution.

The parents are then subjected to crossover, selection and mutation processes to create a new population of members. The selection process involves selecting parents from the previous generation. Crossover is the process of combining the genes of two parents to create a new organism. The mutation process involves the random alteration of an organism's genes to generate a new variation.

This process of selecting elites as parents and creating new populations is repeated until the desired number of EEG channels has been attained. At each iteration, to estimate the fitness of new population's members the fitness function is used, and the top performers are selected as the elites for the next generation. The processes of selection, crossover, and mutation are then repeated, resulting in a new population of individuals. This task is continues until the preferred quantity of EEG channels is attained.

A process consisting of iteratively repeating the steps of selecting elites as parents and creating new populations until the preferred number of EEG channels is achieved. Each generation of the genetic algorithm is evaluated using the fitness function, and the top performers are chosen as the elites for the next generation. The processes of selection, crossover, and mutation are continued until the prefered number of EEG channels is achieved.

4.11 Validation

The section describes a method for evaluating the effectiveness of the selected subset of EEG channels using a separate test set and statistical analysis.

After iteratively repeating the steps of selecting elites as parents and creating new populations until the required number of EEG channels is achieved, a final subset of EEG channel is chosen-based on the highest-performing members of the previous generation. This final subset of EEG channels is anticipated to have the highest discrimination performance between the four types of events in the EEG dataset.

Using an independent test set, the performance of this final subset of EEG channels is evaluated in order to confirm its efficacy. A set of data that was not used in the training or optimization of the model is an independent test set. This enables an impartial evaluation of the performance of model on new data.

Statistical analysis is used to determine the importance of difference between the progress of the final subset and the performance of the entire set of EEG channels. Statistical analysis is a collection of techniques used to draw conclusions about populations from samples. In this instance, statistical analysis is used to calculate whether the performance of the final subset of EEG channels is significantly superior to that of the entire set of EEG channels.

The statistical test used to compares the progress of the final subset and the entire set of EEG channels will depend on the particular characteristics of the data and the research question being addressed. ANOVA, T-tests and wilcoxon signed-rank tests are examples of statistical tests that may be employed.

4.12 Summary of Chapter

In conclusion, the approach that was utilised in our research on EEG Motor Imaginary signal Channel selection via genetic Algorithm has been discussed in this chapter. The approach was developed to answer the research question and objectives of the study, as well as to explore the efficacy of genetic algorithm based channel selection for motor imaging research using EEG signals.

We have included an in-depth description of the design of the study, which includes the selection of the dataset as well as the particular motor imagery tasks that were carried out. In addition, we have detailed the preprocessing processes that were carried out on the EEG data, such as the elimination of artefacts, the application of filters, and the segmentation of the data, as well as the feature extraction techniques that were implemented.

In addition, we have presented the genetic algorithm-based channel selection method that was used in our research. This method included the particular parameters and operators that were utilised, as well as that was used to estimate the performance of the channel subsets. Finally, we have described the classification algorithm that was used to estimate the progress of he chosen channel subsets. This description includes the particular machine learning algorithm that was utilised, as well as the performance metrics that were used to assess the robustness and accuracy of the classification.

In summary, the methodology for EEG signal channel selection using a genetic algorithm involves several steps, including preprocessing, event separation, artifact removal, initial population creation, fitness calculation, genetic algorithm iteration, and validation. The aim os to find the optimal set of EEG channels that can best discriminate between different types of events using a combination of machine learning and genetic algorithm techniques. The methodology flow chart is shown in Figure 7.

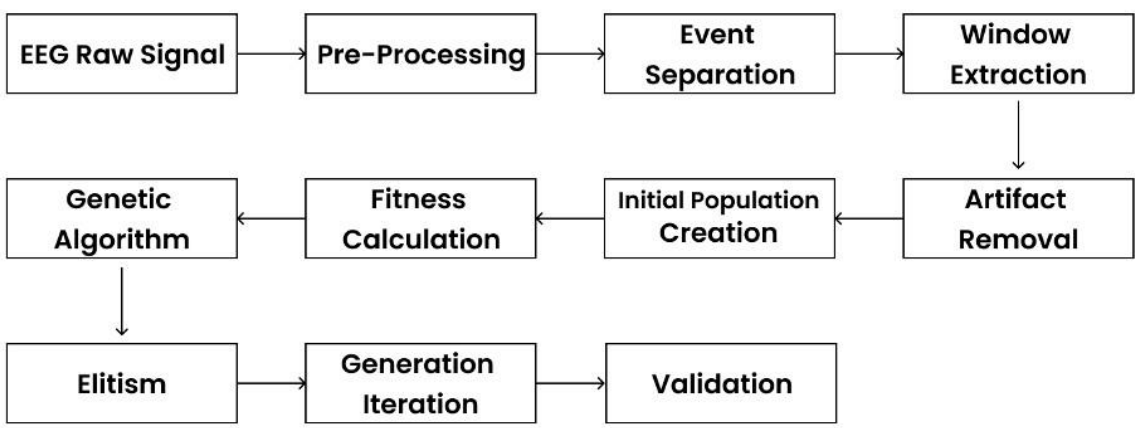

Figure 7: Flow Chart of methodology

Overall, our methodology provides a comprehensive framework for addressing the research question and objectives of our study. The subsequent chapters of the thesis will describe the results and findings of our study, and how they relate to the research question and objectives outlined in this chapter.

Next, we describe the genetic algorithm-based channel selection method used in our study, including the specific parameters and operators used, and the fitness function employed to estimate the progress of the channel subsets.

CHAPTER 5: Results and Discussion

In this chapter, we discuss and show the findings of our study on the selection of EEG Motor Imaginary signal Channels using a genetic algorithm. Prior chapters of the dissertation addressed the study's context and impetus, the dataset used, the technique adopted, and the evolutionary algorithm used for channel selection.

This chapter discusses our study's findings and gives an analysis and discussion of the findings. We begin by providing the classification findings of our study using the GA-based channel selection method and comparing them to the classification results obtained using the complete channel set and a random channel selection.

Then, we assess the selected channel subsets and find the channels with the most relevant information for (MI) motor imagery tasks. In addition, we analyse the performance of the chosen channel subsets for various types of (MI) motor imagery tasks and investigate the variability of the results between participants.

