Journal of Neurodevelopmental Cognition

Journal of Neurodevelopmental Cognition 1 (2019) 16-41
ISSN: 2645-565X
http://www.jncog.sbu.ac.ir

Brain functional connectivity changes in long-term mental stress

Ali Darzi[1], Hamed Azami[2], Reza Khosrowabadi[1]*

[1] *Institute for Cognitive and Brain Science, Shahid Beheshti University, Iran.*
[2]*Institute for Digital Communications, School of Engineering, University of Edinburgh, United Kingdom.*

Abstract

Long-term psychological stress can highly influence brain structure and functions. However, there are only few studies using electroencephalogram (EEG) that have examined this fact. The current study demonstrates a brain-computer interface (BCI) to classify EEG correlates of long-term mental stress in various mental states. The study was performed on 26 healthy right-handed university students and examination period was considered as a long-term mental stressor. Two groups of subjects were selected based on their stress levels evaluated by perceived stress scale (PSS-14). The subjects' EEG data were collected during eyes-open resting state and while they exposed to positive and negative emotional stimuli scored by self-assessment manikin questionnaire (SAM). Several types of features were extracted from EEG data including power spectrum density (PSD), laterality index (LI), correlation coefficient (CC), Canonical correlation analysis (CCA), magnitude square coherence estimation (MSCE), mutual information (MI), phase-slope index (PSI), Granger causality (GC) and directed transfer function (DTF). Subsequently, the extracted features were discriminated using several types of classifiers including k-nearest neighbor (KNN), support vector machine (SVM) and naive Bayesian (NB) classifiers. The proposed BCI was validated by one leave out method and investigation was done in different time windows using low and high frequency resolutions, 7 and 36 frequency bands respectively. The results showed that proposed system can accurately recognize subjects' stress level in various mental states. Moreover, the MI as a functional and DTF as effective connectivity methods yield the highest classification accuracy compared to other feature extraction methods.

Keywords: Long-term mental stress, Electroencephalography, Emotional states, Classification.

* Corresponding author
Email addresses: r_khosroabadi@sbu.ac.ir (Reza Khosrowabadi)

Received: September 2018 *Revised: December 2018*

Introduction

Modern life is full of deadlines, economic problems, work pressures and frustrations. These facts as sources of mental stress can drastically influence the human life if be overloaded or prolonged. Acute (short term) stress is usually not a risk factor for the health, in contrast, long-term stress can lead to serious mental and physical health problems [1,2]. In fact, screening of long-term mental stress is the first step toward its treatment. Therefore, implying a non-invasive method for this purpose is of importance. Physiological markers such as skin conductance (SC), electrocardiogram (ECG) and EEG can potentially be used for this purpose [3,4]. Alteration of heart beat extracted from ECG and changes in skin conductance are good indicators of stress level. Although, they do not directly address effect of stress on brain functioning and could not directly be used for intervention purpose (for instance in a neurofeedback system) [5]. Nevertheless, considering complexity of EEG signal, it may not be easy to recognize the stress level from EEG data. Therefore, it is very important to extract informative features from EEG which significantly correlate with stress level. In fact, the relationship between EEG and stress level could vary in different mental states. Hence, a good recognizer system should be able to identify the stress level in various mental states. In this research, impact of long-term stress on EEG signals is investigated in three emotional states and eyes-open resting state.

In general, mental stress is defined as a feeling of strain in which a person's perceives that a situation is stressful and demands exceed its ability to cope. Since, perceptual responses to stressors are subjective. Therefore, an instrument is required to measure ones' stress level. Vast majority of instruments use self-report questionnaires for this purpose [6] and the Perceived Stress Scale (PSS) is one of them [7] . There are three versions of this questionnaire including 4, 10 and 14 items. The 14-item version is called PSS-14 which measures non-specific perceived stress scale by addressing events experienced in the preceding month. It was designed to quantify how adults find their lives unpredictable, uncontrollable, and overloaded [5]. The questionnaire involves 14 questions which the range of response for each item is from 0 to 4, thus, total possible score in PSS-14 can be from 0 to 56 while a lower result is related to a lower stress level. The questionnaire has very easy and comprehensible questions and the response alternatives are simple to grasp [8]. In this study, the obtained result from PSS-14 shows mean and standard deviation equal to $\mu = 24.5, \sigma = 6.7$. Dealing with the problem as a two states problem, two hard threshold have been considered as 21 ($\mu - \frac{\sigma}{2}$) and 28 ($\mu + \frac{\sigma}{2}$). The subject with PSS-14 score lower than 21 are categorized as mild stress level (stress-free group) and those with scores higher than 28 are considered in concerning situation (stressed group).

We hypothesized that long-term stress alter activity and behavior of the brain and these changes could be recognized in EEG signals. Therefore, EEG data were recorded by an eight-electrode system. Each electrode collects electrical activities of one region of the brain including frontal, central, parietal and occipital regions at left and right sides compared to vertex activity (Cz). EEG signals were gathered in 1 min eyes-open resting state and 3 min while exposed to emotional stimuli (1 min each emotional state).

Since, our investigation was performed in various mental states including emotional states, therefore,

power spectrum density (PSD) of signals and also hemispheric specialization which is a major neurophysiological marker of emotional states [9] were considered as a candidate features in our evaluation. It has been shown that the left hemisphere appears to be more involved in processing of positive emotions, while, the right hemisphere elaborates more in process of negative emotions [10,11]. Therefore, laterality index (LI) was also considered as a potential marker and it was compared with other feature extraction methods. Moreover, studies present the pattern of connectivity between different brain regions as a good indicator of emotional states [12,13,14]. Some of well-known methods of functional connectivity which are used in this study are correlation coefficient (CC), Canonical correlation Analysis (CCA), mean square magnitude square coherence estimation (MSCE), mutual information (MI). Phase-slope index (PSI), Granger causality (GC) and directed transfer function (DTF) are the other three methods which can estimate the direction of information flux in multivariate time series (effective connectivity). Afterward, SVM, KNN and Bayesian classifier were used to classify the data as accurate as possible.

It has been shown that pattern of connectivity between brain regions is frequency specific [2], therefore, for three Fourier transform based methods, PSD, LI and MSCE, two different frequency resolutions, low (7 bands) and high (36 bands) were considered. Since the other methods work on time-domain, different frequency resolution cannot be defined for them. Beside different frequency resolutions, four different time windows of EEG data namely 1s, 2s, 8s, and 56s were compared. Subsequently, the best time window length and frequency resolution that lead to the highest accuracy in classification was highlighted as the best marker for stress identification from EEG signals.

Since, the neurofeedback has successfully been applied for treatment of mental diseases such as anxiety [15], depression [16] and psychosomatic diseases (migraine, tension headache, and hyperpiesia) [17]. Therefore, a neuro-feedback system can potentially be designed based on outcome of this research and subsequently be used for intervention of long-term mental stress. The rest of the paper is organized as follows: Section 2 presents the experimental design and protocol. Sections 3 describes the methods which are proposed to solve the problem. Representing the obtained results and comparing with each other are discussed in Section 5. Finally, Section 5 concludes the proposed methods.

Experimental Design

Identification of long-term mental stress during resting and emotional states is performed based on EEG signals. The procedure of experimental design which contains selecting the stressor, subjects and protocol of data analysis are described in this section.

The mental states in human are typically categorized to four classes based on their underlying time-line. First, attention which lasts for a fraction of second. Second, full blown emotions which usually last more than attention and persist for some seconds to minutes. Next type is known as mood which is usually assigned to a set of emotions which last for minutes to hours. The fourth set which typically is realized as mental disorder can show its signs during a long period such as a year to part of the life [18]. Long-term stress, as a response to psychosocial, physical and physiological problems, has effects on mood.

Considering the time length of long-term mental stress, it could be an underlying state for emotional states. Therefore, any major changes in terms of long-term stress could be recognized in different emotional states. In this regard, our hypothesis related to recognition of long-term mental stress using EEG signal in various emotional states was framed. In fact, many studies have shown that emotional states can be recognized using a short period of EEG signals [19,20] In addition, it is entirely accepted that the brain has some function networks which are active during the resting state. This fact was firstly discovered by Bharat Biswal in 1992. He showed that the brain, even during rest, contains information about its functional organization [21]. By this point of view, if the long-term mental stress influences the brain functions, its effects should also be observable in resting state [22]. Hence, the experiment was designed in a way to recognize the long-term mental stress in both resting and emotional states.

