Spring 2017

A Comprehensive Analysis on EEG Signal Classification Using Advanced Computational Analysis

Kaushik Bhimraj

A COMPREHENSIVE ANALYSIS ON EEG SIGNAL CLASSIFICATION USING ADVANCED COMPUTATIONAL ANALYSIS

by

KAUSHIK BHIMRAJ

(Under the Direction of Rami J. Haddad)

ABSTRACT

Electroencephalogram (EEG) has been used in a wide array of applications to study mental disorders. Due to its non-invasive and low-cost features, EEG has become a viable instrument in Brain-Computer Interfaces (BCI). These BCI systems integrate user's neural features with robotic machines to perform tasks. An important application of this technology is to help facilitate the lives of the tetraplegic through assimilating human brain impulses and converting them into mechanical motion. However, due to EEG signals being highly dynamic in nature, BCI systems are still unstable and prone to unanticipated noise interference. In the initial work, a novel classifier structure is proposed to classify different types of imaginary motions (left hand, right hand, and imagination of words starting with the same letter) across multiple sessions using an optimized set of electrodes for each user. The proposed technique uses raw brain signals obtained utilizing 32 electrodes and classifies the imaginary motions using Artificial Neural Networks (ANN). To enhance the classification rate and optimize the set of electrodes of each subject, a majority voting system combining a set of simple ANNs is used. This electrode optimization technique achieved classification accuracies of 69.83%, 94.04% and 84.56% respectively for the three subjects considered in this work. In the second work, the signal variations are studied in detail for a large EEG dataset. Using the Independent Component Analysis (ICA) with a dynamic threshold model, noise features were filtered. The data was classified to a high

precision of more than 94% using artificial neural networks. A decrease of variance in classification accuracies validated, both, the effectiveness of the proposed dynamic threshold systems and the presence of higher concentrations of noise in data for specific subjects. Nonetheless, based on the variance and classification, subjects were further categorized into two groups. The lower accuracy group was found to have an increased variance in classification accuracies. To confirm these results, a Kaiser windowing technique was used to compute the signal-to-noise ratio (SNR) for all subjects and a low SNR was obtained for all EEG signals pertaining to the group with the poor data classification. This study not only establishes a direct relationship between high signal variance, low SNR, and poor signal classification but also presents classification results that are significantly higher than the accuracies reported by prior studies for the same EEG user dataset.

Index Words: Electroencephalogram, Brain Computer Interface, Independent Component Analysis, Artificial Neural Network

A COMPREHENSIVE ANALYSIS ON EEG SIGNAL CLASSIFICATION USING ADVANCED COMPUTATIONAL ANALYSIS

by

KAUSHIK BHIMRAJ

B.S., Georgia Southern University, 2015

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial Fulfillment

of the Requirement for the Degree

MASTER OF SCIENCE

STATESBORO, GEORGIA

A COMPREHENSIVE ANALYSIS ON EEG SIGNAL CLASSIFICATION USING ADVANCED COMPUTATIONAL ANALYSIS

by

KAUSHIK BHIMRAJ

Major Professor: Rami J. Haddad

Committee: Rocio Alba-Flores

Mohammad Ahad

Electronic Version Approved:

May, 2017

DEDICATION

To my mentor and my parents. I couldn't have done it without you.

ACKNOWLEDGMENTS

First, I would like to express my sincere and heartfelt gratitude to my mentor, Dr. Rami J Haddad for believing in me and giving me an opportunity when most thought otherwise. I would next like to express my deepest gratitude to my parents, for doing everything they could to support me at every step. Finally yet importantly, I would like to thank and appreciate my closest colleague and dear friend, Sylvia Bhattacharya for her valuable forethought and camaraderie.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Non-invasive and cost-effective nature of electroencephalography (EEG) renders it suitable for detecting epilepsy, sleep disorders, brain tumors, and other brain-related conditions. Previous studies have modeled EEG data in brain-computer interface (BCI) systems to perform mechanical motion [1, 2]. Interfacing machines with human cognition provides essential assistance for physically impaired individuals [3]. According to a survey conducted by the Center of Disease Control and Prevention (CDC), 6% of females and 3.5% of males above the age of 18 in the United States suffer from serious physical disability [4]. Amongst these people, some suffer debilitating physical disabilities. Noninvasive Brain Computer Interfacing (BCI) systems have potential as a practical solution to help facilitate these individuals' lives. These systems interface the non-muscular brain signals with a computing machine to process and identify a user's intentions and finally convert them to a controlled artificial motion. But brain signals tend to be dynamic and vary across individuals. This hinders the application of BCI systems for assimilating data and recognizing the user's intended motion becomes a complex task. Using Electroencephalography (EEG), studies have developed interesting methods to test the permanence of identifying people across multiple time-spaced sessions through a comparative analysis using data from multiple classifiers [5]. To recognize and learn such complex patterns of brain activity, a robust classification system is needed.

In 2004, a BCI system was implemented for users to control and guide a mobile robot through a maze using EEG signals [6]. The study outlined a brain-machine interface for future studies and implemented a non-invasive EEG system. The concept of BCI was further explored and used for medical applications wherein a user could control a wheelchair system [7]. The signals were filtered using a bandpass filter and noise was identified to be concentrated below 1 Hz. However, noise was still reported to be persistent within the

signals. BCI systems are relatively easy to implement conceptually, but, when implemented they suffer from low accuracy, high muscle contamination, and noise interference that limit their reliability [8].

Classifiers in BCI bolster user's safety and system reliability through quick and robust classification of data for different neural tasks [9]. Four classifiers namely, support vector machines (SVM), linear discriminant analysis (LDA), statistical classifier (SC), and artificial neural networks (ANN) have proven to be good EEG feature selectors [10]. SVM and LDA classifiers were used with a 10-fold cross validation to classify data from two datasets and resulted in high accuracies with SVM having a higher rate of classification [11]. Here the issue of signal correlation among channels is highlighted which will be further examined in this work. In another study, ANN classification was used along with a fuzzy particle swarm optimization training function to classify EEG signals for 10 subject sample, with 5 subjects being able bodied and the other five suffering tetraplegia [12]. The study provides a comparative analysis for five classifiers and reports its ANN model to have the highest classification accuracy. However, the eye movements were kept to a minimum which does not replicate real world application. In these studies, techniques to classify multidimensional EEG data for BCI systems were explored. However, the scope of these studies have been limited to a few subjects and failed to provide a good reference for reasonable stability in a large scale sample set.

In the initial work, a majority-vote system was implemented to a network of artificial neural networks (ANN) to optimally classify imaginary motions performed by subjects for multiple sessions. The proposed technique optimizes the electrodes used for individual user classification by ranking each electrode's data based on its individual classification accuracy. The best performing electrodes are identified with a rank-based statistical analysis. In the second work, the use of an automatic feature extracting independent component analysis (ICA) system with an ANN classifier that uses the Levenburg-Marquardt training function

to classify a large scale dataset of 105 subjects is proposed. The work validates the relationship between signal to noise ratio (SNR), signal variance across multiple sessions, and signal classification accuracy. The dataset considered for this work was acquired from PhysionNet and consists of both imagined and actual movements performed by 105 subjects [13, 14]. In a different study, wavelet transform features were extracted from the EEG Movement/Imagery dataset and an ANN was used for classification [15]. This study reported a maximum classification accuracy of 68.21%. A phase locking value system (PLV) was used for the same dataset to classify the β (12-30 Hz) and μ (8-12 Hz) rythms for actual movements (78.95% & 63.73%) and imagery tasks (71.55% & 65.55%) in [16]. Another study using two feature selection processes (ICA and frequency band selection), classified the data using an SVM classifier with a Gaussian kernel and reported a high average accuracy of 69% [17]. The average classification accuracy reported is 11% higher than the highest accuracy reported in all previous studies pertaining to the same dataset. However, in this work, details are presented about the data that has not been presented before and propose a robust system that automatically extracts task features using an ICA and classifies them using an ANN classifier.

The remainder of this paper is organized as follows. Chapter 2 provides the literature review for the thesis and introduces key concepts of research. Chapter 3 presents the proposed majority voting technique and provides results from genetic algorithm. Chapter 4 proposes two autonomous noise removal techniques for ICA components and uses an ANN to validate the techniques. Finally, Chapter 5 concludes the paper with a summary of the proposed classification models and future works.

CHAPTER 2

LITERATURE REVIEW

2.1 General Concentrations & Divisions of The Human Brain

The human brain is a vital organ that assists in processing information in order to perform complex day-to-day physical and mental tasks. It is protected within a bone cavity called the cranium. The cranium has several bones that are connected along the suture lines. Pertaining to these suture topographies, the brain regions within the cranium are also classified. The fissures on the grey matter, as shown in Fig 2.1, are divided into several regions, i.e, Frontal lobe, Parietal lobe, Temporal lobe and the Occipital lobe [18]. When observed from the top,

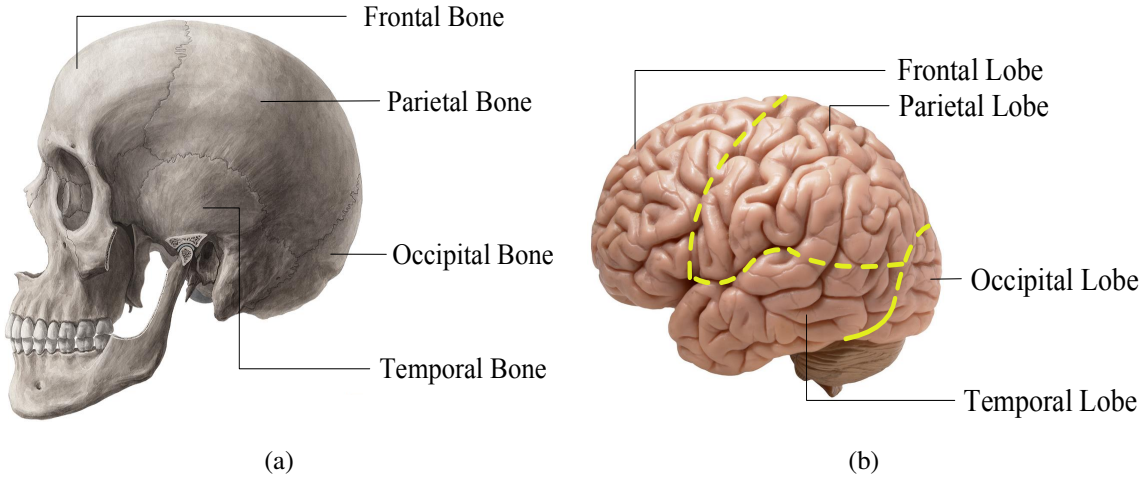


Figure 2.1: a. Bones of the cranium and their suture topography b. Regions of brain pertaining to cranium topography.

the brain seems to be partitioned into left and right hemispheres and the partitioning fissure is called the Great Longitudinal Fissure. The two hemispheres of the brain tend to mirror each other in patterns with little variation but they are still regarded as distinct regions that process information differently [19]. Furthermore, specific regions in the Frontal lobe have been observed to specialize in speech processing, movements of hand, limb, ocular, facial

and body related movements. Regions in the parietal lobe have been observed to integrate sensory inputs with visual information. The temporal region houses the auditory area that processes signals from the ear into meaningful units of information. Lastly, the occipital region is primarly known to process visual information from the optical nerves [20]. Such localizations in the frontal cortex which are responsible for movement of hand and legs are situated in the region between the frontal and parietal region. The region preceding the motor cortex lies within the parietal cortex and is known as the somatosensory region. These regions together process the pressure and intensity of the movement to coordinate movement-related tasks. In such manner, various regions of the brain seamlessly coordinate to provide complex and calculated responses. In Fig 2.3, the motor, somatosenzsory and occipital regions of the brain are shown with emphasis on the visual cortex. To study such co-ordinations across the regions of the brain, researchers have started analyzing data recorded for performed imagery tasks.

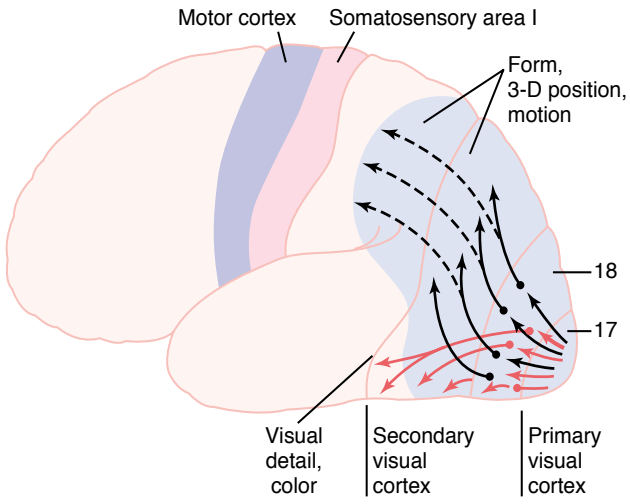


Figure 2.2: Lateral view of the Motor, Somatosensory and Occipital areas of the brain

Using fMRI imaging procedures, data pertaining to constructing and deconstructing mental images were recorded for 15 subjects [21]. The study not only affirms the active involvement of the visual cortex in constructing complex mental images but it also confirms

the co-ordination of other regions of brain in the process. The fMRI results are shown in Fig 2.3. In accordance to the results of this study, this research focuses on analyzing the EEG recordings based on motor imagery.

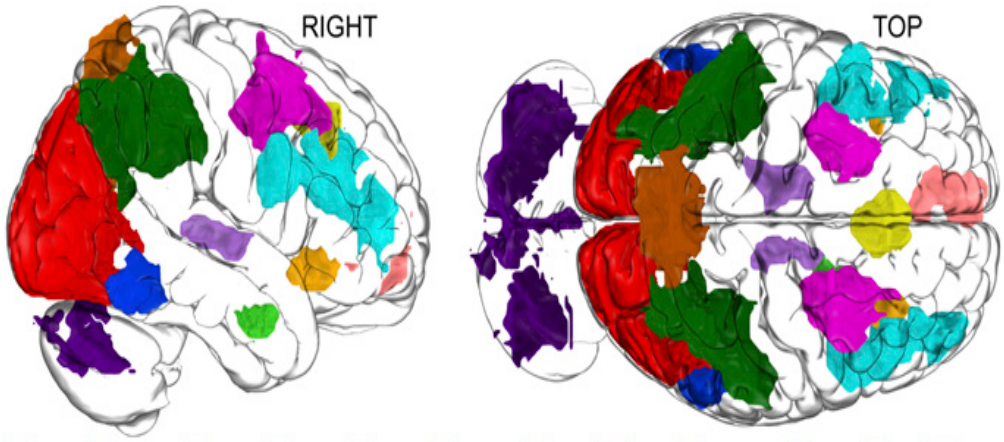
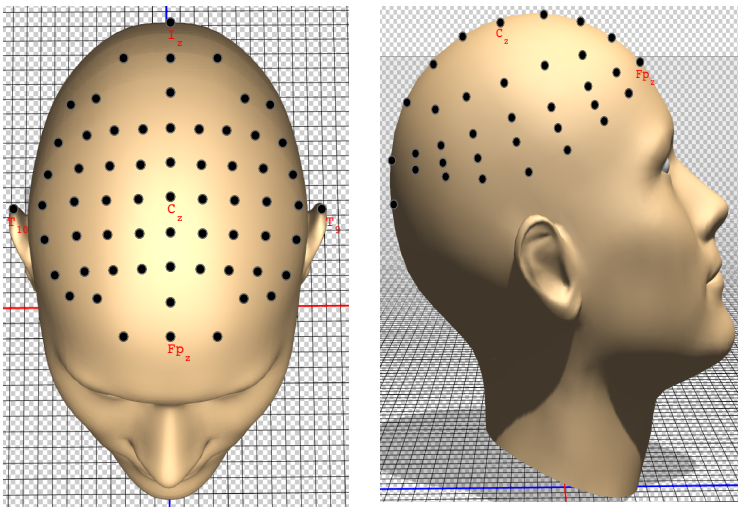


Figure 2.3: fMRI images from a study conducted by researchers are Dartmouth University of the brain when performing imagery tasks [21].