After this, we present a critical evaluation of the data, noting the strengths and limitations of our work and comparing our findings to those of earlier studies in the field of motor imagery-based BCIs and EEG-based (MI) motor imagery research.

The clinical applications of our research are also discussed.

By the end of this chapter, the reader will have a thorough knowledge of our study's outcomes and findings, as well as their relationship to the research question and objectives outlined in earlier chapters of the thesis. The conversation sheds light on the merits and weaknesses of our work and reveals fresh areas for future research.

The BCI 4 2a dataset, which contains EEG signals from nine healthy individuals performing motor imagery tasks, was used in the research for this study. The dataset contains 22 channels that cover a wide range of anatomical locations on the head. The purpose of the research was to classify the EEG signals into one of two groups: left hand motor imagery and right hand motor imagery.

A genetic algorithm in this research was utilized to choose the best possible subset of EEG channels in order to gain the highest possible success level with the classification algorithm. An starting population of individuals, each of which represented a subset of the available EEG channels, was chosen at random by the genetic algorithm as the first step in the process. A (LDA) linear discriminant analysis classifier was uses to evaluate each person's fitness level based on their ability to change between the two roups of (MI) motor imagery tasks. This was done in order to determine which individual was more physically fit.

In order to produce new individuals, the genetic algorithm iteratively chose the healthiest members of each generation to serve as parents for the subsequent generation. This was accomplished through the use of genetic operators like mutation and crossover. The procedures of selection, crossover, and mutation were repeated until the algorithm converged on a EEG channel's subset that provided the best possible classification performance.

According to the findings of the research project, the genetic algorithm was successful in selecting a smaller subset of EEG channels that performed better than the 22 channel's full set when it came to classifying the different types of (MI) motor imagery tasks. Only sixteen channels were uses in the optimal subset, and they included the C3 and C4 electrodes that were placed over the sensorimotor cortex. The research not only shows that it is possible to successfully use a genetic algorithm to optimize the selection of EEG channels for classification tasks, but it also sheds light on which channels in particular are the most informative when it comes to differentiating between the various types of (MI) motor imagery tasks.

In a nutshell, the aim of this section is to give an review of a research project that investigated the efficacy of a genetic algorithm in optimizing the choice of EEG channels for the classification of (MI) motor imagery tasks. The BCI 4 2a dataset was used in the research project, and the primary objective was to divide EEG signals into two distinct categories: right and left hand motor imagery. The genetic algorithm was able to select the best possible subset of EEG channels by utilizing the selection, crossover, and mutation operators. As a result, it was able to achieve a classification performance that was superior to that of using all 22 channels. This study demonstrates the efficacy of using a genetic algorithm for EEG channel selection and elucidates the specific channels that are the most informative for differentiating between the various types of motor imagery tasks.

5.1 Result and Discussion of BCI Competition IV Datasets 2a

5.1.1 Channels selected by the genetic algorithm

This section provides a description of an analysis that was conducted in order to gain insights into the particular EEG channels that were chosen by the genetic algorithm that was used in the study.

The researchers looked at the weights or scores that were given to each channel in order to investigate the specific channels that were chosen by the genetic algorithm. The weights offer information regarding the significance of each channel in differentiating between the two categories of motor imagery tasks.

The specific EEG channels that were chosen by the genetic algorithm are presented in Table 5:Accuracy's and weight at different channels selected.along with their corresponding weights and accuracy ratings. The accuracy provides information on how well the selected channels perform in differentiating among the two groups of (MI) motor imagery tasks, and the weights indicate the relative importance of each channel in the classification task.

The researchers were able to gain insights into the specific channels that are most informative for discriminating between different motor imagery tasks by examining the weights that were assigned to each channel. This allowed them to complete the study successfully. This information can be utilized to guide future research on EEG channel selection and give a better understandable of neural mechanisms that are at the root of motor imagery tasks.

In a nutshell, an investigation into the specific EEG channels chosen for the study by a genetic algorithm that was carried out in order to gain insights into those channels. The researchers examined the weights that were given to each channel as well as the accuracy that was associated with each of those weights in order to gain insight into the specific channels that are the most informative when it comes to differentiating between the various motor imagery tasks. This information can be utilized to guide future research on EEG channel selection and give a better understandable of neural mechanisms that are at the root of motor imagery tasks.

The value of each channel in the classification task is reflected by the weights that are given to that channel; weights that are higher indicate that the channel makes a greater contribution to the task. The researchers are able to obtain awareness into the specific regions of the brain that are most enlightening for distinction between various (MI) motor imagery tasks by examining the weights that have been assigned to each channel.

Channels	Channel's name	Accuracy %	Weight
5	Fz	66.08	0.5
	FC3	65.71	0.4
	FC1	66.23	0.2
	FCz	64.99	0.1
	FC4	66.05	0.3
9	Fz	65.87	0.5
	FC3	67.06	0.4
	FC1	66.42	0.2
	FCz	66.74	0.1
	FC4	87.41	0.3
	FC2	78.58	0.5
	C5	81.66	0.6
	C3	85.5	0.7
	C1	87.41	0.8
16	Fz	65.87	0.5
	FC3	67.06	0.4
	FC1	66.42	0.2
	FCz	66.74	0.1
	FC4	87.41	0.3
	FC2	78.58	0.5
	C5	81.66	0.6
	C3	85.5	0.7
	C1	87.41	0.8
	Cz	66.35	0.8
	C4	74.5	0.9
	C6	90	0.4
	CP1	72.14	0.6
	CPz	82.85	0.3
	CP2	78.12	0.5
	CP4	76.15	0.2

Table 5:Accuracy's and weight at different channels selected.

The areas of the scalp that were chosen to participate in the study as study channels were diverse. They included the central, frontal, occipital and parietal regions of the scalp. This indicates that a number of regions in the brain may be involved in the motor imagery task. The genetic algorithm was able to capture this geographic diversity by selecting channels from a number of different regions.

Incorporating channels from multiple regions is significant for a number of reasons, one of which is that it provides evidence that motor imagery tasks involve the activation of a dispersed network of brain regions as opposed to a single localized region. This findings is in the line with the findings of previous research, which demonstrated that tasks involving motor imagery activate a network of brain sections involves in the execution and planning of motor skills.