Investigation of long-term stress effects requires subject to be exposed to a psychological stressor for a long period. In this study, the examination period was selected as a novel long-term psychological stressor and investigation was performed on a group of undergraduate university students. Encountering to the final exams can be really stressful for students especially those who are young and have little experience. In total 26 right hand student (6 female) in the age range of 18-30 were examined. The subjects stress level was measured by a self-report questionnaire (PSS-14) before EEG recording.

Experimental protocol

The experiment was started with two minutes resting state, one minute with closed eyes and one minute with open eyes and fixed to screen center. After resting states, four emotional stimuli were presented to the subjects, each for one minute in a counterbalanced and random order. After presenting each emotional stimulus, a self assessment Mankin questionnaire (SAM) was being answered by the subject to determine the level of perceived emotion. Figure.2 shows the experimental protocol.

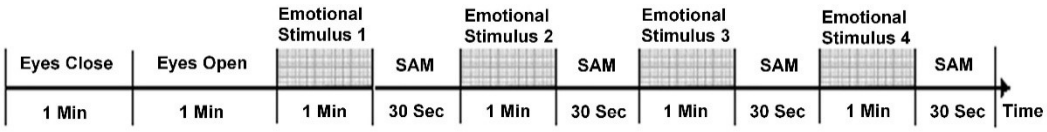

Figure 1. Experimental protocol to collect EEG data

The EEG data were collected while the subject seated at a comfortable chair in a registration room and the experimental procedure was explained to him/her. The EEG was recorded using the BIMEC from Brain-marker BV. The BIMEC has 1 reference channels plus 8 EEG channels with a sampling rate of 250 Hz. The impedance of the Ag/AgCl electrodes was kept below 10 kΩ. Considering the cerebral lateralization during the emotional perception [23], the EEG data was recorded using bilaterally attached electrodes (F3, F4, C3, C4, P3, P4, T7, T8). The10-20 system of electrode placement was implied whereby the C_z was applied as reference channel [24]. The EEG data were collected for a 6-min period of time that comprised of 1 min eyes-close condition, 1 min eyes-open condition, and 1 min for each emotional stimulus. The subject was exposed to 4 sets of different emotional stimuli. The visual stimuli were displayed on a 19 inch monitor position 1 m from the participant's eyes and the audio stimuli were played with speakers with constant output power. The 4 different set of emotional stimuli were presented

randomly and each set of stimulus was only presented once for every subject. A combination of picture and audio stimuli was used to elicit the emotional responses. The pictures were selected form IAPS [22] and audio stimuli were used from synthesized musical experts by Bernard Bouchard [25]. In fact, the perceived emotion may differ from a subject to another; therefore, SAM (Figure 1), was used to measure subject's emotional state.

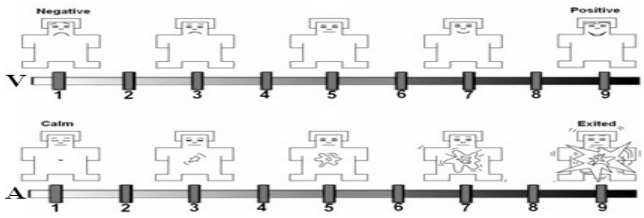

Figure 2. 2D Self assessment Manikin questionnaire [25]

For simplification, four emotional states were defined based on their valence and arousal level. The positive and negative states at calm and exited levels were by boundary lines at score 4 and 6. The answers in border line (between 4 to 6) were eliminated due to ambiguous emotion which perceived. By eliminating the border-line regions, what remains can be easily divided into negative and positive valence at calm and excited level of arousal. However, according to subjects' answers to SAM questionnaire, arousal level of stimuli were not that discriminative. Therefore, our investigation in this paper was performed based on the EEG data labeled only by the valence level.

Methodology

This section describes the procedure of data collection and the analyses performed on the EEG data to differentiate the subjects with different stress level. Discrimination based on EEG signals, needs a good processing technique to be applied. Therefore, a notable feature must be extracted and a suitable classification method should be selected. Figure 3 shows the EEG data processing paradigm used in this study.

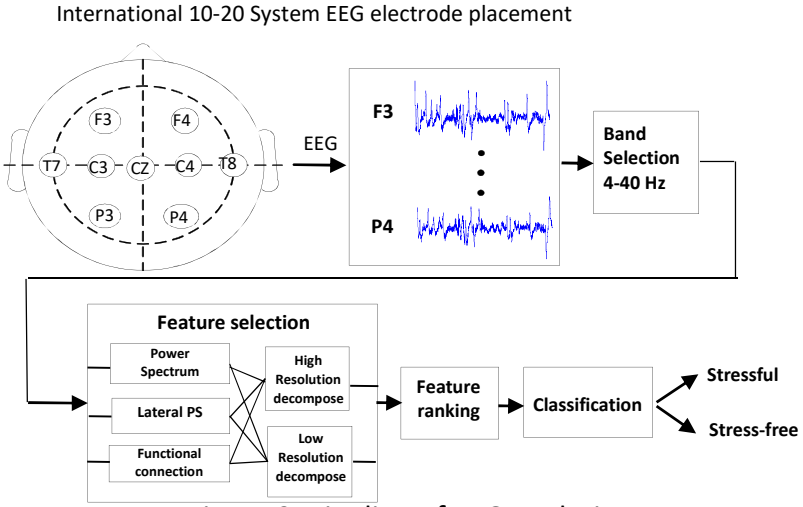

Figure 3. Pipeline of EEG analysis

The mathematical notations used to denote the inputs, outputs, features, related variables and functions are given as follows:

Let $\chi \in \mathbb{R}^{n_s \times n_t \times n_c \times n_p}$ denotes the multi-subject, multi-channel EEG data, where n_s denotes the number of subjects and n_t is the number of trials that each subject has passed and n_c denotes the number of EEG channels. In total, EEG signal of n_s =26 subjects at n_c=8 locations were recorded with sampling rate of 250 Hz. Therefore, by recording EEG data for 60 seconds, n_p=15000 samples were recorded from each channel. Let $\chi = \left[\mathbf{X}_1^1, \mathbf{X}_1^2, \ldots, \mathbf{X}_1^{n_t}, \ldots, \mathbf{X}_s^t, \ldots, \mathbf{X}_{n_s}^{n_t}\right]^T$ whereby $\mathbf{X}_s^t \in \mathbb{R}^{n_p \times n_c}$ denotes the t^{th} single-trial EEG data from the s^{th} subject and $1 \le t \le n_t$, $1 \le s \le n_s$. Let $\mathbf{x}_s^t = \left[\mathbf{x}_{s,1}^t, \mathbf{x}_{s,2}^t, \ldots, \mathbf{x}_{s,i}^t, \ldots, \mathbf{x}_{s,n_c}^t\right]^T$ whereby $\mathbf{x}_{s,i}^t \in \mathbb{R}^{n_p \times 1}$ denotes a vector that represents the time series EEG from the i^{th} channel, $\mathbf{x}_{s,i}^t = \left[x_1, x_2, \ldots, x_n, \ldots, x_{n_p}\right]$ whereby x_n denotes the n^{th} sample. A simplified notation of x_n is used to denote $\mathbf{x}_{s,i}^{t,n}$ in remainder of this paper.

$\gamma \in \mathbb{R}^{n_v \times (n_t n_s)}$ denotes the extracted features that are correlated to $\varphi \in \mathbb{R}^{n_t n_s}$ whereby n_v denotes the number of extracted features for a particular feature extraction method. Let $\gamma = \left[\mathbf{v}_1^1, \mathbf{v}_1^2, \ldots, \mathbf{v}_1^{n_t}, \ldots, \mathbf{v}_s^t, \ldots, \mathbf{v}_{n_s}^{n_t}\right]^T$ whereby $\mathbf{v}_s^t \in \mathbb{R}^{n_v \times 1}$, $\mathbf{v}_s^t = \left[v_1, v_2, \ldots v_k, \ldots v_{n_v}\right]$ where v_k denotes the k^{th} feature.