2.2 Electroencephalography

Electroencephalography has been used to study brain functioning by recording electrical activity of the neural tissue. The electrical impulses generated by the neurons in the brain, permeate the bone tissue to the scalp. These low amplitude signals across the scalp are recorded at specific sampling frequencies of 160 - 512 Hz [22]. Numerous electrodes are placed at specific locations on the scalp based on 10-10 and 10-20 placement systems. The number of electrodes can range anywhere from 1 to 128 electrodes. But most studies use a 32 or 64 electrode setup for study. In Fig 2.4, a 64 electrode EEG setup is shown where the electrodes are placed on scalp using a 10-10 positioning system. EEG signals are used to detect mental abnormalities such as sleep disorders, epilepsy, and paralysis. Due to its noninvasive and cost effective features, EEG has been used in Brain Computer

Interface (BCI) systems. Neural signals recorded by EEG are not only used to detect mental abnormalities but are now being used to transfer human thought into computer enabled action [23].



(a) (b)

Figure 2.4: EEG electrode setup using 10-10 standard system a. Anterior view b. Lateral view

Signals acquired using EEG can also be represented in the frequency domain within four bandwidths. These bandwidths are categorized as alpha, beta, theta and delta. The alpha waves are rhythmic waves that exist between 8 to 13 Hz. These waves originate mostly in adults while awake and in resting state. The waves are mainly concentrated in the occipital region. At normal and attentive state, the beta waves are generated at frequencies between 14 - 80 Hz. These waves predominantly originate in the frontal and parietal regions of the brain. Theta waves are generated during emotional stress, disappointment and frustration. They also occur between 4 - 7 Hz among people suffering from degenerative brain disorders. Lastly, delta waves are generated during deep-sleep state and infancy at 3.5 Hz. These frequency bandwidths are also used to study and analyze the brain.

2.3 Applications of Brain Computer Interface

Brian computer interface (BCI), is an interface that extracts, detects, and translates specific neural generated impulses from the brain into machine defined as actions. BCI systems generally consist of three sections, acquisitioning system, feature extraction system, and signal classification system. This outline of a standard BCI system is shown in Fig 2.5. Within these three sections of the BCI system, EEG has become an effective tool to acquire brain signals from the scalp. In the second section, specific neural features that were performed by the subject triggered by specific visual or vocal queues are identified using feature extraction techniques. Some of the feature extraction techniques transform time based EEG signals to the frequency domain. And data at specific frequencies is extracted and re-transformed back into the time domain.

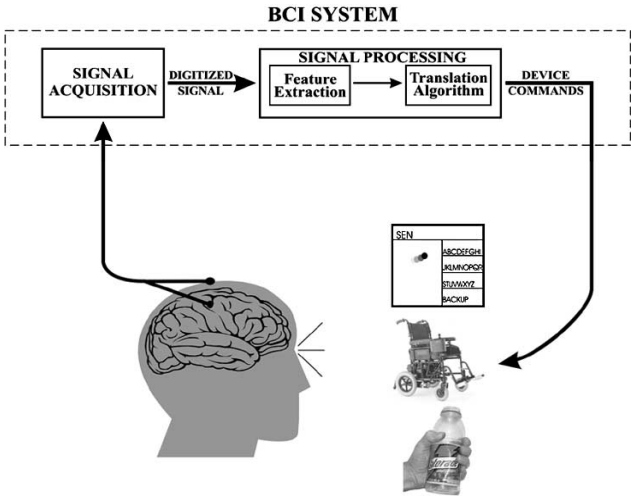


Figure 2.5: Outline of a brain computed interface system with its three main components.

To mitigate noise issues, BCI systems use digital filters or feature selection techniques to remove unwanted artifacts from the EEG data. Sole use of low-pass or band-pass filters, do not remove eye and muscle related contamination present in the signal. However in some EEG recordings for BCI system, users were asked to restrict their eye movements to minimize contamination. This approach is impractical for real world BCI applications. To

remove noise artifacts, blind source separation (BSS) models can be effective. Independent component analysis (ICA) is one such technique which has been successfully used to extract neural related features from contaminated signals [24, 25].

2.4 Artificial Neural Networks

Artificial neural network (ANN) are computation models that mimic the workings of the brain. These models were widely studied in the early eighties and nineties. However, until recently, implementation of this concept had remained a challenge due its computational cost. The ANN models work well for machine learning problems and are better adept to learning and classifying data with complex variations consisting of a large feature sets. The human brain does a marvelous job in learning new things through processing various inputs such as speaking, movement, hearing, sense of touch, sight, listening and complex motor imagery. Scientists hypothesize that brain is highly flexible and can adapt to a new input. In a study, the neural connections between the auditory nerve and the auditory cortex were severed and the optic nerve of the eye was connected to the auditory cortex. Remarkably, the auditory cortex learned to see. In another study, the somatosensroy cortex of the brain which processes the sense of touch was re-wired to receive visual inputs from the optic nerve. The somatosensroy cortex had learned to process the visual signals. These re-wiring experiments establish, that any part of the brain can be used process and learn any kind of input. This encourages researcher to explore a single computational model that can replicate brain's processing capabilities. On a computer, accomplishing such flexibility in learning can seem challenging and to mimic the brain it is essential to introduce the neuron.

2.4.1 Structure of Neuron

The neuron is the main processing unit of brain and there are approximately three trillion neurons in the human brain. The structure of the neuron can be categorized into three parts, the axon, dentrites and the cell body as shown in Fig 2.6. The dentrites connect the cell

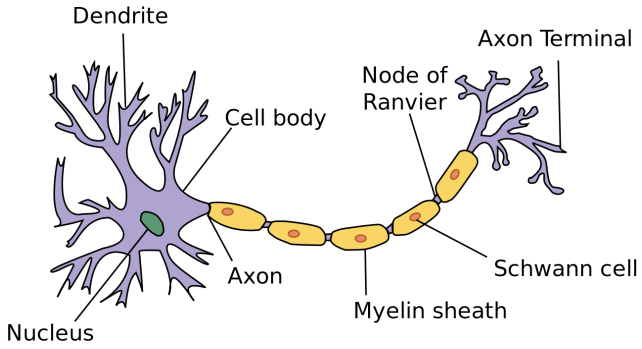
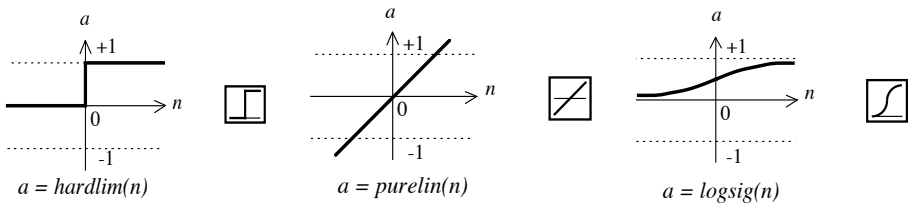


Figure 2.6: Proposed Classification System Model

body to other neurons. There can be numerous dentrites to a neuron and can extend upto 0.5 meters in length. Through dentrites signals are carried into the neuron. The axon is another extension of the cell body that is coated with a protein insulation called the Myelin Sheath. The axon is a longer extension and can measure upto one meters in length. The elongation carries signals out of the cell to the axon terminal.

2.4.2 Brief Overview of Neural Networks

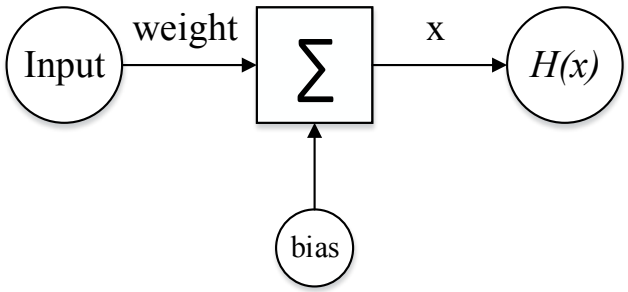
Using the structure of neuron, the computational model of the artificial neural network was designed. In 1943, Warren McCulloch and Walter Pitts created a computational model of the neural network [26]. Backpropagation models were later proposed as addtions to the neural network model in 1975 [27]. Within the ANN, the neuron has one output and one input. In the neuron, the input is multiplied to a weight and fed into a transfer function. Three transfer functions are mainly used in NN systems, i.e, the hard-limit transfer function, linear transfer function and log-sigmoid transfer function. The hard-limit transfer function



(a) Hard-Limit (b) Linear (c) Log-Sigmoid

Figure 2.7: Types of transfer functions used in a neuron [28].

outputs is '0' when the input is less that zero and outputs a '1' when the input is greater than one. Neurons using a linear transfer function use linear approximations or linear filters to provide outputs. Finally, the log-sigmoid transfer function in neurons, takes input values between -∞ and +∞ and provides outputs in a range between 0 and 1. The sigmoid function is frequently used in backpropagation networks due to the differential properties. The outputs for these functions are shown in Fig 2.7. In the figure, the pre-defined MATLAB functions are also shown. Using these functions the neuron computes its output and feeds it forward to the next neuron. Along with inputs and weights, the neuron can also have a bias associated to it. The bias parameter is used to add further emphasis to the neuron's output. The structure of a neuron is shown in Fig 2.8.



$A = f(\text{Input} * \text{weight} + \text{Bias})$

Figure 2.8: Structure of Neuron in an ANN

2.4.3 Feed Forward Network

Basic ANN architecture mainly consists of three types of layers, i.e, input layer, hidden layer and the output layer. As seen in Fig 2.9, a simple ANN consists of three types of layers with several nodes, each designated to a neuron. The computation of in each layer can be

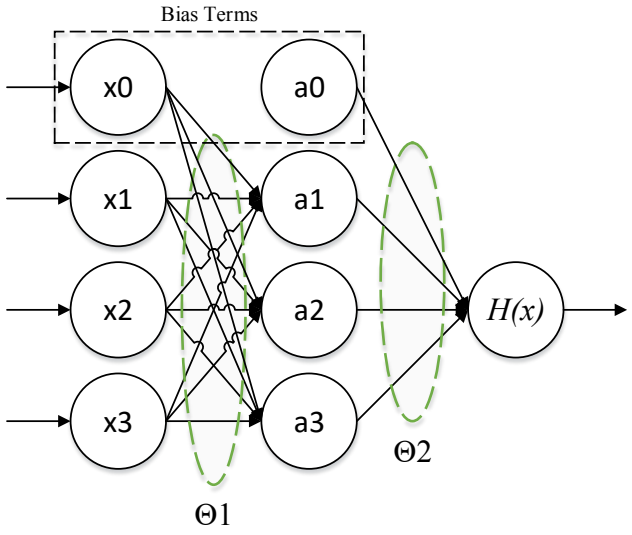


Figure 2.9: Feed Forward Neural Network

mathematically represented as follows.

$$a_1 = g(\theta^1_{10}x_0 + \theta^1_{11}x_1 + \theta^1_{12}x_2 + \theta^1_{12}x_3)$$

$$a_2 = g(\theta^1_{20}x_0 + \theta^1_{21}x_1 + \theta^1_{22}x_2 + \theta^1_{23}x_3)$$

$$a_3 = g(\theta^1_{30}x_0 + \theta^1_{31}x_1 + \theta^1_{32}x_2 + \theta^1_{33}x_3)$$

$$H(x) = g(\theta^2_{10}a_0 + \theta^2_{11}a_1 + \theta^2_{12}a_2 + \theta^2_{13}a_3)$$

where x_0 and a_0 are bias parameters and $g(x)$ is the transfer function. Size of the weight matrix of the layer Θ_1 consists of S_{j+1} x $(S_j + 1)$ values. Here j is the number of layers in the ANN architecture. The weight matrix changes for each layer. In Fig 2.9, the input layer consists of inputs $X = \{x_0, x_1, x_2, x_3\}$, the hidden layer nodes are denoted as $A = \{a_0, a_1, a_2, a_3\}$ and lastly the output layer consists of a single output node with $H(x)$ as

the output. In Eq 2.1, the above equations can be represented a vectorized form,

$$Z = \Theta_1 * X \tag{2.1}$$

where Θ_1 is the weight matrix for the hidden layer and X is the matrix containing the input values. Each of the nodes in the hidden layer compute an output using a transfer function $g(x)$ as shown in Eq 2.2,

$$A = g(Z) \tag{2.2}$$

where A consists output values from all the nodes of the hidden layer. The final output $H(x)$ is computed in Eq 2.3,

$$H(x) = \Theta_2 * A \tag{2.3}$$

where Θ_2 is the weight matrix for the output layer.

2.4.4 Cost Function & Backpropagation

While the feed forward networks contain neurons that calculate outputs based on the given input, weight and bias parameters, the backpropagation model in the neural networks optimizes a cost function and updates the weights in each layer of the neural network. Before the process of backpropagation is explained, it is necessary to introduce the cost function of a neural network model. The function is similar to the cost function of a logistic regression classifier but the cost $J(\Theta)$ in NN models is calculated for all the layers. In Eq 2.4, the cost function is defined as,

$$J(\Theta) = -\frac{1}{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{T} y_j^i log(H(x^i))j + (1 - y_j^i) log(1 - H(x^i))j\right) \tag{2.4}$$

where n is the number of nodes in a layer, T is the number of layers, y_j^i is the expected output of j^{th} node in the i^{th} layer and $H(x^i)$ is the acquired output of the j^{th} node in the i^{th} layer. Using the above cost function, the backpropagation model seek to optimize weights by minimizing $J(\Theta)$.