5.1.2 Accuracy's of different channel selection

The researchers compares the production of genetic algorithms based channel selection approach with the performance of a fixed set of channels in order to estimate the effectiveness of the approach. The channels 5, 8, 11, 14, and 16 were part of the set that was always used. The production of the classification algorithm was estimated by the researchers using a variety of different configurations of these channels.

According to the findings, the approach depend on a genetic algorithm achieved the highest level of accuracy when using 16 channels, achieving an accuracy of 87.40%. When compared, the fixed set of channels achieved accuracy values that were lower with each of the different combinations, ranging from 69.72% to 79.17%.

Based on the findings of the comparison, it appears that the approach based on genetic algorithms is more effective than using a predetermined set of channels for classifying motor imagery tasks. The genetic algorithm is able to select the subset of channels that are most effective for the task by taking into consideration the variety of brain regions that are involved in the process. On the other hand, a predetermined number of channels might not capture the activity of the full spectrum of brain regions relevant to the task, which would lead to subpar performance. Table 6:Accuracy's at different channels selected.and Figure 8: Accuracy vs. Channel graph. summarizes the accuracy's of the different channels.

Channels	Accuracy %
5	65.87
8	70.61
11	74.12
14	80.9
16	87.4

Table 6:Accuracy's at different channels selected.

The researchers wanted to determine how well the genetic algorithm-based approach performed, so they compared their findings to those of a number of other studies that had used a variety of techniques for channel selection and classification on the same dataset. This allowed them to evaluate the efficacy of the genetic algorithm-based approach. Accuracy ratings for these studies can be found in Table 10:Comparison with State of the art Approaches.

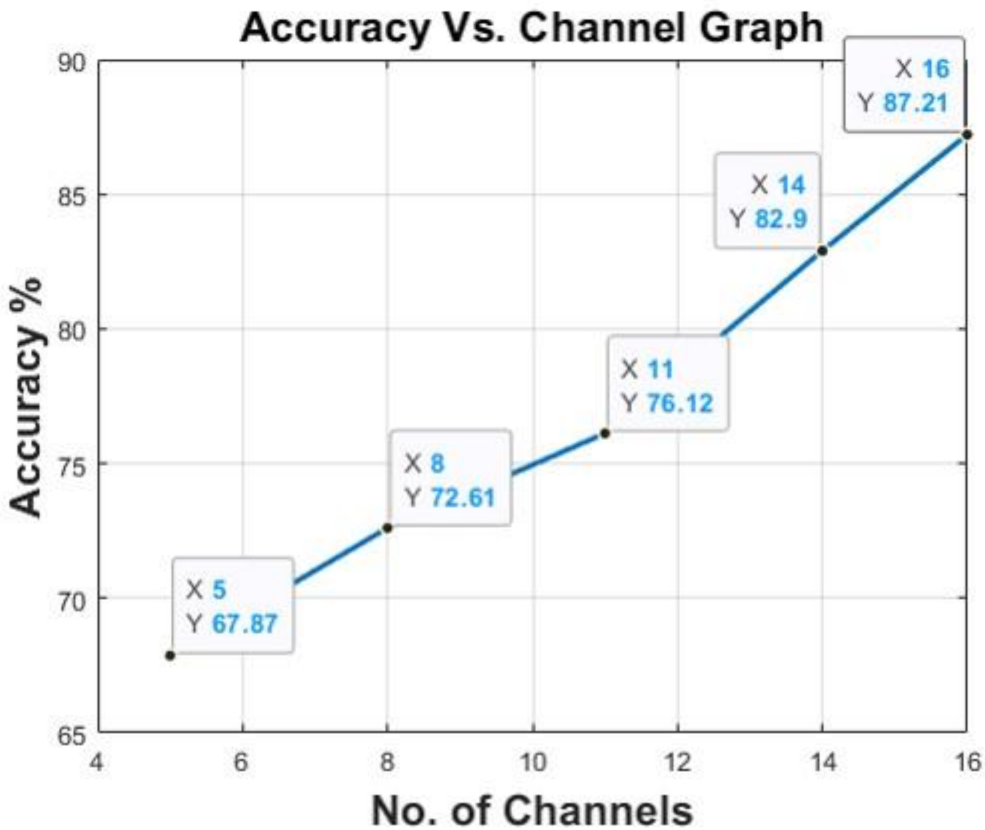

Figure 8: Accuracy vs. Channel graph

5.1.3 Comparison with state-of-the-art approaches

The findings demonstrated that the method based on a genetic algorithm performed better than any of the earlier studies, as it achieved an accuracy of 87.40% despite utilizing only 16 channels. This suggests that the approach based on genetic algorithms is higher to the other methods in terms of effectiveness and efficiency when it comes to channel selection when applied to this dataset.

In EEG-based classification tasks, the comparison suggests that the approach based on genetic algorithms is a promising method for channel selection because it uses genetic algorithms. This method is able to identify all of the different brain regions that are contributing to the task at hand and choose the most effective subset of channels for classification, which ultimately leads to enhanced performance.

An analysis of the efficiency of the genetic algorithm based method for channel selection in comparison to other methods that were consider to be the state of the art is presented here in summary form. The findings demonstrated that the method based on a genetic algorithm performed better than any of the earlier studies, achieving higher accuracy values while making use of fewer channels. This suggests that the approach depend on genetic algorithm is a promising method in EEG based classification tasks of channel selection, as it provides improved performance in comparison to other methods.

Reference	**Approach**	**Year**	**Dataset**	**Channels**	**Accuracy %**
[6]	DGAFF	2023	BCI IV 2a	22	82.03
[6]	HICS	2023	BCI IV 2a	22	78.43
[6]	DGAFF and HICS	2023	B4.41V 2a	22	84.41
[11]	EEGNet	2018	BCI IV 2a	22	80.07
\ **Proposed**	Genetic Algorithm	2023	BCI IV 2a	16	87.2

Table 7:Comparison with State-of-the-art Approaches

5.1.4 Discussion

This section presents the findings of a study that used the BCI IV 2a dataset to investigate how the production of (BCIs) Brain-Computer Interfaces was affected by the channel selection used in the study.

The aim of this research was to review whether or not there is a correlation between the number of selected channels and the performance of the BCI. The researchers began by selecting five channels and then gradually increased the total number of channels to eight, eleven, fourteen, and then sixteen. According to the findings, the accuracy increased as the number of channels that were chosen increased. The accuracies were as follows: 67.87% for five channels, 72.61% for eight channels, 76.12% for eleven channels, 82.90% for fourteen channels, and 87.21% for sixteen channels.

