A Fuzzy Approach Towards Cognitive Load Detection Using Function Near-Infrared Spectroscopy

A Thesis

submitted in partial fulfillment of the requirements for

the Degree of

Master of Electronics & Tele-Communication Engineering

by

Mainak Dan

Registration No. 111857 of 2010-2011
Examination Roll No. M4ETC1611

under the guidance of

Prof. Amit Konar
DEPARTMENT OF ELECTRONICS & TELECOMMUNICATION ENGINEERING
JADAVPUR UNIVERSITY, KOLKATA-700032.

May, 2016

FACULTY OF ENGINEERING AND TECHNOLOGY
JADAVPUR UNIVERSITY

CERTIFICATE

This to certify that the thesis entitled "**A Fuzzy Approach Towards Cognitive Load Detection Using Functional Near-Infrared Spectroscopy**" has been carried out by **Mainak Dan** (Examination Roll No. **M4ETC1611**, Registration No. 111857 of 2010-11) under my guidance and supervision and be accepted in partial fulfillment of the requirement for the degree of Master of Electronics Telecommunication Engineering.

Supervisor

____

Dr. Amit Konar
Professor
Department of Electronics and Tele-
Communication Engineering
Jadavpur University
Kolkata-700032

Head of the Department

____

Dr. Palaniandavar Venkateswaran
Professor
Department of Electronics and Tele-
Communication Engineering
Jadavpur University
Kolkata-700032

____

Prof. Sivaji Bandyopadhyay
Dean, Faculty of Engineering and Technology
Jadavpur University
Kolkata – 700032.

FACULTY OF ENGINEERING AND TECHNOLOGY
JADAVPUR UNIVERSITY

CERTIFICATE OF APPROVAL*

The foregoing thesis is hereby approved as a creditable study of an engineering subject and presented in a manner satisfactory to warrant its acceptance as prerequisite to obtain the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn there in but approve the thesis only for which it is submitted.

Committee on Final Examination
for the Evaluation of the Thesis

____________________________ ____________________________

Signature of the Supervisor Signature of the Examiner

*Only in the case the thesis is approved

FACULTY OF ENGINEERING AND TECHNOLOGY
JADAVPUR UNIVERSITY

DECLARATION OF ORIGINALITY AND COMPLIANCE OF ACADEMIC THESIS

I hereby declare that this thesis entitled "A Fuzzy Approach Towards Cognitive Load Detection Using Functional Near-Infrared Spectroscopy" contains literature survey and original research work by the undersigned candidate, as part of his Degree of Masters of Electronics and Tele-Communication Engineering.

All information have been obtained and presented in accordance with academic rules and ethical conduct.

I also declare that, as required by these rules and conduct, I have completely cited all the materials and results that are not original to this work.

Mainak Dan
Examination Roll No.: **M4ETC1611**
Registration No.: **111857 of 2010-11**

Signature of the Candidate
Date ....

Acknowledgements

First and foremost, I would like to express my earnest gratitude and heartfelt indebtedness to my supervisor, Prof. Amit Konar, Department of Electronics and Tele-Communication Engineering, for the privilege and pleasure, of allowing me to work under him towards my Degree of Masters of Electronics and Tele-Communication Engineering. This work would not have been materialized, without his whole-hearted help and support.

I am thankful to Prof. Iti Saha Mishra and Prof. P. Venkateswaran who have been acting as Head of the Department of Electronics and Tele-Communication Engineering during the tenure of my studentship. I would also like to show my gratitude to Anuradha Saha, Research Scholar, Department of Electronics and Tele-Communication Engineering for her continuous support and valuable advices during the entire course of this thesis work.

I am indebted to a number of my classmates and friends Ranodip Das, Srishti Srivastava, Rimita Lahiri and Ranita Saha for their constant support, good wishes and for stimulating a fulfilled atmosphere during my stay at Jadavpur University. Lastly, I would like to thank my parents and brother for their constant support and encouragement.

Mainak Dan

Dedicated to My Parents…

Abstract

.

The main notion of this thesis is to identify the cognitive load during a mental arithmetic task experiment using fNIRS signals. The first objective is to classify the difficulty level and the state of inactivity during the given task. To identify the classes, the feature vectors have to undergo all the possible steps of a pattern classification problem. In this paper, we have developed a novel Feature Selection technique to reduce the dimension of the feature vectors by omitting the redundant features. For this purpose, an objective function depending upon the class density or likelihood functions is optimized using the well-known Differential Evolution algorithm. General type-2 fuzzy classifier is used for subsequent classification step. The general type-2 fuzzy classifier is used, as it provides additional degrees of freedom in designing membership functions and in modelling uncertainties than the normal type-1 fuzzy system, which relies on only one membership function. The proposed Feature Selection technique gives a satisfactory accuracy results over principal component analysis. The fuzzy classifier outperforms the other well-known classifier like support vector machine, k-nearest neighborhood. Experimental result reveals that the proposed likelihood-based FS induced type 2 fuzzy classifier attains the highest classification accuracy (above 90% in each case) as compared to its standard competitors. The load of a subject undergoing the experiment is measured at a particular class relying upon the mean type- 1 fuzzy value of all feature entities. A clear discrimination in concentration level from 16 channels has been observed for each distinct feature set.

CONTENTS

LIST OF PUBLICATIONS

[1] Dan, Mainak, Anuradha Saha, Amit Konar, Anca L. Ralescu, and Atulya K. Nagar. "A Type-2 Fuzzy Approach Towards Cognitive Load Detection Using fNIRS Signals." In *Fuzzy Systems Conference, 2016. FUZZ-IEEE 2016. IEEE International*. (Accepted)

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 Introduction

This chapter introduces basic structure of a Brain-Computer interfacing system. In section 1.1, BCI systems and different brain signal acquisition techniques are discussed. The main goal of the thesis work is to detect cognitive load and hence, in section 1.2, the problem definition along with the usefulness of such detection is outline. In subsequent sections, the motivation and the contribution to the thesis are discussed. The chapter is concluded with describing the organization of the thesis.

1.1 Brain-Computer Interfacing System

In recent years, development of brain computer interfacing (BCI) [1]-[2] systems has gained a considerable amount of importance from scientific communities. BCI researchers have put emphasis on modeling and developing systems which will be used to create a direct communication medium between brain and outside world without involving muscles or peripheral nervous system [3]. In BCI, external visual or auditory stimuli are provided to the user and with user's intentions; brain signals in terms of electric potential or concentration change of hemoglobin in blood vessels of inner tissue are generated. These signals are applied for

- biomedical engineering such as wheel chair [3]-[5], mind-driven motion of robots [6]-[8], thought-controlled driving [9]-[10], prosthetic devices [11]-[12] etc.,
- sleep disorders, neurological diseases and attention monitoring,
- correlating observable behaviour with recorded brain signals.

A complete BCI system consists of several steps - a brain-signal acquisition, signal pre-processing to remove any noise or artifacts, classification of the signal to understand the user's intention. At last, classifier output drives external motor or devices such that it can work with user's command. This type of control systems may be open loop or closed loop in nature in which visual or auditory signals are often used as feedback to the user for precise operation, for example in motor imagery BCI system [13]. There are two types of brain signal acquisition techniques, namely invasive and non-invasive.

1.1.1 Invasive Brain Signal Acquisition

Invasive BCI [14] techniques are mostly used for repairing damaged sight within the brain. As the devices are implanted within the grey matter of the brain by neurosurgical approach, it is able to produce high quality brain signals. But due to scar-tissue build up as the brain matters start to react with the externally implanted devices, the signals sometimes become weaker or non-existent in nature [15]. To avoid this problem, the partially invasive BCI techniques are often used, in which the devices are implanted under the cortex, but outside the grey matter of human brain. Electrocorticography (ECoG) [16] is one of the partially invasive BCI techniques by which better resolution electrophysiological signals than non-surgical or non-invasive BCI can be accumulated.

1.1.2 Non-Invasive Brain Signal Acquisition

To avoid surgical approach and high cost, nowadays, non-invasive techniques like Electroencephalogram (EEG) [17], functional Near-Infrared Spectroscopy (fNIRS) [18], functional Magnetic Resonance Imaging (fMRI) [19] are preferred over invasive techniques such as to collect brain signals.

In EEG, the electrophysiological activities are measures by recording the potentials at different electrodes above the scalp due to underlying neuronal activities [17]. fMRI is a functional neuro-imaging technique, in which it is assumed that neural activities have direct impact over cerebral blood flow. The brain activity is measured by detecting the blood-oxygen-level dependent (BOLD) signal [19] which is related to energy use by brain signals. In case of fNIRS signal, the change of the concentration of oxy-hemoglobin (HbO) and deoxy-hemoglobin (HbR) in the blood vessels of the underlying tissues is captured [18]. These non-invasive technique not only help in developing BCI systems to assist physically challenged people, but also help in developing systems that can be used to measure cognitive load for several applications like lie detection [20], attention measurement [21], detection of mental state [22] etc.

In this thesis work, fNIRS-BCI techniques is used due to the device's portability, low-expensive, easy to use, susceptibility to noise etc. unlike EEG and MEG or any other techniques which depend upon the electrophysiological signals, it is susceptible to electrical noises as it is an optical signal based imaging method. In fNIRS technique [18], the external infrared (IR) light sources are used to pass the photons through several layer of human brain. As blood contains haemoglobin which is carrier for the oxygen in our blood-system, some of the photons are back-reflected from the blood vessels of underlying tissues and are captures at particularly placed IR detectors above the human head. The amount of these photons is dependent upon the concentration of oxy and de-oxy haemoglobin. The detail mathematical relation between the amount of photons in terms of intensity and the concentration measurements are discussed in the subsequent chapters. There exist several literatures, which discussed about the promising outcomes of fNIRS BCI systems

1.2 Cognitive Load Detection Problem

According to cognitive psychology, cognitive load is defined as the total mental effort used during completion of a task. Cognitive load theory was developed out of the study of problem by John Sweller in the late 1980s [23]."Cognitive load theory has been designed to provide guidelines intended to assist in the presentation of information in a manner that encourages learner activities that optimize intellectual performance" [24]. Sweller's theory employs aspects of information processing theory to emphasize the inherent limitations of concurrent working memory load on learning during instruction. It makes use of the schema as primary unit of analysis for the design of instructional materials [24]. There exist several literatures in the field of psychology to measure the load during metal task [25]. The advantages of such load detection is that it can help in lie detection [20], attention detection [21] etc. In this thesis, an approach is introduced to measure the cognitive load during a mental arithmetic task [26] experiment from pattern recognition's perspective. There are namely two objectives to be completed.

i. The first objective is to classify the difficulty level and the state of inactivity during the given task.
ii. Secondly, the load of a subject undergoing the experiment is measured.

To identify the classes, the feature vectors have to undergo all the possible steps of a pattern classification problem. The work is discussed with sufficient details in subsequent chapters.

1.3 Motivation Leading to Thesis

S cientist communities have devoted their research minds in understanding human brain since past several decades. It is very hard to understand properly how human brain works, interpret what model to learn, how to learn it, when to learn that model etc. Though several learning strategies have already been proposed to mimic the learning procedure of human brain, it is still an open problem to scientific society. The main motivation for starting working in brain signal processing is to understand the human brain precisely, to understand the behavior of brain of an Alzheimer patient and thus, will be helpful for human society. Also some people having been suffering from cerebral palsy, epileptic seizure, improper motor function problem, therefore developing BCI systems at low cost, but with faster response and portability will help them in several ways for smoother life style. It is also a challenging problem in cognitive load theory to detect cognitive load properly as instructional design helps learners to reduce excessive amount of mental load during a task [24]. To deal with a complex part of human body, to understand its functionality, and in subsequent stages to use this knowledge in developing systems for social welfare led me to work in cognitive load detection problem from the perspective of pattern recognition.

1.4 Contribution to the Thesis

This thesis work aims for the detection of cognitive load on the basis of the fuzzified output of the feature vectors. The difficulty level is classified using general type-2 [27]-[29] classifier based on the assumption that difficulty level of problems may seem different to different subjects. Though the difficulty level of the problems are defined pre-experimentation, there exists some uncertainties in finding out appropriate difficulty level for each problem for each subject taking part in the experiment. The general type-2 fuzzy classifier is used, as it provides additional degrees of freedom in designing membership functions and in modelling uncertainties than the normal type-1 fuzzy system, which relies on only one membership function [27]-[29].

The second novelty lies in the feature selection (FS) step, where likelihood based objective function is designed with necessary mathematical approach. This selection technique identifies the most appropriate features from the original feature vector depending upon the class conditional density functions by optimizing a cost function using Differential Evolution (DE) [30] algorithm.

1.5 Organization of the Thesis

Chapter 2: Functional Near-Infrared Spectroscopy

This chapter introduces us to the fundamentals of fNIRS system, its mathematical background and data acquisition technique. Also it describes thoroughly the pre-processing techniques effectively used to remove noise and artifacts from fNIRS raw signals for further subsequent processing during classification problem.

Chapter 3: Classification and Load Detection

The main focus of this chapter is to provide a thorough description of the newly proposed feature selection technique and Fuzzy type-2 classifier for completion of the objective of this thesis work. Also, some well-known dimension reduction technique like Principal Component Analysis (PCA) and well-known classifiers like Support Vector

Machine (SVM), *k*-Nearest Neighborhood (*k*-NN), and Linear Discriminant Analysis (LDA) are discussed in context of classification of fNIRS signal classification.

Chapter 4: Experiment and Analysis

This chapter deals with all the experiments done in the laboratory. Some new features are also proposed and their discriminative natures are discussed with some supporting figures. The chapter concludes with the classification accuracy of the proposed classifier and comparison of this accuracy with other well-known classifiers and their statistical significance.

Chapter 5: Conclusion

This thesis work is concluded at chapter 5 by summarizing the complete work and some future research scopes to improve the cognitive load detection from pattern recognition's perspective.

Appendix A: MATLAB Source Codes

Here, the necessary source codes that can be executed in MATLAB environment for the proposed methods are given in tabular form.

References

[1] Dornhege, Guido. *Toward brain-computer interfacing*. MIT press, 2007.

[2] Nijholt, Anton, and Don FD Tan. "Brain-computer interfacing for intelligent systems." *Intelligent Systems, IEEE* 23, no. 3 (2008): 72-79.

[3] Long, Jinyi, Yuanqing Li, Hongtao Wang, Tianyou Yu, Jiahui Pan, and Feng Li. "A hybrid brain computer interface to control the direction and speed of a simulated or real wheelchair." *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* 20, no. 5 (2012): 720-729.

[4] Scherer, Reinhold, Felix Lee, Alois Schlögl, Robert Leeb, Horst Bischof, and Gert Pfurtscheller. "Toward self-paced brain–computer communication: navigation through virtual worlds." *Biomedical Engineering, IEEE Transactions on* 55, no. 2 (2008): 675-682.

[5] Galán, Ferran, Marnix Nuttin, Eileen Lew, Pierre W. Ferrez, Gerolf Vanacker, Johan Philips, and J. del R. Millán. "A brain-actuated wheelchair: asynchronous and non-invasive brain–computer interfaces for continuous control of robots." *Clinical Neurophysiology* 119, no. 9 (2008): 2159-2169.

[6] Millan, Jd R., Frédéric Renkens, Josep Mouriño, and Wulfram Gerstner. "Noninvasive brain-actuated control of a mobile robot by human EEG."*Biomedical Engineering, IEEE Transactions on* 51, no. 6 (2004): 1026-1033.

[7] Bell, Christian J., Pradeep Shenoy, Rawichote Chalodhorn, and Rajesh PN Rao. "Control of a humanoid robot by a noninvasive brain–computer interface in humans." *Journal of neural engineering* 5, no. 2 (2008): 214.

[8] Chae, Yongwook, Jaeseung Jeong, and Sungho Jo. "Toward brain-actuated humanoid robots: asynchronous direct control using an EEG-based BCI."*Robotics, IEEE Transactions on* 28, no. 5 (2012): 1131-1144.

[9] Curran, Eleanor, Peter Sykacek, Maria Stokes, Stephen J. Roberts, Will Penny, Ingrid Johnsrude, and Adrian M. Owen. "Cognitive tasks for driving a brain-computer interfacing system: a pilot study." *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* 12, no. 1 (2004): 48-54.

[10] Lin, Fu-Chang, Li-Wei Ko, Chun-Hsiang Chuang, Tung-Ping Su, and Chin-Teng Lin. "Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system." *Circuits and Systems I: Regular Papers, IEEE Transactions on* 59, no. 9 (2012): 2044-2055.

[11] Serruya, Mijail D., Nicholas G. Hatsopoulos, Liam Paninski, Matthew R. Fellows, and John P. Donoghue. "Brain-machine interface: Instant neural control of a movement signal." *Nature* 416, no. 6877 (2002): 141-142.

[12] Grigorescu, Sorin M., Thorsten Lüth, Christos Fragkopoulos, Marco Cyriacks, and Axel Gräser. "A BCI-controlled robotic assistant for quadriplegic people in domestic and professional life." *Robotica* 30, no. 03 (2012): 419-431.

[13] Schlögl, Alois, Felix Lee, Horst Bischof, and Gert Pfurtscheller. "Characterization of four-class motor imagery EEG data for the BCI-competition 2005." *Journal of neural engineering* 2, no. 4 (2005): L14.

[14] Gulati, Tanuj, Seok Joon Won, Dhakshin S. Ramanathan, Chelsea C. Wong, Anitha Bodepudi, Raymond A. Swanson, and Karunesh Ganguly. "Robust neuroprosthetic control from the stroke perilesional cortex." *The Journal of Neuroscience* 35, no. 22 (2015): 8653-8661.

[15] Millán, José del R., and J. Carmena. "Invasive or noninvasive: understanding brain-machine interface technology." *IEEE Engineering in Medicine and Biology Magazine* 29, no. EPFL-ARTICLE-150426 (2010): 16-22.

[16] Salanova, V., F. Andermann, A. Oliver, T. Rasmussen, and L. F. Quesney. "Occipital lobe epilepsy: electroclinical manifestations, electrocorticography, cortical stimulation and outcome in 42 patients treated between 1930 and 1991." *Brain* 115, no. 6 (1992): 1655-1680.