Assigning an output class (stressful and stress-free) to each raw EEG signal needs a pipeline processing consists of 5 main steps consist of band selection, feature extraction, feature ranking, signal decomposition and classification methods. Some basic and well-known functions were used in this study such as, rectangular windowing, discrete Fourier transform, power spectral density, cross power spectral density and cross correlation. Each section is described completely with the mathematical notations of these functions.

EEG signal decomposition

One of the main factors in EEG signal analysis is artifact and noise removal. Many well-known artifacts such as breathing, heart rate and eye movement pollute the EEG signal, mostly in low frequencies. There are also other artifacts which will affect the high frequency bands such as electrical power line (50/60 Hz). Hence, selecting a working band which has the most information and least artifacts is very important. Here, a band of 4-40 Hz was chosen. Subsequently, analyses were performed using two types of bands decomposition.

Band decomposition

Each frequency band of brain electrical activity is thought to have correlation with specific brain functions. For instance, the alpha waves at occipital areas represent the visual cortex activities as well as they have active role in network coordination and communication [26]. Therefore, it is important to decompose EEG signal into some specific frequencies at the first step of EEG data processing. In this study, first, the EEG data were decomposed to trivial Quantitative EEG (QEEG) bands including theta (4-8), alpha1 (8-10 Hz) and alpha2 (10-12 Hz), beta1 (12-16 Hz), beta2 (16-20 Hz), beta3 (20-30 Hz) and gamma1 (30-40 Hz).

Subsequently, another type of decomposition with high resolution characteristics (1 Hz) is implemented to have more accurate information about oscillations in different frequencies. It should be noticed that three out of seven feature extraction methods consider frequency-domain thus, different frequency decomposition can be taken into account.

Time windowing

Alongside of different resolutions in frequency domain, to cope the BCI scopes which need high classification rate in minimum time period, different window sizes in time domain were also investigated. Let $\mathbf{W} \in \mathbb{R}^{N_w \times N_c}$, $\mathbf{W} = \left[\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_i, \ldots, \mathbf{w}_{n_c}\right]^T$ where $\mathbf{w}_i \in \mathbb{R}^{N_w \times 1}$, denotes a vector that represents the windowed time series EEG signal from the i^{th} channel. Let $\mathbf{w}_i = \mathrm{f}_{\mathrm{w}}\left(\mathbf{x}_{s,i}^t, s_t, N_w\right)$ denotes the rectangular window function on $\mathbf{x}_{s,i}^t$ given as:

$$\mathrm{f}_{\mathrm{w}}\left(\mathbf{x}_{s,i}^t, s_t, N\right) = \left[\mathbf{w}_i^1, \mathbf{w}_i^2, \ldots, \mathbf{w}_i^n, \ldots, \mathbf{w}_i^{N_w}\right] \quad (1)$$

where N_w denotes the number of windows that can be exist, s_t denotes the considered start point, and $\mathbf{w}_i^n$ denotes the n^{th} sample of the windowed signal that is the ($s_t + n$)th sample of the i^{th} channel of the EEG data. Four time windows including N=250, 500, 2000, 14000 standing for 1, 2, 8 and 56 seconds period of time were evaluated in this study. In each windowing process, the first and last 500 samples were not taken into account to avoid the effect of baseline.

Feature Extraction

There are numbers of techniques for features extraction from EEG signal [27,28,29,30,31,32]. Several methods such as event-related synchronization or desynchronization, event-related potentials, visual-evoked potentials, and features from quantitative EEG [33, 34, 35, 36] have been evaluated. In this study PSD, LI and MSCE which are estimated by power spectrum density of signals as well as MI, GC, PSI and DTF as functional and effective connectivity estimation methods are implemented. In this section, the feature extraction methods are discussed briefly.

Power Spectral Density

Today, it has been proved that brain signals work in some special frequency bands. Therefore, investigation on power of signal in frequency domain can reveal lots of information. The simplest method to investigate the power a signal is through power spectrum of a signal which is equal to Fourier transform of the signal in power of two or in other hand, the Fourier transform of autocorrelation of the signal. The Fourier transform maps a signal from time domain into frequency domain. In fact, the Fourier transform measures the correlation between the signal and the sinusoid signals with different wavelength (frequencies). Thus, an overview can be obtained by tracking the quantity of correlation in each frequency. The higher correlation means the more activation on the frequency.

Let $\mathbf{z}_i = f_{DFT}(\mathbf{w}_i)$ denotes the discrete Fourier transform (DFT) whereby $\mathbf{z}_i \in \mathbb{C}^{N\times 1}$ is a transform of the windowed EEG signal $\mathbf{w}_i$ from the time domain to the frequency domain of the *i*th channel of the EEG data. Let $\mathbf{z}_i^f$ denotes the *f*th sample of $\mathbf{z}_i$ given by:

$$\mathbf{z}_i^f = \sum_{n=0}^{N-1} \mathbf{w}_i^n e^{\frac{-J2\pi fn}{N}} \quad (1)$$

Where $f = [0,1,...,N-1]\times\left(\frac{F_s}{2\times N}\right)$, and J is the imaginary unit.

Let $\mathbf{s}_{ii} = N \times f_{DFT}(\mathbf{r}_{ii})$ denotes the PSD of $\mathbf{w}_i$ whereby $\mathbf{s}_{ii} \in \mathbb{C}^{N\times 1}$ is the DFT of the autocorrelation function $\mathbf{r}_{ii}$. Let $\mathbf{s}_{ij} = N \times f_{DFT}(\mathbf{r}_{ij})$ denotes the cross power spectral density (CPSD) of $\mathbf{w}_i$ and $\mathbf{w}_j$ whereby $\mathbf{s}_{ij} \in \mathbb{C}^{N\times 1}$ is the DFT of the cross correlation function $\mathbf{r}_{ij}$. The PSD and the CPSD functions are given by

$$f_{DFT}(\mathbf{r}_{ii}) = \frac{1}{N}\mathbf{z}_i\mathbf{z}^*_i \quad (2)$$

$$f_{DFT}(\mathbf{r}_{ij}) = \frac{1}{N}\mathbf{z}_i\mathbf{z}^*_j \quad (3)$$

After computing the PSD value $\mathbf{c}_{ij}$ for all the EEG channels, the feature vector $\mathbf{v}_s^t \in \mathbb{C}^{n_v\times 1}$ is given by $\mathbf{v}_s^t = \left[\mathbf{c}_{11}, \mathbf{c}_{12}, \ldots, \mathbf{c}_{ij}, \ldots, \mathbf{c}_{n_c n_f}\right]$ where the number of extracted features is $n_v = n_c \times n_f$. The n_f is the number of frequency bands (36/frequency resolution) which is equal to 7 and 36 for low and high frequency resolution, respectively. According to the number of channels which is equal to 8, the number of extracted features is equal to 56 and 288.

Lateral Index based on Power Spectral Density

Although, the left and right lobes of brain have similar parts which are functionally related but the activation of the same parts in the two lobes are not analogous. Therefore, if the activation of one region on the cortex has nothing to tell us, maybe the relation of the two symmetric regions can reveal a significant answer about the differences between the stressful and control groups. A number of researches have reported the benefit of this features for mental state categorization [10,11]. The lateral signals which have been used in this study are calculated as below:

$$Lateral_S = \frac{S_{left} - S_{right}}{S_{left} + S_{right}} \quad (4)$$

Which the *S* is the PSD of signals and S_{left}, S_{right} stand for the PSD of the signal from left and right side of brain, respectively. The left and right lobes relationships were calculated at frontal, central, temporal and parietal regions. In fact, the laterality index of EEG signals monitors the differences between brain activities at the left and right regions.

By computing the Lateral index $\mathbf{c}_{ij}$ for all the EEG lateral channels, the feature vector $\mathbf{v}_s^t \in \mathbb{R}^{n_v \times 1}$ is given by $\mathbf{v}_s^t = \left[\mathbf{c}_{11}, \mathbf{c}_{12}, \ldots, \mathbf{c}_{ij}, \ldots, \mathbf{c}_{n_m n_f}\right]$ where the number of extracted features is $n_v = n_f \times n_c / 2$. In the formula the n_m is considered as $n_m = n_c / 2$ due to lateral view to the channels. Considering n_f equal to 7 and 36 for low and high resolution of frequency decomposition, 28 and 144 features are extracted, respectively.