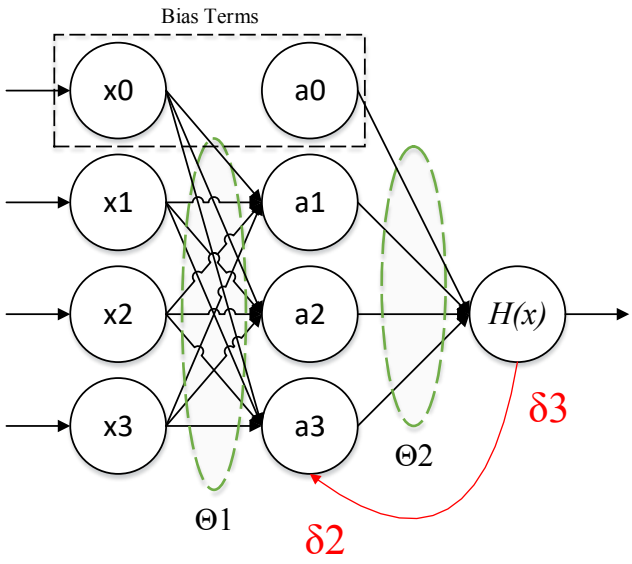


Figure 2.10: Backpropagation in a Feed Forward NN

In Eq 2.5, using the cost function an error δ_j^i in each layer is calculated as,

$$\delta_j^i = a_j^i - y_j^i \tag{2.5}$$

where a_j^i is the computed output, y_j^i is the expected output at j^{th} node in the i^{th} layer. The error in each layer is summed for all the nodes and then multiplied to the error of the previous layer. The errors are from the output layer to the input layer in a backward direction as shown in Fig 2.10. For such network, the backpropagation process is shown in as,

$$\delta_j^3 = \delta_j^i = a_j^4 - y_j$$
$$\delta_j^2 = (\Theta^2 \delta_j^3) * a_j^2(1 - a_j^2)$$

Using the computed errors the cost $J(\Theta)$ is minimized by,

$$\frac{\delta J(\Theta)}{\delta \Theta} = a_T^n \delta_T^n \tag{2.6}$$

The key computational aspects of the research have been explained in this chapter. The proposed techniques using EEG, augmented BCI with ANN classification is explored in the next chapters.

CHAPTER 3

OPTIMIZATION OF EEG-BASED IMAGINARY MOTION CLASSIFICATION USING MAJORITY-VOTING

3.1 Overview

In this chapter, a majority-vote system was added to a network of artificial neural networks (ANN) to optimally classify imaginary motions performed by subjects for multiple sessions. The proposed technique optimizes the electrodes used for individual user classification by ranking each electrode's data based on its individual classification accuracy. The best performing electrodes are identified with a rank-based statistical analysis. The remainder of this chapter is organized as follows. Section 3.2 discusses the EEG data used in the experiments conducted to validate the proposed method. Section 3.3 illustrates the proposed method. Section 3.4 reports the results of the proposed method. Section 3.5 introduces the genetic algorithm to optimize electrodes and seeks to further improve classfication accuracy, the findings are also compared for improvements. Finally, the results and findings obtained from these experiments are summarized in the conclusion in section 3.6.

3.2 Dataset Description

Dataset V from the BCI Competition III was used for validation. The data was recorded by the Istituto Dalle Molle di Intelligenza Artificiale Percettiva (IDIAP) research institute in Switzerland [29]. The dataset consists of brain activity pertaining to three specific tasks from three healthy male subjects. These brain signals were recorded in repeated sessions over a course of a single day and the subject's imagination of three types of exercises included - left hand movement, right hand movement, and imagination of words starting with the same letter. Each of these three sessions lasted for 4 minutes with brief intervals ranging between 5-10 minutes. During each session, the subjects imagined a specific task

Table 3.1: Electrodes & Their Alloted Channel Numbers

Channels	Electrode	Channels	Electrode
1	F_{P1}	17	O_2
2	AF_3	18	PO_4
3	F_7	19	P_4
4	F_3	20	T_6
5	FC_1	21	CP_6
6	FC_5	22	CP_2
7	T_3	23	C_4
8	C_3	24	T_4
9	CP_1	25	FC_6
10	CP_5	26	FC_2
11	T_5	27	F_4
12	P_3	28	F_8
13	P_Z	29	AF_4
14	PO_3	30	FP_2
15	O_1	31	F_Z
16	O_Z	32	C_Z

for a time period of 15 seconds and tasks were changed as indicated by the operator. The voltages across the scalp were measured using the Biosemi portable EEG machine with a sampling frequency of 512 Hz. Thirty two electrodes were used to acquire data and electrodes were placed at specific locations on the scalp using the IEEE 10-20 standard system. Electrodes were assigned a number for ease of reference and the numbers relating to the electrode names are listed in Table 4.2. The raw data from these 3 subjects over the 3 sessions comprise of $3 \times 3 \times 32 = 288$ total data channels.

3.3 ANN Structure & Majority Voting System

The prime focus of this research is to implement a Majority Voting (MV) system to optimize the classification accuracy for each subject. It's main contribution lies in the unique structure

for this classifier that utilizes a network of relatively simple ANNs with a majority vote circuit. Instead of using a single complicated ANN with 32 inputs, using majority vote simplified the ANN structure. Before the voting system is explained, it is essential to first describe the system structure. The EEG recordings for each of the three subjects contain columns of data acquired from each of the 32 channels. The data from each column is input into an artificial neural network (ANN). These neural networks use back propagation to train, test, and validate the data. In this experiment, 70% of the data was used for training, and 30% was used for testing and validation. The ANN structure has a single neuron assigned for input, while the hidden layer consists of 10 neurons in its first layer and 20 neurons in the second. The output layer was allotted 3 neurons. The outputs acquired from the neural networks were categorized by the MV system into 3 imaginary task categories i.e, class 7, class 2, and class 3.

The process is visually illustrated in Figure 3.1. The acquired ANN data is split into three columns for each of the 32 channels. Across each row of the three columns, the maximum value is replaced with a '1', and the rest are changed to '0'. Next, these binary values are added to the binary values acquired from other channel's ANN outputs. This process repeated for each row results in a single output that consists of the majority voted data from all of the specified channels.

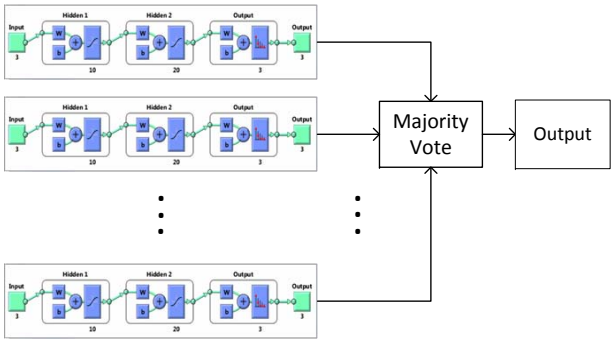


Figure 3.1: Proposed Classification System Model

Algorithm 1 Majority-Vote Model of ANN Outputs.

Precondition: For each subject, $\{x_1, x_2, ...x_n\}$ is data from n channels.

function MAJORITY VOTE($x_1, x_2, ..., x_n$)
  for $k \leftarrow 1$ to n **do** ▷ Number of iterations
    $[A_k, y_k] \leftarrow ANN(x_k)$
  end for ▷ Y contains y_k outputs and A contains A_k classification accuracies from ANN.
  $\{z_1, z_2, ...z_n\} \leftarrow rank(Y)$ based on A_n. ▷ Z has z_n ranked channels.
  for $i \leftarrow 1$ to N **do** ▷ N ranked outputs considered.
    $z_i \leftarrow \{a_i, b_i, c_i\}$
{$\{a_i, b_i, c_i\}$ are 3 imaginary tasks contained in z_i.}
    $[rows, columns] \leftarrow size(z_i)$
    for $j \leftarrow 1$ to $rows$ **do**
      $z_i^* \leftarrow max(\{a_{ij}, b_{ij}, c_{ij}\}) = 1$ and
      $z_i^* \leftarrow nonmax(\{a_i, b_i, c_i\}) = 0$
{Row max for 3 tasks replaced with 1, others with 0.}
    end for
    $\delta_i \leftarrow sum(z_i^*)$
  end for
  $\delta \leftarrow \{\delta_1, \delta_2, ..., \delta_N\}$
  for $m \leftarrow 1$ to N **do**
    $MajorityVoted \leftarrow max(\delta_m) = 1$ and
    $MajorityVoted \leftarrow nonmax(\delta_m) = 0$
  end for
end function

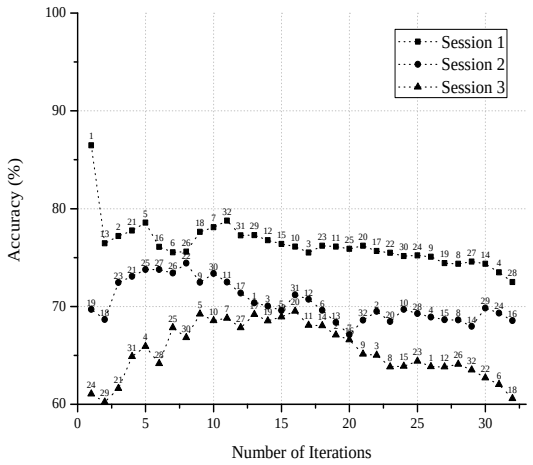
Using the MV system, the overall aggregated classification accuracy of all 32 channels was evaluated to obtain a baseline accuracy. Next, all 32 channels were ranked based on their individual classification accuracies and then sorted in a descending order. By using the obtained ranked and sorted sequence, the channels were majority voted in an iterative process starting from the most to the least accurate channel. This process of majority voting is detailed in Algorithm 1. The process continues until the overall aggregated classification accuracy of the majority vote is optimized.

Using this process, the set of electrodes used to generate the maximum classification accuracy was considered as the optimal set for that user. To quantify the effectiveness of the proposed approach, the classification accuracies of the optimized set of electrodes for each user and the overall classification accuracy across all users were obtained and analyzed in Section 3.4.

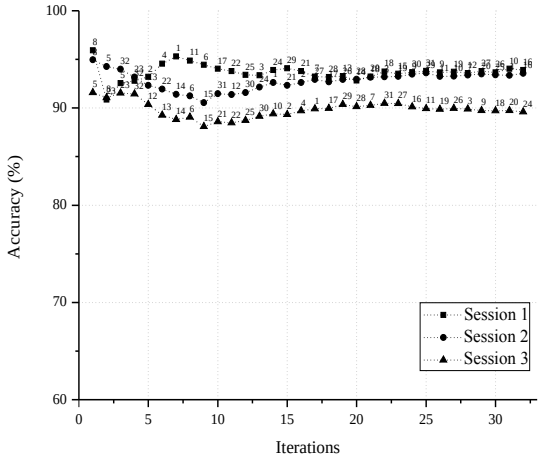
3.4 Results & Data Analysis

Data for each subject contains 3 sessions of repeated imaginary tasks. Using the majority voting system, channel data for each session was ranked and sorted according to the individual ANN channel accuracies. In the channel optimization process, the majority voting process was used twice. The first time, channel optimization was performed for each session separately. Using this method, wide variations were noticed across the three sessions among the optimized channel sets for each subject. This was due to the presence of time and magnitude related variance in the input data itself. The MV accuracies across the 32 iterations are presented in Figures 3.2-(a,b,c) for all three subjects in the dataset. In these plots, the x-axis does not represent 32 separate channels but rather 32 iterations of sequential majority voting of ranked channels based on their accuracies.

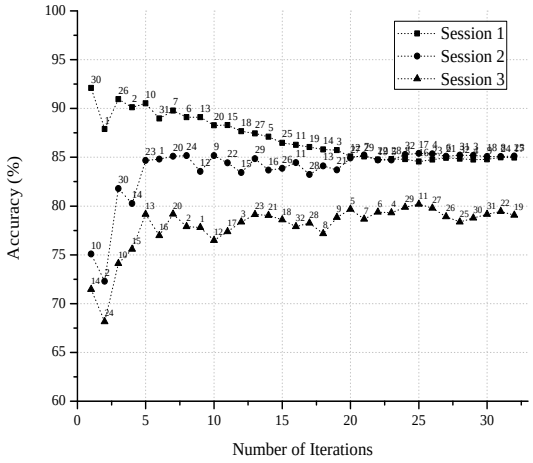
In order to mitigate the fluctuations across the subject sessions and stabilize the optimization process, the majority voting process was used again. However, this time the



(a) Subject 1



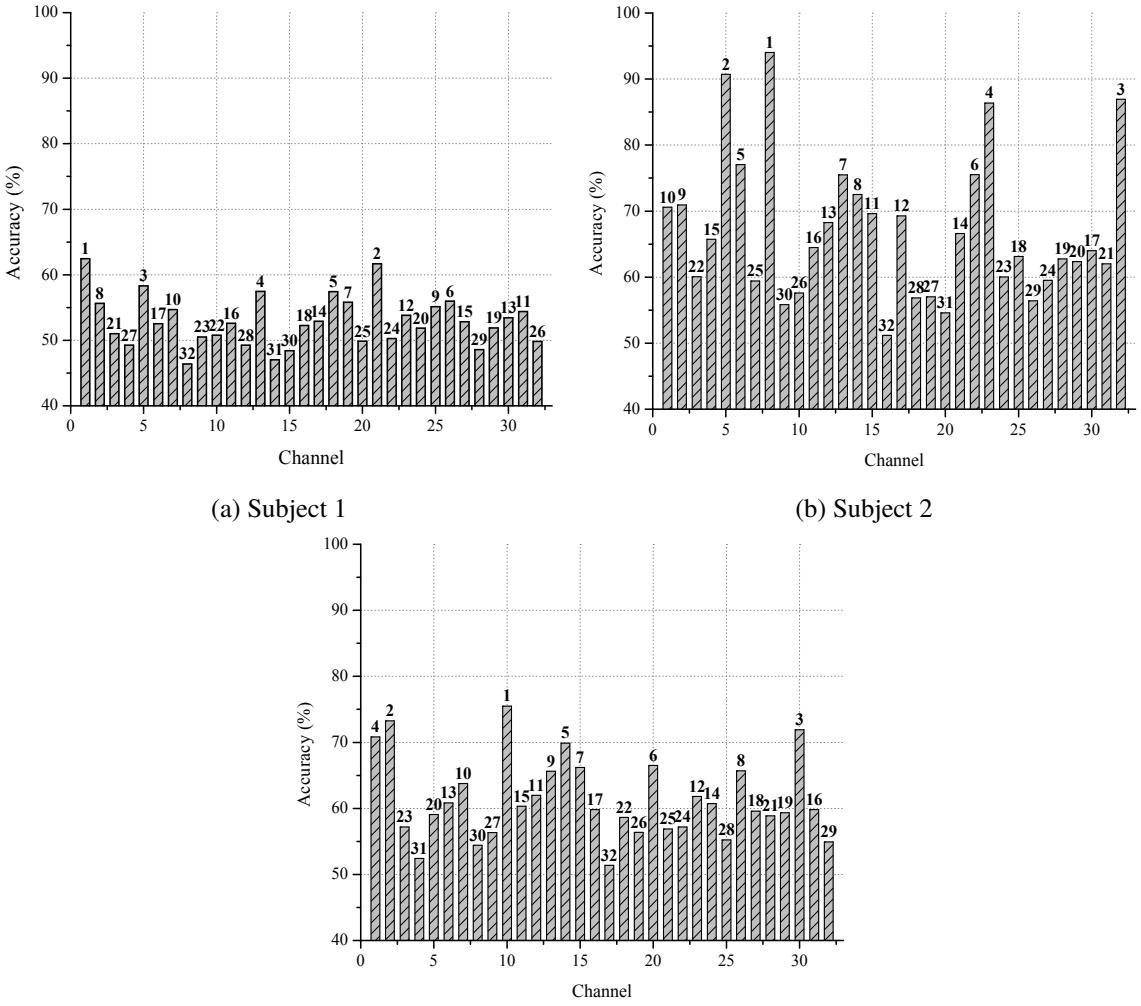
(b) Subject 2



(c) Subject 3

Figure 3.2: Accuracies for Sequential Majority Voted Channel Combinations based on Ranked Individual Channel Accuracies for each of the three Sessions.

individual channel accuracies for each session were averaged and then sorted in decreasing order. The channel sequence obtained was input into the majority voting system. The averaged ANN accuracies for each channel is shown in Figures 3.3-(a,b,c) for the three subjects. Their ranks are also specified in the figure. Using these ranks, the MV system evaluated the classification accuracies for each session. The resulting MV accuracies from the 32 iterations were averaged to get the final individual accuracies for each of the subjects. The



(a) Subject 1

(b) Subject 2

(c) Subject 3

Figure 3.3: Individual Channel Classification Accuracy Averaged Across all 3 Sessions for each Subject and their Respective Rank.

highest accuracy value was considered and the channel combination sequence associated with that accuracy was considered as the optimized channel set for the specific subject. Using the above process, optimized channel sets were obtained for all three subjects.