[17] Sanei, Saeid, and Jonathon A. Chambers. *EEG signal processing*. John Wiley & Sons, 2013.

[18] Coyle, Shirley M., Tomás E. Ward, and Charles M. Markham. "Brain–computer interface using a simplified functional near-infrared spectroscopy system." *Journal of neural engineering* 4, no. 3 (2007): 219.

[19] Weiskopf, Nikolaus, Klaus Mathiak, Simon W. Bock, Frank Scharnowski, Ralf Veit, Wolfgang Grodd, Rainer Goebel, and Niels Birbaumer. "Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI)." *Biomedical Engineering, IEEE Transactions on*51, no. 6 (2004): 966-970.

[20] Ding, Xiao Pan, Xiaoqing Gao, Genyue Fu, and Kang Lee. "Neural correlates of spontaneous deception: A functional near-infrared spectroscopy (fNIRS) study." *Neuropsychologia* 51, no. 4 (2013): 704-712.

[21] Negoro, Hideki, Masayuki Sawada, Junzo Iida, Toyosaku Ota, Shohei Tanaka, and Toshifumi Kishimoto. "Prefrontal dysfunction in attention-deficit/hyperactivity disorder as measured by near-infrared spectroscopy."*Child Psychiatry & Human Development* 41, no. 2 (2010): 193-203.

[22] Herff, Christian, Dominic Heger, Ole Fortmann, Johannes Hennrich, Felix Putze, and Tanja Schultz. "Mental workload during n-back task-quantified in the prefrontal cortex using fNIRS." *Frontiers in human neuroscience* 7, no. 1 (2013): 935-940.

[23] Sweller, John. "Cognitive load during problem solving: Effects on learning."*Cognitive science* 12, no. 2 (1988): 257-285.

[24] Sweller, John, Jeroen JG Van Merrienboer, and Fred GWC Paas. "Cognitive architecture and instructional design." *Educational psychology review* 10, no. 3 (1998): 251-296.

[25] Paas, Fred GWC, and Jeroen JG Van Merriënboer. "The efficiency of instructional conditions: An approach to combine mental effort and performance measures." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 35, no. 4 (1993): 737-743.

[26] Tanida, Masahiro, Kaoru Sakatani, Ruriko Takano, and Keiko Tagai. "Relation between asymmetry of prefrontal cortex activities and the autonomic nervous system during a mental arithmetic task: near infrared spectroscopy study." *Neuroscience letters* 369, no. 1 (2004): 69-74.

[27] Karnik, Nilesh N., Jerry M. Mendel, and Qilian Liang. "Type-2 fuzzy logic systems." *Fuzzy Systems, IEEE Transactions on* 7, no. 6 (1999): 643-658.

[28] Karnik, Nilesh N., and Jerry M. Mendel. "Introduction to type-2 fuzzy logic systems." In *Fuzzy Systems Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*, vol. 2, pp. 915-920. IEEE, 1998.

[29] Karnik, Nilesh N., and Jerry M. Mendel. "Operations on type-2 fuzzy sets."*Fuzzy sets and systems* 122, no. 2 (2001): 327-348.

[30] Storn, Rainer, and Kenneth Price. "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces." *Journal of global optimization* 11, no. 4 (1997): 341-359..

CHAPTER 2

Functional NIRS

The main focus of this chapter is to introduce fNIRS method to the reader. The fundamental and mathematical background of fNIRS systems are discussed followed by signal acquisition technique. Some previous related works using fNIRS signal are also described in brief. The chapter ended with the discussion of various noises encountered during fNIRS signal processing and some techniques to remove those artifacts.

2.1 Fundamentals of fNIRS

Functional near-infrared spectroscopy (fNIRS) is a neuro-imaging technique, in which the change in the optical properties of the underlying tissues of human brain due to change in the concentration of oxy-haemoglobin (HbO) and de-oxy-haemoglobin (HbR) [1] in the blood vessels of the tissues is assumed to be coupled with the brain activity. fNIRS signals have taken a significant amount of importance due to its electrical noise susceptibility (as it is an optical signal) in developing BCI systems using motor imagery signals taken from motor cortex (for example, driving simulation) to cognitive load detection with the help of mental arithmetic (MA) [2] task signals collected from pre-frontal cortex as it provides us with both spatial and temporal information unlike EEG [3] and fMRI [4], which have high values in only one of them.

The near-infrared spectrum corresponds to the optical wavelength window of 650-900 nm can penetrate the outer surface of the human brain, including the cranium and the various meninges and fluids surrounding the brain [1]. The photons traversed through the different layers of scalp, skull, cortical regions and brain fluids undergo several optical phenomena like diffusion, scattering, which create a random walking path of them within the human head. Some photons are absorbed while a significant amount are back reflected from the blood vessels of the tissues, as the concentration change of oxy-haemoglobin and de-oxy-haemoglobin take place. The back-reflected photons can travel up to several centimetres from the original point source location above the human cortical surface and are spotted out by suitably placed IR detectors. The changes in the emitted light intensity due to change in optical properties of the tissues can be calculated using

Beer-Lamberts Law [1], which measures optical density or attenuation (*OD*) in terms of chromophomore concentration (c). The corresponding mathematical relation [5] is described as,

$$OD = \log_{10}\left(\frac{I_0}{I}\right) = \alpha cLB + G, \tag{1}$$

where, where OD is attenuation, I_0 is the incident light intensity (mW), I is the transmitted light intensity (mW), α is the specific extinction coefficient (mol^{-1} m^{-1}), c is the concentration of the chromophomore (mol), L is the distance between the source and the detector (m), B is the differential path length factor and G is a term to account for scattering losses.

Differences in the absorption spectra of HbO and HbR allow the measurement of relative changes in haemoglobin concentration through the use of light attenuation at multiple wavelengths. Two or more wavelengths are selected, with one wavelength above and one below the isosbestic point of 810 nm at which HbO and HbR have identical absorption coefficients. The corresponding modified Beer-Lambert's Law [6] for measuring change of attenuation is expressed linearly in term of change in concentration of HbO and HbR as,

$$\Delta OD = \left(\alpha_{HbO}\Delta c_{HbO} + \alpha_{HbR}\Delta c_{HbR}\right) BL, \tag{2}$$

where, Δ is used to indicate the change in the particular parameter value.

2.2 fNIRS Signal Acquisition

BCI uses brain signals to collect information on the user's intensions. The first step in developing an fNIRS-BCI system is to acquire suitable brain signals. The two most common brain areas from where the signals are acquired are the primary motor cortex and the prefrontal cortex. The signals that are responsible for motor imagery and motor execution tasks are collected from motor cortex; prefrontal cortex are used in collecting signals are related to mental arithmetic, mental games, landscape imagination etc. In spite of the fact that, there exist several different emitter-detector configurations used in these two areas, the emitter-detector distances are designed to keep within a typical range due to their significant role in fNIRS measurement. For example, an increase in emitter-detector distance corresponds to an increase in imaging depth [7]. To measure blood oxygenation signals from the cortical areas, 3 cm separation between emitter and detector was proposed [8]. A separation of less than 1 cm might contain only skin-layer contribution, whereas that of more than 5 cm might result in weak, noisy signals [9] which are not at all useful for further processing. A typical emitter-detector configuration on the head and the paths traveled by light to reach two detectors are shown in Figure 2. A suitable number of emitter/detector pairs for adequate extraction of neuronal activity vary depending on the type of brain signals that are used for BCI purpose. For the prefrontal cortex, 3 emitters and 8 detectors are suitable for adequately acquiring most brain signals corresponding to prefrontal tasks [10]-[16]. The path between emitter and detector is known the channel. For brain activities corresponding to motor cortex tasks, 6 emitters and 6 detectors can cover the entire motor cortex. In the previous studies, 4 emitters and 4 detectors [17], 6 emitters and 6 detectors [18], and 5 emitters and 4 detectors have been applied to acquire motor-cortex activities.

2.3 Previous Related Works

This section deals with the previous related works using fNIRS signals. The discussion has been extended from motor imagery signals taken from motor cortex to the cognitive tasks signal for which mainly pro-frontal cortex is responsible.

2.3.1 Motor Cortex Activity

Brain signal collection from primary motor cortex is a natural means of providing a strong BCI control over external devices for motor intention execution. Moreover, these might also be an effective data acquisition from the perspective of neuro-rehabilitation to support the people with motor-neuron diseases. Motor execution and motor imagery are the two most common activities that are gathered from the motor cortex of human brain.

2.3.1.1 Motor Execution

The motor execution task activates the motor cortex due to the movement of the human body parts, which involves the muscular activities and muscular tension due to physical work. Due to involvement of muscular contraction in motor execution task, motor execution-based BCIs are affected by somatic sensory feedback from contracting muscles and, therefore, the neuronal modulation may not be solely from the central nervous system. Several motor execution tasks including finger tapping [19]-[21], hand tapping [16], [22], arm lifting [23], knee extension [23] and hand grasping/gripping [24], [25] have been used in the previous studies.

2.3.1.1 Motor Imagery

Motor imagery can be characterized as a covert mental process of envisioning of physiological movement of one's own body part without inclusion of muscular tension, contraction or flexion. Since the main goal of BCI is to develop a communication channel for motor-disabled people, motor imagery is one of the most useful tasks in fNIRS-BCI. The motor imagery tasks specify imagination of the squeezing of a soft ball, covert imagery of a simple or complex sequence of finger tapping, imagination of feet tapping, imagination of hand grasping/gripping, imagination of wrist flexion, imagination of flexion and extension of elbow, and folding and unfolding of specific fingers [26]. Unlike motor execution tasks, the motor imagery signals are free of somatic sensory feedback.

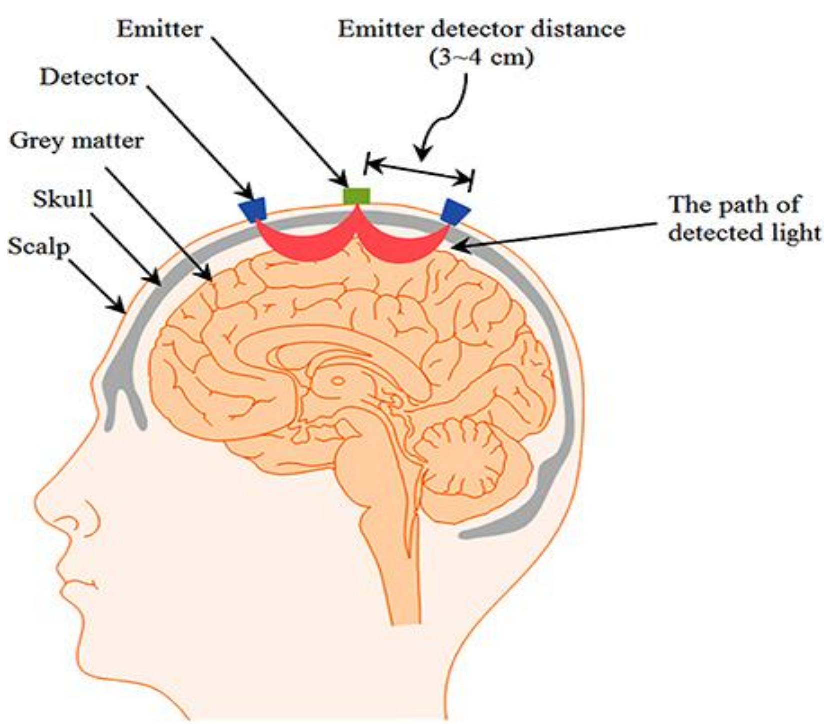


Figure 1 fNIRS Signal Acquisition Process

2.3.2 Prefrontal Cortex Activity

As fNIRS signals are more disturbed and attenuated by the motion artifacts due to the slippage in hairs, the activities in the prefrontal cortex are also a good choice for fNIRS-BCI. Also, they are likely to be more effective in the case of motor-function related disability. Given these advantages, most studies have used the pre-frontal activities showing promising results. Some of the commonly utilized prefrontal activities for fNIRS-BCI are mental arithmetic, music imagery, mental counting, and landscape imagery [26].

2.3.2.1 Mental Arithmetic

Mental arithmetic (sometimes called mental calculation) means performing secret calculation using the brain without any help in the form of paper, pen, calculator, computer, etc. It activates the prefrontal cortex. Since it does not include any muscle movement, it is widely used for fNIRS-BCI. Various studies have effectively shown its possibility as a mental assignment for BCI. Mental arithmetic entails mental multiplication or other arithmetic tasks. However, the most commonly utilized mental arithmetic is backwards subtraction, which involves subtraction of a small number (for example, a two-digit number) from a large number (for example, a three-digit number) with successive subtraction of a randomly appearing small number from the result of the previous subtraction [26].

2.3.2.2 Music Imagery

Music imagery (also called mental singing) consists of organizing and analyzing music in the brain without any external auditory stimulus. It has been successfully demonstrated music imagery as a brain activity that can be effectively used for fNIRS-BCI.

Besides mental arithmetic and music imagery, various other tasks in the prefrontal cortex have been demonstrated with promising outcomes. These involve mental counting, landscape imagery, mental character writing, object rotation, change-detection tasks, labyrinth tasks, and emotion-induction tasks. Some studies have demonstrated direct decoding of neural correlates corresponding to subjective preferences, deception, visual stimuli, and others [26].

2.4 fNIRS Signal Processing

While conducting the experiment, it is certain that various kinds of noises cause disturbance in the data stream of fNIRS. These noises are generated as the ideal condition for experiment is difficult to create. These noises can be divided into three categories in broad sense – instrumental noise, experimental noise and physiological noise. The first two categories are not at all related to any physiological activities. Thus it is better to remove these noises before removing the physiological noises from the output of the modified Beer-Lambert’s law [6].

2.4.1 Instrumental Noise Removal

Instrumental Noises are generated due to instrumental degradation or by variation of surround light intensity. These disturbances create a constant high frequency noises and can be easily removed by low pass filter with appropriately chosen cut-off frequencies (say, 3-5 Hz) [26].

2.3.2. Experimental Noise Removal

Experimental noise or error is generated from the motion artifacts such as head motion which results in the dislocation of optodes. These can be visualized in the fNIRS data as spike artifacts due to abrupt change in light intensity as the optodes change their positions. There exist several filtering methods for removing motion artifacts and other experimental errors such as the Wiener filtering-based method, eigenvector-based spatial filtering, wavelet-analysis-based methods, Savitzky-Golay type filters, and others [26]. For comparison of various techniques and their advantages and disadvantages, the readers can go through the literature [27].

2.4.3. Physiological Noise Removal

Physiological noise includes several kinds of artifacts due to heartbeat (1-1.5 Hz), blood pressure fluctuations or Mayer wave (around 0.1 Hz), respiration (0.2 – 0.5 Hz) [21] etc. There are several well-known techniques such as band-pass filtering [1], [10]-[12], ICA [28], ARMAx method, advance filtering technique [27] to remove physiological noise. The techniques that are used in the thesis are described with sufficient details.

2.4.3.1. Band Pass Filtering

Band-pass filtering is an effective way to remove physiological noises as the frequency range is known a priori. Some fNIRS-BCI studies have shown promising results using a simple low-pass, or a high-pass or a band-pass filtering to remove physiological noises. Various cut-off frequencies for band-pass filtering have been used for decades to sample out the noises: For example, [10], [12], [29] and [30] have used the frequency bands of 0.01~0.8 Hz, 0.1~0.5 Hz, 0.01 ~ 0.2 Hz and 0.1 ~ 0.5 Hz, respectively. In general, the band of 0.1 ~ 0.4 Hz is used in this thesis work as it can eliminate significantly a large portion of physiological noises such as heart beat and Mayer waves [26] without disturbing the fNIRS signal of interest elicited by a task of 10 seconds period. The types of band-pass filtering include Butter worth filters, Elliptic filters, and Chebyshev filters. However, there is no study which signifies advantage of a particular filtering method over others yet.

2.4.3.2. Advance Filtering Method

If the frequency band of a physiological noise, for example due to respiration, overlap with frequency band of the hemodynamic response, then band-pass filtering is not an effective means for removing physiological noises. Therefore, other methods, such as adaptive filtering [27], PCA [31] and ICA [29] have found significant application area in eliminating physiological noise. To account for physiological clamors, additional noise-related elements can be included to the regression model. In addition to modeling the canonical functional response, a series with adaptive amplitudes and phase components in order to model specific physiological noise contribution from heartbeat, respiration, and blood pressure can be incorporated. The autoregressive moving average with exogenous signals (ARMAx) model-based approach incorporating physiological signals as exogenous signals can be used to predict the brain state during a particular cognitive task. The fNIRS signal at each channel can be regarded as an output from a linear combination of various components. The components include the dynamical characteristics of the HbO and HbR changes in a specific brain region (the influence from the current/previous stimuli), the physiological signals, the baseline fluctuation, and other noises.

Independent Component Analysis

According to A.Hyvarinen *et al.* [28] Independent Component Analysis (ICA) is a *method for finding underlying factors or components from multivariate (multi-dimensional) statistical data*. ICA is a special case of Blind Source Separation (BSS) problem in which it is assumed that the sources are statistically independent and non-Gaussian in nature. Using ICA, one can separate multiple source signals into additive subcomponent or target signals. ICA tries to find a direction in vector space along which the source components are independent in nature. In our experiment, 16 channel data have been collected using the fNIRS device from pre-frontal cortex. Though band pass filtering is used to remove several artefacts, still there is a scope of noise interference between these channels and also from several physiological disturbances which results in low SNR rate at each channel. Thus, use of ICA is necessary as an advance filtering method to deal with such kind of nuisance.

Let us assume that the data (observed from 16 channels of the fNIRS device used) are represented by $x(t) = \left(x_1(t), x_2(t), \ldots, x_m(t)\right)^T$ and the sub-components are denoted as $s(t) = \left(s_1(t), s_2(t), \ldots, s_m(t)\right)^T$. The objective of ICA is to transform the observed data x into maximum number of independent components s by a linear transform W dependent upon a function $F(s_1, s_2, \ldots, s_m)$ of independence. So the task of ICA is to find both the target variable $s(t)$ and the matrix W while only the variable $x(t)$ is observable. The mathematical derivation of ICA is based on two principles,

i. The additive components s_i and s_j are uncorrelated.

ii. Also the non-Gaussianity of the transform $s = Wx$ is locally maximum under the constraint that the variance of x is constant.