Functional Connectivity

In general, brain connectivity is referred to three types of connections between distinct units within a nervous system: structural (anatomically), functional and effective connections [37,38]. Anatomic connection stands for the connection of each neuron of nervous system or structural network of the cerebral cortex [39]. The other two types of connections stand for coincident firing of two regions. Functional connectivity is defined as the temporal correlations between spatially remote neurophysiological events [40]. Point to point functional connectivity is defined in terms of observed correlations or covariance without considering the anatomical connectivity. In this study, the MSCE [41] and MI [40] are used as functional connectivity methods and the extracted features are compared with the other methods.

Correlation Coefficient

Correlation between two signals shows how dependent they are to each other. One of the widely used method to calculate the level of dependency is Pearson method. This method which is based on statistic parameters like covariance and standard deviation, is defined as follow:

$$\rho_{XY} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} \qquad (5)$$

Where the cov and σ stand for covariance and standard deviation. The Pearson's method is good in finding linear correlation between two variables. The output value, ρ, can be any real number in the range of -1 to 1 which 1 and -1 mean the highest zero phase and 180 degree dependency, respectively. The extracted feature vector from CC, $\mathbf{v}_s^t \in \mathbb{R}^{n_v \times 1}$ is given by $\mathbf{v}_s^t = \left[\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{n_m}\right]$ where the number of extracted features is $n_v = n_m \times 1$. The n_m is equal to $n_m = n_c(n_c - 1)/2$ due to the nature of functional connection which is bilateral. For instance the correlation of C3-T7 is not different with T7-C3 correlation. Thus, 28 correlation coefficients are obtained. The CC feature extraction method works on time domain of signal, therefore the n_f, number of frequency windows, is considered equal to one.

Canonical Correlation Analysis

Consider $X = (x_1, x_2, \ldots, x_n)^T$ and $Y = (y_1, y_2, \ldots, y_m)^T$ of random variables with covariance matrix as $\sum XY = \mathrm{cov}(X,Y)$. The canonical correlation finds linear combination of x_i and y_i in a way that produce the highest correlation [42]. The concept of CCA is close to principle component analysis (PCA) [43] with

the difference of finding similar components in two groups rather one group of data.

Using CCA, the extracted feature vector is $\mathbf{v}_s^t \in \mathbb{R}^{n_v \times 1}$ is given by $\mathbf{v}_s^t = \left[\mathbf{c}_{11}, \mathbf{c}_{12}, \ldots, \mathbf{c}_{ij}, \ldots, \mathbf{c}_{n_m n_o}\right]$ where $n_v = n_f \times n_m \times n_o$. The n_o, number of components in CCA, is equal to minimum order of X and Y series which is Considered to be $n_c / 2$. The n_m is equal to n_c and as same as CC method the n_f is propounded to be 36 and 7 due to high and low frequency resolution. Therefore, the number of obtained features is $n_v = n_f \times n_c^2 / 2$. Considering n_c equal to 8, the number of obtained features for high and low frequency resolution is 1152 and 224, respectively.

Mean Square Coherence Estimate

The MSCE is a function of frequency that gives a value $\mathbf{c}_{ij} \in \mathbb{R}^{n_{f1} \times 1}$ to indicate how well the *i*th channel $\mathbf{w}_i$ corresponds to the *j*th channel $\mathbf{w}_j$ at each frequency [57]. The MSCE value $\mathbf{c}_{ij}$ where $0 \le \mathbf{c}_{ij} \le 1$ is estimated using the Welch's averaged method. The MSCE value $\mathbf{c}_{ij}$ is a function of $\mathbf{s}_{ii}$, $\mathbf{s}_{jj}$, and $\mathbf{s}_{ij}$ given by:

$$\mathbf{c}_{ij}^f = \frac{\left|\mathbf{s}_{ij}^f\right|}{\mathbf{s}_{ii}^f \mathbf{s}_{jj}^f} \qquad (7)$$

Where $\mathbf{s}_{ii}^f, \mathbf{s}_{ij}^f, \mathbf{c}_{ij}^f$ denote the *f* th samples of $\mathbf{s}_{ii}, \mathbf{s}_{ij}, \mathbf{c}_{ij}$ respectively.

The cross spectral density and the power spectral densities are estimated using equations (2) and (3) respectively. A Hamming window of size 256 with overlap of 12.5% (n_{f_1} =32) were considered for all the considered time resolutions (window size) except for 1 second window which Hamming window with size of 128 samples and overlap of 12.5% (n_{f_1} =16) was used.

The extracted feature vector from MSCE, $\mathbf{v}_s^t \in \mathbb{R}^{n_v \times 1}$ is given by $\mathbf{v}_s^t = \left[\mathbf{c}_{11}, \mathbf{c}_{12}, \ldots, \mathbf{c}_{ij}, \ldots, \mathbf{c}_{n_m n_f}\right]$ where the number of extracted features is $n_v = n_f \times n_c (n_c - 1) / 2$. By two decomposition methods, low and high resolution in frequency domain, 196 and 1008 features are obtained, respectively.

Mutual Imformation

Mutual dependency of one signal to another can be measured by MI. Unlike the correlation coefficient or coherency, the MI is not limited to real-valued variable and can be implemented on variety of variable types. The basic idea behind the MI is determining the similarity of joint probability distribution function (PDF) of two signals and the product of their marginal PDFs. The MI method is defined as below:

$$MI(X;Y)=\sum_{y\in Y}\sum_{x\in X}p(x,y)\log(\frac{p(x,y)}{p(x)p(y)}) \quad (8)$$

Where $p(x,y)$, $p(x)$ and $p(y)$ are the joint PDF of x and y and marginal PDF of x and y, respectively. Dependency can be interpreted as influence of one signal over another, thus, there should not be any mutual information between two independent signals. This can be obtained from equation (8). Let the extracted feature vector from MI be $\mathbf{v}_s^t \in \mathbb{R}^{n_v\times 1}$ which $\mathbf{v}_s^t=\left[\mathbf{c}_1,\mathbf{c}_2,\ldots,\mathbf{c}_i,\ldots,\mathbf{c}_{n_v}\right]$. Number of extracted features is $n_v=n_m\times n_f$ where $n_m=n_c(n_c-1)/2$ and n_f is depend on the way of frequency windowing which can be either 36 or 7. In summary, 1008 and 196 features are calculated for high and low frequency decomposition.

Effective connectivity

Effective connectivity is defined as the influence of one neuronal system exerts over another either on synaptic or cortical level [56]. Effective connectivity method estimates the level and direction of information flow from signal x to signal y. In this study three effective connectivity methods, GC, SPI and DTF are implemented. A brief introduction to these methods is mentioned in this section.

Granger Causality

Granger causality is a statistical test to investigate if one signal can be useful to forecast the other one. The GC is based on linear autoregressive (AR) models. Consider a bivariate linear AR model of two variables X_1 and X_2 as follow:

$$\begin{aligned}X_1(t)&=\sum_{i=1}^{N}a_{11}(i)X_1(t-i)+\sum_{i=1}^{N}a_{12}(i)X_2(t-i)+E_1(t)\\X_2(t)&=\sum_{i=1}^{N}a_{21}(i)X_1(t-i)+\sum_{i=1}^{N}a_{22}(i)X_2(t-i)+E_2(t)\end{aligned} \quad (9)$$

The equation can be rewritten as

$$\begin{pmatrix}A_{11}(t) & A_{12}(t)\\ A_{21}(t) & A_{22}(t)\end{pmatrix}\begin{pmatrix}X_1(t)\\ X_2(t)\end{pmatrix}=\begin{pmatrix}E_1(t)\\ E_2(t)\end{pmatrix} \quad (10)$$

Where the $A_{11}(0)=1$, $A_{22}(0)=1$ and $A_{nm}(t)|_{t\neq 0}=-a_{nm}(t)$. In this case by taking Fourier transform and inverting the matrix, equation 10 is driven.

$$\begin{pmatrix}X_1(f)\\ X_2(f)\end{pmatrix}=\begin{pmatrix}H_{11}(f) & H_{12}(f)\\ H_{21}(f) & H_{22}(f)\end{pmatrix}\begin{pmatrix}E_1(f)\\ E_2(f)\end{pmatrix} \quad (11)$$