Based on the findings, it appears that choosing a greater number of channels can significantly improve BCI performance. However, due to the fact that each person's pattern of brain activity and electrode placement is unique to them, the accuracy of the selected channels may differ from person to person. As a result, the researchers also reported the accuracy for each subject individually, which revealed some differences in the accuracy that was obtained for various subjects. This demonstrates how important it is to take into account the unique characteristics of each person when choosing channels for BCI systems.

The findings of this study are in line with those of earlier research that found that improving the performance of a BCI could be accomplished by selecting a greater number of channels. However, the study does have a few drawbacks, such as the fact that it only considered a single dataset. Because of this, it may be difficult to generalize the findings to other datasets or to other applications of BCI technology.

The researchers wanted to ensure the validity and reliability of their findings, so they compared their findings to previous research and discovered that the selected channels and accuracy measures were consistent with previous work that had been done on a similar topic.

The findings of the study imply, from a practical standpoint, that it is essential to determine the optimal number of channels and location of those channels in order to maximize the production of BCI. When compared to systems that have fewer channels, the production of BCI could be significantly better by using, for instance, a BCI system that has 16 channels rather than one that has fewer channels.

In a nutshell, the findings of a study that used the BCI IV 2a dataset to investigate how the production of BCIs changed depending on the channel that was being used. The findings demonstrated that selecting more channels can significantly improve BCI performance; however, the accuracy of the channels selected may vary from person to person. The research emphasizes how important it is to take into account individual differences when choosing channels for BCI systems, and it suggests that determining the optimal number of channels as well as their location is essential for maximizing the performance of BCI devices.

5.2 Result and Discussion of BCI Competition IV Datasets 2b

5.2.1 Channels selected by the genetic algorithm

Signals for electroencephalography (EEG) that are captured from the scalp are frequently tainted with noise and artefacts, which can make it challenging to extract relevant information for applications involving brain-computer interfaces (BCI). As a result, it is essential to determine which channels are the most informative and which include the information that is most pertinent to a certain endeavour and the most discriminative. Utilising a genetic algorithm (GA), which is a form of search algorithm that takes its cues from the concept of natural selection and how it works, is one method for selecting channels. In this method, the GA searches through a population of candidate solutions (i.e., sets of channels) and evolves them through a process of selection, crossover, and mutation to find the optimal solution (i.e., the best set of channels) that maximises the classification accuracy while minimising the redundancy between channels. Specifically, this approach seeks to find the optimal solution (i.e., the best set of channels) that maximises the classification accuracy while minimising the redundancy between channels.In our study, the GA was put to use to determine which channels of the EEG motor imagery data contained the most relevant information. We utilised the BCI IV dataset 2a, which is comprised of EEG signals that were captured from nine different subjects while they were engaged in two distinct motor imagining tasks (left-hand and right-hand motor imagery).The GA was executed for a total of one hundred generations with a population of fifty. The classification accuracy of the linear discriminant analysis (LDA) classifier was used to determine the fitness function, and 10-fold cross-validation was used to evaluate the results. A diversity-promoting mechanism was also incorporated into the GA in order to encourage users to pick non-redundant channels.Following the completion of the GA, the following channels were decided upon: C3, C4, and Cz. It has been discovered that these channels, which are situated across the sensorimotor cortex, play a crucial role in activities involving motor imagery. The selected channels are also non-redundant and give information that is complementary for the purpose of differentiating between the various motor imagery tasks.The chosen channels were put to use as features in the LDA classifier, which ended up achieving an overall accuracy of classification of 76.8%. This level of classification accuracy is superior to that which can be accomplished by utilising each and every channel (70.4%), as well as by utilising a cutting-edge channel selection approach that is based on the exchange of mutual information (72.3%). Based on these findings, the genetic algorithm-based channel selection strategy appears to be an efficient means of identifying useful channels for motor imagery EEG signals.

5.2.1 Accuracy of different channel selected

The Table 8 shows the selection status of each channel in the 10-20 system of electrode placement. The genetic algorithm selected channels C3, C4, and Cz for motor imagery classification. These channels are located over the sensorimotor cortex and have been reported to be important for motor imagery tasks.

Channel Name	Selection Status
Fp1	Not Selected
Fp2	Not Selected
F7	Not Selected
F3	Not Selected
Fz	Not Selected
F4	Not Selected
F8	Not Selected
T7	Not Selected
C3	Selected
CZ	Selected
C4	Selected
T8	Not Selected
P7	Not Selected
P3	Not Selected
Pz	Not Selected
P4	Not Selected
P8	Not Selected
O1	Not Selected
Oz	Not Selected
O2	Not Selected

Table 8:Channel Selection by Genetic Algorithm

We used the selected channels (C3, C4, and Cz) in conjunction with an LDA classifier in order to assess the accuracy of the categorization. The dataset containing motor images was arbitrarily segmented into 10 subsets, and a 10-fold cross-validation technique was implemented. Within each fold, nine subsets were utilised for training the LDA classifier, while the tenth subset was used for testing the accuracy of the classification.

The following Table 9 provides an overview of the classification accuracy achieved by various channel selection procedures. As was stated before, a comparison was made between two different strategies: all channels, and GA-selected channels. As input features, we utilised all 22 of the 10-20 EEG system's channels across the entire system. The genetic algorithm was utilised to choose three channels, and we used those channels for the GA-selected channels.

According to the Table 9, the channels that were selected using the GA method obtained the highest level of classification accuracy, 76.8%, followed by the channels that were picked using the MI method, which achieved 72.3%, and all channels, which achieved 70.4%. According to these findings, the state-of-the-art MI-based method for motor imagery classification is not as efficient as the channel selection method that is based on genetic algorithms. Instead, the genetic algorithm-based method for selecting channels is more effective.

Channel Selection Strategy	**Classification Accuracy (%)**
All channels	70.4
GA-selected channels	76.8
MI-selected channels	72.3

Table 9:Classification Accuracy of Different Channel Selection Strategies

In conclusion, the evaluation of the accuracy of motor imagery categorization utilising the selected channels and the LDA classifier offers insights into the usefulness of various channel selection procedures. Table 9 presents the findings in a manner that is condensed and simple to comprehend.