Based on these two principles, let us discuss the generative model for linear noiseless ICA is developed mathematically. After removing the noises, signals classified through regular pattern recognition steps which are discussed in the next chapter.

References

[1] Coyle, Shirley M., Tomás E. Ward, and Charles M. Markham. "Brain–computer interface using a simplified functional near-infrared spectroscopy system." *Journal of neural engineering* 4, no. 3 (2007): 219.

[2] Tanida, Masahiro, Kaoru Sakatani, Ruriko Takano, and Keiko Tagai. "Relation between asymmetry of prefrontal cortex activities and the autonomic nervous system during a mental arithmetic task: near infrared spectroscopy study." *Neuroscience letters* 369, no. 1 (2004): 69-74.

[3] Sanei, Saeid, and Jonathon A. Chambers. *EEG signal processing*. John Wiley & Sons, 2013.

[4] Weiskopf, Nikolaus, Klaus Mathiak, Simon W. Bock, Frank Scharnowski, Ralf Veit, Wolfgang Grodd, Rainer Goebel, and Niels Birbaumer. "Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI)." *Biomedical Engineering, IEEE Transactions on*51, no. 6 (2004): 966-970.

[5] Villringer, Arno, and Britton Chance. "Non-invasive optical spectroscopy and imaging of human brain function." *Trends in neurosciences* 20, no. 10 (1997): 435-442.

[6] Ayaz, Hasan, Patricia A. Shewokis, Adrian Curtin, Meltem Izzetoglu, Kurtulus Izzetoglu, and Banu Onaral. "Using MazeSuite and functional near infrared spectroscopy to study learning in spatial navigation." *J. Vis. Exp* 56, no. 3443 (2011): 10-3791.

[7] McCormick, Patrick W., Melville Stewart, Gary Lewis, Manuel Dujovny, and James I. Ausman. "Intracerebral penetration of infrared light: technical note."*Journal of neurosurgery* 76, no. 2 (1992): 315-318.

[8] Gagnon, Louis, Meryem A. Yücel, Mathieu Dehaes, Robert J. Cooper, Katherine L. Perdue, Juliette Selb, Theodore J. Huppert, Richard D. Hoge, and David A. Boas. "Quantification of the cortical contribution to the NIRS signal over the motor

cortex using concurrent NIRS-fMRI measurements."*Neuroimage* 59, no. 4 (2012): 3933-3940.

[9] Gratton, Gabriele, Carrie R. Brumback, Brian A. Gordon, Melanie A. Pearson, Kathy A. Low, and Monica Fabiani. "Effects of measurement method, wavelength, and source-detector distance on the fast optical signal." *Neuroimage* 32, no. 4 (2006): 1576-1590.

[10] Luu, Sheena, and Tom Chau. "Decoding subjective preference from single-trial near-infrared spectroscopy signals." *Journal of neural engineering* 6, no. 1 (2008): 016003.

[11] Power, Sarah D., Tiago H. Falk, and Tom Chau. "Classification of prefrontal activity due to mental arithmetic and music imagery using hidden Markov models and frequency domain near-infrared spectroscopy." *Journal of neural engineering* 7, no. 2 (2010): 026002.

[12] Power, Sarah D., Azadeh Kushki, and Tom Chau. "Towards a system-paced near-infrared spectroscopy brain? computer interface: differentiating prefrontal activity due to mental arithmetic and mental singing from the no-control state." *Journal of neural engineering* 8, no. 6 (2011): 066004.

[13] Power, Sarah D., Azadeh Kushki, and Tom Chau. "Automatic single-trial discrimination of mental arithmetic, mental singing and the no-control state from prefrontal activity: toward a three-state NIRS-BCI." *BMC research notes*5, no. 1 (2012): 141.

[14] Power, Sarah D., Azadeh Kushki, and Tom Chau. "Intersession consistency of single-trial classification of the prefrontal response to mental arithmetic and the no-control state by NIRS." *PloS one* 7, no. 7 (2012): e37791.

[15] Khan, M. Jawad, Melissa Jiyoun Hong, and Keum-Shik Hong. "Decoding of four movement directions using hybrid NIRS-EEG brain-computer interface."*Front. Hum. Neurosci* 8, no. 244 (2014): 10-3389.

[16] Naseer, Noman, Melissa Jiyoun Hong, and Keum-Shik Hong. "Online binary decision decoding using functional near-infrared spectroscopy for the development

of brain–computer interface." *Experimental brain research* 232, no. 2 (2014): 555-564.

[17] Sitaram, Ranganatha, Haihong Zhang, Cuntai Guan, Manoj Thulasidas, Yoko Hoshi, Akihiro Ishikawa, Koji Shimizu, and Niels Birbaumer. "Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain–computer interface." *NeuroImage* 34, no. 4 (2007): 1416-1427.

[18] Naseer, Noman, and Keum-Shik Hong. "Classification of functional near-infrared spectroscopy signals corresponding to the right-and left-wrist motor imagery for development of a brain–computer interface." *Neuroscience letters* 553 (2013): 84-89.

[19] Cui, Xu, Signe Bray, and Allan L. Reiss. "Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics." *Neuroimage*49, no. 4 (2010): 3039-3046.

[20] Cui, Xu, Signe Bray, and Allan L. Reiss. "Speeded near infrared spectroscopy (NIRS) response detection." *PLoS one* 5, no. 11 (2010): e15474.

[21] Seo, Young-Wook, Seung-Duk Lee, Dal-Kwon Koh, and Beop-Min Kim. "Partial least squares-discriminant analysis for the prediction of hemodynamic changes using near infrared spectroscopy." *Journal of the Optical Society of Korea* 16, no. 1 (2012): 57-62.

[22] Hai, Nguyen Thanh, Ngo Q. Cuong, Truong Q. Dang Khoa, and Vo Van Toi. "Temporal hemodynamic classification of two hands tapping using functional near-infrared spectroscopy." *Front. Hum. Neurosci* 7, no. 516 (2013): 10-3389.

[23] Shin, Jaeyoung, and Jichai Jeong. "Multiclass classification of hemodynamic responses for performance improvement of functional near-infrared spectroscopy-based brain–computer interface." *Journal of biomedical optics* 19, no. 6 (2014): 067009-067009.

[24] Nagaoka, Takafumi, Kaoru Sakatani, Takayuki Awano, Noriaki Yokose, Tatsuya Hoshino, Yoshihiro Murata, Yoichi Katayama, Akihiro Ishikawa, and Hideo Eda.

"Development of a new rehabilitation system based on a brain-computer interface using near-infrared spectroscopy." In *Oxygen transport to tissue XXXI*, pp. 497-503. Springer US, 2010.

[25] Fazli, Siamac, Jan Mehnert, Jens Steinbrink, Gabriel Curio, Arno Villringer, Klaus-Robert Müller, and Benjamin Blankertz. "Enhanced performance by a hybrid NIRS–EEG brain computer interface." *Neuroimage* 59, no. 1 (2012): 519-529.

[26] Naseer, Noman, and Keum-Shik Hong. "fNIRS-based brain-computer interfaces: a review." *Frontiers in human neuroscience* 9 (2015).

[27] Cooper, Robert J., Juliette J. Selb, Louis Gagnon, Dorte Phillip, Henrik W. Schytz, Helle K. Iversen, Messoud Ashina, and David A. Boas. "A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy." (2012).

[28] Hyvarinen, Aapo, and J. Karhunen. "Oja. Independent component analysis."*John Wiley & Sons* (2001).

[29] Hu, Xiao-Su, Keum-Shik Hong, and Shuzhi Sam Ge. "fNIRS-based online deception decoding." *Journal of neural engineering* 9, no. 2 (2012): 026012.

[30] Tomita, Yohei, Francois-Benoit Vialatte, Gérard Dreyfus, Yasue Mitsukura, Hovagim Bakardjian, and Andrzej Cichocki. "Bimodal BCI using simultaneously NIRS and EEG." *Biomedical Engineering, IEEE Transactions on* 61, no. 4 (2014): 1274-1284.

[31] Zhao, Haitao, Pong Chi Yuen, and James T. Kwok. "A novel incremental principal component analysis and its application for face recognition."*Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*36, no. 4 (2006): 873-886.

CHAPTER 3

Classification and Load Detection

This chapter deals with classification of fNIRS signal after pre-processing and feature extraction. In section 3.1, the problem solving approach is discussed followed by Feature Selection and Dimension Reduction technique in section 3.2. This section demonstrates the proposed likelihood based DE-driven FS technique. In the next section (3.3) classification process is discussed by describing some well-known classifiers and also elaborating the proposed type-2 fuzzy classifier.

3.1 Problem Solving Approach

This chapter is dedicated to the classification of the fNIRS signal and cognitive load detection of individual subjects. One of the objectives is to find the state of activity and also the difficulty level of the problems that have been come across by the subjects during the experiment. The purpose of this experiment is to find the cognitive load encounter in the subjects during this mental arithmetic cognitive task. To find the state of inactivity and difficulty level, we are entirely depending upon the nature of the fNIRS signal. Thus classification of these signals is primarily required. As it is considered as a classification problem, the solution of this problem consists of several well-known steps – data acquisition, pre-processing, feature extraction (FE), feature selection (FS) and dimension reduction (DR), classification. Some well-known classifiers have been described with sufficient mathematical background. In this chapter, the proposed classifier based on type-2 fuzzy [1]-[5] logic is proposed for classification. Also, a FS tool based on the likelihood estimation [6] is also proposed for better performance. Data acquisition and pre-processing steps are described in chapter 2 and the feature extraction and classifier performance will be discussed in the next chapter. These approaches are described with sufficient details in the subsequent sections.

3.2 Feature Selection

The feature selection (FS) technique is used in the learning algorithm to choose a smaller subset of features without any loss of information from the larger feature set when there exist a few samples for training but the original feature set is of high dimension [7]. The reasons behind the essentiality of FS technique are as follows.

i. There exist a number of redundant features in the original feature set. When the number of data samples for training is lower with respect to the dimension of the feature vector, there is a possibility of overtraining the classifier. To get rid of overtraining, it is necessary to choose a subset of features such that no redundancy is present.

ii. If the dimension of the feature vector reduces, it is obvious that the training phase will take shorter time.

iii. The complexity of the model reduces as the dimension of the feature vector reduces.

In this work, a new feature selection strategy is proposed based on the likelihood functions or class density functions. The technique is discussed with sufficient mathematical description in the following subsection.

3.2.1 Proposed Likelihood Based Feature Selection

Let $\mathbf{X}_{N\times D}=\{\vec{X}_1,\vec{X}_2,\ldots\vec{X}_N\}$ be a set of N pattern vectors or data points, each having D features. Given such $\mathbf{X}_{N\times D}$ matrix, the object is to find a feature matrix $\mathbf{X}_{N\times n}$, where $n\leq D$. The notion of this FS technique is to select the features which involve in significant increment of likelihood of a feature vector as a whole within its own class and decrement of likelihood in other classes. The probability of the feature vector $\vec{X}_i^c$ with prior information that class c has already appeared in the experiment is given by $p(\vec{X}_i^c/c)$. Similarly, the probability of the feature vector $\vec{X}_i^c$ with prior information that class d has already appeared in the experiment is given by $p(\vec{X}_i^c/d)$. Now, for an classification problem with K number of classes, if an iterative optimization technique is used, objective is to maximize the difference between intra and inter-class likelihood function which can be expressed mathematically as,

$$p(\vec{X}_i^c / c) - \tfrac{1}{K-1} \sum_{\forall d, d \neq c} p(\vec{X}_i^c / d) \tag{3}$$

Thus, the overall objective function considering all the feature vectors of all the classes can be written as,

$$J = \sum_{\forall c} \sum_{\forall i, i \in c} \left(p(\vec{X}_i^c / c) - \tfrac{1}{K-1} \sum_{\forall d, d \neq c} p(\vec{X}_i^c / d) \right) \tag{4}$$

Let, $\vec{X}_i^c$ be the i^{th} vector belong to class c, $\vec{\mu}^c$ be the mean vector of class c, Σ^c be the co-variance matrix of class c and $p(\vec{X} / c) \sim N(\vec{\mu}^c, \Sigma^c)$. By modifying equation (4) and neglecting higher order terms, we have,

$$J = L_1 - \lambda . L_2 \tag{5}$$

where, scale factor $\lambda \in (0,10]$

$$L_1 = \sum_{\forall c} \sum_{\forall i, i \in c} \left(\frac{1}{K-1} \sum_{\forall d, d \neq c} L_1^d - L_1^c \right), \tag{6}$$

$$L_2 = \sum_{\forall c} \sum_{\forall i, i \in c} \left(\frac{1}{K-1} \sum_{\forall d, d \neq c} \frac{1}{(2\pi)^{\frac{n}{2}} \left| \Sigma^d \right|^{\frac{1}{2}}} - \frac{1}{(2\pi)^{\frac{n}{2}} \left| \Sigma^c \right|^{\frac{1}{2}}} \right), \tag{7}$$

$$L_1^d = \frac{1}{2.(2\pi)^{\frac{n}{2}} \left|\Sigma^d\right|^{\frac{1}{2}}} (\vec{X}_i^c - \vec{\mu}^d)^T \Sigma^{d^{-1}} (\vec{X}_i^c - \vec{\mu}^d), \tag{8}$$

$$L_1^c = \frac{1}{2.(2\pi)^{\frac{n}{2}} \left|\Sigma^n\right|^{\frac{1}{2}}} (\vec{X}_i^c - \vec{\mu}^c)^T \Sigma^{c^{-1}} (\vec{X}_i^c - \vec{\mu}^c). \tag{9}$$

Equation (5) is maximized using Differential Evolution (DE) algorithm. To initialize population vectors, string vectors of dimension *D* consisting of '1' (true) and '0' (false) are constructed. The true value at a position indicates the inclusion of the feature value of that particular index into the lower dimensional feature vectors. These low dimensional feature vectors are used to obtain the value of the cost function of equation (5) assuming that their original class definitions have not been changed. The string vectors are updated iteratively using DE/rand/1/bin [7] technique. The scale factor λ is chosen experimentally by taking the ratio of L_1 and L_2, and average them over a number of these ratios to maintain a compatibility between these two values. After optimization the best string vectors found so far is used for reduction of the original feature vectors. Now, when a new unknown feature vector comes into sight for testing, this string vector helps the recognizer to down select the features. The corresponding algorithm using DE [7] as an optimizer is explained in table I.

Table I Algorithm for Likelihood Based DE-Driven Feature Selection

Algorithm 1: Likelihood Based DE-Driven Feature Selection

1. **Initialization:** Declare *D*-dimensional population vectors of size *NP* $\vec{Z}_i, i = 1 \quad to \quad NP$, where each vector is a binary string with elements '0' or '1'. Set crossover rate, *Cr = 0.7* and scale factor *F = 0.8.*

2. **Mutation:** For each target vector $\vec{Z}_i$ select random mutually distinct $\vec{Z}_j, \vec{Z}_k, \vec{Z}_l$ to create the donor vector using the equation

$$\vec{V}_i = \vec{Z}_j + F.(\vec{Z}_k - \vec{Z}_l)$$

   Also the elements of donor vectors are rounded to nearest integer from the set {0,1}.

3. **Recombination/ Crossover:** Create an offspring vector $\vec{U}_i$ for each target vectors using the vectors $\vec{Z}_i$ and $\vec{V}_i$ where the *j*th of the offspring vector is obtained by the following rule,

$$\boldsymbol{u_{i,j}} = \begin{cases} \boldsymbol{v_{i,j}}, & \boldsymbol{if\ a\ randomly\ selected\ number\ r, between\ 0\ and\ 1 \leq Cr} \\ \boldsymbol{z_{i,j}}, & \boldsymbol{otherwise} \end{cases}$$

4. Selection: Find out the indices of target vector and offspring vectors with entries '1'. Select the corresponding features from the feature set. Compute the objective value using equation (5). Find out the binary string vector which is able to produce minimum objective function value. Translate that vector to the next generation or iteration

5. Repeat from step 2 until the termination criterion is not reached.

6. Output the best fit string vectors after the termination criterion is achieved. The indices with '1' values are the features of a feature vector to be selected for subsequent stages.

3.2.2 Dimension Reduction Using PCA

Principle Component Analysis [8] is a statistical method that linearly transforms the correlated set of observed variables into linearly uncorrelated set of variables in least square sense, also known as principle components. The number of uncorrelated variables is less than or equal to the number of observed variables. It is orthogonal transformation used to project higher dimensional data into lower dimension space onto a line through the sample mean in the direction along the Eigen vectors of the largest Eigen values of the scatter matrix of the observed set of data as shown in the figure 2. Let us discuss the mathematical explanation in detail from the perspective of fNIRS signal processing such that the readers get a comprehensive overview of using PCA in fNIRS.

Let us consider *n* *d*-dimensional data points or feature vectors $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ obtained from the pre-processed fNIRS signal. The objective is to find a vector $\vec{x}_o$ such that sum of the squared distance between $\vec{x}_o$ and various $\vec{x}_k$'s is minimum. The objective function can be constructed as,

$$J_o\left(\vec{x}_o\right) = \sum_{k=1}^{n} \left\| \vec{x}_o - \vec{x}_k \right\|^2 \tag{10}$$

Taking the derivative of $J_o\left(\vec{x}_o\right)$ with respect to $\vec{x}_o$, it is found that

$$\vec{x}_o = \frac{1}{n} \sum_{k=1}^{n} \vec{x}_k = \vec{m} \tag{11}$$

Where $\vec{m}$ is the sample mean.

Therefore, for dimension reduction, we have to project the feature vectors along the direction of the sample mean $\vec{m}$. Now from vector algebra, any vector can be written in terms of other vector according to the law of the triangle.