Where the H matrix is known as transfer matrix. Using the transfer matrix the spectral matrix, S, can be defined as:

$$S(f)=\langle X(f)X^*(f)\rangle=\langle H(f)\sum H^*(f)\rangle \quad (12)$$

Where <.> denotes expectation value and the asterisk signifies matrix transposition
complex conjugation, Σ is the covariance matrix of the residuals E(t), and H is the transfer matrix. The spectral G-causality from j to i is then

$$I_{j\to i}(f) = -\ln\left(1 - \frac{(\sum_{jj} - \frac{\sum_{ij}^{2}}{\sum_{ii}})\,|H_{ij}(f)|^{2}}{S_{ii}(f)}\right) \qquad (13)$$

Where the 'I' stands for the flow of information form point j to point i. The obtained feature vector form GC is like $\mathbf{v}_s^t \in \mathbb{R}^{n_v \times 1}$ which $\mathbf{v}_s^t = \left[\mathbf{c}_{11}, \mathbf{c}_{12}, \ldots, \mathbf{c}_{ij}, \ldots, \mathbf{c}_{n_m n_f}\right]$. Number of extracted features is $n_v = n_m \times n_f$ where $n_m = n_c^{\,2}$ due to information flow difference from node j to i and i to j. Number of windows in frequency domain, n_f, is equal to 36 and 7 according to resolution of decomposition. Thus, number of all obtained features is equal to 2304 and 448 for high and low resolutions, respectively.

Directed Transfer function

Directed transfer function is an effective connectivity method which is based one Granger causality [44]. This method considers all the channels into account simultaneously and makes a profit of transfer matrix, H, that is calculated in GC.

Consider multivariate Autoregressive (MVAR) of r^{th} order for multi-channel EEG signal. Using equations (9-11), the transfer matrix, H, can be calculated. By definition of DTF, the casual influence from j^{th} channel to i^{th} channel for a specific frequency interval is defined as equation (14).

$$(I_{j\to i}(f))^2 = |H_{ij}(f)|^2 \qquad (14)$$

Where $H_{ij}(f)$ is the i^{th} row and j^{th} column element of the transfer matrix. The normalized DTF is given as equation (15).

$$(\gamma_{j\to i}(f))^2 = |H_{ij}(f)|^2 \Big/ \sum_{j=1}^{n_c} |H_{ij}(f)|^2 \qquad (15)$$

Let consider $\mathbf{v}_s^t \in \mathbb{R}^{n_v \times 1}$ to be the obtained feature vector form DTF is like which $\mathbf{v}_s^t = \left[\mathbf{c}_{11}, \mathbf{c}_{12}, \ldots, \mathbf{c}_{ij}, \ldots, \mathbf{c}_{n_m n_f}\right]$. Number of extracted features is $n_v = n_m \times n_f$ where $n_m = n_c^{\,2}$ and $n_f = 36 \,\&\, 7$ due to methods of frequency decomposition. Overall, 2304 and 448 features are extracted for two frequency resolution.

Phase-Slope Index

Most of effective connection estimator like Granger causality are depend on the asymmetry of two channels. The asymmetry between two channels can happen not just because of the information flow but also for several factors such as independent background activity. To solve the problem of independent background activity which may cause asymmetry between the nodes, PSI method is proposed.
Consider two measures signals, $y_i(t)$ for i=1,2 consist of two independent sources $x_i(t)$ for i=1,2.

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \beta\eta(t) \quad (16)$$

Where $\eta(t)$ is a set of M independent noise sources which are added into the measurement channels by an unknown $2 \times M$ matrix B. PSI is based on the slope of cross-spectra phase of two signals i and j. The cross spectra between two signals is as follow:

$$S_{ij}(f) = \left\langle Y_i(f) * Y_j(f) \right\rangle \quad (17)$$

Where $\langle . \rangle$ denotes expectation value and the asterisk signifies matrix transposition and complex conjugation. Considering the cross spectra concept, the PSI is defined as

$$\psi_{ij} = \xi(\sum_{i=1}^{F} C(f) * C(f + \Delta f)) \quad (18)$$

Where C_{ij} is complex coherency and it is defined as $C_{ij}(f) = S_{ij}(f) \Big/ \sqrt{S_{ii}(f) S_{jj}(f)}$, Δf is the frequency resolution, F is the frequency interval which slope is summed, IM is imaginary operator and S is cross spectra. Finally, normalized output is always better to deal with. Thus, the last step is to make the output of SPI normalized.

$$\psi_{ij} = \psi_{ij} \Big/ \sigma_{\Psi} \quad (20)$$

With std being estimated by the jackknife method. Consider the obtained feature vector form PSI be like $\mathbf{v}_s^t \in \mathbb{R}^{n_v \times 1}$ which $\mathbf{v}_s^t = \left[\mathbf{c}_{11}, \mathbf{c}_{12}, \ldots, \mathbf{c}_{ij}, \ldots, \mathbf{c}_{n_m n_f} \right]$. Number of extracted features is $n_v = n_m \times n_f$ where $n_m = n_c^{\,2}$. Number of windows in frequency domain is equal to 36 and 7 for high and low resolution of decomposition which leads to 2304 and 448 features, respectively.

Feature selection

Combination of signal decomposition techniques and feature selection methods leads to hundreds of features. Although each feature contains neuropsychological data in time and frequency domain but the most informative features should be selected as the best features. The most informative feature is defined as the one which can discriminate the EEG data of the two groups with highest accuracy. For this purpose, a parameter should be defined to assess the obtained feature and demonstrate the level of efficiency. Therefore, a procedure based on two-sample t-test is used. T-test is a statistical examination of two population means. A two-sample t-test examines whether the means of two sets of samples are

different [45]. The two-sample t-test between target and control groups is performed for each feature separately. The significance level of each feature is then presented by p-value which less p-value means more significance difference between the two groups. Subsequently, all of features with P-value less than 0.01 are selected and also sorted form least to most P-values. Then, a simple forward feature selection method is used to find the best subset of features with highest classification accuracy. After all, the selected features are validated by one-leave-out method as it is mentioned in the next section.

Classification Methods

By selecting the best features, now, it is the time to find out how accurate these features could separate stressful group from the stress-free group. One of the well-known, fast and simple classifiers that have been used in a large number of signal and image applications is Bayesian classifier. This classifier is particularly used when the dimensionality of the inputs is large [46]. The classifier will model the distribution of each class, and then based on each model, will predict the probability of belonging a data to the classes. Hence, it tries to minimize the misclassification rate as much as possible by adjusting the model of distribution of classes [61]. *K*-nearest neighbor (KNN) is another well-established and simple classifier method. The algorithm is based on *K* closest training examples in the feature space and works well when there is a little or no previous information about the distribution of data [46,47]. In this study the *K* is considered to be one. The other classification algorithm used in this study is support vector machine (SVM). A SVM classifies the data by finding the best hyperplane that separates all data points of one class from those of the other class. The best hyperplane for an SVM means the one with the largest margin between the two classes. In this study, a linear SVM has been used to avoid the setting of parameters in SVM with kernels such as sigma in radial basis function (RBF). The classifiers parameters were tuned so that lead to the highest accuracy rate. An overview of the whole procedure is presented in Figure 4.