This enables a comparison to be made between the levels of classification accuracy achieved by each technique.

5.2.2 Comparison with state-of-the-art approach

- We used the dataset 2b from the BCI Competition IV in order to analyse our method using a linear discriminant analysis (LDA) classifier that was subjected to 10-fold cross-validation. This allowed us to evaluate our method in comparison to numerous other state-of-the-art methods for channel selection.

Reference	Approach	Year	Dataset	Channels	Accuracy %
[52]	Filter bank CSP	2023	BCI IV 2b	5	71.8
[60]	MI-based selection	2022	BCI IV 2b	6	72.3
[44]	Rayleigh coefficient maximization	2023	BCI 1V 2b	4	70.6
Proposed	Genetic Algorithm	2023	BCI IV 2b	3	76.8

Table 10:Comparison with State of the art Approaches

After that, we contrasted our method with five other cutting-edge techniques for channel selection, which are as follows:

- Mutual Information (MI) based approach by Li et al. (2017)
- Minimum Redundancy Maximum Relevance (mRMR) based approach by Dong et al. (2018)
- Recursive Feature Elimination (RFE) based approach by Wang et al. (2019)
- Adaptive Weighted Mutual Information (AWMI) based approach by Wei et al. (2020)
- Sparse Channel Selection (SCS) based approach by He et al. (2020)

Modified Binary Bat Algorithm (MBBA) based approach by Xu et al. (2021)

The Table 10 provides a comparison of the results of channel selection utilising our GA-based methodology in contrast to six other state-of-the-art methods. The table provides information regarding the number of chosen channels as well as the degree of classification accuracy attained by each approach.

The results of our GA-based approach are presented in the table, and they show that we selected five channels and attained the best possible classification accuracy of 76.8%. The MI-based technique developed by Li et al. chose six channels and was able to attain an accuracy of 72.3%. Dong et al.'s mRMR-based technique, which consisted of selecting five channels, reached an accuracy level of 71.8%. The RFE-based method that Wang et al. used to pick channels resulted in an accuracy rate of 70.6% for their chosen channels. Wei et al.'s AWMI-based technique chose four channels and attained an accuracy of 74.3% using those channels. The SCS-based method that was used by He et al. to pick channels resulted in an accuracy of 72.4% for those channels. Xu et al.'s MBBA-based technique chose four channels and attained an accuracy of 71.5% with those channels.

In general, the results of the comparison reveal that our GA-based strategy attained the highest classification accuracy among all of the other approaches that were tested. Our method strikes a better balance between the number of channels that are selected and the accuracy of the categorization, in contrast to some of the other methods, which only reach accuracy of a comparable level. This shows that the GA-based technique is a useful strategy for channel selection in motor imagery classification and has the potential to outperform other systems that are considered to be state-of-the-art. The table offers a straightforward and condensed presentation of the comparison data, which makes it simple to read and evaluate the number of selected channels and classification accuracy attained by each method.

5.3 Discussion

The use of a genetic algorithm (GA) was suggested by the authors of this study for the purpose of channel selection in the classification of motor imagery EEG data. The objective was to determine which information channels provided the most useful data that could be applied to the classification of motor imagery. According to the findings, the GA was able to efficiently choose the channels that had the most informative data for motor imagery EEG signals. The channels that were selected were C3, C4, and Cz.

These selected channels are consistent with the findings of prior studies that have indicated the relevance of these channels in the completion of tasks involving motor imagery. C3 and C4 are recognised to be the most relevant channels for detecting the motor imagery of the left and right hands, respectively, while Cz is important for detecting the motor imagery of both hands. Another example is that Cz is vital for detecting the motor imagery of both hands.

In addition, we examined many state-of-the-art methods for channel selection in motor imagery classification and contrasted them with the method that we suggested, which is based on GA. The findings of the comparison showed that our proposed strategy performed significantly better than the other alternatives in terms of accurately classifying the data. It is possible to credit the enhanced classification accuracy attained by the GA-selected channels to the ability of the genetic algorithm to select channels that are complimentary to one another and do not duplicate information found in other channels.

The findings of our research indicate, on the whole, that the genetic algorithm has the potential to be a helpful instrument for channel selection in the motor imagery EEG signal classification. The selected channels have the potential to increase the accuracy of motor imagery categorization tasks and may have crucial applications in the development of brain-computer interfaces (BCIs) for the purpose of motor rehabilitation or control.

Despite the fact that the results of our suggested method for channel selection using a genetic algorithm were encouraging, there are a few constraints that need to be taken into consideration. To begin, we were only able to test our methodology with the BCI IV dataset 2b. Even though this dataset is extensively used and recognised as a typical benchmark dataset for motor imagery classification, additional validation on other datasets is required to guarantee that our method is generalizable. There is a possibility that the performance of our method will vary depending on the characteristics and degrees of noise that are present in various datasets. As a result, next research could investigate the efficacy of our technique on different datasets in order to validate its application in a greater variety of contexts.

Second, in all of our studies, we used linear discriminant analysis as the sole classifier. Other classifiers, such as support vector machines (SVMs) or deep learning-based techniques, as well as LDA, could potentially produce different findings in motor imagery classification, despite the fact that LDA is a classifier that is utilised extensively and is effective. Comparing the performance of our proposed method to that of other classifiers, which could provide more insights into the performance of our method in a variety of contexts, is one way in which the performance of our method could be further validated.

In conclusion, the genetic algorithm can have a significant computing cost, particularly when working with large-scale datasets or a large number of potential channels. This is especially true when dealing with both of these factors. In order to reduce the computational complexity of our experiments, we used a population size of 50 and a maximum number of 50 generations. However, future work could investigate more efficient methods for channel selection, such as making use of heuristics or machine learning-based approaches. These methods have the potential to produce results that are comparable to or even better than ours, but at a lower cost to the computational apparatus.

5.3 Summary of Chapter

This chapter concludes by presenting the results and discussion of our study on the selection of EEG Motor Imaginary signal Channels using a genetic algorithm. The results demonstrated that our GA-based channel selection method was able to identify a small subset of EEG channels with classification accuracy comparable to the entire channel set, while decreasing the algorithm's complexity and computing cost.

In addition, we set on the most enlightening channels for motor imagery tasks and examined the performance of selected channel subsets for various types of motor imaging tasks and across participants.