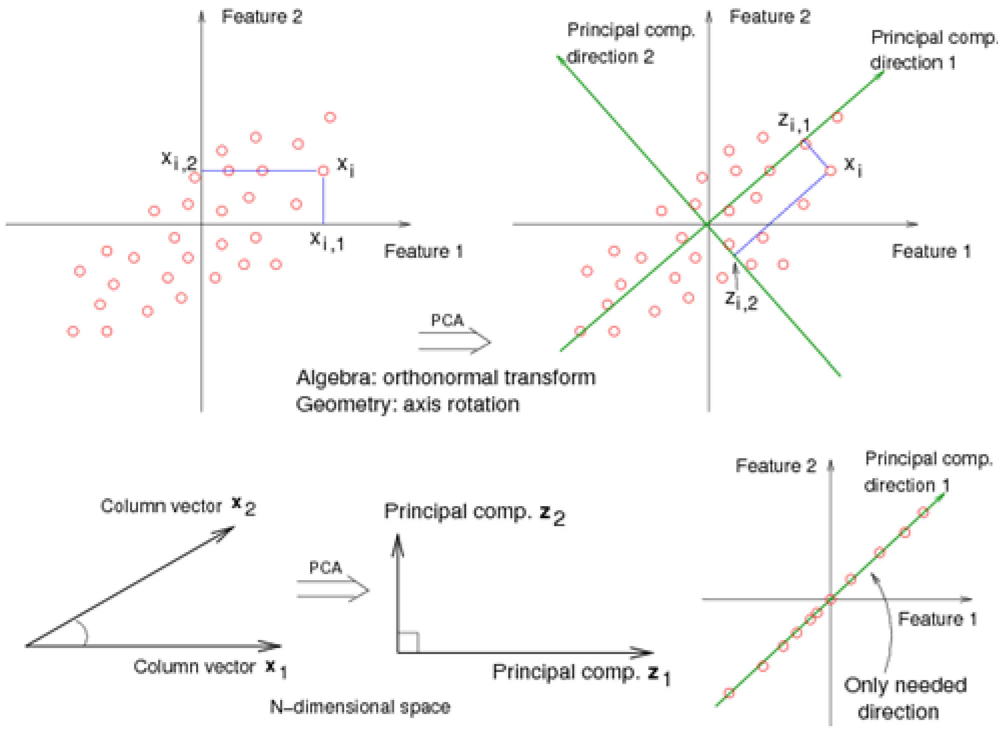


Figure 2 Principal Compnent Analysis in 2D feature space

Thus,

$$\vec{x}_k = \vec{m} + a_k \vec{e} \tag{12}$$

where, $\vec{e}$ is the unit vector from $\vec{m}$ to $\vec{x}_k$. To find the optimal set of coefficients a_k's, an objective function is to be minimized, mathematically expressed as,

$$J_1(a_1, a_2, ..., a_n, \vec{e}) = \sum_{k=1}^{n} \left\| (\vec{m} + a_k \vec{e}) - \vec{x}_k \right\|^2 \tag{13}$$

By taking proper mathematical approach, the value is found to be $a_k = \vec{e}^T (\vec{x}_k - \vec{m})$. Thus the set of coefficients can be found by projecting the vectors onto the line in the direction of the unit vector that passes through the sample mean. The next step is to find the value of the unit vectors.

Let us define the scatter matrix **S** as,

$$\mathbf{S} = \sum_{k=1}^{n} (\vec{x}_k - \vec{m})(\vec{x}_k - \vec{m})^T \tag{14}$$

To find the unit vectors let us modify the previous objective function in term of the unit vector which is,

$$J_1(\vec{e}) = -\vec{e}^T \mathbf{S}\vec{e} + \sum_{k=1}^{n} \|\vec{x}_k - \vec{m}\|^2 \tag{15}$$

To minimize $J_1(\vec{e})$ we need to maximize $\vec{e}^T \mathbf{S}\vec{e}$. A constrained optimization problem can be designed which as follows.

$$maximize \quad \vec{e}^T \mathbf{S}\vec{e}, \ \text{subject} \quad \text{to} \ \|\vec{e}\|^2 = 1 \tag{16}$$

By using Lagrange multiplier, it can be transformed into an unconstrained optimization problem by defining a function *u*. Mathematically, it can be written as,

$$u = \vec{e}^T \mathbf{S}\vec{e} - \lambda(\vec{e}^T \vec{e} - 1) \tag{17}$$

Now,

$$\begin{aligned}&\frac{\partial u}{\partial \vec{e}} \overset{\Delta}{=} 0\\&or, 2\mathbf{S}\vec{e} - 2\lambda\vec{e} = 0\\&or, \mathbf{S}\vec{e} = \lambda\vec{e}\end{aligned} \tag{18}$$

Thus, it is found that the principle components are the directional vectors along the direction of the Eigen vectors of the scatter matrix. The problem can be extended for *d'*-dimensional projection. The steps involved are enumerated below.

i. Find the scatter matrix **S** of the observed data vectors.
ii. Find the Eigen values and corresponding Eigen vectors of the scatter matrix.
iii. Arrange the Eigen values in ascending order.
iv. Find the difference between two consecutive Eigen values starting from the largest one.
v. Consider the Eigen values after which the difference gets a higher increment.
vi. Construct a matrix **A,** whose each column corresponds to the Eigen vectors of the corresponding Eigen values survived during step v.
vii. Find $\vec{y} = \mathbf{A}^T\vec{x}$ where, $\vec{y}$ is a reduced dimensional orthogonally projected feature vector.

It is evident from the mathematical derivation of PCA that

i. Scatter matrix is real and symmetric and hence, the Eigen vectors are orthogonal,
ii. Eigen vectors are the principles axes of the hyper ellipsoidal cloud of the observed data vectors,
iii. PCA reduces the dimension of the feature vectors by restricting attention to the directions along which the scatter of the cloud of data points are greatest.

Though PCA does not find any subset of the original feature vector, it reduces the dimensionality by orthogonally transforming the feature vectors and creates a new subset. In this work, we also compare the classifier accuracy using PCA and the proposed FS technique individually.

3.3 Classification and Cognitive Load Detection

The final step of any pattern recognition problem is to classify the pattern vectors. By classification we mean to identify the group or class of an object or pattern vectors. There is a multiple learning strategy such as Supervised Learning, Unsupervised Learning, and Reinforcement Learning etc. In this work, supervised learning strategy is used and hence, it consists of two phases – training and testing. In BCI, several literatures use the standard classifier like Support Vector Machines (SVM) [9], Artificial Neural Network (ANN) [10], k-Nearest Neighbourhood (k-NN) [11]. Also type-2 based fuzzy [1]-[5] classifier is proposed. The classifier descriptions with necessary supporting figures are discussed in this section for the convenience of the reader.

3.3.1 Support Vector Machine

Support Vector Machines (SVM) [9] is a maximum margin classifier which tri39es to maximize the distance between the hyper plane and nearest training data points or also known as support vectors.

Let us consider two classes which are levelled as +1/-1. The samples denoted by $Z=\{x^t, r^t\}$ where, x^t are the observed data points and r^t are the corresponding class level (+1/-1). The objective or training phase of the SVM classifier is to find an optimal hyper plane which best separates the training vector and also maximizes the margin between hyper plane and nearest training data points (support vectors) on the either side of the plane for better generalization. Let the hyper plane is denoted by,

$$\vec{w}^T \vec{x} + w_o \tag{19}$$

In training phase it is necessary to find the weight vector $\vec{w}$ and the bias w_0 such that

$$\begin{aligned} \vec{w}^T \vec{x} + w_o &\geq +1, \quad if \quad r = +1 \\ \vec{w}^T \vec{x} + w_o &\leq -1, \quad if \quad r = -1 \end{aligned} \tag{20}$$

Figure 3 shows the classification strategy of SVM classifier with linear hyper plane for features vectors with two feature entities for two-class problem. The data points on the dotted lines are called support vectors. Distance between two dotted lines is the margin. For multi-class problem, "*one verses all*" strategy is used and the number of hyper plane obtained is same as the number of the class present in the training dataset. The following points are to be noted while using SVM classifier is used to classify the fNIRS signal.

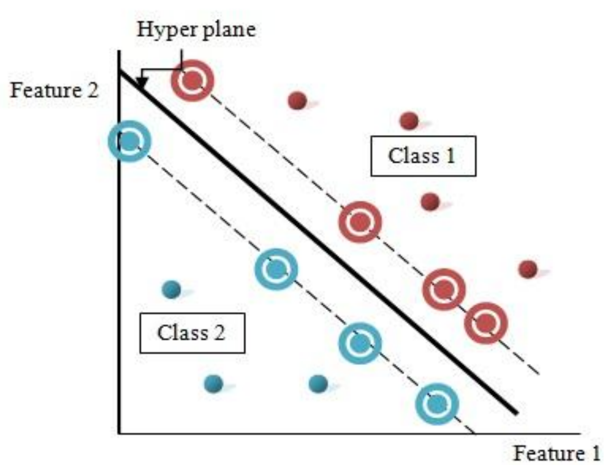


Figure 3 SVM Classifier

1. During solving the constrained optimization problem in an unconstrained manner, the equations are converted to dual form such that the complexity depends on the number of the fNIRS training samples not on the dimension of the feature vectors.
2. If the complete dataset is not linearly separable or some of the points are misclassified during training phase, which is quite obvious for physiological signals like fNIRS, one can use soft margin method by defining slack variable which will store the value of the deviation met by some of the observed variables. There are two kind of deviations – the observed data may lie on the wrong side of the hyper plane so that it is misclassified, and another one is that the data point is on the right side of the hyper plane but in the margin i.e. it is not at a sufficient distance from the hyper plane.
3. If the training set is not linearly separable at all, one can use kernel functions to project the dataset in a higher dimensional space so that they can be linearly separated, without using a nonlinear function to create a hyper plane. This is known as kernel tricks (for example, Radial Basis Function (RBF)). In this situation, the dual form optimization problem, as mentioned in point 1, is useful.

In our study we use both linear kernel and RBF kernel based SVM to classify the feature vectors. The accuracies are discussed in the next chapter. For classification of unknown feature vector obtained from unlabeled fNIRS signal, one only needs to find the sign of the outcome of the hyper plane equation.

3.3.2 *k*-Nearest Neighborhood Classifier

k-nearest neighborhood [11] is a non-parametric method for classification used in pattern recognition problem. The input is only consists of k closest examples in the feature space. An object is classified by a majority vote of its neighbours, with the object being assigned to the class most common among its *k* nearest neighbours (*k* is a positive integer, typically small). If $k = 1$, then the object is simply assigned to the class of that single nearest neighbour. The steps of the classification algorithm are enumerated below.

i. Find the distance between the unknown feature vector and the observed data points with known class label. There exist several distance metric between two vectors $\vec{x}$ and $\vec{y}$ with covariance matrix **S**, such as,

   a. Euclidean distance : $d_E(\vec{x},\vec{y}) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$,

   b. Manhattan distance : $d_{TC}(\vec{x},\vec{y}) = \sum_{i=1}^{n}|x_i - y_i|$,

   c. Mahalanobis distance : $d_M(\vec{x},\vec{y}) = \sqrt{(\vec{x}-\vec{y})^T \mathbf{S}^{-1}(\vec{x}-\vec{y})}$.

ii. Find out *k* nearest data points from the test point and their corresponding class level.

iii. Take majority voting to find out the class of the unknown test feature vector.

In this work, Mahalanobis distance metric is used, as though it is computationally complex than other two metrics, it uses the spread function in term of covariance matrix, which produce effective and better accuracy than the other two distance metrics.

3.3.3 Proposed Fuzzy Type-2 Classifier

The notion of a *type-2 fuzzy set* was brought into light by Zadeh as an extension of the concept of a normal or type-1 fuzzy set. In type-2 fuzzy set, the membership functions of the variables are itself type-1 fuzzy set. This is effective in those cases where there exist some uncertainties in determining the exact type-1 membership function. In that case, if the knowledge base itself carries so many uncertainties, these lead to rules having uncertain antecedents and/or consequents, which in turn translate into uncertain antecedent and/or consequent membership functions. In this problem, though the difficulty level of each problem is pre-determined, it is obvious that the difficulty levels vary from subject to subject, and thus, accurate rules or knowledge base is very tedious to construct. Before describing the type-2 classifier in details, I like to introduce the necessary definitions [1]-[5] to the readers and advantage of type-2 fuzzy logic to validate the effectiveness in this classification problem.

Definition 1: A T2 Fuzzy set, denoted as $\tilde{A}$ is characterized by a T2 MF $\mu_{\tilde{A}}(x,u)$, wherew $x \in X \quad and \quad u \in J_X \subseteq (0,1]$ i.e.,

$$\tilde{A} = \left\{ \left((x,u), \mu_{\tilde{A}}(x,u) \right) \mid \forall x \in X, \forall u \in J_X \subseteq (0,1], 0 \le \mu_{\tilde{A}}(x,u) \le 1 \right\} \tag{21}$$

Definition 2: The *domain* of a secondary MF is called the *primary membership* of *x*. Thus, in the equation (21), J_x is the primary membership function of *x*.

Definition 3: The *amplitude* of a secondary MF is called a *secondary grade*.

Definition 4*:* Uncertainty in the primary memberships of a T2 FS consists of a bounded region that we call the *footprint of uncertainty* (FOU). It is the union of all primary memberships i.e.,

$$FOU\left(\tilde{A}\right)=\bigcup_{x\in X} J_x \tag{22}$$

Definition 5: The *upper membership function* (UMF) and *lower membership function* (LMF) of $\tilde{A}$ are two T1 MFs that bound the FOU. UMF is the upper bound of FOU and LMF is associated with the lower bound of the FOU, and they are denoted by $\bar{\mu}_{\tilde{A}}\left(x,u\right)$ and $\underline{\mu}_{\tilde{A}}\left(x,u\right)$, i.e., $\forall x$

$$\begin{aligned}\bar{\mu}_{\tilde{A}}\left(x,u\right) &\equiv \overline{FOU\left(\tilde{A}\right)}\\ \underline{\mu}_{\tilde{A}}\left(x,u\right) &\equiv \underline{FOU\left(\tilde{A}\right)}\end{aligned} \tag{23}$$

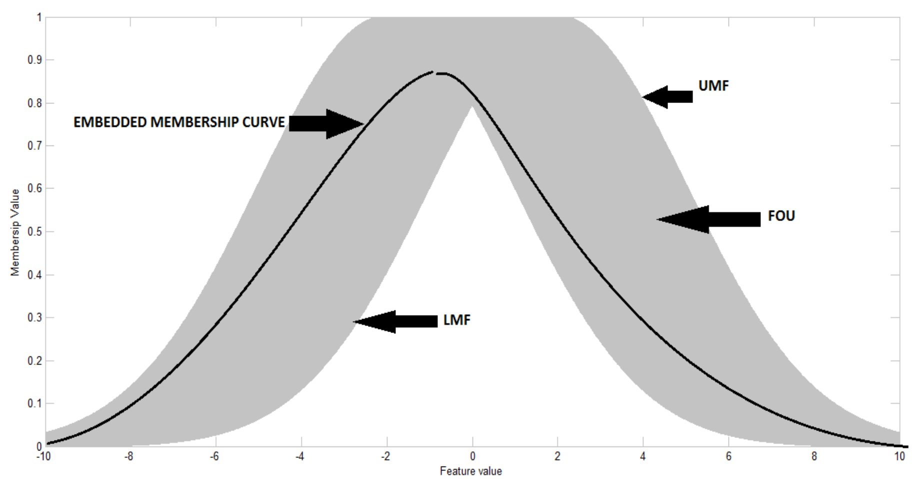


Figure 4 UMF, LMF and FOU of T2 FS

The general type-2 fuzzy classifier is used, as it provides additional degrees of freedom in designing membership functions and in modelling uncertainties than the normal type-1 fuzzy system, which relies on only one membership function. The Gaussian fuzzy membership functions for each feature entity are developed for different subjects for a particular class (Figure 6). If there is *S* number of subjects available for the experiment, then we obtain *S* number of Gaussian membership curves for a particular feature entity in a class. We also consider the secondary membership to be Gaussian in nature where the mean and variance are computed using the information from upper membership (UMF) and lower membership function (LMF) (Figure 5).