In Figure 3.4, the highest accuracy can be observed at the 20th iteration, and at this iteration, the first 20 channels among the ranked channels were deemed optimal based on their individual classification accuracies. The classification accuracy of 69.83% was

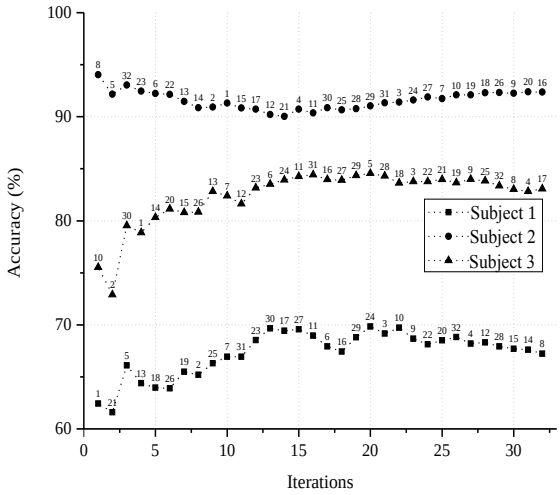
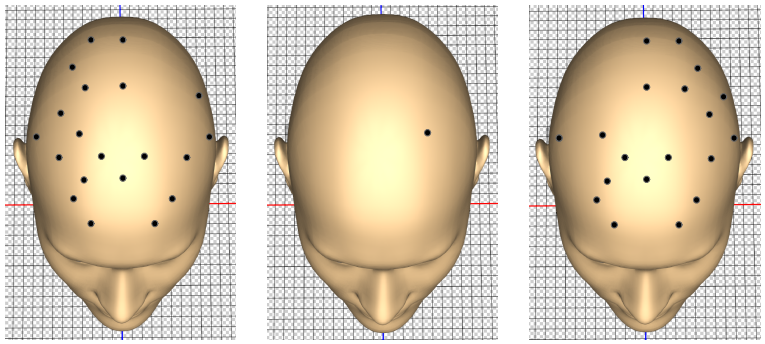


Figure 3.4: Subject Optimized Channel Accuracies over 32 Iterations



(a) Subject 1 (b) Subject 2 (c) Subject 3

Figure 3.5: Optimized electrode sets for each subject

achieved. However for Subject 2, a single electrode (C3) was deduced as optimal by the MV system to perform all three tasks and the classification accuracy was 94.04%. This was due to a significantly higher individual channel classification accuracy for that particular channel. Finally, electrodes for subject 3 were optimized to a set consisting of 14 electrodes with a classification accuracy of 84.56%. Table 3.2 summarizes the optimized classification accuracies of the three subjects. The reported overall accuracy calculated by averaging the classification accuracy of all subjects was 82.81% which is higher than the 79.96% accuracy obtained by optimizing a set of electrodes for multiusers [30].

Table 3.2: Electrode Optimization

Subject	Classification Accuracy
1	69.83%
2	94.04%
3	84.56%
Average	**82.81%**

Figures 3.5-(a,b,c) illustrate the placement of these optimized electrodes for subject 1, subject 2, and subject 3, respectively. It is obvious from these results, that each subject utilized a different set of electrodes to obtain the optimal classification accuracy. For subjects 1 and 3, both sides of the brain worked together through corpus callosum with a certain side of the brain being more dominant than the other, while subject 2 was left lateralized. Based on these results, it is evident that subject 1 has prominent activity in the right side of the brain while subject 3 has shown more activity in the left part. Therefore, it can be concluded that people with certain hand affinity have elevated activity in respective brain regions. Figures 3.5-(a,c) illustrate the left-hand dominance and right-hand dominance for subject 1 and subject 3, respectively.

3.5 Further Optimization Using Genetic Algorithm

The genetic algorithm (GA) is a computational model that seeks to find a global minimum of complex function with given parameters and a user-defined cost function. Genetic algorithm has been used in a wide array of applications. In a study genetic algorithm has been used to extract complex features from EEG for BCI application and two classifier were used to classify the data to validate the effectiveness of feature extraction [31]. Features of five mental tasks extracted in this study. In another study, epileptic seizures in EEG signals were classified using wavelet transform and genetic algorithm [32]. They acquired a high

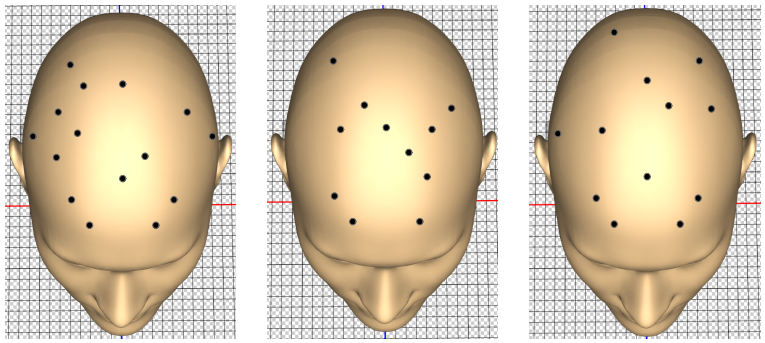
classification accuracy of 94.3% and 98% for normal and epileptic features respectively. Using genetic algorithm in optimizing EEG related data classification has always provided higher classification. To further enhance the classification accuracy of this study, the genetic algorithm was employed to the majority voting procedure. Before the process is explain in much detail, the parameters of the genetic algorithm need to be introduced.

Table 3.3: Genetic Algorithm Parameters

GA Options	Description	Values
Creation Function	Creates the initial population. In this case, a random 32-bit was created that was used as initial chromosomes.	[0,1]
Crossover Fraction	Fraction by which the population characteristics are carried to the next generation.	0.8
Crossover Function	Used to create new chromosomes from the parent chromosomes for each generation.	Default
Elite Count	Positive integer value that guarantees the amount of chromosomes survive to the next generation	0.05
Max Generations	Maximum number of generations the GA can run	30
Population Size	Size of the population	64
Mutation Function	This function adds mutations for the GA to come out of local minima	Default

The parameters in in Table 3.3, the parameters that were focused are listed along with their description. The algorithm had a creation function that created 64-bit long strings of rendom '0' and '1'. Thirty strings were created for the initial population. These strings were used to choose the electrode data to be used in the majority voting procedure. The majority-voting was used as a cost function. Based on the output accuracy of the majority vote, the chromosomes in the GA were chosen in each generation. The genetic algorithm was set to stall after 30 generations. In this manner, the genetic algorithm was used to optimize the data for all the three subjects in the dataset. From the acquired results of the output, it was noticed that the GA was able to reduce number of electrodes required for

computation by more than 50% in subject 2 and 3. The optimized electrode positions are shown in Fig 3.6 for all the three subjects.



(a) Subject 1 (b) Subject 2 (c) Subject 3

Figure 3.6: Optimized electrode sets for each subject

For subject 1, electrodes were reduced by 50% and classification accuracy was improved by an average of 2.51% across all three sessions. Electrode for subject 2, were also reduced by 65.62%. This percent was computed based on the 32 electrodes and not on the results acquired through majority voted results. Accuracy was improved by 0.218% across all the three sessions. Finally for subject 3, the electrodes were optimized to a total of 12 electrodes which is a 62.50% reduction in the number of electrodes used. The classification accuracy was improve by an average of 1.9%. The optimized electrode sets and classification accuracies for each subject can be seen in Table 3.4.

Table 3.4: Weight Factors & Overall Classification Accuracy of the High & Low Groups

	Optimized Electrode Set	GA - Optimized Accuracy (%)	Majority-Vote Accuracy (%)	Improvement (%)	Channel Reduction
Subject 1	1 2 5 7 10 13 18 19 20 21 23 24 25 29 30 31	72.35%	69.83%	2.51%	20%
Subject 2	1 4 5 8 10 18 22 23 29 30 32	94.25%	94.04%	0.218%	-9.09%
Subject 3	1 2 9 10 13 14 17 23 24 29 30 31	86.414%	84.56%	1.9%	40%
	Average	84.34%	82.81%	1.56%	

3.6 Summary

In this chapter, majority vote system was applied to a system of neural networks in order to optimally classify three imaginary tasks performed by three subjects. This proposed approach optimized the electrodes for each individual user by ranking the electrodes based on their individual classification accuracies. Averaging the electrode accuracies reduced the variance across the sessions and by implementing the majority vote the best performing electrodes are shortlisted for each user. It was observed that using a network of simple neural networks along with the majority vote system improved the classification accuracy of each subject significantly. In addition, the averaged overall classification accuracy of all three subjects increased from 79.96% to 82.81%. Using the genetic algorithm, the acquired accruacy from the majority vote was further optimized to 84.34% which is 1.56% greater. The number of electrodes used in the computation was also reduced by more than 50% with this process. In this work it can concluded that not only accuracy was improved but also the computation time was effective reduced through electrode optimization by using majority voting and genetic algorithm. By this it is evident that classification accuracy is user-dependent and hence each user has a different set of optimal electrodes. To further validate and investigate this technique, the majority voting system should be tested on a much larger EEG dataset.

CHAPTER 4

A COMPREHENSIVE STUDY OF MOTOR IMAGERY EEG-BASED CLASSIFICATION USING INDEPENDENT COMPONENT ANALYSIS AND ARTIFICIAL NEURAL NETWORKS

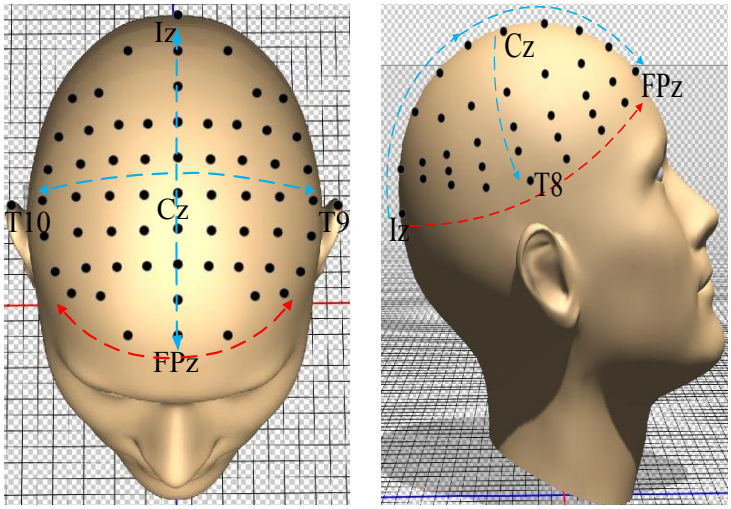
4.1 Overview

In this chapter, robust technique is proposed that combines the use of an automatic feature extracting independent component analysis (ICA) system with an ANN classifier that uses the Levenburg-Marquardt training function to classify a large scale dataset of 105 subjects. The work validates the relationship between signal to noise ratio (SNR), signal variance across multiple sessions, and signal classification accuracy. The dataset considered for this work was acquired from PhysionNet and consists of both imagined and actual movements performed by 105 subjects. In a different study, wavelet transform features were extracted from the EEG Movement/Imagery dataset and an ANN was used for classification. This study reported a maximum classification accuracy of 68.21%. A phase locking value system (PLV) was used for the same dataset to classify the β (12-30 Hz) and μ (8-12 Hz) rythms for actual movements (78.95% & 63.73%) and imagery tasks (71.55% & 65.55%). Another study using two feature selection processes (ICA and frequency band selection), classified the data using an SVM classifier with a Gaussian kernel and reported a high average accuracy of 69%. The average classification accuracy reported is 11% higher than the highest accuracy reported in all previous studies pertaining to the same dataset. However, in this work, details are presented about the data that has not been presented before and propose a robust system that automatically extracts task features using an ICA and classifies them using an ANN classifier. The remainder of this chapter is organized as follows. Section 4.2 details of the dataset. Section 4.3 presents the feature extraction technique and the ANN architecture. Section 4.5 presents the experimental results and

findings of this work. Finally, Section 4.6 concludes the chapter with a summary of the proposed classification method and findings.

4.2 Dataset Description

The EEG Motor Movement/Imagery dataset acquired from PhysioNet website consisted data obtained from 109 subjects. This data was acquired at a sampling frequency of 160 Hz. Sixty-four electrodes were arranged based on the international 10-10 standard system as shown in Figure 4.1-(a,b). Subjects performed two tasks during each recording and the recordings comprised of actual movement of a task along with its imaginary motion. The tasks performed were opening and closing of left fist, opening and closing of right fist, opening and closing of both fists, and opening and closing of both feet. These tasks were performed in 3 sessions that were recorded during the same day. Two of the four tasks were performed in each recording with intermittent resting in between each task. The tasks were performed for 4 seconds and the total duration of the entire recording was about 2 minutes. The actual motions for these tasks were first recorded then followed by the imaginary motions. The data was provided with event related annotations. However, information regarding age, gender, and handedness for the subjects was not available. For this work, the 4 tasks including the brain rest state were arranged into two separate sets (Set 1 & Set 2) as shown in Table 4.1. Set 1 contained data pertaining to imagined opening and closing left fist and opening and closing right fist respectively. Set 2 contained data pertaining to imagined opening and closing both fists and onset of motion of both feet. The focus of this work is to only classify imaginary motion, actual movements performed were ignored. Also, data from 4 subjects among the 109 subjects had data sampled at a different frequency (128 Hz) and were not considered as well. Information regarding age, gender, and handedness for the subjects was acquired from Dr. Peter Brunner at Albany Medical College, Department of Neurology. Among the 105 subjects, 99 had their gender specified



(a) Front view. (b) Side view.

Figure 4.1: Arrangement of 64 electrodes on a simulated human scalp using the 10-10 international system.

Table 4.1: Imagery Task Categorization & Subject Details for Dataset

	Task 1	Task 2	Task 3
Set 1	Rest	Left Fist	Right Fist
Set 2	Rest	Both Fists	Both Feet
	Gender	Age	Handedness
	40 Male	19-67	6 Left 34 Right
	59 Female	19-58	4 Left 55 Right
	6 Unknown	–	–

of which 40 were male and 59 were female. The age group considered was widespread and consistent among both genders. The handedness among the subjects was mostly right handed with only 10 subjects being left handed. In each session, almost half of the recording contained data pertaining to mental resting while the other half was divided between two tasks. Before using this dataset, it was first pre-processed to remove noise artifacts.