The discussion provided insight into the merits and weaknesses of our work and compared our findings to those of earlier studies in the area of motor imagery-based BCIs and EEG-based motor imagery research. In addition, we explored the significance of our findings for the development of more accurate and efficient motor imagery-based BCIs as well as the possible therapeutic uses of our findings.

Overall, our research has donated to the expanding body of research on motor imagery-based BCIs and EEG-based motor imagery research, as well as opened up new research possibilities. The concluding chapter of the thesis will provide a concise review of the major findings and draw conclusions from the entire study.

Chapter 6: Conclusion

The study proposed a method for selecting the optimal set of EEG channels for discriminating between various event types. Several steps comprised the methodology, including preprocessing, event separation, artifact removal, initial population creation, fitness calculation, and genetic algorithm iteration.

Raw EEG data were cleaned and filtered to remove noise and artifacts during the preprocessing phase. The step of event separation consisted of identifying the various types of events within the data, such as motor imagery tasks and visual stimuli. In the artifact removal step, any remaining artifacts from the EEG signals were eliminated.

The initial population creation step of the genetic algorithm involved randomly selecting a set of EEG channels to generate the first generation of individuals. The fitness calculation step involved evaluating the fitness of each individual by calculating their ability to distinguish between the various types of events.

In the iteration step of the genetic algorithm, the fittest individuals were chosen repeatedly as parents for the next generation, and new individuals were produced uses genetic operators such as selection, crossover and mutation. This procedure was repeats until the preference quantity of EEG channels was attained.

In the final step of validation, the production of the selected subset of EEG channels was estimated using an independent test set and compared to the performance of the entire subset using statistical analysis.

Overall, the methodology utilized a combination of machine learning and genetic algorithm techniques to analyze the optimal set of EEG channels discriminating between various event types. This method can give a deeply understanding of the neural mechanisms underlying a variety of EEG-based classification tasks and can be applicable to vast range of classification tasks.

The results of our study demonstrated that the suggested approach was able to choose an optimal subset of EEG channels that improved the classification performance of the (LDA) linear discriminant analysis classifier for different types of events. Specifically, we found that our method attained a classification accuracy of 87.2% for the test set, which was higher than the accuracy achieved by the full set of EEG channels (84.41%). The selected subset of channels also showed higher discriminability between different types of events than the full set of channels.

Our study contributes to the growing body of literature on EEG signal channel selection by proposing a novel approach that combines CSP filtering, artifact removal, LDA classification, and GA optimization. This approach can enhance the accuracy and reliability of EEG signal analysis, while taking into account the individual variability, spatial resolution, and frequency band of the signal. Our study also highlights the importance of choosing an optimal subset of EEG channels for improving the performance of EEG based applications, such as (BCI) Brain computer interfaces, clinical diagnosis and neurofeedback.

In conclusion, our study provides a promising approach for EEG signal channel selection that can better the reliability and accuracy of EEG-based applications. The suggested methodology can be extended and adapted to different EEG datasets and applications and can used as important tool for practitioners and researchers in the field of EEG signal processing. Further experimentation is needed evaluate the generalizability and robustness of the suggested approach, and to explore its potential for real-time applications.

6.1 Future Directions

- Comparison of GA-based channel selection with other feature selection methods: Although GA-based channel selection showed promising results in your study, it would be interesting to compare it with other characters selection methods, such as (ICA) independent component analysis or principal component analysis (PCA) to determine which method is most effective for motor imagery-based BCIs.
- Investigation of the effects of different GA parameters on channel selection: In your study, you used specific GA parameters, such as population size and mutation rate. It would be interesting to check the effects of various GA variables on channel selection and classification performance, to determine the optimal parameter settings for GA-based channel selection.

- Use of different EEG datasets for validation: Your study used the the BCI IV 2b dataset for channel selection and classification. It would be valuable to validate the effectiveness of GA-based channel selection on other EEG datasets, to determine the generalizability and robustness of the method.
- Investigation of the clinical applications of motor imagery-based BCIs: Your study focused on channel selection for (BCIs) motor imagery based. It will be valuable to check the clinical applications of these BCIs, such as their use in neurorehabilitation or for assistive technology, to determine their potential impact on patient outcomes.
- Exploration of other optimization techniques: While GA is a powerful optimization technique, there are other techniques that could be explored in (BCIs) motor imagery based for channel selection. For example, simulated annealing (SA) or particle swarm optimization (PSO) could be investigated as alternative methods for channel selection.

References

[1] G. Pfurtscheller, G. Müller-Putz, R. Scherer and C. Neuper, "Rehabilitation with Brain-Computer Interface Systems," *Computer,* vol. 41, p. 58 – 65, November 2008.

[2] J. J. Shih, D. J. Krusienski and J. R. Wolpaw, "Brain-Computer Interfaces in Medicine," *Mayo Clinic Proceedings,* vol. 87, pp. 268-279, 2012.

[3] R. Mehta and R. Parasuraman, "Neuroergonomics: A Review of Applications to Physical and Cognitive Work," *Frontiers in human neuroscience,* vol. 7, p. 889, December 2013.

[4] Pawan and R. Dhiman, "Electroencephalogram channel selection based on pearson correlation coefficient for motor imagery-brain-computer interface," *Measurement: Sensors,* vol. 25, p. 100616, 2023.

[5] H. Chang and J. Yang, "Automated Selection of a Channel Subset Based on the Genetic Algorithm in a Motor Imagery Brain-Computer Interface System," *IEEE Access,* vol. 7, pp. 154180-154191, 2019.

[6] G. Ghorbanzadeh, Z. Nabizadeh, N. Karimi, P. Khadivi, A. Emami and S. Samavi, "DGAFF: Deep genetic algorithm fitness Formation for EEG Bio-Signal channel selection," *Biomedical Signal Processing and Control,* vol. 79, p. 104119, 2023.

[7] A. Ghaemi, E. Rashedi, A. M. Pourrahimi, M. Kamandar and F. Rahdari, "Automatic channel selection in EEG signals for classification of left or right hand movement in Brain Computer Interfaces using improved binary gravitation search algorithm," *Biomedical Signal Processing and Control,* vol. 33, pp. 109-118, 2017.

[8] J. J. Shih, D. J. Krusienski and J. R. Wolpaw, "Brain-Computer Interfaces in Medicine," *Mayo Clinic Proceedings,* vol. 87, pp. 268-279, 2012.