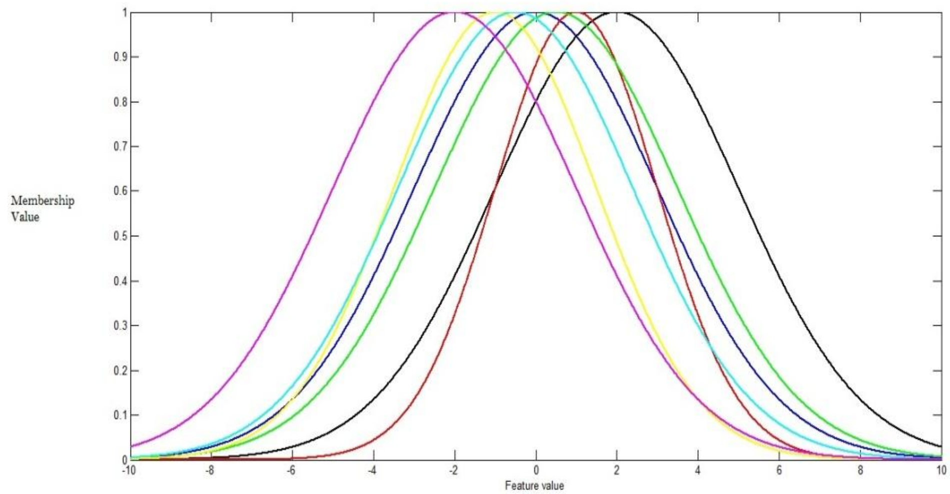


Figure 5 Gaussian T1 Membership Curve of a Feature Entity for Different Subject

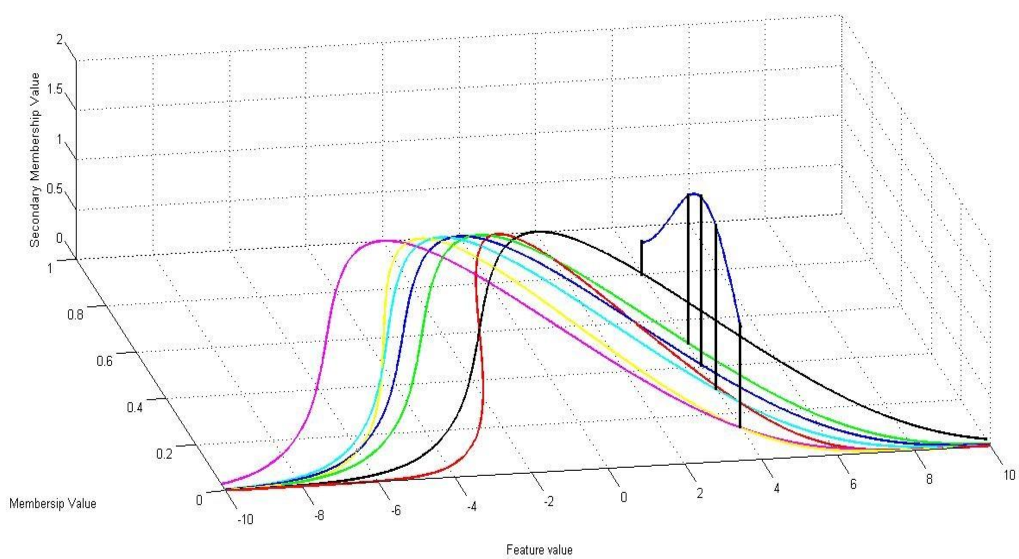


Figure 6 Gaussian T2 Membership Curve of a Feature Value x = 4

There exists different kind of rules for type reduction [1]-[5] of a general type-2 fuzzy set. When a feature vector with unknown class label is given to the system, we compute the type-1 membership value of a feature entity considering the vertical slice at that particular feature value (Figure 7) by the following mathematical expression,

$$\mu'_F(x) = \frac{\sum\limits_{\forall \mu_F(x)} \mu_{\tilde{F}}(x, \mu_F(x)).\mu_F(x)}{\sum\limits_{\forall \mu_F(x)} \mu_{\tilde{F}}(x, \mu_F(x))} \tag{24}$$

where, $\mu_F(x)$ is the primary membership value of feature F at a value x and $J(x, \mu_F(x))$ is the secondary membership value of the primary membership $\mu_F(x)$. The summation is taken over all the primary membership values found from the S number of subjects between UMF and LMF. We find all of the membership values of all the features for all the classes and store them in matrix $\mathbf{F}$ whose rows are indicating the membership values of the features for a class and columns are indicating the membership value of a feature entity in all the classes.

After getting the matrix $\mathbf{F}$, we find the average membership value in a particular class by summing along the column of a matrix, expressed as,

$$\bar{\mu} = \frac{\sum\limits_{\forall F} \mu'_F}{No.\quad of \quad features} \tag{25}$$

The class with highest average membership value is the class corresponds to the unknown feature vector. The corresponding algorithm for Type-2 fuzzy classification is given in table II.

Table II Type-2 Fuzzy Classification

Algorithm 2: Fuzzy Type- 2 Classification

1. Find the Gaussian Membership Function of a feature entity for a particular subject whose feature vectors are in training set by computing the mean and variance of that feature entity from the set of training vectors corresponding to that subject.
2. Repeat step 1 for all the feature entities.
3. Repeat steps 1 and 2 for all the subjects in the training set. Also obtain the same for different classes of training vectors.
4. In the classification stage, take vertical slices at the feature value of a particular feature entity in the fuzzy space for different classes.
5. Compute the mean and the variance of the T1 membership values corresponding to the vertical slices.
6. Compute the secondary Gaussian membership values at that particular feature value along the vertical slices.
7. Compute the weighted type-2 membership value using equation (24).
8. Repeat steps 4-7 for all other unlabeled feature entity values.
9. Store the weighted membership values for different classes of particular feature entity in a column vector and merge the column vectors vertically to get a matrix **F** of dimension $\boldsymbol{K} \times \boldsymbol{d}$, where K is the number of classes and d is the dimension of the feature vectors.
10. Obtain the class of the unlabeled feature vector using equation (25).

After determining the class, we now compute the load on each subject by the value of the average membership in a class, which is discussed in details in the next chapter.

References

[1] Mendel, Jerry M., and Robert I. Bob John. "Type-2 fuzzy sets made simple."*Fuzzy Systems, IEEE Transactions on* 10, no. 2 (2002): 117-127.

[2] Karnik, Nilesh N., and Jerry M. Mendel. "Operations on type-2 fuzzy sets."*Fuzzy sets and systems* 122, no. 2 (2001): 327-348.

[3] Mendel, Jerry M., Robert I. John, and Feilong Liu. "Interval type-2 fuzzy logic systems made simple." *Fuzzy Systems, IEEE Transactions on* 14, no. 6 (2006): 808-821.

[4] Karnik, Nilesh N., and Jerry M. Mendel. "Centroid of a type-2 fuzzy set."*Information Sciences* 132, no. 1 (2001): 195-220.

[5] Karnik, Nilesh N., Jerry M. Mendel, and Qilian Liang. "Type-2 fuzzy logic systems." *Fuzzy Systems, IEEE Transactions on* 7, no. 6 (1999): 643-658.

[6] Bishop, Christopher M. "Pattern Recognition." *Machine Learning* (2006).

[7] James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*. Vol. 112. New York: springer, 2013.

[8] Zhao, Haitao, Pong Chi Yuen, and James T. Kwok. "A novel incremental principal component analysis and its application for face recognition."*Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*36, no. 4 (2006): 873-886.

[9] Distante, Cosimo, Nicola Ancona, and Pietro Siciliano. "Support vector machines for olfactory signals recognition." *Sensors and Actuators B: Chemical* 88, no. 1 (2003): 30-39.

[10] Hecht-Nielsen, Robert. "Theory of the backpropagation neural network." In*Neural Networks, 1989. IJCNN., International Joint Conference on*, pp. 593-605. IEEE, 1989.

[11] Song, Yang, Jian Huang, Ding Zhou, Hongyuan Zha, and C. Lee Giles. "Iknn: Informative k-nearest neighbor pattern classification." In *Knowledge Discovery in Databases: PKDD 2007*, pp. 248-264. Springer Berlin Heidelberg, 2007.

CHAPTER 4

Experiment and Analysis

This chapter is dedicated toward the detail description of fNIRS experiments done in Artificial Intelligence Lab of Jadavpur University. In Section 4.1, the experimental set up is described along with the detail timeline of the stimulus used for this purpose. In Section 4.2, features are discussed with their significance. The pre-processing steps are discussed in the next section. Section 4.4 gives a detail description and comparison of the outcomes of the classifiers with statistical significance. The chapter is concluded with the load detection for individual subjects in different problems of different difficulty level.

4.1 Experiment

So far, we have discussed about the theoretical background necessary for working with fNIRS systems and also the proposed feature selection and classification techniques for cognitive load detection. This chapter includes a brief description of fNIRS device used, experimental setup along with subject and stimuli details, and experiments: to i) extract fNIRS features, ii) select the most significant features from a large pool of extracted features, iii) detect cognitive load for easy, moderate and hard mental tasks, iv) compare the classifier performance and v) to validate classifier performance using McNemar's statistical test.

The experiment has been performed at Artificial Intelligence Lab, Jadavpur University, where the brain response of human subjects is captured using a popular brain-imaging device called fNIRS (Figure 8). This device has been selected for the present problem because of its non-invasiveness, capability to localize and measure oxy-haemoglobin and de-oxy-haemoglobin, low-cost and portability [1]. It can be shown form Figure 8(b) that fNIRS band which has been used in this experiment, has 4 infra-red (IR) sources and 10 IR detectors. The path from IR source to detector of an fNIRS device is termed as channel, which provides measurements of oxy-hemoglobin (HbO) and de-oxy-hemoglobin (HbR) blood concentration. Therefore, the present fNIRS band captures brain images from 16 channels when the band is attached to the forehead of human subjects during the experiment.

Figure 9 shows an experimental framework; where a subject is asked to perform mental mathematical operations only by pointing out the correct answer from a set of four options. fNIRS band is placed on her forehead to measure cerebral blood oxygen concentration during the trial as pre-frontal cortex data are useful for this type of experiment. Eight such healthy subjects of ages between 21 and 27 years are selected to perform the experiments. They are instructed to restrict their movement in order to avoid unwanted movement-related artifacts to some extent.

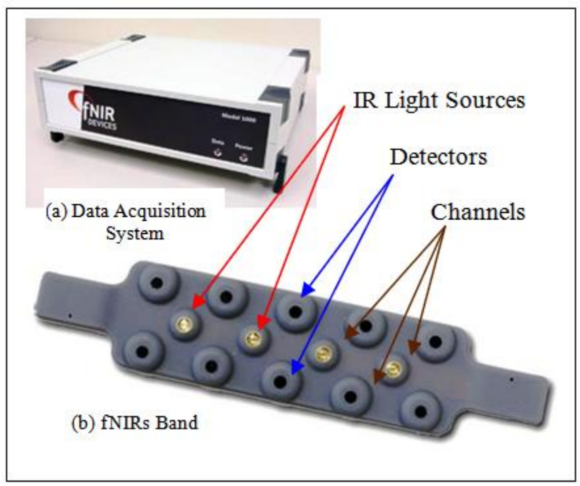


Figure 7 fNIRS device to capture brain images during cognitive tasks

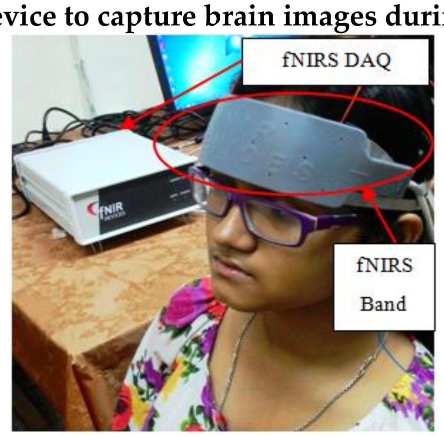


Figure 8 A subject is performing cognitive tasks while brain images are captured using fNIRS device

An experimental trial contains five mental mathematical problems, each of 20 seconds and a 5-second rest between each problem. The stimulus began with "start" instruction followed by a slide containing the instructions for the subjects for subsequent tasks and a 5 second of rest time respectively. Figure 10 shows the timeline of the stimulus used for performing the tasks. For the present problem, three kinds of experimental trial have been prepared based on three difficulty levels: easy (E), moderate (M) and hard (H). Each subject has to perform each of these three trials for 7 times, resulting in 35 experimental instances for each difficulty level. The cognitive load of the subjects is also classified into three distinct levels including low, medium and high, depending on the difficulty levels: E, M and H respectively.

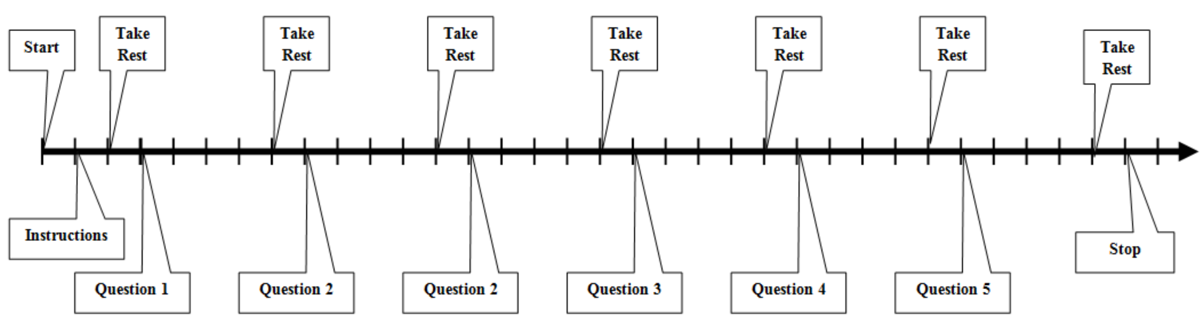


Figure 9 Time Flow of the stimulus Used. The scale is divided into 5sec segments

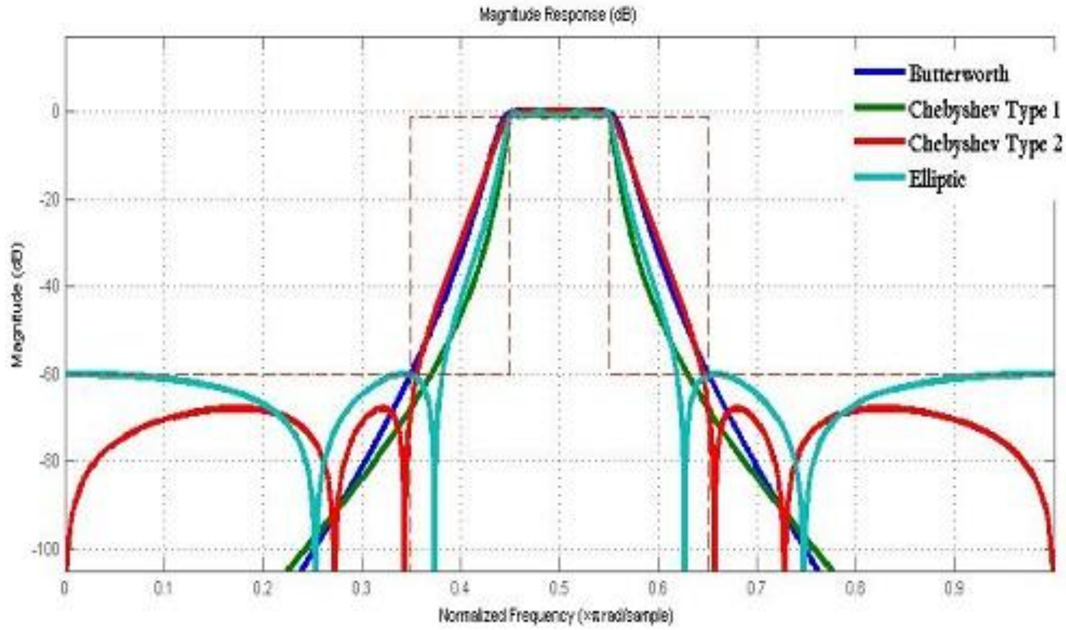


Figure 10 Frequency response of a Butterworth, Chebyshev-I, Chebyshev-II, and Elliptic Filters

4.2 Experiment 1: Preprocessing

Before further processing like feature extraction and selection, classification etc, we need to eliminate any artifacts present in the signal, else it will create disturbance and results in faulty classification of signals. The signals first undergo ICA [2] technique for removing physical noises and also for reducing the interference of the signals from other channels by extracting the independent components of 16 channels. In next step, these signals are band-passed using a Chebyshev type 1 IIR filter of order 10. Though several literatures uses different band-pass filters like Chebyshev type 1, type 2, Butterworth, Elliptic; in this literature, Chebyshev type 1 filter is used. Figure 11 shows that the sharp roll-off and good attenuation in stop-band and low ripples are obtained for Chebyshev type 1 filter than others keeping the order fixed. The two cut-off frequencies are defined as 0.1 Hz and 0.4 Hz for band-pass filtering. Figures 12-19 show the pre-processing output of oxy-haemoglobin (Figure 12-15) and de-oxy-haemoglobin (Figure 16-19) after raw signals are provided from randomly chosen 4 channels.