Table 4.2: Electrodes & Their Alloted Channel Numbers

Channels	Electrode	Channels	Electrode	Channels	Electrode	Channels	Electrode
1	FC_5	17	CP_1	33	F_1	49	P_3
2	FC_3	18	CP_z	34	F_z	50	P_1
3	FC_1	19	CP_2	35	F_2	51	P_z
4	FC_z	20	CP_4	36	F_4	52	P_2
5	FC_2	21	CP_6	37	F_6	53	P_4
6	FC_4	22	F_{P1}	38	F_8	54	P_6
7	FC_6	23	F_{Pz}	39	FT_7	55	P_8
8	C_5	24	F_{P2}	40	FT_8	56	PO_7
9	C_3	25	AF_7	41	T_7	57	PO_3
10	C_1	26	AF_3	42	T_8	58	PO_z
11	C_z	27	AF_z	43	T_9	59	PO_4
12	C_2	28	AF_4	44	T_{10}	60	PO_6
13	C_4	29	AF_8	45	TP_7	61	O_1
14	C_6	30	F_7	46	TP_8	62	O_z
15	CP_5	31	F_5	47	P_7	63	O_2
16	CP_3	32	F_3	48	P_5	64	I_z

4.3 Proposed Noise Extraction Model using Fixed/Variable ICA Thresholding

The data acquired from each subject was pre-processed using independent component analysis (ICA). The algorithm is closely related to the blind source separation (BSS) model and has been examined in detail in [24, 25]. In ICA, the input signals are reconstructed into statistically independent components (ICs), wherein neural source signals are highlighted and noise features such as electrooculographic (EOG), electromyography (EMG) and other such bodily interferences are supressed. Each IC needs to be visually inspected and the noise features need to be filtered manually. In a study, ICA was used to extract eye related noise artifacts using heuristic techniques that identified blink related components and contamination were validated using scalp topographies [33]. These heuristic techniques

require prior knowledge about the noise related signal patterns. This renders the use of ICA as impractical for BCI systems. To circumvent the issue of filtering the signals using a manual procedure, various autonomous techniques have been proposed to automate the detection of specific contamination within the ICA reconstructed EEG signal. One such method uses the spectral gradient of data. The gradient is calculated from the power spectral density estimate for the ICs [34]. The slope is observed on a log-log scale and a threshold parameter is used to eliminate heavily contaminated data. This technique is simple to implement and remove high muscular contamination. Another method uses the local maxima of the ICs and zeroes the entire width of artifacts that are higher than the computed threshold parameter. The method was validated across two datasets [35]. These studies perceive noise to be specific to a certain contamination and propose techniques to filter them out. This is prevalent, as characteristics of noise generated from eye and muscular movements are peculiar in signal shape. It requires experience and prior knowledge of the property of noise to recognize and filter out such artifacts. Autonomous techniques focused on filtering noise from sources, such as, eye, muscular and other related interference have not yet been effective to drastically mitigate such interference. So in this study, an autonomous extraction model is proposed that computes a dynamic threshold to compensate for noise contamination from all interference and filter them out.

4.3.1 Proposed ICA Model

The data acquired from each subject was pre-processed using independent component analysis (ICA). The algorithm is closely related to the blind source separation (BSS) model and has been examined in detail in [24, 25]. In ICA, the input signals are reconstructed into statistically independent components (ICs), wherein neural source signals are highlighted and noise features such as electrooculographic (EOG), electromyography (EMG) and other such bodily interferences are supressed. Each IC needs to be visually inspected and the

noise features need to be filtered manually. In a study, ICA was used to extract eye related noise artifacts using heuristic techniques that identified blink related components and contamination were validated using scalp topographies [33]. These heuristic techniques require prior knowledge about the noise related signal patterns. This renders the use of ICA as impractical for BCI systems. To circumvent the issue of filtering the signals using a manual procedure, various autonomous techniques have been proposed to automate the detection of specific contamination within the ICA reconstructed EEG signal. One such method uses the spectral gradient of data. The gradient is calculated from the power spectral density estimate for the ICs [34]. The slope is observed on a log-log scale and a threshold parameter is used to eliminate heavily contaminated data. This technique is simple to implement and remove high muscular contamination. Another method uses the local maxima of the ICs and zeroes the entire width of artifacts that are higher than the computed threshold parameter. The method was validated across two datasets [35]. These studies perceive noise to be specific to a certain contamination and propose techniques to filter them out. This is prevalent, as characteristics of noise generated from eye and muscular movements are peculiar in signal shape. It requires experience and prior knowledge of the property of noise to recognize and filter out such artifacts. Autonomous techniques focused on filtering noise from sources, such as, eye, muscular and other related interference have not yet been effective to drastically mitigate such interference. So in this study, an autonomous extraction model is proposed that computes a dynamic threshold to compensate for noise contamination from all interference and filter them out.

4.3.2 Proposed Autonomous ICA Model

In data, 64 electrodes positioned using the 10-10 system provide potentials attributing to specific tasks. However, these electrodes tend to record data from other regions of the brain, resulting in added interference and noise artifacts in the signal. This attributes associate

to the cocktail party problem. The ICA algorithm, with no advance knowledge regarding signal sources, finds a matrix that untangles the 64 signals into independent components which would reflect the original source data.

Let Z be the input matrix containing subject related EEG data in a [p x q] arrangement, p and q be the number of channels and the number of samples, respectively. The average of channel p, μ_p, is evaluated as follows:

$$\mu_p = \frac{1}{q}\sum_{j=1}^{q} Z_{pj} \tag{4.1}$$

Let μ be a vector of all channel averages $[\mu_1, \mu_2, ..., \mu_p]$, $J_{1,q}$ be a vector $[1, 1, ..., 1]$ of size (1xq). To zero-mean the channel data, a centered input matrix, Z^c, is generated as follows:

$$Z^c = Z - \mu_p^T J_{1,q} \tag{4.2}$$

The centered input matrix, Z^c, is used to evaluate the covariance matrix, ξ, as follows:

$$\xi = \frac{(Z^c)(Z^c)^T}{q-1} \tag{4.3}$$

The singular value decomposition (SVD) is used to decompose the covariance matrix ξ into three separate matrices as follows:

$$\xi = UDV^T \tag{4.4}$$

where U and V are orthonormal matrices that are mutually orthogonal. The columns of matrices U and V are referred to as the left and right singular values, respectively. The left and right singular values represent the eigenvectors of $\xi\xi^T$ and $\xi^T\xi$, respectively. D is a diagonal matrix of singular values $\sigma_{ij} = 0$ except if $i = j$ which results in $\sigma_{11} \geq \sigma_{22} \geq \cdots \geq 0$. The squares of $D's$ diagonal singular values (σ_{ii}^2) represent the eigenvalues of $\xi\xi^T$ and $\xi^T\xi$ which are organized in a descending order starting with the most dominant singular values.

Before applying the ICA algorithm, the centered input matrix, Z^c, is whitened or decorrelated by using the eigen-value decomposition (EVD) of the covariance matrix. The whitening process transforms the original matrix into a new matrix Z_w^c that all its components are uncorrelated and have a variance equal to 1. The whitened matrix Z_w^c is evaluated as follows:

$$Z_w^c = U\frac{1}{\sqrt{D}}U^T Z^c \tag{4.5}$$

where $\frac{1}{\sqrt{D}}$ is basically evaluated by taking square-root of the inverse of each element in the D diagonal matrix.

To apply the ICA algorithm, the Gaussian neg-entropy model is used. This model seeks to minimize the Gaussianity of the signal through maximizing the distance between a Gaussian function with an exact variance and Z_w^c. The Gaussian function is defined as follows:

$$G = kZ_w^c e^{-0.5(kZ_w^c)^2} \tag{4.6}$$

where k is the weight which is updated based on the departure measure from Gaussian entropy. Using Eq. (4.8), the weights for the model are updated as follows:

$$k = \frac{1}{q}\sum_{j=1}^{q} GZ_w^c + \left(\frac{1}{q}\sum_{j=1}^{q} kZ_w^c G\right)k_L \tag{4.7}$$

where k_L contains weight values from the previous iteration. The weight matrix k contains weights of all the input samples. The SVD is also used to decompose the weight matrix into three separate matrices as follows:

$$k = QSR^T \tag{4.8}$$

where Q and R represent the left and right singular values, respectively. S is the

diagonal matrix of singular values. After the SVD, the weight matrix k is whitened or decorrelated by using EVD as follows:

$$k_D = Q\frac{1}{\sqrt{S}}Q^T k \tag{4.9}$$

where the weight matrix k_D contains the decorrelated values. To generate a stopping criteria ε for the weights, a minimizing error function is used as follows:

$$\varepsilon = max(1 - \langle k_D, k_L \rangle) \leq \varepsilon_{th}, \quad \varepsilon \, \epsilon \, [0 \; 1] \tag{4.10}$$

The weights are updated across multiple iterations till the maximum error value is minimized to a threshold value ε_{th}. In Eq. (4.11), the weights with the minimized error is multiplied with the centered and whitened input matrix Z_w^c,

$$Z_{ICA} = k_D Z_w^c \tag{4.11}$$

where Z_{ICA} contains reconstructed independent components associated with their source EEG signals. Using the above algorithm, the source signals obtained from the electrodes were reconstructed to emphasize the neural activities of the imaginary motions and suppressed the noise features. In order to extract noise related features from Z_{ICA}, an autonomous threshold algorithm was used where the local maxima for IC signals from Z_{ICA} were recorded. Noise within data manifests as voltage spikes created through linear combinations of imaginary motion data with unwanted eye, muscle and or other noise related artifacts. Using the recorded local threshold values, the highest and the lowest maxima for each IC are recognized. These maximum and minimum values are averaged across all the ICs and are used as limits to compute the threshold parameter. Using Eq. (4.12), threshold J^{th} is calculated as follows:

$$J^{th} = \hat{h} - \eta(\hat{h} - \hat{l}) \quad \eta \, \epsilon \, [0 \; 1] \tag{4.12}$$

where $\hat{h}$ and $\hat{l}$ are the averages of the highest and lowest local maxima across all ICs. Between these two values, the threshold parameter is controlled using η. After a thorough investigation, it was determined that η=0.5 was an optimal value for the threshold to effectively filter noise features. The computed threshold value varied for each subject across ICs. However, this threshold model does not account for fluctuations in the noise contamination across individual ICs. Therefore, a novel autonomous threshold parameter model is proposed. This model adds a variable scaling factor ρ_k to increase the threshold versatility. The autonomous threshold J_k^{th} is calculated in Eq. 4.13 as follows:

$$J_k^{th} = \hat{h} - \rho_k \eta(\hat{h} - \hat{l}) \quad \eta \, \epsilon \, [0 \; 1] \tag{4.13}$$

where ρ_k is a variable scaling factor for the k^th IC which is a ratio between average of all samples of the k^th IC and the average of all samples from all the ICs. This parameter indicates the degree of contamination within a specific IC in comparison with the overall contamination across all other ICs. The ratio for ρ_k is defined as follows:

$$\rho_k = N \frac{\sum_{j=1}^{n} u_{kj}}{\sum_{i=1}^{N} \left(\sum_{j=1}^{n} u_{ij} \right)} \tag{4.14}$$

where u_{kj} is the k^{th} IC samples while u_{ij} is the i^{th} IC samples. N and n represent the number of ICs and the number of samples in each IC, respectively. The noise features were extracted using both thresholds (J^{th} & J_k^{th}) and the data for both the processes were recorded for further analysis. The EEG data, using both thresholding (fixed and variable) techniques was filtered for noise features and the results acquired from both techniques were saved for comparison. Filtered data from both processes was compared and validated using classification accuracies from an artificial neural network. Classification accuracy was considered as the main criteria of improvement. The thresholding process for both

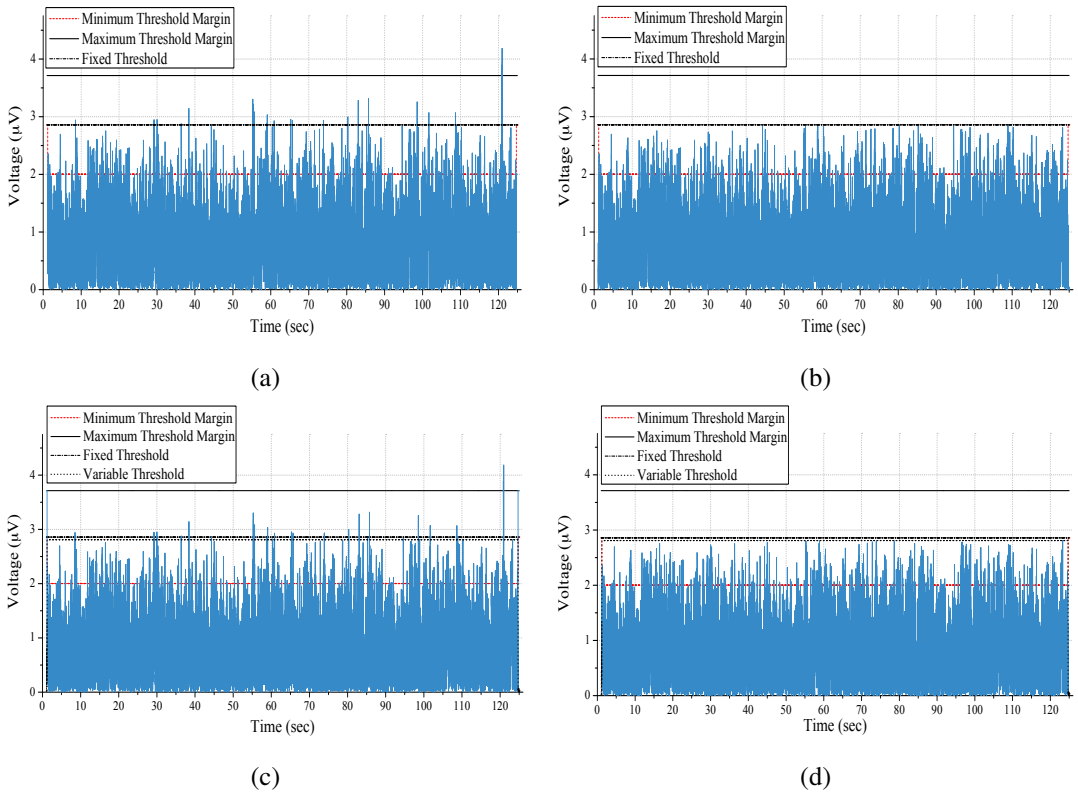


Figure 4.2: Autonomous noise artifact filtering for EEG signals a) Threshold is computed to be 2.85μV for data from electrode at P_5. b) Signal above the threshold J^{th} is zeroed c) Threshold J_k^{th} is computed to be 2.80μV. d) All values above the threshold are zeroed.

techniques is shown in Figure 4.2, for single IC channel named P_5 which was chosen from one of the three sessions for a single subject. For this signal, the threshold was computed to be 2.85μV. In Figure 4.2-b, the voltage spikes in the EEG signal exceeding the threshold value were removed by zeroing them out. This process was implemented for all the ICs across 105 subjects. However, due to threshold values being constant across all the ICs for a subject, the threshold was not able to compensate for variations in noise contamination in some ICs. Few ICs contained high concentrations of noise while few had very low amounts of visually noticeable noise fluctuations and using a fixed threshold was not adequate. To mitigate the issue, a variable threshold was implemented that changed in accordance with

the noise level for each of the ICs. This was realized by adding another tuning factor ρ_k to change Eq. (4.12) and make the threshold value tunable. In Figure 4.2-c, the same EEG channel named P_5 is plotted with the new threshold at 2.80μV and values above J_k^{th} were zeroed as shown in Figure 4.2-d. The EEG data, using both thresholding (fixed and variable) techniques was filtered for noise features and the results acquired from both techniques were saved for comparison. Filtered data from both processes was compared and validated using classification accuracies from an artificial neural network. Classification accuracy was considered as the main criteria of improvement.