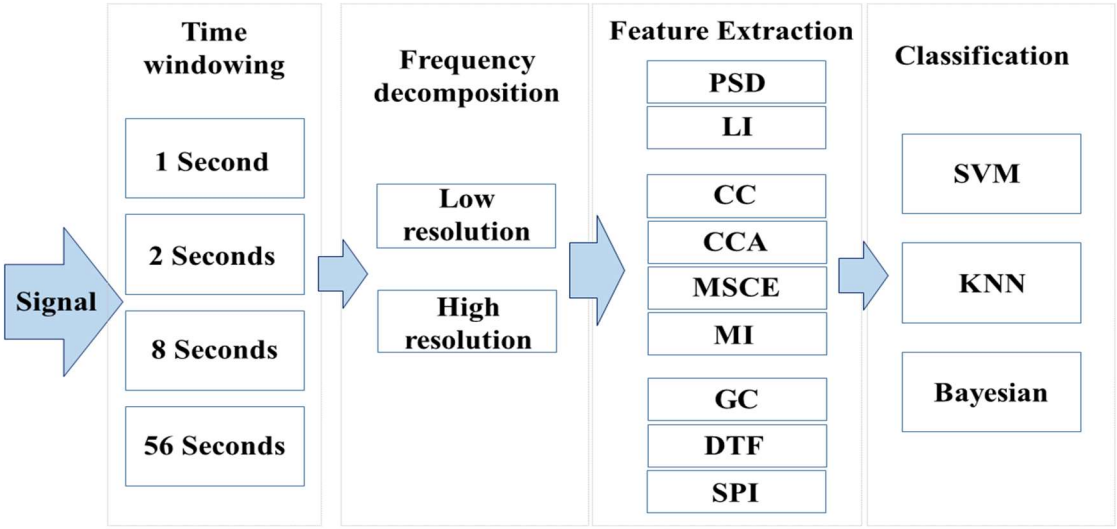

Figure 4. The overview of the classification paradigm

Results

The goal of this study is to identify subjects' stress level using their EEG data. Hence, the subjects' stress level is measured by PSS-14 questionnaire and their EEG data were collected in various mental states including the rest and four emotional states. Subjects were labeled to two classes of stressful and stress-free based on their answers to PSS-14. Each subject was exposed to four emotional stimuli and the induced emotional states (arousal and valance levels) were measured by SAM questionnaire. The emotional states were categorized by considering boundary for valance and arousal levels. Four classes of emotions including two valence states (positive and negative) and two arousal states (calm and excited) were identified. The numbers of subjects identified in each emotional states are presented in Tabel 1. The investigation in different arousal level was eliminated from consideration due to the unbiased and small number of subjects in each group.

Table 1. Number of valid subjects for each emotional states

Emotional state / Group	Arousal < 4	Arousal > 6	Valence <4	Valence >6
Stressful (PSS-14 ≥ 28)	5	10	13	13
Stress-free (PSS-14 ≤ 21)	15	8	12	16

Thus, determination of stress level is limited to high and low valence categories which in this paper is referred as negative and positive emotion states. Beside positive and negative emotional states, the eyes open resting state is also investigated. Three supervised classification methods including SVM, KNN and NB are used to classify each subjects into stressful or stress-free classes. Tables 2 to 5 present the results of validated classification accuracy by considering the fact that features with P-value less than 0.01 are taken to account. Hachure in few boxes in these tables is representative for not finding any features with P-value less than 0.01. Tables 2 and table 3 present the classification accuracy for identification of stress level during positive (Pos) and negative (Neg) emotional states. Results of all feature extraction methods with high resolution decomposition (36 bands each of them with 1Hz bandwidth) are shown in Table2 and results for low resolution decomposition (7 bands of theta, alpha, beta and gamma) are shown in Table3. Subsequently, results of eyes-open resting state are also shown in Tables 4 and 5. The results for all extracted features with high resolution of frequency decomposition are presented in Table 4 and with low resolution one are shown in Table 5.

Table 2: Classification accuracy for high resolution decomposition during positive (Pos) and Negative (Neg) emotional states

Features	Window size	1sec		2sec		8sec		56sec	
	State / Classifier	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos
PSD	**SVM**								
	KNN								
	Bayesian								
LI	**SVM**		80.00				77.1429		80.00
	KNN		77.143				82.857		71.429
	Bayesian		77.143				77.143		77.143
MSCE	**SVM**								
	KNN								
	Bayesian								
CC	**SVM**			100	97.142	100	85.714	90	82.857
	KNN			100	97.142	100	97.142	83.333	82.857
	Bayesian			96.666	94.285	100	97.142	80	80
CCA	**SVM**	100	100	100	100	100	94.285	86.666	97.142
	KNN	100	94.285	93.333	97.142	96.666	88.571	80	88.571
	Bayesian	100	94.285	100	97.142	100	94.285	86.666	97.142
MI	**SVM**	0	0	0	0	0	0	0	0
	KNN	0	0	0	0	0	0	0	0
	Bayesian	0	0	0	0	0	0	0	0
GC	**SVM**	100	100	100	100	93.333	97.142	93.333	94.285
	KNN	100	100	100	100	96.666	94.285	83.333	91.428
	Bayesian	100	100	100	97.142	96.666	100	90	94.285
SPI	**SVM**	100	97.142	100	100	90	94.285	76.666	71.428
	KNN	93.333	97.142	100	100	96.666	91.428	80	62.857
	Bayesian	93.333	97.142	100	100	90	91.428	83.333	71.428
DTF	**SVM**	100	100	100	100	100	100	86.666	100
	KNN	100	97.142	100	100	96.666	100	80	94.285
	Bayesian	100	100	100	94.285	100	100	86.666	97.142

Table 3: Classification accuracy for low resolution decomposition during positive (Pos) and Negative (Neg) emotional states

Features	Window size	1sec		2sec		8sec		56sec	
	State / Classifier	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos
PSD	**SVM**	70.00							
	KNN	73.33							
	Bayesian	73.33							
LI	**SVM**	76.667	74.285	86.667	74.286	83.333	91.429	90.00	71.428
	KNN	70.00	68.571	80.000	68.571	86.667	94.285	83.333	82.857
	Bayesian	70.00	74.286	90.000	74.286	86.667	94.285	86.667	68.571
MSCE	**SVM**						94.285	73.333	74.285
	KNN						91.429	66.667	54.285
	Bayesian						88.571	70.00	74.285
CC	**SVM**	100	100	100	100	80	82.857	73.333	80
	KNN	100	97.142	96.666	94.285	83.333	88.571	66.666	65.714
	Bayesian	100	94.285	100	100	76.666	82.857	76.666	77.142
CCA	**SVM**	100	91.428	100	97.142	93.333	91.428	73.333	82.857
	KNN	93.333	94.285	100	91.428	86.666	88.571	83.333	57.142
	Bayesian	100	97.142	100	94.285	90	91.428	80	80
MI	**SVM**	0	0	0	0	0	0	0	0
	KNN	0	0	0	0	0	0	0	0
	Bayesian	0	0	0	0	0	0	0	0
GC	**SVM**	100	100	100	97.142	100	97.14286	83.333	88.571
	KNN	86.666	94.285	93.333	94.285	93.333	97.142	80	85.714
	Bayesian	100	100	100	100	96.666	94.285	80	94.285
SPI	**SVM**	96.666	100	100	100	86.666	97.142	80	80
	KNN	90	91.428	86.666	97.142	76.666	85.714	73.333	82.857
	Bayesian	83.333	94.285	90	94.285	96.666	94.285	73.333	80
DTF	**SVM**	100	100	100	100	100	100	86.666	82.857
	KNN	96.666	97.142	100	100	100	100	80	88.571
	Bayesian	100	100	93.333	100	96.666	97.142	83.333	88.571

Table 4. Classification accuracy based on high resolution decomposition during eyes-open resting state

Features	Window size / Classifier	1sec	2sec	8sec	56sec
PSD	SVM	94.444	100.00	66.666	72.222
	KNN	88.888	88.888	66.666	94.444
	Bayesian	94.444	94.444	77.777	83.333
LI	SVM	100.00	100.00	77.777	
	KNN	100.00	100.00	61.111	
	Bayesian	100.00	100.00	77.777	
MSCE	SVM	100.00	100.00	100.00	77.778
	KNN	100.00	100.00	100.00	72.222
	Bayesian	94.444	94.444	100.00	77.777
CC	SVM	88.888	94.444	100	
	KNN	100	100	100	
	Bayesian	100	100	100	
CCA	SVM	100	100	100	100
	KNN	100	100	100	83.333
	Bayesian	100	100	100	94.444
MI	SVM	0	0	0	0
	KNN	0	0	0	0
	Bayesian	0	0	0	0
GC	SVM	100	100	100	88.888
	KNN	100	100	100	94.444
	Bayesian	100	100	94.444	94.444
SPI	SVM	100	100	100	83.333
	KNN	100	88.888	88.888	83.333
	Bayesian	100	100	100	94.444
DTF	SVM	100	100	100	88.888
	KNN	100	100	100	88.888
	Bayesian	94.444	100	100	88.888

Table 5. Classification accuracy based on low resolution decomposition during eyes-open resting state