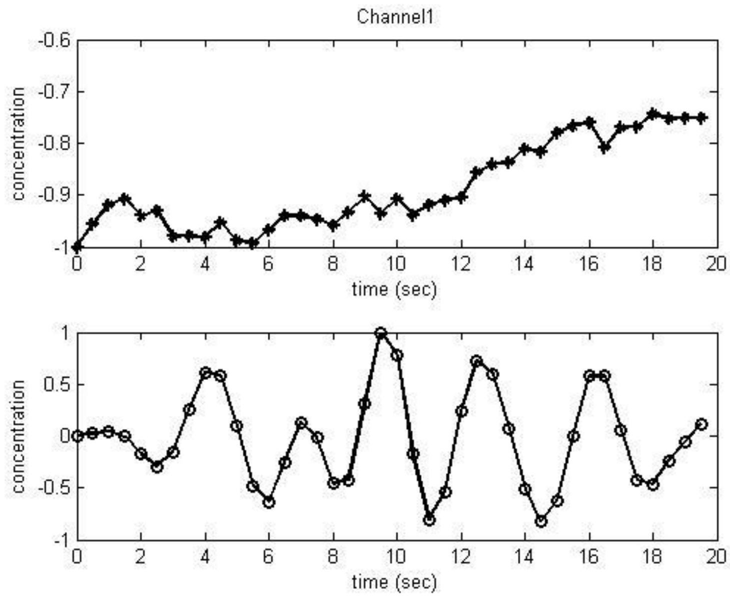


Figure 11 Pre-processing output of oxy-heamoglobin captured in Channel 1

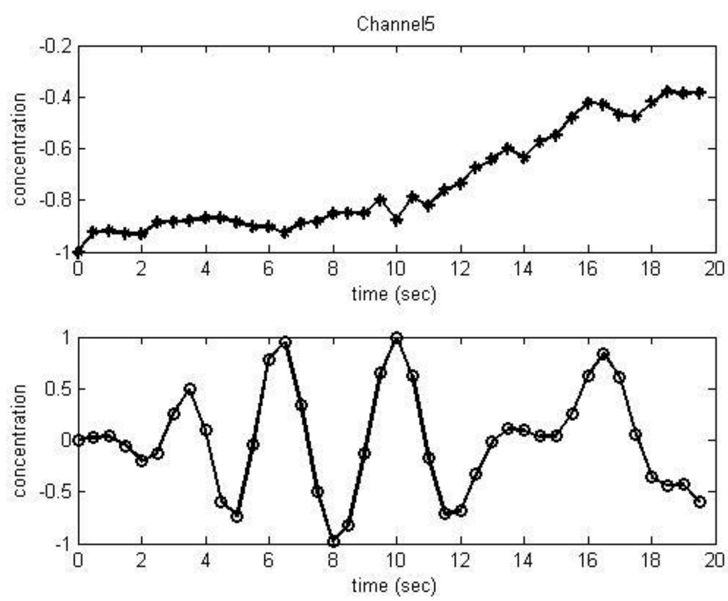


Figure 12 Pre-processing output of oxy-heamoglobin captured in Channel 5

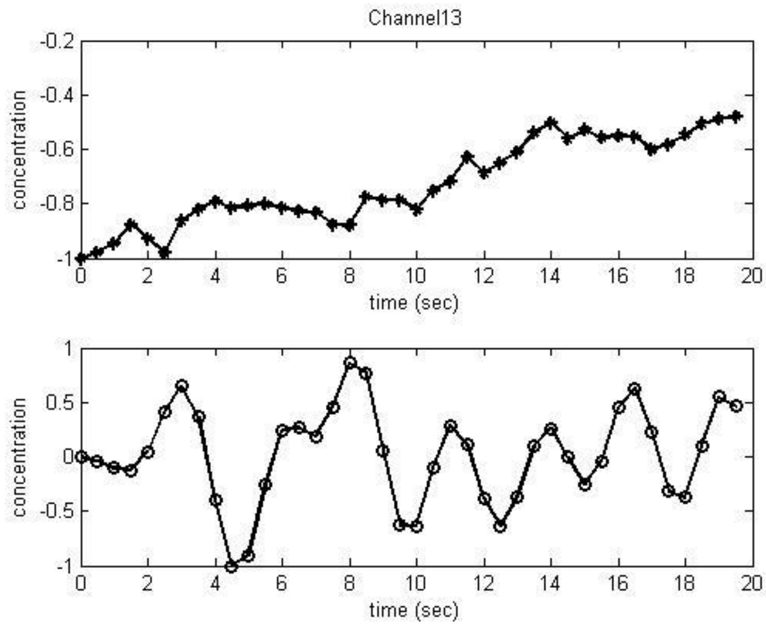


Figure 13 Pre-processing output of oxy-heamoglobin captured in Channel 13

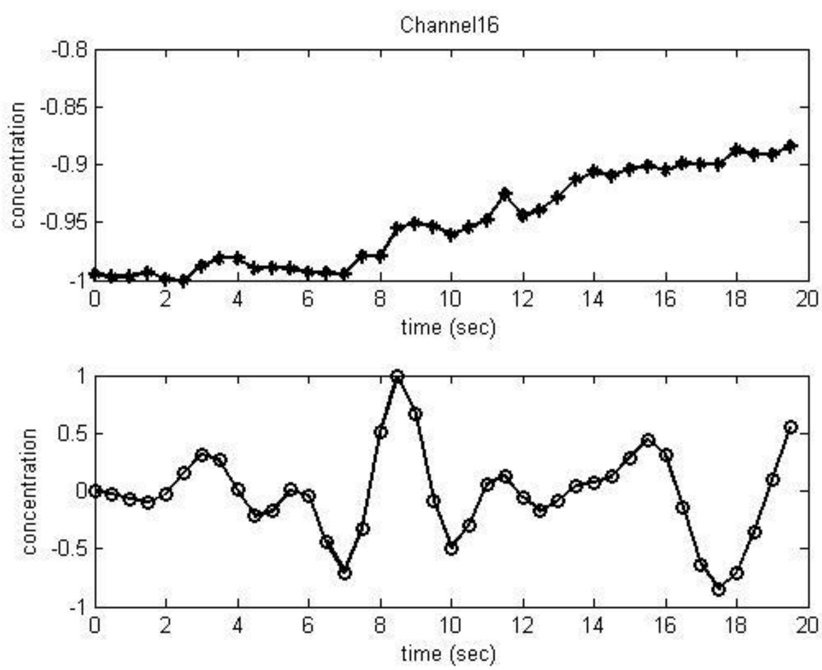


Figure 14 Pre-processing output of oxy-heamoglobin captured in Channel 16

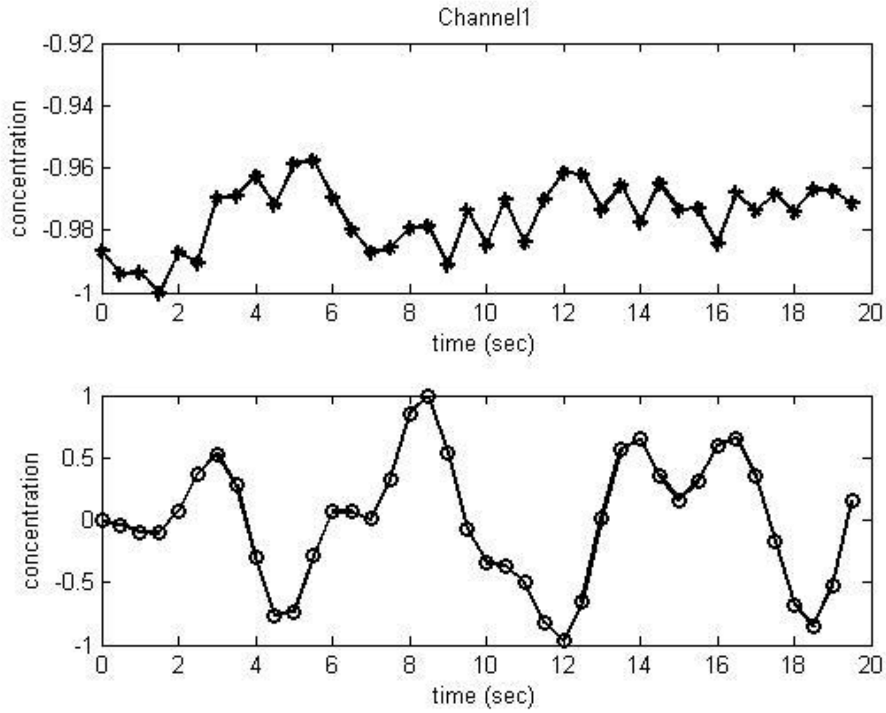


Figure 15 Pre-processing output of de-oxy-heamoglobin captured in Channel 1

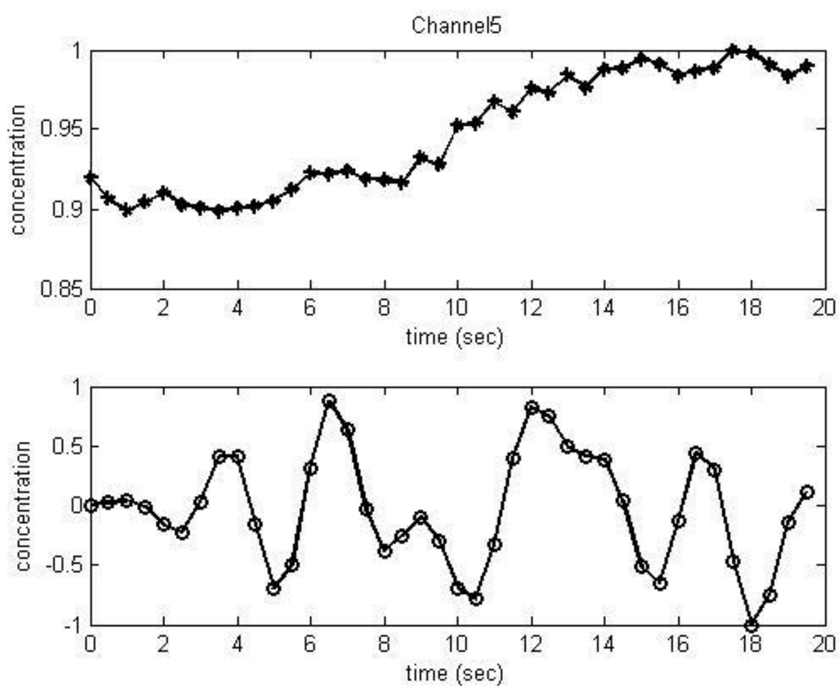


Figure 16 Pre-processing output of de-oxy-heamoglobin captured in Channel 5

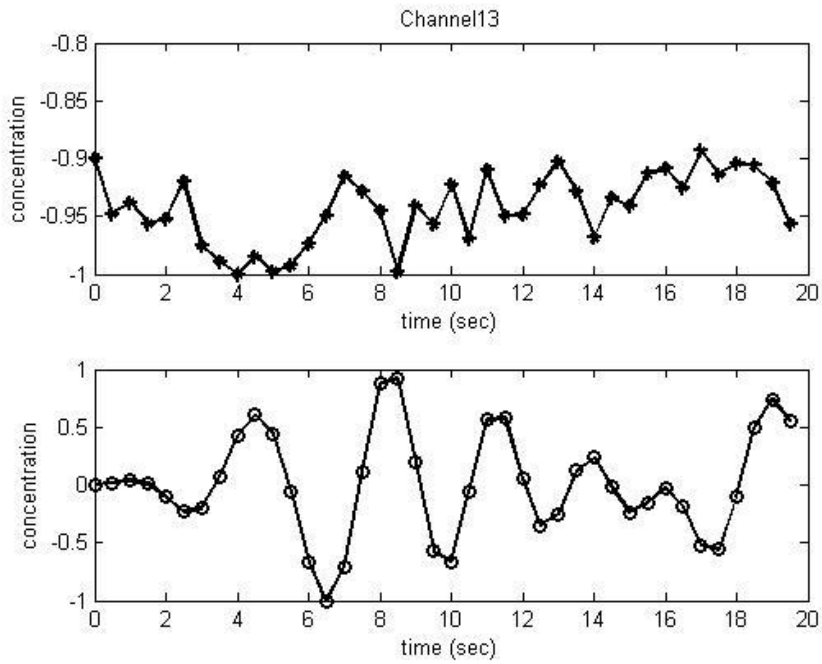


Figure 17 Pre-processing output of de-oxy-heamoglobin captured in Channel 13

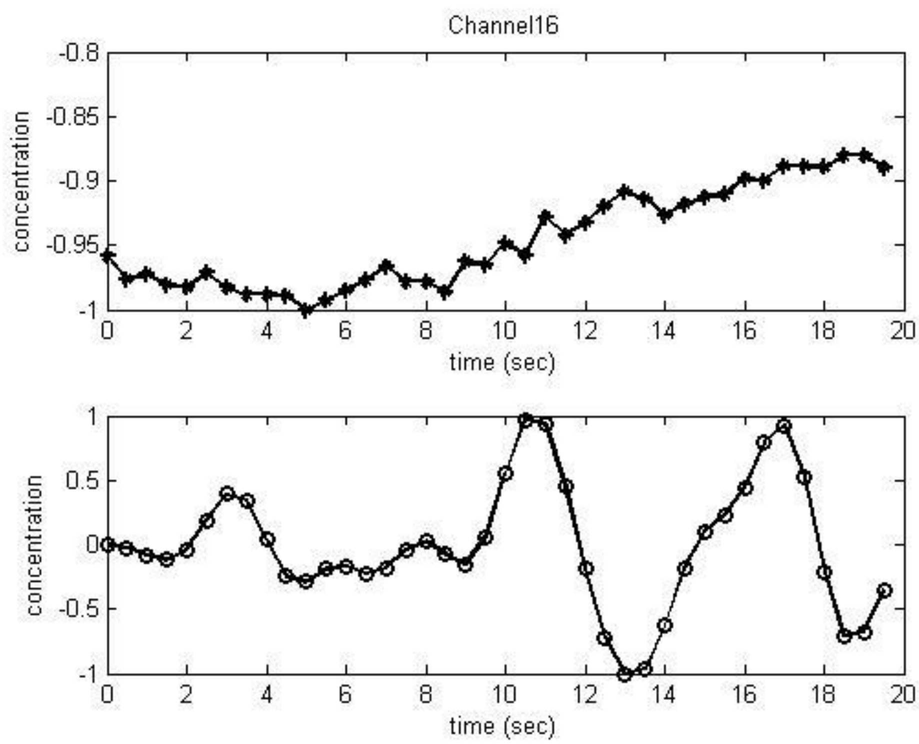


Figure 18 Pre-processing output of de-oxy-heamoglobin captured in Channel 16

4.3 Experiment 2: Feature Extraction

Feature extraction (FE) is the second and one of the most important steps in pattern classification problem, since feature can best describe the pattern itself which is hidden within a raw signal. In this experiment, we have selected a list of features that can be used to extract necessary information from fNIRS data to detect the cognitive load of human subjects. We consider seven following feature sets.

i. F_1 : Mean value of HbO concentration.
ii. F_2 : Mean value of HbR concentration.
iii. F_3 : Mean value of HbO+HbR concentration.
iv. F_4 : Variance of HbO+HbR concentration.
v. F_5 : Mean of HbO-HbR concentration.
vi. F_6 : Variance of HbO-HbR concentration.
vii. F_7 : Average slop of HbO-HbR concentration.

To perform FE, we start with fNIRS data acquired for three different levels of cognitive tasks: E, M and H. For each kind of difficulty level, we obtain HbO and HbR data, each having dimension of 35×16×40, where, 35 represents the number of experimental instances, 16 represents the number of channels of fNIRS and 40 represents the samples. To extract first feature set $F_{1,}$ we take mean of HbO data across samples recorded by each channel and obtain 16 features for each of 35 instances. Therefore, F_1 contains mean-HbO features having dimension of 35×16. In similar fashion, feature sets F_2-F_7 is prepared by determining their respective 16 features for 35 instances, which finally present a feature matrix of 35×112 dimension for each difficulty level for each subject.

A clear discrimination in concentration level from 16 channels has been observed for each feature set. For better understanding, we provide feature level discrimination of 7 features determined from randomly chosen 4 channels, while the 3 subjects are performing cognitive tasks during one experimental trial. Figures 20-23, 24-27, 28 present the feature level discrimination of seven features between 3 subjects determined from 4 channels for easy, medium and hard levels respectively. It is clearly from these figures that the features are quite similar between these subjects for a particular difficulty level task.

In a second measure, we have also shown that the features are quite discriminative in nature for different difficulty levels taken from 4 channels for a particular subject. Figures 22-25 are showing these characteristics in the feature set. Thus, choice of these features is quite effective in terms of classification.