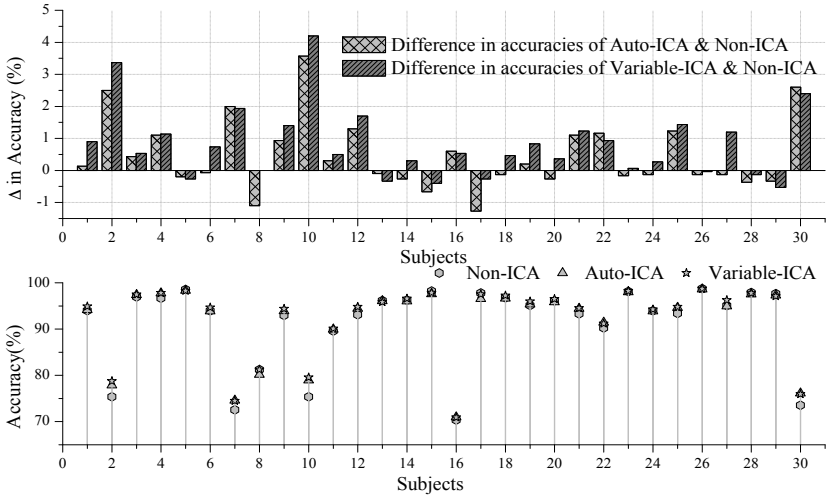
4.3.3 Preliminary Results

Among the 105 subjects, 15 males and 15 females were randomly chosen for analysis. The chosen male subjects varied in age from 19 to 67 years and females subjects varied in age from 22 to 58 years respectively. The handedness for the subject sample size comprised of 6 left handed and 24 right handed subjects. In this study, the subjects that did not have any genders specified (6 unknowns) were not considered in the random selection process.

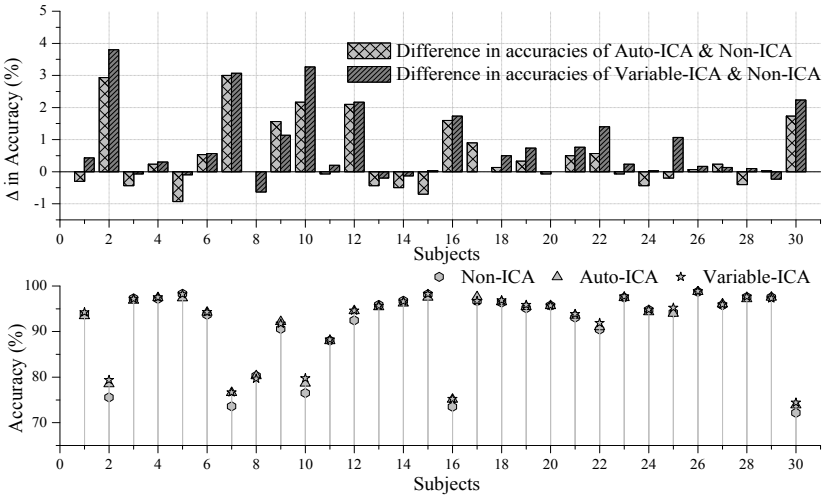
For these selected subjects, the ICA algorithm along with the autonomous noise extracting model were used to acquire ICs with removed noise artifacts. These ICs are input into a neural network structure for classification. The classification was performed using a three layered neural network containing input, hidden and output layers. The input layer consisted 64 neurons for 64 inputs. The hidden layer for the system consisted of 10 neurons and used a tan-sigmoid transfer function. Finally, the output layer was assigned one neuron and used a linear transfer function. Half of the data was used for training while the other half was used for validation and testing. In this neural network, Levenburg-Marquardt (LM) optimization was used as training function. The reasons behind choosing the LM optimization will expliained in the following subsection.

Subjects performed six imagery tasks and the data was recorded over three sessions

within the time span of a single day. The six imagery tasks were categorized into two sets and Eq. (4.10-4.11) were used separately to extract noise artifacts from data. This validation process seeks to highlight differences between the two equations wherein the process involves the use of a fixed threshold technique called fixed threshold and an adaptive threshold technique called variable threshold technique. The fixed threshold technique calculates the threshold using local maxima from all ICs for a subject. In the technique, the threshold remains unchanged when removing noise artifacts. The variable threshold technique considers the threshold acquired using the fixed threshold technique and adjusts it based the noise variations in the ICs. In this technique, every IC uses a distinct threshold to filter noise variations and results in better noise artifact removal. Data from both sets were input into the ANN and their classification accuracies were saved and averaged among three sessions. Accuracies from the ANN were noticed to be significantly higher than reported in previous studies that used the same dataset. Analysis in this work reports an 11.79% improvement in classification accuracy from latest results reported in [36]. Even without using autonomous noise extraction systems with ICA, high classification accuracies were achieved. Baseline overall accuracies of 91.30% and 91.24% for data from Set 1 and Set 2 were acquired. To improve this accuracy further, autonomous noise feature removal techniques were used. Accuracies obtained using fixed threshold data were subtracted from the baseline accuracies to quantify the improvement. The same process was repeated for data acquired using the variable threshold. In Figure 4.3-(a,b), accuracies for Set 1 and Set 2 were acquired using both noise extraction techniques are plotted with baseline accuracy values. The improvements in classification per-subject are highlighted in the upper graphs of the figure. Using the fixed threshold, an overall improvement compared to baseline classification was observed to be 0.46% and 0.47% for Set 1 and Set 2 respectively. Similarly using variable threshold technique for Set 1 and Set 2, a slightly higher improvements were observed at 0.82% and 0.76% respectively.



(a)



(b)

Figure 4.3: Difference in ANN accuracy for fixed threshold and variable threshold techniques. a) Classification results of data from Set 1 for 30 subjects. b) Classification results of data from Set 2 for 30 subjects.

In Table 4.3, subject related information is highlighted along with overall classification accuracies of Set 1 and Set 2 data. The accuracies are reported for data acquired using no threshold, fixed threshold, and variable threshold. There was a slight improvement in classification when using both autonomous noise extracting features. For the 30 subjects,

Table 4.3: Classification Results of the Neural Network

Male	Female	Male (Age)	Female (Age)	Handedness	Method	Set	Task 1 (avg)	Task 2 (avg)	Task 3 (avg)	Overall (avg)	Variance (σ^2)	Δ w.r.t No Threshold
					No Threshold	Set 1	95.01%	87.39%	87.63%	91.30%	2.51	–
						Set 2	94.90%	87.07%	87.87%	91.24%	2.16	–
15	15	19-67	22-58	6 Left, 24 Right	Fixed Threshold	Set 1	95.38%	88.07%	87.99%	91.76%	2.07	0.46%
						Set 2	95.23%	87.81%	88.42%	91.71%	1.86	0.47%
					Variable Threshold	Set 1	95.52%	88.65%	88.60%	92.11%	2.32	0.81%
						Set 2	95.36%	88.19%	88.93%	91.99%	1.49	0.75%

the average accuracy for each specific single task, the overall average accuracies, and the variance of the overall accuracies for each set using all three methods no threshold, fixed threshold, and variable threshold are listed in Table 4.5. The variances listed in table are the averaged variances of accuracies across the 3 sessions for all the subjects. Also, the average difference in the accuracy using the fixed threshold and variable threshold technique is presented. Across both sets, differences were relatively higher and mostly positive for the variable threshold data.

4.4 Various Training Functions and ANN Architecture

ANN has been used to effectively classify real world applications concerning biological data [37, 38]. Figure 4.4 illustrates a typical artificial neural network structure. ANN models are used to learn and classify mental abnormalities, speech and face patterns [39]. Neural networks are used to model non-linear data and tends to provide predictions or classification outputs with reduced error rates. However, ANNs have their own challenges. Prior to implementation, numerous parameters require tuning and the networks are prone to over-fitting problems. Also, neural networks are computationally expensive. Despite all this, ANN models show great potential especially when efficient training algorithms are used. Among the different types of training functions for ANN classifiers, some high performing algorithms were identified. These algorithms are classified into two categories. The first

utilizes heuristic techniques and the other uses numerical optimization techniques [28]. In this study, one heuristic technique and seven numerical optimization techniques were investigated during initial experimentation. The classifiers were tested using data from a session recorded for a single subject performing three tasks. The accuracy, time, mean squared error (MSE), gradient factor, and the number of epochs were observed and listed in Table 4.4. The Polak-Ribiere conjugate gradient, BFGS Quasi-Newton, and the

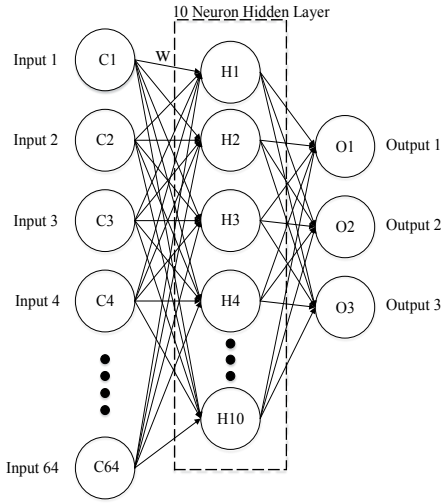


Figure 4.4: ANN Architecture

Scaled conjugate gradient classifiers provided high classification accuracies. However, their training functions resulted in longer training period and required high number of epochs to converge. The variable learning rate technique using the heuristic technique provided the lowest classification accuracy (60.41%) while still converging relatively fast. In addition, its mean-squared error (MSE) and gradient descent were also observed to be the highest. Finally, the Levenberg-Marquardt (LM) within the ANN classifier provided the highest classification accuracy with a relatively fast classification time of 117 seconds and was also able to reduce the MSE error significantly. The LM-ANN classifier uses the Levenberg-Marquardt optimization to update weights and biases for the neural network. The optimization minimizes the least square problem which is shown in Eq. (4.15). In this

Table 4.4: Back-Propagation Perceptron Classifiers and Classification Information.

Training functions used in the neural network	Accuracy (%)	Time (sec)	Minimized MSE	Gradient	Epochs (max 1000)
Conjugate Gradient with Beale-Powell Restarts	80.61	60	0.115	0.009	175
BFGS Quasi-Newton	88.43	141	0.078	0.005	480
Resilient Back-propagation	80.20	306	0.115	0.01	1000
Scaled Conjugate Gradient	88.39	191	0.076	0.004	617
Fletcher-Powell Conjugate Gradient	80.32	85	0.109	0.012	248
Polak-Ribiere Conjugate Gradient	88.92	246	0.079	0.008	697
One Step Secant	75.45	50	0.130	0.023	146
Variable Learning Rate Back-propagation	60.41	52	0.173	0.116	166
Levenberg Marquardt	93.46	117	0.057	0.014	59

work, Y is considered to be the least squared error and is defined as:

$$Y = \sum_{i=1}^{n} (y_i - f(x_i, \alpha))^2 \tag{4.15}$$

where n is number of data points, x_i is the input value and y_i is output value and $f(x_i, \alpha)$ is the prediction function. The LM algorithm optimizes the weights to minimize Y and has been previously used in research studies for signal classification. The LM based neural network classifier was used to classify thumb, index and middle finger movements which was used to control a bionic arm very effectively [40]. Another study used the LM-ANN in combination with the genetic algorithm to optimize weights to improve classification accuracy and classification time for a quicker detection of hypoglycemia [41]. In this study, LM-ANN is used to classify data from 105 subjects that had performed six different tasks. The neural network architecture used in this study composed of 64 neurons assigned to the input layer. Ten neurons were assigned to the hidden layer that used a tan-sigmoid transfer function. Three neurons were assigned to the output layer that used a linear transfer function. 50% of the data was used for training, 25% was used for validation, and the remainder 25% was used for testing purposes.

4.5 Data Classification Results & Analysis

After preliminary results were studied closely, the threshold techniques were tested for all 105 subjects. To test these techniques, data were first filtered using both thresholds and then classified using the ANN. The three recording sessions were classified for Set 1 and Set 2 imaginary motions. The resulting classification accuracies for these sessions were averaged for each user. Variance of the classification accuracies across these sessions were also recorded. To visualize the data classification improvement using the proposed technique, the classification accuracies of the dataset without using a threshold (baseline), with using a fixed threshold, and with using an variable threshold were plotted in Figures 4.5-(a,b) for both sets. These figures indicate that, in general, using a fixed threshold improves the classification accuracy compared to not using a threshold. However, using the variable threshold provided the highest classification accuracy. Figures 4.5-(c,d) illustrate the improvements of the classification accuracies of both fixed threshold and variable threshold compared to not using any threshold or any noise extraction technique. To quantify the overall improvement due to pre-processing (de-noising) the independent components using threshold techniques, the overall mean of the sessions' averaged accuracies of both noise filtering techniques were computed. An overall improvement of 0.43% and 0.42% was observed for Set 1 and Set 2 pre-processed using fixed threshold. The accuracies were further improved using the variable threshold with an overall improvement of 0.75% and 0.65% for Set 1 and Set 2 respectively. Accuracy for each specific task was calculated using the ratio number of samples accurately classified and the total number of samples in for the task. In this study this ratio is referred to as the weight factor. By multiplying these weight factors with their respective task-specific accuracies and summing the product, the overall weighted classification accuracy is acquired.