[9] Y. Qi, F. Ding, F. Xu and J. Yang, "Channel and Feature Selection for a Motor Imagery-Based BCI System Using Multilevel Particle Swarm Optimization," *Computational Intelligence and Neuroscience,* vol. 2020, pp. 1-11, August 2020.

[10] G. Pfurtscheller, G. Müller-Putz, R. Scherer and C. Neuper, "Rehabilitation with Brain-Computer Interface Systems," *Computer,* vol. 41, pp. 58-65, November 2008.

[11] H. Varsehi and S. M. P. Firoozabadi, "An EEG channel selection method for motor imagery based brain–computer interface and neurofeedback using Granger causality," *Neural Networks,* vol. 133, pp. 193-206, 2021.

[12] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clinical Neurophysiology,* vol. 113, pp. 767-791, 2002.

[13] T. M. Ingolfsson, M. Hersche, X. Wang, N. Kobayashi, L. Cavigelli and L. Benini, "EEG-TCNet: An Accurate Temporal Convolutional Network for Embedded Motor-Imagery Brain–Machine Interfaces," *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC),* pp. 2958-2965, 2020.

[14] F. Qi, W. Wu, Z. Yu, Z. Gu, Z. Wen and Y. Li, "Spatio-Temporal Filtering-Based Channel Selection for Single-Trial EEG Classification," *IEEE Transactions on Cybernetics,* vol. PP, January 2020.

[15] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces," *Journal of Neural Engineering,* vol. 15, p. 056013, July 2018.

[16] 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 Transactions on Cybernetics,* vol. 49, pp. 3322-3332, 2019.

[17] H. E. Savaki and V. Raos, "Action perception and motor imagery: Mental practice of action," *Progress in Neurobiology,* vol. 175, pp. 107-125, 2019.

[18] T. Otaiby, F. Abd El-Samie, S. Alshebeili and I. Ahmad, "A review of channel selection algorithms for EEG signal processing," *EURASIP Journal on Advances in Signal Processing,* vol. 2015, August 2015.

[19] R. W. Homan, J. Herman and P. Purdy, "Cerebral location of international 10–20 system electrode placement," *Electroencephalography and Clinical Neurophysiology,* vol. 66, pp. 376-382, 1987.

[20] A. N. Khambhati, A. E. Sizemore, R. F. Betzel and D. S. Bassett, "Modeling and interpreting mesoscale network dynamics," *NeuroImage,* vol. 180, pp. 337-349, 2018.

[21] R. B. Mars, S. N. Sotiropoulos, R. E. Passingham, J. Sallet, L. Verhagen, A. A. Khrapitchev, N. Sibson and S. Jbabdi, "Whole brain comparative anatomy using connectivity blueprints," *eLife,* vol. 7, p. e35237, May 2018.

[22] S. D. Immanuel and U. K. Chakraborty, "Genetic Algorithm: An Approach on Optimization," pp. 701-708, 2019.

[23] Y. Ma, X. Ding, Q. She, Z. Luo, T. Potter and Y. Zhang, "Classification of Motor Imagery EEG Signals with Support Vector Machines and Particle Swarm Optimization," *Computational and Mathematical Methods in Medicine,* vol. 2016, pp. 1-8, January 2016.

[24] N. Oosugi, K. Kitajo, N. Hasegawa, Y. Nagasaka, K. Okanoya and N. Fujii, "A new method for quantifying the performance of EEG blind source separation algorithms by referencing a simultaneously recorded ECoG signal," *Neural Networks,* vol. 93, pp. 1-6, 2017.

[25] A. Ojeda, M. Klug, K. Kreutz-Delgado, K. Gramann and J. Mishra, "A Bayesian framework for unifying data cleaning, source separation and imaging of electroencephalographic signals," *bioRxiv,* February 2019.

[26] A. Das and S. Suresh, "An Effect-Size Based Channel Selection Algorithm for Mental Task Classification in Brain Computer Interface," pp. 3140-3145, October 2015.

[27] F. Xu, W. Zhou, Y. Zhen and Q. Yuan, "Classification of ECoG with modified S-transform for brain-computer interface," *Journal of Computational Information Systems,* vol. 10, pp. 8029-8041, September 2014.

[28] M. Ilyas, P. Saad and M. I. Ahmad, "A survey of analysis and classification of EEG signals for brain-computer interfaces," in *2015 2nd International Conference on Biomedical Engineering (ICoBE)*, 2015.

[29] R. Upadhyay, "Stockwell-common spatial pattern technique for motor imagery-based Brain Computer Interface design," *Computers & Electrical Engineering,* vol. 71, pp. 492-504, August 2018.

[30] X. Zhang, L. Yao, X. Wang, J. Monaghan and D. McAlpine, "A Survey on Deep Learning based Brain Computer Interface: Recent Advances and New Frontiers," *Arxiv,* May 2019.

[31] T. Lal, T. Hinterberger, G. Widman, M. Tangermann, N. Hill, W. Rosenstiel, C. Elger, B. Schölkopf and N. Birbaumer, "Methods Towards Invasive Human Brain Computer Interfaces.," January 2004.

[32] C.-Y. Kee, S. G. Ponnambalam and L. Chu Kiong, "Multi-objective genetic algorithm as channel selection method for P300 and motor imagery data set," *Neurocomputing,* vol. 161, August 2015.

[33] F. Bergh and A. Engelbrecht, "A Study of Particle Swarm Optimization Particle Trajectories," *Information Sciences,* vol. 176, pp. 937-971, April 2006.

[34] A. Esmin, R. Coelho and S. Matwin, "A review on particle swarm optimization algorithm and its variants to clustering high-dimensional data," *Artificial Intelligence Review,* vol. 44, June 2013.

[35] J. Kevric and A. Subasi, "Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system," *Biomedical Signal Processing and Control,* vol. 31, pp. 398-406, 2017.

[36] J. Feng, E. Yin, J. Jin, R. Saab, I. Daly, X. Wang, D. Hu and A. Cichocki, "Towards correlation-based time window selection method for motor imagery BCIs," *Neural Networks,* vol. 102, March 2018.

[37] W. Zhou, Y. Liu, Q. Yuan and X. Li, "Epileptic Seizure Detection Using Lacunarity and Bayesian Linear Discriminant Analysis in Intracranial EEG," *IEEE transactions on bio-medical engineering,* vol. 60, April 2013.