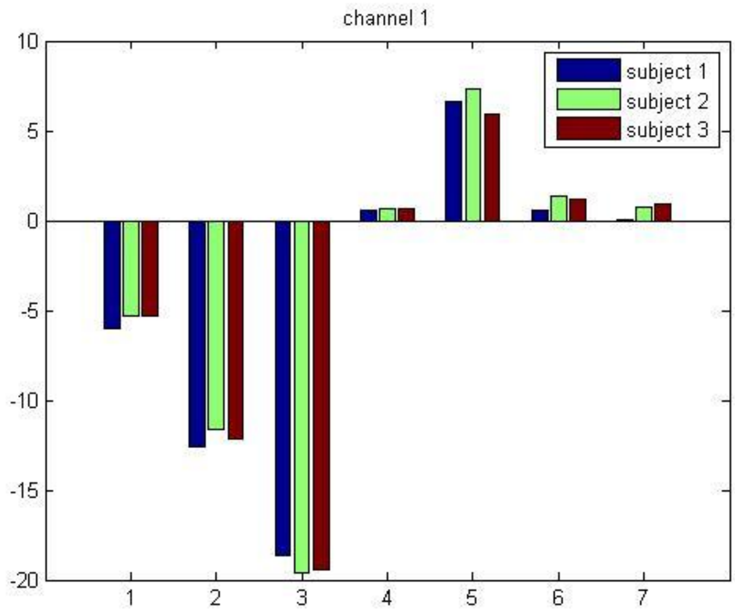


Figure 19 Feature Level Discrimination for Easy Task captured in channel 1

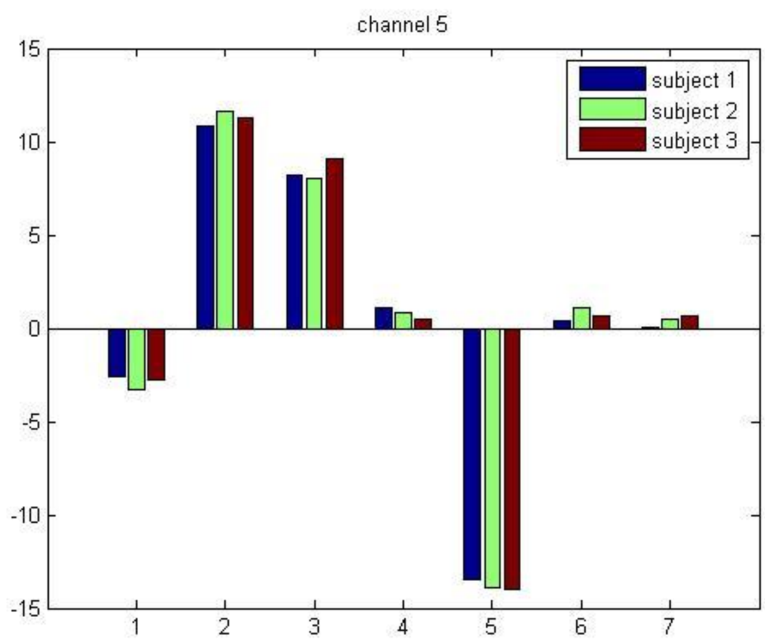


Figure 20 Feature Level Discrimination for Easy Task captured in channel 5

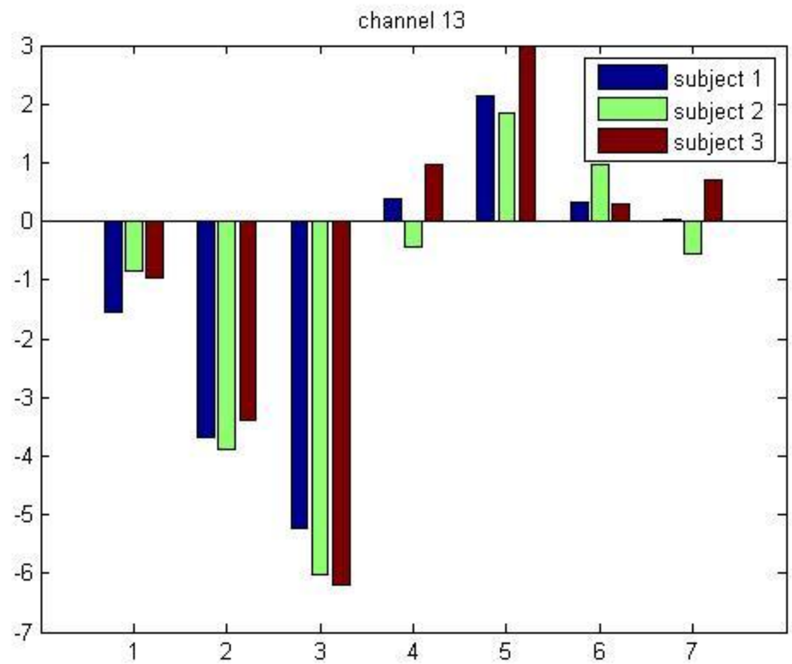


Figure 21 Feature Level Discrimination for Easy Task captured in channel 13

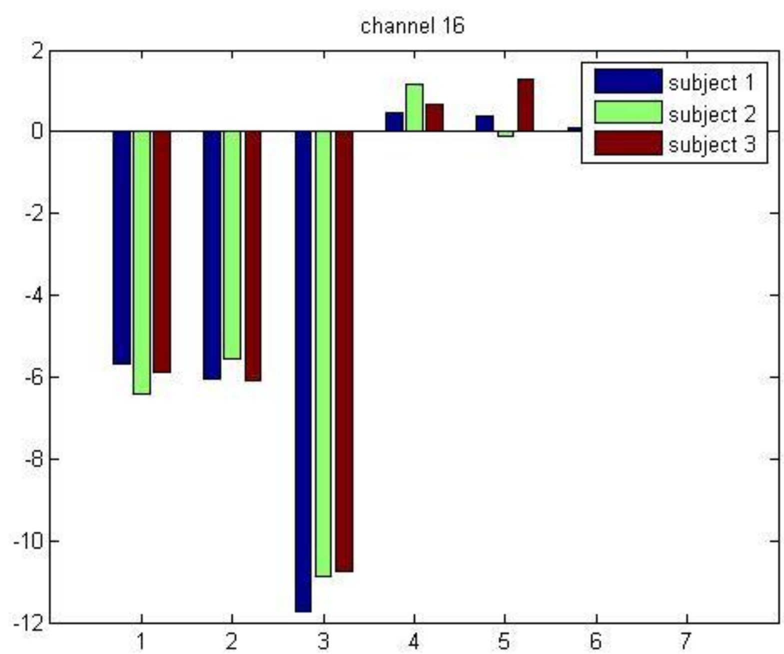


Figure 22 Feature Level Discrimination for Easy Task captured in channel 16

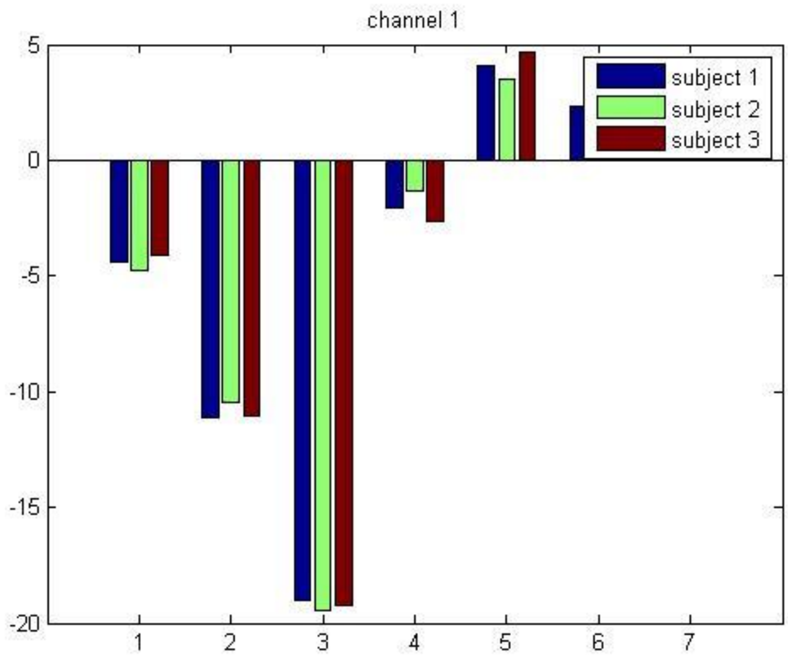


Figure 23 Feature Level Discrimination for Medium Task captured in channel 1

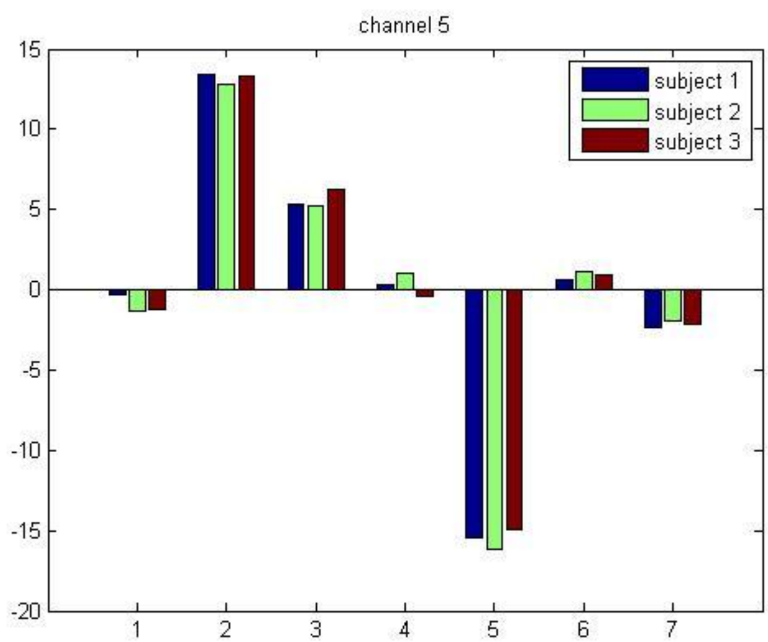


Figure 24 Feature Level Discrimination for Medium Task captured in channel 5

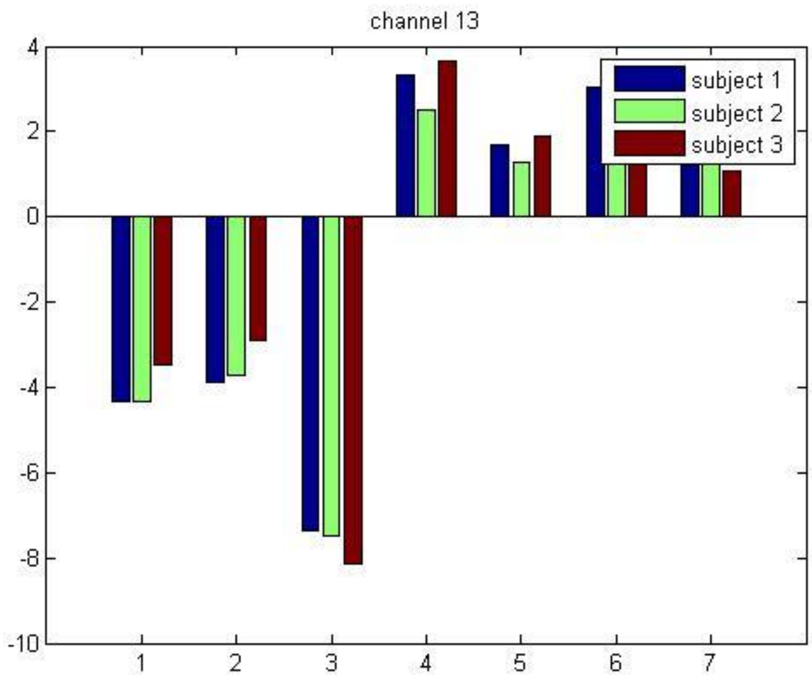


Figure 25 Feature Level Discrimination for Medium Task captured in channel 13

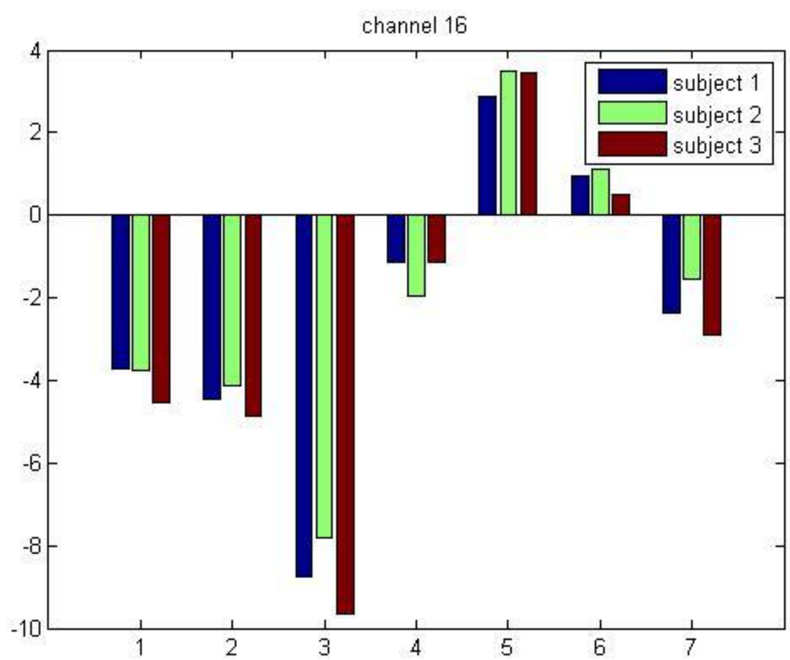


Figure 26 Feature Level Discrimination for Medium Task captured in channel 16

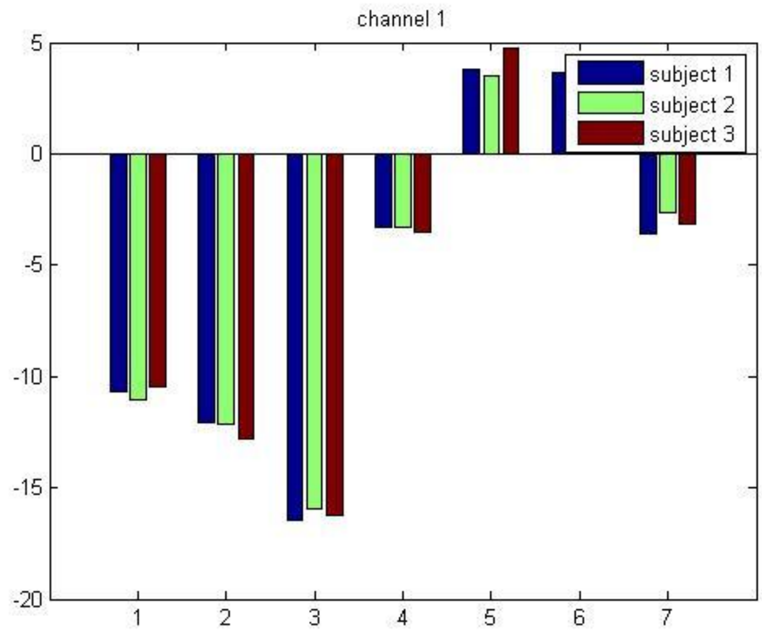


Figure 27 Feature Level Discrimination for Hard Task captured in channel 1

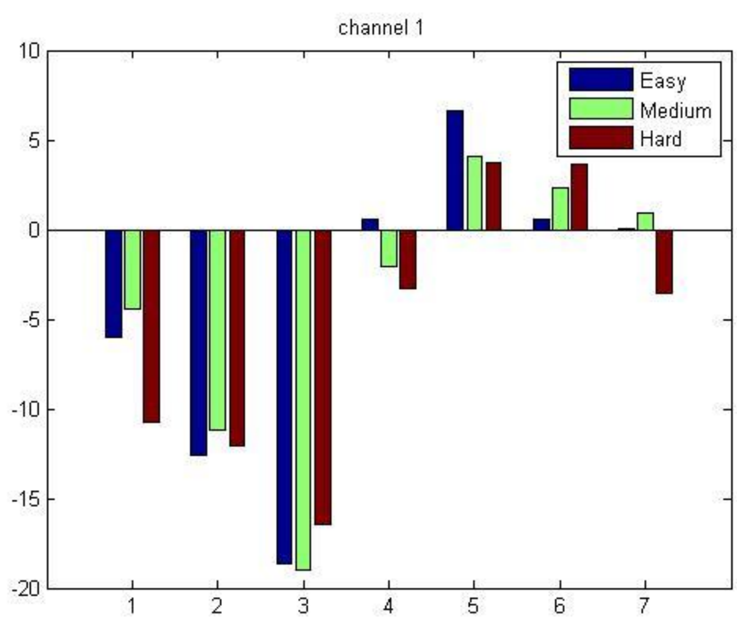


Figure 28 Feature Level Discrimination for Different Task captured in channel 1

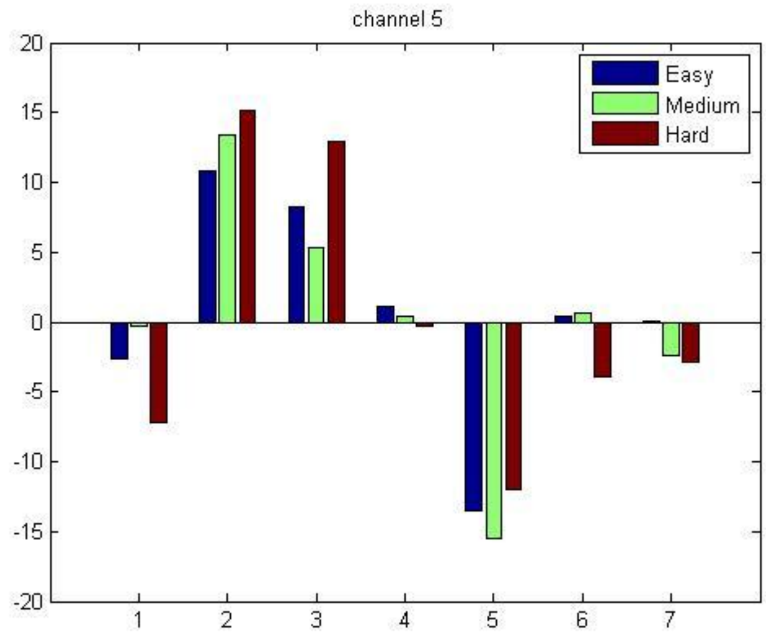


Figure 29 Feature Level Discrimination for Different Task captured in channel 5

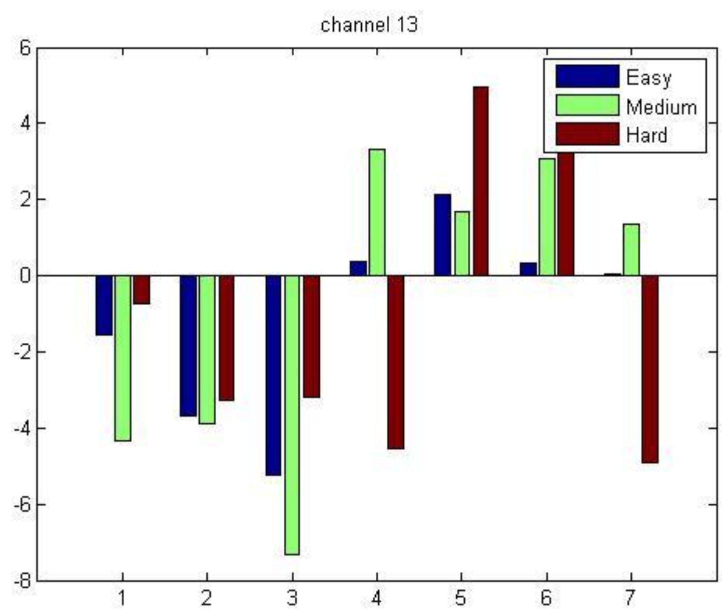


Figure 30 Feature Level Discrimination for Different Task captured in channel 13

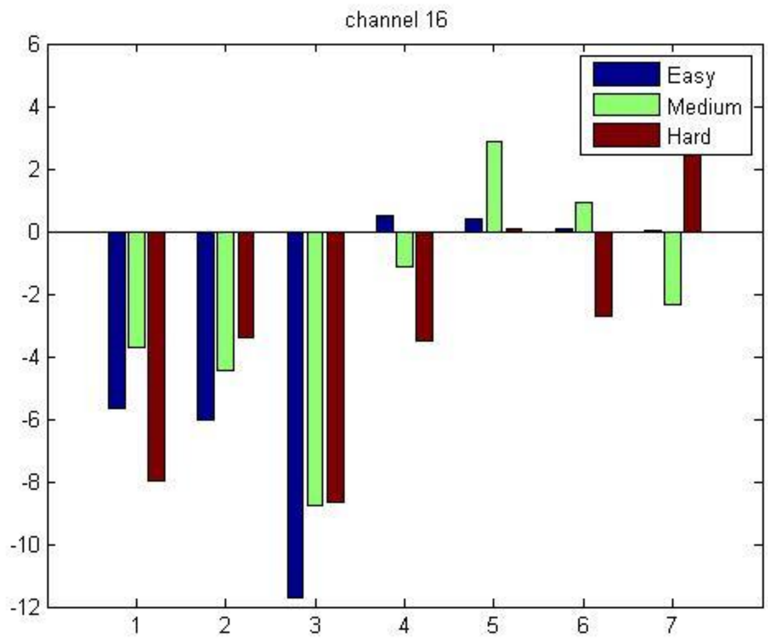
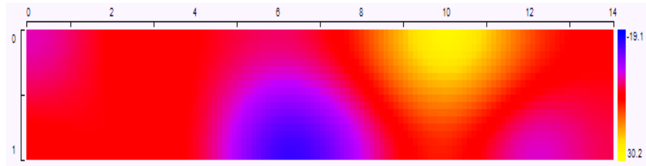


Figure 31 Feature Level Discrimination for Different Task captured in channel 16

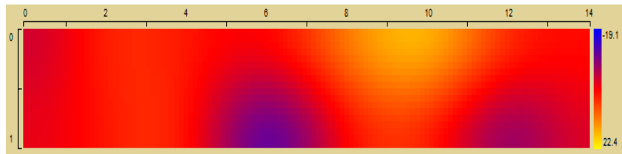
4.4 Experiment 3: Feature Selection

Selection of appropriate features is necessary to classify any pattern recognition problem, if the extracted feature set is found significantly large and the no of samples available for training is small in number. Having a large dimensional feature set (here, 112), this experiment also utilizes a DE-induced likelihood-based feature selection (FS) technique that optimally selects 25 most significant features that are able to correctly represent the pattern vectors.