Task accuracies, weight factors, and the overall classification accuracies using both (fixed threshold & variable threshold) techniques along with results for no threshold data

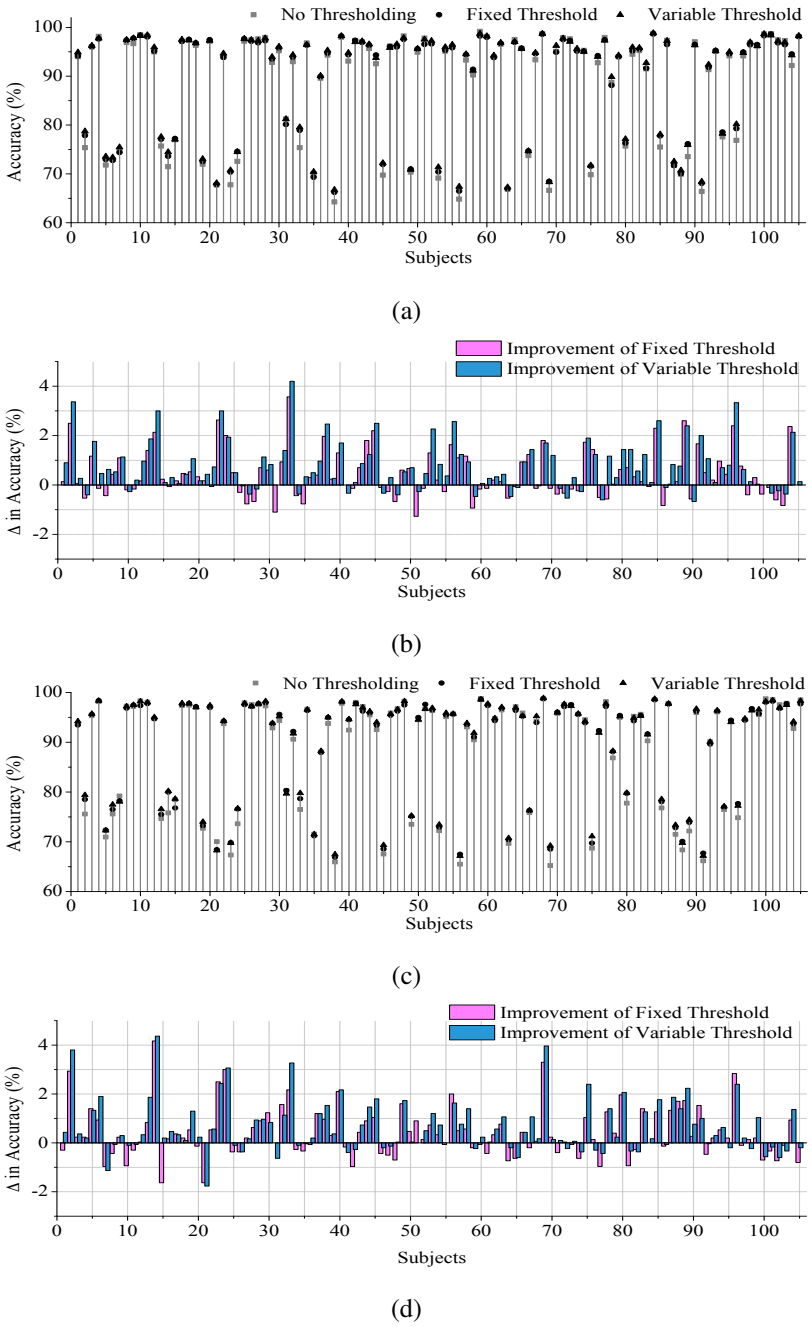


Figure 4.5: Comparative performance analysis of no, fixed, and variable threshold techniques, a) Classification accuracies of Set 1 tasks, b) Improvement of classification accuracies of Set 1 tasks, c) Classification accuracies of Set 2 tasks, d) Improvement of classification accuracies of Set 2 tasks.

are presented in Table 4.5. In addition to the classification accuracies, the variance of the classification accuracies were quantified. It was noticed that classifications with low accuracies resulted in higher variance across their different sessions compared to classifications with high accuracies. This relation between high variance and low classification accuracy was further explored for data acquired using the variable threshold technique as illustrated in Figure 4.6.

Figures 4.6-(a,b) highlight the classification accuracies variance of Set 1 and Set 2 for all the 105 subjects using the LM-ANN classifier with variable threshold. The corresponding classification accuracies of Set 1 and Set 2 for all the 105 subjects are presented in Figures 4.6-(c,d). It can be depicted from these figures that the low classification accuracies, in general, have a corresponding high variance. Among all the 105 subjects, 31 are identified to have a low classification accuracy with a corresponding high variance. Therefore, the subjects were categorized into two groups, the low and high accuracy groups. The low accuracy group had 31 subjects while the rest were categorized as the high accuracy group.

Table 4.5: Task-Specific Classification Results for Fixed Threshold & Variable Threshold

Thresholding Method	Set	Task 1 (avg)	Weight Fac 1	Task 2 (avg)	Weight Fac 2	Task 3 (avg)	Weight Fac 3	Overall (avg)	Variance (σ^2)	Δ w.r.t No Threshold
Non	Set 1	94.03%	0.502	83.18%	0.251	83.42%	0.246	88.73%	3.03	–
	Set 2	94.02%	0.502	83.26%	0.249	83.87%	0.249	88.85%	3.71	–
Fixed	Set 1	94.26%	0.502	84.13%	0.251	83.70%	0.246	89.16%	2.95	0.43%
	Set 2	94.27%	0.502	83.94%	0.249	84.35%	0.249	89.27%	3.33	0.42%
Variable	Set 1	94.49%	0.502	84.62%	0.251	84.08%	0.246	89.48%	2.73	0.75%
	Set 2	94.44%	0.502	84.21%	0.249	84.69%	0.249	89.50%	3.15	0.65%

In Table 4.6, the breakdown of the accuracies based on the task performed for the two subject groups are presented along with gender, age, and hand dominance information. For the high accuracy group, Task 1 was classified at an average overall accuracy of 97.54% and

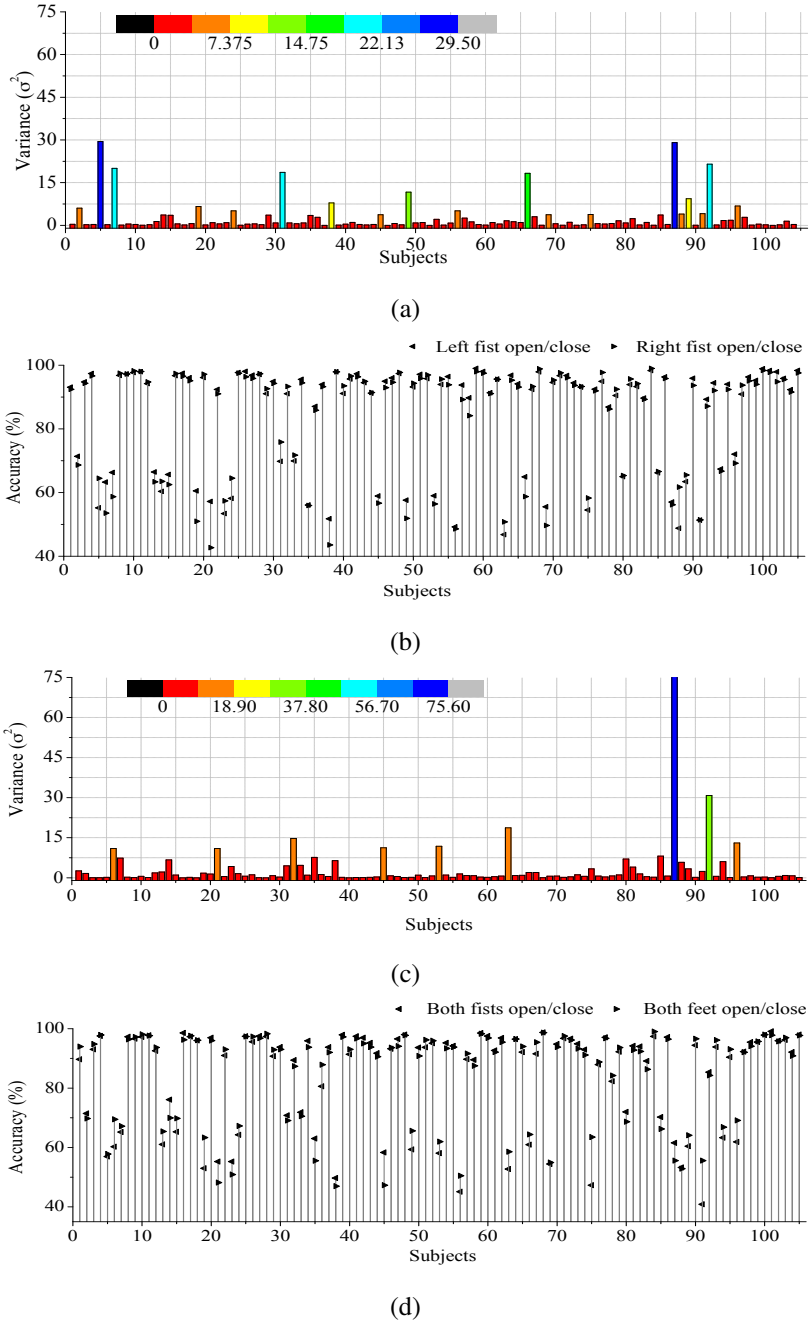


Figure 4.6: Imaginary tasks classification of 105 subjects using the LM-ANN classifier using variable threshold, a) Variance of classification accuracies of Set 1 tasks, b) Classification accuracies of Set 1 tasks, c) Variance of classification accuracies of Set 2 task, d) Classification accuracies of Set 2 tasks.

97.41% for Set 1 and Set 2 respectively. Task 2 and Task 3 from both sets were classified with accuracies greater than 94%. For the low accuracy group, Task 1 was classified at an average overall accuracy of 87.19% and 87.34% for Set 1 and Set 2 respectively. Task 2 and Task 3 were classified at a relatively lower classification that didn't exceed 62%. Task 1, which represented 50% of entire dataset (brain rest state), was classified with relatively high accuracy in both sets. However, Tasks 2 and 3 accuracies were more than 30% lower in the low accuracy group compared to the high accuracy group.

Since the tasks are not equally distributed, weight factors are used to evaluate the overall accuracy as listed in Table 4.7. In the table, the overall accuracies along with their corresponding variances are also specified for the two groups. The overall accuracy for the high accuracy group was above 95% and classification for the lower accuracy group was below 75%. Also from Table 4.7, it evident that the variance is in fact higher for data classified at a lower accuracy.

Table 4.6: Categorized High & Low Groups and their Averaged Classification Accuracies for Performed Tasks

Groups	Male	Female	Unknown (Gender)	Male (Age)	Female (Age)	Handedness	Set	Task 1 (avg)	Task 2 (avg)	Task 3 (avg)
High Accuracy	27	41	6	19-67	22-58	6 Left 62 Right	Set 1 Set 2	97.54% 97.41%	94.89% 94.36%	94.56% 94.40%
Low Accuracy	13	18	–	20-59	20-51	4 Left 27 Right	Set 1 Set 2	87.19% 87.34%	60.11% 59.98%	59.07% 61.51%

For the 31 subjects that were classified at a lower accuracy, certain factors were identified to be potential leading causes for such low classification accuracy. The source of these signals predominantly corresponds to the motor cortex regions and the somatosensory cortex regions in the parietal lobe. These regions are responsible for processing and coordinating motor movements [42]. Low accuracies are usually attributed to low frequency

Table 4.7: Weight Factors & Overall Classification Accuracy of the High & Low Groups

Groups	Set	Weight factor 1	Weight factor 2	Weight factor 3	Overall (avg)	Variance (σ^2)
High Accuracy	Set 1	0.502	0.252	0.245	96.16%	0.93
	Set 2	0.502	0.249	0.249	95.92%	1.19
Low Accuracy	Set 1	0.501	0.250	0.249	73.54%	7.03
	Set 2	0.501	0.250	0.248	74.18%	7.82

noise artifacts below 9 Hz caused by impedance variations due to movement of electrodes specifically near the parietal region [43]. Therefore, to clarify that the underlying issue was indeed caused due to unanticipated interference, the signal-to-noise ratio (SNR) for each of the channels was calculated using an estimated power spectral density (PSD) created using a Kaiser windowing technique. The windowing technique uses a modified Bessel function of the first kind $\zeta_v(x)$ that is defined as follows:

$$\zeta_v(x) = \left(\frac{1}{2}x\right)^v \sum_{k=1}^{q} \frac{(\frac{1}{4}x^2)^k}{k!\Gamma(v+k+1)} \tag{4.16}$$

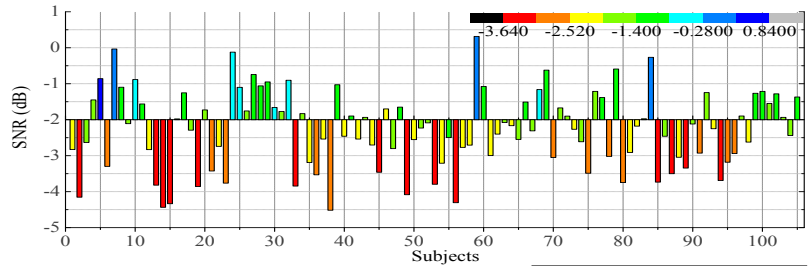
where q is the number of samples, v is the degree of the function, $\Gamma(x)$ is the gamma function. For the zeroth order Bessel function, $\zeta_0(x)$, Eq. (4.16) is modified as follows:

$$\zeta_0(x) = \sum_{k=1}^{q} \frac{(\frac{1}{4}x^2)^k}{k!} \tag{4.17}$$

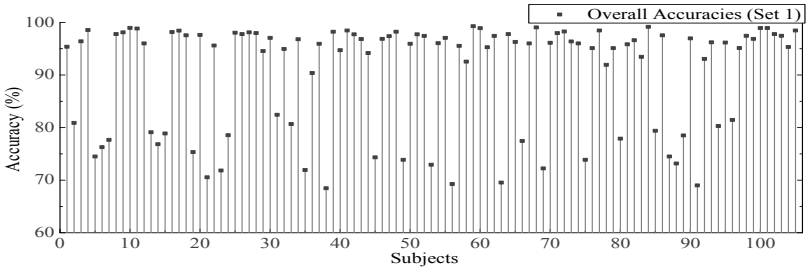
This zeroth order Bessel function is used in the Kaiser windowing technique. The Kaiser window, $w[n]$, is defined as follows:

$$w[n] = \begin{cases} \dfrac{\zeta_0\left(\pi\beta\sqrt{1-\left(\frac{2n}{N-1}-1\right)^2}\right)}{\zeta_0(\pi\beta)}, & 0 < n \leq N-1 \\ 0, & \text{otherwise.} \end{cases} \tag{4.18}$$

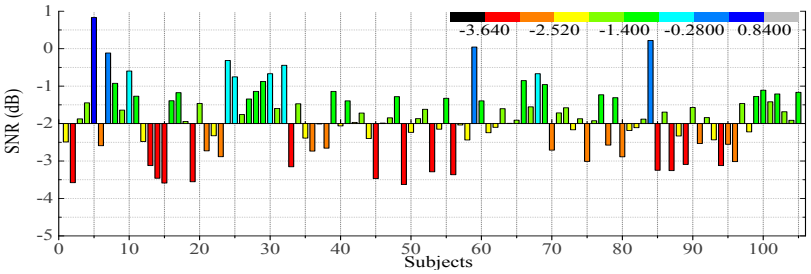
where N is the window size, β is a non negative real number that determines the shape of the window.



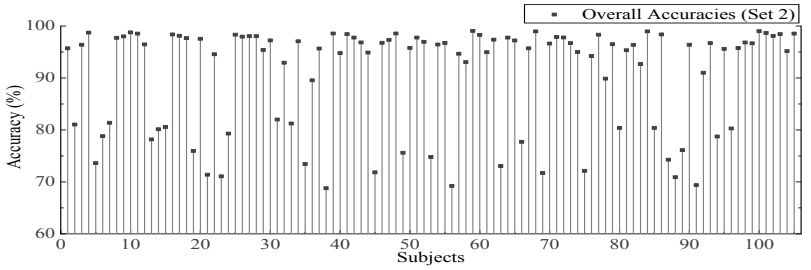
(a)



(b)



(c)



(d)

Figure 4.7: Signal-to-noise ratio and overall classification accuracy from 105 subjects where, a) SNR values are plotted for data pertaining to Set 1 tasks, b) Classification accuracies of Set 1 tasks, c) SNR values are plotted for data pertaining to Set 2 tasks, d) Classification accuracies of Set 2 tasks.