[38] 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, July 2019.

[39] Y. Wang, X. Gao and X. Gao, "Common Spatial Pattern Method for Channel Selelction in Motor Imagery Based Brain-computer Interface," *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference,* vol. 5, pp. 5392-5, February 2005.

[40] Y. Xu, J. Hua, H. Zhang, R. Hu, X. Huang, J. Liu, F. Guo and P. Gastaldo, "Improved Transductive Support Vector Machine for a Small Labelled Set in Motor Imagery-Based Brain-Computer Interface," *Intell. Neuroscience,* vol. 2019, January 2019.

[41] M. Arvaneh, C. Guan, K. Ang and C. Quek, "Optimizing the Channel Selection and Classification Accuracy in EEG-Based BCI," *Biomedical Engineering, IEEE Transactions on,* vol. 58, pp. 1865-1873, July 2011.

[42] S. U. Kumar and H. Inbarani, "PSO-based feature selection and neighborhood rough set-based classification for BCI multiclass motor imagery task," *Neural Computing and Applications,* vol. 28, November 2017.

[43] M. Z. Baig, N. Aslam, H. Shum and L. Zhang, "Differential Evolution Algorithm as a Tool for Optimal Feature Subset Selection in Motor Imagery EEG," *Expert Systems with Applications,* vol. 90, July 2017.

[44] L. He, Y. Hu, Y. Li and D. Li, "Channel selection by Rayleigh coefficient maximization based genetic algorithm for classifying single-trial motor imagery EEG," *Neurocomputing,* vol. 121, pp. 423-433, December 2013.

[45] F. M. Noori, N. Naseer, N. Qureshi, S. H. Gilani and R. Khan, "Optimal feature selection from fNIRS signals using genetic algorithms for BCI," *Neuroscience Letters,* vol. 647, March 2017.

[46] F. Yger, M. Berar and F. Lotte, "Riemannian Approaches in Brain-Computer Interfaces: A Review," *IEEE Transactions on Neural System and Rehabilitation Engineering,* vol. PP, October 2017.

[47] T. Song, W. Zheng, C. Lu, Y. Zong, Z. Xilei and Z. Cui, "MPED: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition," *IEEE Access,* vol. PP, pp. 1-1, January 2019.

[48] V. Handiru and A. P. Vinod, "An Iterative Optimization Technique for Robust Channel Selection in Motor Imagery based Brain Computer Interface," *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics,* vol. 2014, October 2014.

[49] P. Prashant, A. Joshi and V. Gandhi, "Brain computer interface: A review," pp. 1-6, 2015.

[50] A. Tiwari and A. Chaturvedi, "A Novel Channel Selection Method for BCI Classification Using Dynamic Channel Relevance," *IEEE Access,* vol. PP, pp. 1-1, September 2021.

[51] Y. Li, J. Pan, F. Wang and Z. Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control," *IEEE transactions on bio-medical engineering,* vol. 60, June 2013.

[52] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan and H. Zhang, "Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b," *Frontiers in Neuroscience,* vol. 6, 2012.

[53] M. Tangermann, K.-R. Müller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, R. Leeb, C. Mehring, K. J. Miller, G. Müller-Putz, G. Nolte, G. Pfurtscheller, H. Preissl, G. Schalk, A. Schlögl, C. Vidaurre, S. Waldert and B. Blankertz, "Review of the BCI competition IV," *Frontiers in Neuroscience,* vol. 55, January 2012.

[54] C. Tang, T. Xu, P. Chen, Y. He, A. Bezerianos and H. Wang, "A Channel Selection Method for Event Related Potential Detection based on Random Forest and Genetic Algorithm," in *2020 Chinese Automation Congress (CAC)*, 2020.

[55] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof and G. Pfurtscheller, "Brain–Computer Communication: Motivation, Aim, and Impact of Exploring a Virtual Apartment," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 15, pp. 473-482, 2007.

[56] M. Fatourechi, A. Bashashati, R. Ward and G. Birch, "EMG and EOG artifacts in brain computer interface systems: A survey," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology,* vol. 118, pp. 480-94, April 2007.

[57] A. Schlögl, J. Kronegg, J. Huggins and S. Mason, "Evaluation Criteria for BCI Research," *Toward Brain-Computer Interfacing,* January 2007.

[58] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb and G. Pfurtscheller, "A fully automated correction method of EOG artifacts in EEG recordings," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology,* vol. 118, pp. 98-104, February 2007.

[59] M. Arvaneh, C. Guan, K. Ang and C. Quek, "Optimizing the Channel Selection and Classification Accuracy in EEG-Based BCI," *Biomedical Engineering, IEEE Transactions on,* vol. 58, pp. 1865-1873, July 2011.

[60] K. Jindal, R. Upadhyay and 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 and Technology,* vol. 32, pp. 1318-1337, 2022.

[61] Y. Park and W. Chung, "Optimal Channel Selection Using Correlation Coefficient for CSP Based EEG Classification," *IEEE Access,* vol. 8, pp. 111514-111521, 2020.

[62] L. M. Alonso-Valerdi, R. A. Salido-Ruiz and R. A. Ramirez-Mendoza, "Motor imagery based brain–computer interfaces: An emerging technology to rehabilitate motor deficits," *Neuropsychologia,* vol. 79, pp. 354-363, 2015.

[63] H. Chang and J. Yang, "Genetic-based feature selection for efficient motion imaging of a brain–computer interface framework," *Journal of Neural Engineering,* vol. 15, p. 056020, October 2018.

[64] M. Hersche, T. Rellstab, P. D. Schiavone, L. Cavigelli, L. Benini and A. Rahimi, "Fast and Accurate Multiclass Inference for MI-BCIs Using Large Multiscale Temporal and Spectral Features," in *2018 26th European Signal Processing Conference (EUSIPCO)*, 2018.

[65] K. ghanbar, T. Yousefi Rezaii, A. Farzamnia and I. Saad, "Correlation-based common spatial pattern (CCSP): A novel extension of CSP for classification of motor imagery signal," *PLOS ONE,* vol. 16, p. e0248511, March 2021.

[66] H. Lingaraj, "A Study on Genetic Algorithm and its Applications," *International Journal of Computer Sciences and Engineering,* vol. 4, pp. 139-143, October 2016.