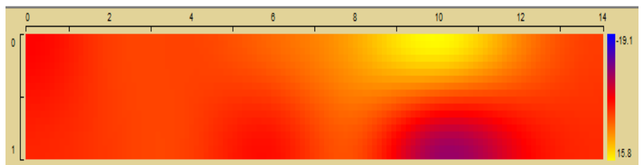
The performance of the proposed FS technique is compared with the well-known principal component type-2 analysis (PCA), when the selected features are fed to the fuzzy classifier. Result is given in Table III, wherefrom it can be concluded from the table that the proposed FS technique provides better classification accuracy than PCA [4].



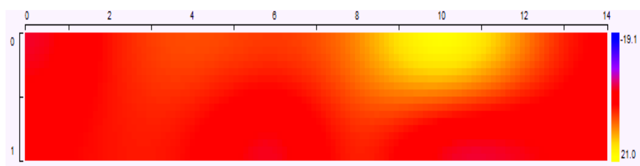
(a)



(b)



(c)



(d)

Figure 33. Oxygenation level recorded from 14 sensor-detector pair. (a): inactive condition, (b): easy task, (c): moderate task, (d): hard task

Table III Comparison between Mean and Standard Deviation of Fuzzy Type-2 Classifier with PCA Based and Proposed Likelihood Based FS Technique

Features Dimensions	Average Classifier Accuracy (in %)		Statistical Significance
	PCA + Fuzzy Type-2 Classifier	Proposed likelihood + Fuzzy Type-2 Classifier	
112 (Reduced to 25)	88.791 (0.01407)	**92.597 (0.01207)**	+

4.6 Experiment 5: Cognitive Load Detection for Different Difficulty Level

This experiment deals with classification of cognitive load into three classes: low, medium and hard and also determine subjective load for E, M, and H levels of cognitive tasks. For this we determine fuzzy membership values for each level (class) of cognitive load: low, medium and hard. For a particular class, we quantified the range of fuzzy memberships (0,1] in three levels as shown in the Table IV.

Table IV Measurement of Cognitive Load According To Average Fuzzy Membership Values

Range of Fuzzy Membership Values	Cognitive Load
0.0-0.3	Low
0.3-0.7	Medium
0.7-1.0	High

The membership ranges, as mentioned in Table IV, are used to detect subjective load while the particular subject is offered easy, moderate and hard cognitive tasks one by one. The cognitive load for randomly selected four subjects are presented in Table V, which gives a clear indication that using the proposed technique, cognitive load differs subject to subject for the same degree of complexity.

In order to detect cognitive load distribution of the subjects with increasing difficulty level, oxygenation level is captured from fNIRS and presented in Figure 33 (a)-(d). From Figure 33 (a)-(d), we can observe the change in oxygenated blood volume of a subject with the increase in difficulty level of a task. From the color bar shown in figure, it is clear that there is a high rise of brain activity with the increase in oxygenation, the highest of which is represented by yellow color, whereas the lowest brain activation is represented by blue color. Now, from figure 33 (a), it is evident that there requires less oxygenated blood during inactive (rest) situation. In addition, it is also prominent from figure 33 (a) that in inactive situation, brain activity becomes low, which results in the appearance of blue region. Figure 33 (b) shows the oxygenation level during easy task performance, where blue regions as well as its intensity tend to get lower. In figure 33 (c), difficulty level of the task becomes moderate, where the human brain needs to perform more activity, thus flow of oxygenated blood tends to get higher, so more reddish/yellowish region appears, whereas blue region decreases significantly. Lastly, Figure 33 (d) shows the oxygenation for hard task. This task requires the highest brain activation, which is evident from only the red/yellow regions near the forebrain, and no de-oxygenation takes place.

Table V Subject Wise Cognitive Load Detection for 4 Subjects Corresponds to Table IV

Class Subject	Load (In Alphabetic Form)		
	Easy	Moderate	Hard
1	Low	Medium	Medium
2	Medium	High	High
3	Low	Low	Medium
4	Medium	High	High

4.7 Experiment 6: Proposed Fuzzy Type-2 classifier Performance

We have compared type-2 fuzzy classifier performance using proposed likelihood-based FS algorithm with linear support vector machine (LSVM) [5], k-nearest neighbor (kNN) with Euclidean distance metric [6] and support vector machine with radial basis function (SVM-RBF) [8] classifiers (Table VI). The classifier performance of kNN is averaged over three different values of *k* (3, 5, 7). Table VI indicates that the proposed likelihood-based FS induced type 2 fuzzy classifier attains the highest classification accuracy (above 90% in each case) as compared to its standard competitors. The last column of the table represents statistical significance of the difference of the means of best two algorithms. Positive (+) significance means that if we use two-tailed test, then the t value of 49 degrees of freedom becomes significant at a 0.05 level of significance. Negative (-) significance indicates not statistically significant and 'NA' represents the cases for which two or more algorithms best accuracy results.

Table VII provides the individual class performance by using confusion matrix of four different classes (easy, moderate, hard and inactivity) while implementing type-2 fuzzy classifier and proposed likelihood-based FS technique. Table VII indicates that the classification accuracy for the individual class is high, over 90%

Table VI Mean and Standard Deviation of Classifier Accuracy Using Proposed Likelihood Based FS Technique

Features Dimensions	Difficulty Level / Class	Average Classifier Accuracy (In %)				Statistical Significance
		L-SVM	KNN	SVM-RBF	Type-2 Fuzzy Classifier	
112 (Reduced to 25)	Easy	75.156 (0.01936)	77.431 (0.0)	83.174 (0.01493)	**93.797 (0.01949)**	+
	Medium	72.309 (0.01584)	76.139 (0.01894)	80.197 (0.01346)	**92.647 (0.01719)**	+
	Hard	71.473 (0.01575)	73.164 (0.01795)	76.197 (0.01976)	**90.794 (0.01019)**	+
	Inactivity	74.794 (0.02941)	76.981 (0.01349)	76.197 (0.01976)	**93.147 (0.02009)**	+

4.8 Experiment 7: McNemar’s Statistical Test

McNemar’s test [9] is one popular statistical test to compare the relative performance of the proposed algorithm with existing standard techniques. Here, we too apply McNemar’s test to compare performance of the proposed likelihood-based feature selection induced type-2 fuzzy classifier with the above mentioned three standard classifiers including LSVM, kNN and SVM-RBF. The results of Mcnemar’s test, as reported in Table VIII depends on the value of the parameter *p*, where *p* indicates the estimated probability of rejecting the null hypothesis of a study question when that hypothesis is true. Table VIII confirms that the proposed classifier outperforms all its competitors by a wider margin.

Table VIII Confusion Matrix of Four Different Classes Using Fuzzy Type-2 Classifier and Proposed Likelihood-Based FS Technique

Predicted Class Actual Class	Easy	Moderate	Hard	Inactivity
Easy	**93.797**	2.458	1.0977	2.6473
Moderate	2.598	**92.647**	1.913	2.842
Hard	3.892	3.197	**90.794**	2.117
Inactivity	3.414	1.941	1.498	**93.147**

Table VII Statistical Test

Classifier algorithm used for comparison (features:25)	REFERENCE ALGORITHM: LIKELIHOOD-BASED FS INDUCED TYPE-2 FUZZY CLASSIFIER	
	Z	p
L-SVM	36.257	p<0.00001
kNN	21.145	p<0.00001
SVM-RBF	7.145	p<0.00001

References

[1] Seraglia, Bruno, Luciano Gamberini, Konstantinos Priftis, Pietro Scatturin, Massimiliano Martinelli, and Simone Cutini. "An exploratory fNIRS study with immersive virtual reality: a new method for technical implementation." *Front. Hum. Neurosci* 5, no. 176 (2011): 10-3389.

[2] Hyvärinen, Aapo, Juha Karhunen, and Erkki Oja. *Independent component analysis*. Vol. 46. John Wiley & Sons, 2004.

[3] Naseer, Noman, and Keum-Shik Hong. "fNIRS-based brain-computer interfaces: a review." *Frontiers in human neuroscience* 9 (2015).

[4] Zhao, Haitao, Pong Chi Yuen, and James T. Kwok. "A novel incremental principal component analysis and its application for face recognition."*Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*36, no. 4 (2006): 873-886.

[5] Saha, Simanto, Soupayan Datta, Amit Konar, and Ramadoss Janarthanan. "A study on emotion recognition from body gestures using Kinect sensor." In *Communications and Signal Processing (ICCSP), 2014 International Conference on*, pp. 056-060. IEEE, 2014.

[6] Song, Yang, Jian Huang, Ding Zhou, Hongyuan Zha, and C. Lee Giles. "Iknn: Informative k-nearest neighbor pattern classification." In *Knowledge Discovery in Databases: PKDD 2007*, pp. 248-264. Springer Berlin Heidelberg, 2007.

[7] Hecht-Nielsen, Robert. "Theory of the backpropagation neural network." In*Neural Networks, 1989. IJCNN., International Joint Conference on*, pp. 593-605. IEEE, 1989.

[8] Distante, Cosimo, Nicola Ancona, and Pietro Siciliano. "Support vector machines for olfactory signals recognition." *Sensors and Actuators B: Chemical* 88, no. 1 (2003): 30-39.

[9] Saha, Anuradha, Amit Konar, Avhishek Chatterjee, Anca Ralescu, and Atulya K. Nagar. "EEG analysis for olfactory perceptual-ability measurement using a recurrent neural classifier." *Human-Machine Systems, IEEE Transactions on* 44, no. 6 (2014): 717-730.

CHAPTER 5

Conclusion and Future Scope

The thesis is concluded with this chapter. In section 5.1, the conclusion and fruitfulness of the work is discussed briefly. In the next section, an outline is provided for future research works that can be done in this domain.

5.1 Feature Selection

This paper proposes an interesting approach to detect cognitive load of human subjects using fNIRS signal. Seven different feature sets are utilized here to identify right features to classify cognitive load. Type-2 fuzzy classifier is used as it gives additional degree of freedom to model the uncertainties a dataset carries within than the ordinary type-1 fuzzy logic. Without depending upon a single rule-base, it takes into consideration a perturbation in an uncertainty to create a more flexible rule-based for fuzzy reasoning. The fuzzy membership ranges have been defined to detect subjective load while the particular subject performs easy, moderate and hard cognitive tasks. Cognitive load distribution of the subject with increasing difficulty level is analyzed from the oxygenation level recorded using fNIRS device. It is clear from the experimental results that the proposed likelihood-based FS induced type 2 fuzzy classifier attains the highest classification accuracy (above 90% in each case) in comparison to its standard competitors including LSVM, kNN, BPNN and SVM-RBF. Experiment also reveals that individual class performance (easy, moderate, hard and inactivity) by using type-2 fuzzy classifier and proposed likelihood-based FS technique is high, over 90%.

5.2 Scope of Future Research

Though, a lot of works has been done cognitive load detection and also on the pre-frontal cortex brain signal using fNIRS device for mental arithmetic stability, there lay a lot of scopes in working in this domain.

According to free lunch theorem, no classification algorithm can be termed as best. Thus, the first scope lies in designing better classifier for dealing with this imagery cognitive task related problems, where the job is more difficult (say, combination of mental arithmetic, landscape, music imagery).

Cognitive load has been measured using the quantification technique. Here lies a scope on developing a better metric to measure the load, and mapping it with the spatial load distribution image taken from the fNIRS to verify the effectiveness of the metric.

APPENDIX A

MATLAB Source Codes

This section deals with the necessary source codes developed on a MATLAB environment. The codes are easy to use and have been organized in a tabular form with comments for the ease of use.

Table IX Noise Elimination from fNIRS Signal

Source Code 1: Noise Removal

```
r = 16; %#independent components = #channnels
n = size(X,1);
[Xica, A, T, mu] = myICA(X', r);
t = 0 : 0.5 : size(Xica,2)/2-0.5;

for i = 1 : size(X,2)
   figure;
   subplot(2,1,1), plot(t,X(:,i)/max(abs(X(:,i))));

   subplot(2,1,2), plot(t,Xica(i,:)/max(abs(Xica(i,:))));
end

for i = 1 : size(Xica,1)
   figure;
   subplot(2,1,1), plot(t,Xica(i,:)/max(abs(Xica(i,:))));
   x = Xica(i,:)';
   BandPassSpecObj = fdesign.bandpass('N,Fp1,Fp2,Ap',10,0.1,0.4,1,1.96);
   BandPassFilt = design(BandPassSpecObj, 'cheby1');
   y = filter(BandPassFilt, x);
   M(i,:) = y';
   subplot(2,1,2), plot(t,y/max(abs(y)));
end
```

Table X Proposed Feature Selection Method

Source Code 2: Likelihood Based DE-Driven Feature Selection

```
function [ X_red ] = feature_selection( Xfea )

%%
%This function tries to optimize the feature vector using differential
%evolution technique.
%%
%Parameter definitions

MAXiter = 500;     %maximum no of iteration used
NP = 50 ;      %number of vectors in a pool of population
F = 0.8;
CR = 0.7;
lambda = rand(1)*10;
Xdim = size(Xfea,2);
Xmut = zeros(NP,Xdim);
fvalue = zeros(NP,1);
%%
%Initialization
Xpop = (ones(NP,Xdim) - rand(NP,Xdim))>0.5;

for i = 1 : NP
   Xindex = [];
   Xindex = find(Xpop(i,:)>0);
   fvalue(i,:) = obj_func(Xfea(:,Xindex),y, lambda);
end
[~ ,minindex] = min(fvalue); %value and index of the minimum cost function
best = Xpop(minindex,:);      %best feature vector of the generation

%%
%Generation

for iter = 1 : MAXiter
   for i = 1 : NP

       %Select the three random vectors for vector i
       r1 = floor(rand(1)*NP)+1;
       r2 = floor(rand(1)*NP)+1;
       r3 = floor(rand(1)*NP)+1;
```

```
    while(r1==i)
       r1 = floor(rand(1)*NP)+1;
    end
    while(r2 == i||r2 == r1)
       r2=floor(rand(1)*NP)+1;
    end
    while(r3 == i|| r3 == r1 || r3 == r2)
       r3 = floor(rand(1)*NP)+1;
    end
    %%
    %Mutation
    %formation of mutant vector

    Xmut(i,:) = (Xpop(r1,:) + F*(Xpop(r2,:) - Xpop(r3,:)))>0.5;

    for j = 1 : Xdim
       if rand(1) < CR
          Xu(i,j) = Xmut(i,j);
       else
          Xu (i,j) = Xpop(i,j);
       end
    end
    Xuindex = [];
    Xuindex = find(Xu(i,:)==1);

    fvalue_u = obj_func(Xfea(:,Xuindex),y, lambda);

    if fvalue_u < fvalue(i,1)
       fvalue(i,1) = fvalue_u;
    end
  end
    [~ ,minindex] = min(fvalue);
    best = Xpop(minindex,:);
end
    X_red = Xfea(:,find(best>0));
end
```

Table XI Likelihood-Based Objective Function

Source Code 3: Likelihood Based Objective Function

```
function [ f ] = obj_func( X , y, lambda )

NC = max(y); % total number of classes present
Xnew = zeros (size(X,1),size(X,2),NC);

for i = 1: size(X,1)
   for j = 1 : NC
      if y(i,1) == j
         Xnew (i,:,j) = X(i,:);
      end
   end
end
mean_val = zeros(NC,size(X,2));
cov_val = zeros(size(X,1),size(X,2),NC);
inv_cv = zeros(size(X,1),size(X,2),NC);
for i = 1 : NC
   mean_val(i,:) = mean(Xnew(:,:,i));
   addi = zeros(size(Xnew,1),size(Xnew,2));
   for j = 1 : size(Xnew,1)
      addi = addi + (Xnew(j,:,i)-mean_val(i,:))'*(Xnew(j,:,i)...
                                    -mean_val(i,:));
   end
   cov_val(:,:,i) = addi/(Size(Xnew,1)-1);
   inv_cv(:,:,i) = pinv(cov_val(:,:,i));
end

term1 = zeros(NC,size(Xnew,1),NC);
for i = 1 : NC
   for j = 1 : size(Xnew,1)
      for k = 1 : NC
         term1(k,j,i) = (Xnew(j,:,i)-mean_val(k,:))*inv_cv(:,:,k)*...
                  (Xnew(j,:,i)-mean_val(k,:))'...
                  ./sqrt(det(cov_val(:,:,k)));
      end
   end
end
term2 = 0;
for i = 1 : NC-1
   for j = i+1 : NC
```

```
    term2 = term2 + (1/sqrt(det(cov_val(:,:,j))))- ...
          (1/sqrt(det(cov_val(:,:,i))));
  end
end
L1 = 0;
for i = 1 : NC
  L1 = L1 + 4*sum(sum(term1(i,:,i)))-sum(sum(term1(:,:,i)));
end

f = L1 + lambda*term2;
end
```

Table XII Proposed Type-2 Fuzzy Classifier

Source Code 4: Type 2 Fuzzy Classifier

```
function [ t1red ] = t2fuzzclassifier(Xtest,m,s)
% This function is used to generate the type 2 gaussian type membership
% function of the test input variables and subsequent reduction of type 2
% into type 1 by using weighted average method.
r = size(m,1);
c = size(m,2);
max_chan = size(Xtest,1);
fgauss = zeros(max_chen,r,c);
for i = 1 : max_chan
  for j = 1 : r
    for k = 1 : c
      fgauss(i,j,k) = gaussmf(Xtest(i,k),[s(j,k) m(j,k)]);
    end
  end
end
t1red = zeros(max_chan,c);
for i = 1 : max_chan
  for j = 1 : c
    sd = std(fgauss(i,:,j));
    mn = mean(fgauss(i,:,j));

    f2gauss = gaussmf(fgauss(i,:,j) , [sd mn]);
    t1red(i,j) = sum(f2gauss.*fgauss(i,:,j))/sum(f2gauss);
  end
end
end
```