To calculate the SNR of EEG signals, a modified periodogram was generated using the Kaiser windowing function in Eq. (4.18) with $\beta = 38$. The generated periodogram was used to compute the SNR values. From the periodogram, the power of the fundamental frequency represented the desired signal while the rest of the periodogram excluding the first 6 harmonics represented the undesired noise. These SNR values were averaged across all the 64 channels for each subject. The average SNR values were observed to be consistent across the 3 sessions for both Set 1 and Set 2. The subjects within the low classification accuracy group, obtained SNR values that are lower compared to the high classification accuracy group. The average SNR of both the low and high classification accuracy groups were -2.81dB and -1.8dB respectively. In a recent study, lower classifications of EEG signals were associated with lower SNR and high variance across the channel sets [44]. The SNR values were compared with the averaged overall accuracies for all the 105 subjects. In Figure 4.7-(a,b), the overall classification accuracies for all the subjects are compared to their respective SNR values in both sets. From this figure, it is noticed that the SNR values are higher for subjects with high data classification, which validates the increase in the noise variance among the low accuracy groups.

The findings establish that high variance and low SNR can be used to gauge the noise level in EEG signals, which affects the classification accuracies. However, without prior knowledge, identifying the true source of these noise features would be extremely difficult. Therefore, the speculated sources of noise are categorized as noise artifacts created due to movement of body, uncertainty in electrode placement, facial and ocular movements, tense body posture, anxieties, etc. Nevertheless, these circumstances resemble real-world settings and the findings of this study would greatly assist in providing direction to future studies.

4.6 Summary

In this paper, using two automated feature extraction procedures, the data was extracted. With artificial neural networks, data acquired from each extraction procedure was classified. The classification accuracies for each subject was compared for both extraction techniques and it was found that variable threshold technique was able to compensate more effectively to changing noise concentrations within the EEG signal. This resulted in further increase in classification by an average of 0.75% and 0.65% for data in both sets. It was also noticed that the average variance across the data decreased with increase in classification accuracy. The next step involved focusing on the variance in data and establishing a correlation between classification and variance. The variance in classification accuracies among three sessions was plotted for each subject in comparison to their respective classification accuracies. Users with higher accuracies had an average variance of 0.93 and 1.19 for both sets respectively. Users with lower accuracies had an average variance of 7.03 and 7.82 for both sets. So users with low classification had a relatively higher variance, which confirmed that users with poor classification had a higher variance. Reasons for poor classification among users were still unclear and further analysis was needed. The signal to noise ratio was computed for each subject using the Kaiser windowing technique and the SNR results showed a higher signal to noise ratio of -6.64dB for subjects with high classification accuracies. Users with low accuracies were observed to lower SNR of -8.97dB. This procedure clearly validated a wide presence of high noise levels in EEG signals pertaining to specific subjects. These findings provide a good understanding about poor classification, high variance and low SNR. Speculating a few reasons for high noise interference in specific user data, this work seeks to motivate more research to find reasons for such noise.

CHAPTER 5

CONCLUSION & FUTURE WORK

The majority voting system was applied to a system of neural networks in order to optimally classify three imaginary tasks performed by three subjects. This proposed approach optimized the electrodes for each individual user by ranking the electrodes based on their individual classification accuracies. Averaging the electrode accuracies reduced the variance across the sessions and by implementing the majority vote the best performing electrodes were shortlisted for each user. It was observed that using a network of simple neural networks along with the majority vote system improved the classification accuracy of each subject significantly. In addition, the averaged overall classification accuracy of all three subjects increased from 79.96% to 82.81%. Using the genetic algorithm, the acquired accruacy from the majority vote was further optimized to 84.34% which is 1.56% greater. The number of electrodes used in the computation was also reduced by more than 50% with this process. Therefore, it is concluded that classification accuracy is user-dependent and hence each user has a different set of optimal electrodes. To further validate and investigate this technique, the majority voting system should be tested on a much larger EEG dataset. In the following work, using two automated feature extraction procedures, large scale data was extracted. With artificial neural networks, data acquired from each extraction procedure was classified. The classification accuracies for each subject was compared for both extraction techniques and it was found that Variable-ICA technique was able to compensate more effectively to changing noise concentrations within the EEG signal. This resulted in further increase in classification by an average of 0.75% and 0.65% for data in both sets. It was also noticed that the average variance across the data decreased with increase in classification accuracy. The next step involved focusing on the variance in data and establishing a correlation between classification and variance. The variance in classification accuracies among three sessions was plotted for each subject in comparison to their respective clas-

sification accuracies. Users with higher accuracies had an average variance of 0.93 and 1.19 for both sets respectively. Users with lower accuracies had an average variance of 7.03 and 7.82 for both sets. So users with low classification had a relatively higher variance, which confirmed that users with poor classification had a higher variance. Reasons for poor classification among users were still unclear and further analysis was needed. The signal to noise ratio was computed for each subject using the Kaiser windowing technique and the SNR results showed a higher signal to noise ratio of -6.64dB for subjects with high classification accuracies. Users with low accuracies were observed to lower SNR of -8.97dB. This procedure clearly validated a wide presence of high noise levels in EEG signals pertaining to specific subjects. These findings provide a good understanding about poor classification, high variance and low SNR. Speculating a few reasons for high noise interference in specific user data, this work seeks to motivate more research to find reasons for such noise. In future work, the genetic algorithm will be used to optimize the electrodes of the ANN results. This analysis is underway and the research will be updated with results in the neat future.

REFERENCES

[1] J. d. R. Millán, F. Galán, D. Vanhooydonck, E. Lew, J. Philips, and M. Nuttin, "Asynchronous Non-Invasive Brain-Actuated Control of An Intelligent Wheelchair," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2009, pp. 3361–3364.

[2] G. Pires, M. Castelo-Branco, and U. Nunes, "Visual P300-Based BCI to Steer a Wheelchair: A Bayesian Approach," in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2008, pp. 658–661.

[3] "Self-Paced (Asynchronous) BCI Control of a Wheelchair i."

[4] C. for Disease Control and Prevention, "Summary Health Statistics: National Health Interview Survey, 2014 Table A-10a." [Online]. Available: http://ftp.cdc.gov/pub/health_statistics/nchs/nhis/shs/2014_shs_table_a-10.pdf

[5] E. Maiorana, D. La Rocca, and P. Campisi, "On the Permanence of EEG Signals for Biometric Recognition," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 1, pp. 163–175, 2016.

[6] J. R. Millan, F. Renkens, J. Mouriño, and W. Gerstner, "Noninvasive Brain-Actuated Control of A Mobile Robot By Human EEG," *IEEE Transactions on biomedical Engineering*, vol. 51, no. 6, pp. 1026–1033, 2004.

[7] K. Tanaka, K. Matsunaga, and H. O. Wang, "Electroencephalogram-Based Control of An Electric Wheelchair," *IEEE transactions on robotics*, vol. 21, no. 4, pp. 762–766, 2005.

[8] D. P. Subha, P. K. Joseph, R. Acharya, and C. M. Lim, "EEG Signal Analysis: A Survey," *Journal of medical systems*, vol. 34, no. 2, pp. 195–212, 2010.

[9] S. M. Grigorescu, T. Lüth, C. Fragkopoulos, M. Cyriacks, and A. Gräser, "A BCI-Controlled Robotic Assistant for Quadriplegic People in Domestic and Professional Life," *Robotica*, vol. 30, no. 03, pp. 419–431, 2012.

[10] L. Bi, X.-A. Fan, and Y. Liu, "EEG-Based Brain-Controlled Mobile Robots: A Survey," *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 2, pp. 161–176, 2013.

[11] S. Siuly and Y. Li, "Improving the Separability of Motor Imagery EEG Signals Using A Cross Correlation-Based Least Square Support Vector Machine for Brain–Computer Interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 4, pp. 526–538, 2012.

[12] R. Chai, S. H. Ling, G. P. Hunter, Y. Tran, and H. T. Nguyen, "Brain–Computer Interface Classifier for Wheelchair Commands Using Neural Network With Fuzzy Particle Swarm Optimization," *IEEE journal of biomedical and health informatics*, vol. 18, no. 5, pp. 1614–1624, 2014.

[13] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A General-Purpose Brain-Computer Interface (BCI) System," *IEEE Transactions on biomedical engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[14] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, Physiotoolkit, and Physionet Components of A New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[15] M. Tolić and F. Jović, "Classification of Wavelet Transformed EEG Signals with Neural Network for Imagined Mental and Motor Tasks," *Kineziologija*, vol. 45, no. 1, pp. 130–138, 2013.

[16] A. Loboda, A. Margineanu, G. Rotariu, and A. M. Lazar, "Discrimination of EEG-Based Motor Imagery Tasks by Means of A Simple Phase Information Method," *International Journal of Advanced Research in Artificial Intelligence (IJARAI)*, vol. 3, no. 10, 2014.

[17] J. Sleight, P. Pillai, and S. Mohan, "Classification of Executed and Imagined Motor Movement EEG Signals," *Ann Arbor: University of Michigan*, pp. 1–10, 2009.

[18] F. H. Netter, *Atlas of Human Anatomy, Professional Edition: including NetterReference. com Access with Full Downloadable Image Bank*. Elsevier Health Sciences, 2014.

[19] T. Lombronzo, "The truth about the left brain / right brain ... - npr." [Online]. Available: https://tinyurl.com/ozfmp2v

[20] S. Standring, *Gray's Anatomy: The Anatomical Basis of Clinical Practice*. Elsevier Health Sciences, 2015.

[21] A. Schlegel, P. J. Kohler, S. V. Fogelson, P. Alexander, D. Konuthula, and P. U. Tse, "Network Structure and Dynamics of The Mental Workspace," *Proceedings of the National Academy of Sciences*, vol. 110, no. 40, pp. 16 277–16 282, 2013.

[22] V. L. Towle, J. Bolaños, D. Suarez, K. Tan, R. Grzeszczuk, D. N. Levin, R. Cakmur, S. A. Frank, and J.-P. Spire, "The Spatial Location of EEG Electrodes: Locating The Best-Fitting Sphere Relative to Cortical Anatomy," *Electroencephalography and clinical neurophysiology*, vol. 86, no. 1, pp. 1–6, 1993.

[23] F. Mormann, K. Lehnertz, P. David, and C. E. Elger, "Mean Phase Coherence as A Measure for Phase Synchronization and its Application to The EEG of Epilepsy Patients," *Physica D: Nonlinear Phenomena*, vol. 144, no. 3, pp. 358–369, 2000.

[24] A. Hyvarinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis," *IEEE transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

[25] A. Hyvärinen and E. Oja, "Independent Component Analysis: Algorithms and Applications," *Neural networks*, vol. 13, no. 4, pp. 411–430, 2000.

[26] W. S. McCulloch and W. Pitts, "A Logical Calculus of The Ideas Immanent in Nervous Activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[27] T. Kohonen, "An Introduction to Neural Computing," *Neural networks*, vol. 1, no. 1, pp. 3–16, 1988.

[28] M. H. Beale, M. T. Hagan, and H. B. Demuth, "Neural Network ToolboxâĎć UserâĂŹs Guide," in *R2016a, The MathWorks, Inc., 3 Apple Hill Drive Natick, MA 01760-2098,, www. mathworks. com*. Citeseer, 2016.

[29] J. R. Millan, "On The Need for On-Line Learning in Brain-Computer Interfaces," in *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, vol. 4. IEEE, 2004, pp. 2877–2882.

[30] S. Bhattacharya, R. Haddad, and M. Ahad, "A Multiuser EEG Based Imaginary Motion Classification Using Neural Networks," in *IEEE-Southeast Conf 2016. Conference of the*, 2016.

[31] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of Linear, Nonlinear, and Feature Selection Methods for EEG Signal Classification,"

IEEE Transactions on neural systems and rehabilitation engineering, vol. 11, no. 2, pp. 141–144, 2003.

[32] H. Ocak, "Optimal Classification of Epileptic Seizures in EEG Using Wavelet Analysis and Genetic Algorithm," *Signal processing*, vol. 88, no. 7, pp. 1858–1867, 2008.

[33] T.-P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Removal of Eye Activity Artifacts from Visual Event-Related Potentials in Normal and Clinical Subjects," *Clinical Neurophysiology*, vol. 111, no. 10, pp. 1745–1758, 2000.

[34] S. Fitzgibbon, D. DeLosAngeles, T. Lewis, D. Powers, T. Grummett, E. Whitham, L. Ward, J. Willoughby, and K. Pope, "Automatic Determination of EMG-Contaminated Components and Validation of Independent Component Analysis Using EEG During Pharmacologic Paralysis," *Clinical Neurophysiology*, vol. 127, no. 3, pp. 1781–1793, 2016.

[35] S. Çınar and N. Acır, "A Novel System for Automatic Removal of Ocular Artefacts in EEG by Using Outlier Detection Methods and Independent Component Analysis," *Expert Systems with Applications*, vol. 68, pp. 36–44, 2017.

[36] V. S. Handiru and V. A. Prasad, "Optimized bi-objective eeg channel selection and cross-subject generalization with brain–computer interfaces," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 6, pp. 777–786, 2016.

[37] G. J. Jiang, S.-Z. Fan, M. F. Abbod, H.-H. Huang, J.-Y. Lan, F.-F. Tsai, H.-C. Chang, Y.-W. Yang, F.-L. Chuang, Y.-F. Chiu *et al.*, "Sample Entropy Analysis of EEG Signals Via Artificial Neural Networks to Model PatientsâĂŹ Consciousness Level Based on Anesthesiologists Experience," *BioMed research international*, 2015.

[38] I. Omerhodzic, S. Avdakovic, A. Nuhanovic, and K. Dizdarevic, "Energy Distribution of EEG Signals: EEG Signal Wavelet-Neural Network Classifier," *arXiv preprint arXiv:1307.7897*, 2013.

[39] R. P. Lippmann, "Pattern Classification Using Neural Networks," *IEEE communications magazine*, vol. 27, no. 11, pp. 47–50, 1989.

[40] J. Zhao, Z. Xie, L. Jiang, H. Cai, H. Liu, and G. Hirzinger, "Levenberg-Marquardt Based Neural Network Control for A Five-Fingered Prosthetic Hand," in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. IEEE, 2005, pp. 4482–4487.

[41] L. B. Nguyen, A. V. Nguyen, S. H. Ling, and H. T. Nguyen, "Combining Genetic Algorithm and Levenberg-Marquardt Algorithm in Training Neural Network for Hypoglycemia Detection Using EEG Signals," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2013, pp. 5386–5389.

[42] J. E. Hall, *Guyton and Hall Textbook of Medical Physiology*. Elsevier Health Sciences, 2015.

[43] D. L. Schomer and F. L. Da Silva, *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 2012.

[44] M. Motro, "Searching for Causes of Poor Classification Performance in A Brain-Computer Interface Speller," 2014.