Features	Window size / Classifier	1sec	2sec	8sec	56sec
PSD	SVM	83.333	83.333		72.222
	KNN	72.222	72.222		55.555
	Bayesian	83.333	83.333		72.222
LI	SVM	100.00	100.00	100.00	
	KNN	100.00	100.00	100.00	
	Bayesian	94.444	94.444	88.888	
MSCE	SVM	100.00	100.00	100.00	83.333
	KNN	100.00	94.444	100.00	55.555
	Bayesian	100.00	100.00	88.888	77.777
CC	SVM	100	100	88.8889	

	KNN	100	100	88.8889	
	Bayesian	100	88.888	88.8889	
CCA	SVM	100	100	94.444	77.777
	KNN	94.444	100	94.444	72.222
	Bayesian	100	100	94.444	77.777
MI	SVM	0	0	0	0
	KNN	0	0	0	0
	Bayesian	0	0	0	0
GC	SVM	100	100	100	88.888
	KNN	94.444	100	100	83.333
	Bayesian	100	88.888	100	83.333
SPI	SVM	100	100	100	94.444
	KNN	88.8889	94.444	100	88.888
	Bayesian	94.444	100	100	83.333
DTF	SVM	100	100	100	83.333
	KNN	100	100	100	94.444
	Bayesian	88.888	100	100	83.333

According to the tables 2 to 5, it is hard to select the best feature extraction method, best classification procedure and also the best resolution for time-wise and frequency-wise decomposition. Still, responding to the main question is challengeable that if emotional states and also resting state can be good modalities to predict the level of stress. Although, these questions cannot be responded flawlessly but some statistical analysis are done to clarify the results. For instance, the mean and standard deviation of classification accuracy for emotional states and resting state is presented in figure 5. Three states (Negative and positive emotion and resting state) are divided by two categories of high and low resolution frequency decomposition. According to figure 5, the eyes open resting state is a good modality to find stress level and between the emotional states, the positive emotions can be a better modality than negative ones. Form the fact that positive emotions can be a better modality, it can be interpreted that long term mental stress can influence positive emotions more than negative ones.

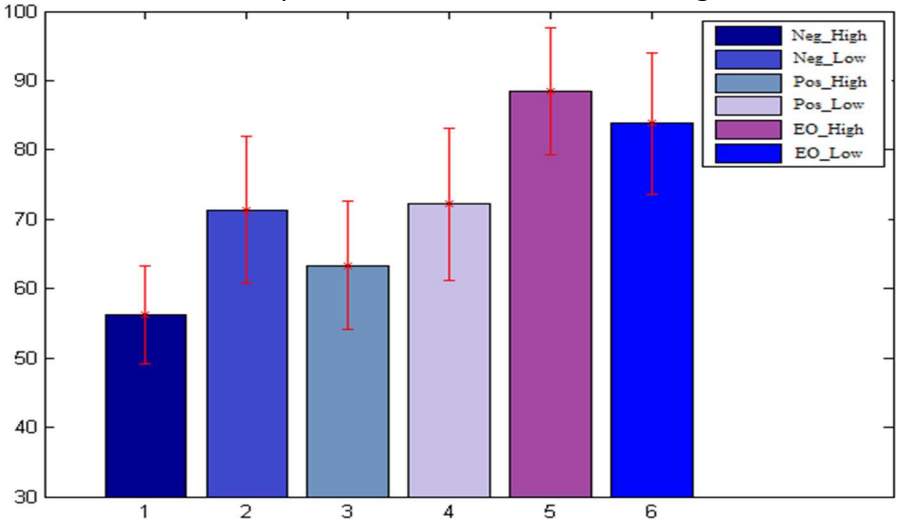

Figure 5. Mean and standard deviation of classification accuracy for different states.

In figure 5 the 'Neg', 'Pos' and 'EO' are stand for negative and positive emotions and eyes open states, respectively. The 'High' and 'Low' terms represent high and low resolution of frequency decomposition. Figure 6 assesses the effect of different time domain decomposition approaches on EEG signal. Same as previous figure, height of each bar refers to mean of accuracy and the red error line on top of each bar shows the standard deviation of the category. As it can be observed, using whole time interval of signal to develop the BCI can lead to lower efficiency of the BCI. Windowing of signal by 8 second intervals can provide us the highest accuracy but it should be noticed that the variance of this category is much higher than 2-second one.

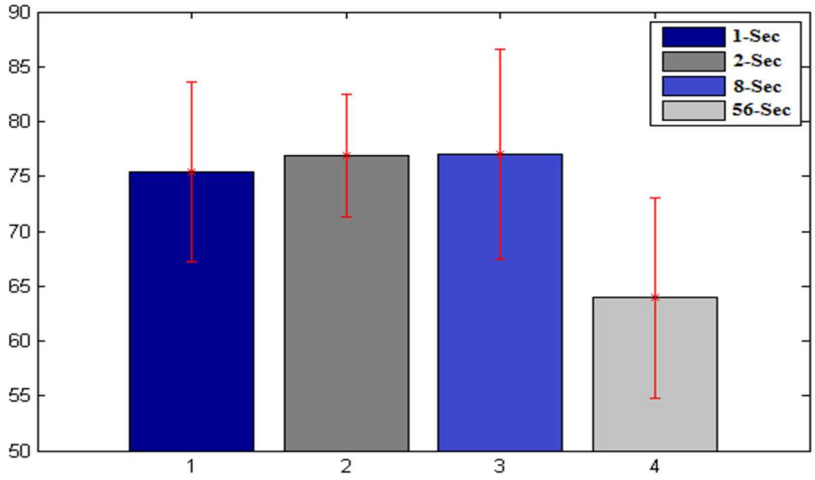

Figure 6. The mean of classification accuracy for different time windowing approaches.

To find out which classifier performs better than the other two, the performance of the three classifiers are investigated. Figure 7 shows the mean value and standard deviation of classification accuracy for each classifier.

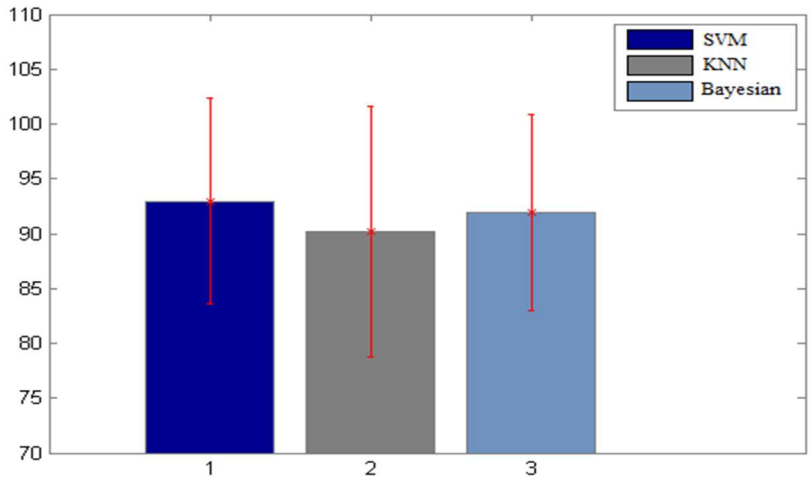

Figure 7. The performance of classifiers.

According to figure 7, the SVM classifier shows superiority to other classification methods used in this study. Although, the SVM shows higher classification rate but the running time of SVM is about 5 times longer than Bayesian and KNN. Thus, in online tasks, the Bayesian can be better choice in expense of decreasing in classification rate. The efficiency of feature extraction methods is another critical issue in the BCI design and brain signal processing. Presenting the performance of feature extraction methods, figure 8 shows the performance of implemented methods. As it can be inferred form figure 8, the basic methods based on power spectrum of signals cannot be a proper choice to classify the EEG data. Obviously, the functional and effective connection methods show superiority to PSD-based methods and between all of proposed methods, the DTF has the best accuracy and shows the highest performance.

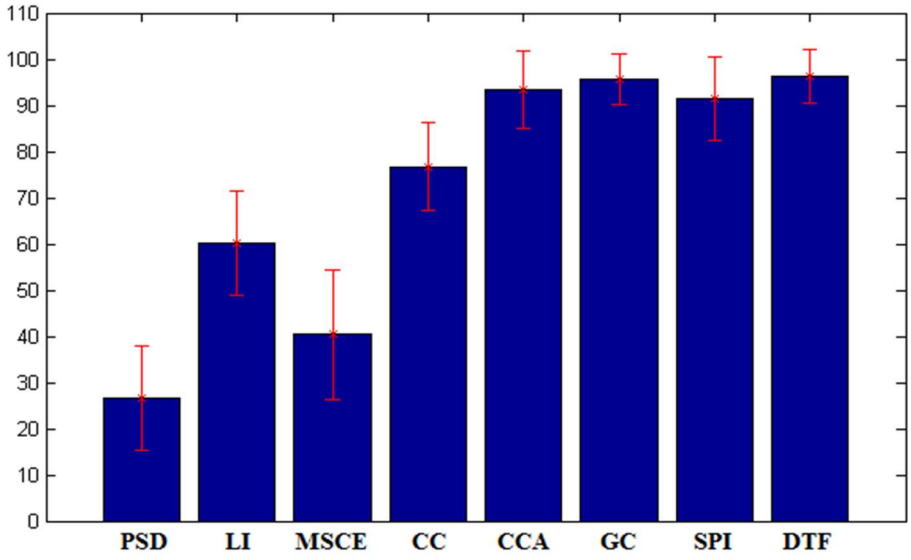

Figure 8. The performance of each feature extraction method.

Table. The most effective connections by each methods.

Modality	Negative Emotion Feature/ Frequency band/ Time	Positive Emotion Feature / Frequency band / Time	Eyes open resting state Feature / Frequency band / Time
PSD	$F_3 / \theta / t = 7\,\text{sec}$	---	$P_3 / \theta / t \in [20,22]$ $F_3 / \beta_2 / t$ □ 10 $F_3 / \alpha_1 / t = 17$
LI	$T / \gamma_2 / t \in \{15,22,43\}$	$T / \gamma_2 / t = 32$ $C / \theta / t \in [4,8]$	$C / \theta \,\&\, \alpha_1 / t < 10$ $C / \beta_1 / t \in [20,30]$
MSCE	$P_4 - T_8 / \beta_1 / t < 5$	$F_3 - P_3 / \beta_2 / t < 15$ $F_3 - P_4 / \gamma_2 / t < 10$ $P_4 - T_8 / \beta_2 / t = 25$ $P_4 - T_8 / \gamma_1 / t < 10$	$T_7 - T_8 / \beta_2 / t \in [25,30]$ $P_3 - P_4 / \gamma_2 / t \in [10,15]$
CC	$F_3 - F_4 / \beta_1 / t = all$ $F_3 - F_4 / \gamma_1 / t \in [16,24]$	$F_3 - F_4 / f = all / t = all$	Not a particular feature has found
GC	$C_4 \to F_3 / \gamma_1 \,\&\, \gamma_2 / t = 25$ $P_3 \to T_7 / f = all / t \in [15,30]$	$P_3 \to T_7 / f = all / t < 10$ $C_3 \to T_7 / \beta_2, \gamma_1 \,\&\, \gamma_2 / t < 10$ $C_4 \to T_8 / f = all / t \in [20,30]$	$C_3 \to T_7 / \beta_2 \,\&\, \gamma_2 / t < 10$ $C_3 \to T_8 / \alpha_2 \,\&\, \gamma_1 / t < 10$ $C_3 \to F_3 / f = all / t \in [30,40]$
PSI	$F_3 \to F_4 / \alpha_1, \gamma_1 \,\&\, \gamma_2 / t = all$ $C_3 \to F_4 / \beta_2 / t = all$ $F_3 \to T_7 / \gamma_1 \,\&\, \gamma_2 / t < 10 \,\&\, t > 50$	$C_4 \to F_3 / \beta_2 \,\&\, \gamma_1 / t = all$ $C_3 \to T_7 / \beta_2 \,\&\, \gamma_1 / t \in [15,25]$ $C_4 \to T_7 / \theta \,\&\, \alpha_1 / t < 10$	$C_3 \to F_4 / \beta_2 / t \in [8,16]$ $F_3 \to T_8 / \beta_2 \,\&\, \gamma_2 / t \in [40,50]$
DTF	$C_4 \to F_3 / f = all / t > 50$ $F_4 \to F_3 / \beta_1, \beta_2, \gamma_1 \,\&\, \gamma_2 / t \in [20,30]$	$C_4 \to F_3 / \alpha_1, \alpha_2, \beta_1 \,\&\, \beta_2 / t \in [20,40]$ $F_3 \to F_4 / \theta \,\&\, \gamma_1 / t > 40$	$C_4 \to F_4 / \beta_2 / t = all$ $T_8 \to T_7 / \theta \,\&\, \alpha_1 / t \in [20,25]$

Centro-temporal: benign Rolandic Epilepsy (Several references are available) GC modality shows C-T connections.

Parieto-Temporal: WE have found it before and the same relation has found in schizophrenia and autism spectrum disorder (several references are mentioned in the red paragraph)
This novel tip (MSCE feature for long-term stress identification) can be used for intervention purpose such as in a neurofeedback using the most significant patterns. Since MSCE features extracted in trials with length of 56 seconds show the highest classification accuracy using SVM classifier, therefore, the most informative features of the this scenario could be selected as a therapy marker. Table 6 presents the most significant connections that are changed due to long-term mental stress.

The most significant feature which produces high level of discrimination between stressful and stress-free subjects is the functional connection between left parietal (P3) and right Temporal (T8) regions in gamma band (P_value=2.69×10^{-4}). In average, the synchronization between the above mentioned regions

is decreased about 55 percent in stressful subjects. The alteration of cortical synchrony between temporal and parietal regions has been reported in some mental disorders such as autism (ref) and schizophrenia [48]. For instance, it has been shown that resting state functional connection decrease between parietal and temporal regions and between the temporal cortices bilaterally in schizophrenic patients [48]. The attenuation of short-range connectivity within bilateral temporal and left parietal regions has also been reported in adults with autism spectrum disorder (ASD) [49]. It is interesting to know that adults with ASD show significantly higher subjective stress and poorer ability to cope with stress in everyday life, as compared to healthy controls [50] . Any conclusion can be made but what if we consider the long-term stress as a facilitating factor for mental disorders which influence the functional connection (FC) between temporal and parietal regions. Therefore, this facts highlight the importance of this FC which needs to be more taken care.

It should be noted that long-term stress does not only influence temporal-parietal connection but this alteration happens together with other changes such as variation of FC at beta band between left central and left temporal connection (C3 and T7). The high level of identification of stress level has been achieved using all of these features.

There are some suggestions for future works in this area. First of all, each scientific research will bring more reliable results by having more subjects. In this research, we only could recruit 26 young adults which can be improved in a more general study. Not only the number but age and gender of subjects are also very important. Therefore, a study on a group of subjects with a wider range of age and investigation of effect of gender on the results is suggested for future studies. In addition, using an EEG recorder with a better spatial resolution (64 or 128 channels) could uncover more details.

Conclusion

In this article, effect of long-term mental stress on brain functioning is investigated using EEG signal. The study is conducted using various mental states including the rest (eyes-open) and two emotional states (positive and negative valance). An EEG based BCI is then proposed to discriminate the stressful subjects from stress-free ones. Effectiveness of various feature extraction methods such as PSD, LI, MSCE, CC, CCA, MI, GC, DTF and SPI is evaluated in different time and frequency resolutions. The extracted features from each time and frequency resolution are utilized for classification purpose. Three classification algorithms namely KNN, NB and SVM were applied and the results were validated by one-leave out method. The obtained results showed the superiority of functional connectivity method (MSCE) to others. Consequently, use of the most significant discriminating features, working on enhancement of FC between temporal and parietal region was suggested as potential way to cope with long-term mental stress.

Refereces:

[1] U. Lundberg, "Stress and (Public) Health," *in International Encyclopedia of Public Health, Oxford: Academic Press*, pp. 241-250, 2008.

[2] C.R. Noback, R.J. Demarest, "The Nervous System, Introduction and Review," *annals of neurology McGraw-Hill Book Company*, pp. 323, 1986.

[3] O. Kohlish and F. Schaefer, "Physiological changes during computer task: responses to mental load or to motor demands," *Ergonomics*. vol. 39, no. 2, pp. 213-224, 1996.

[4] A.Gevins, M.E. Smith, H. Leong, L. McEvoy, S. Whitfield, R. Du, and G. Rush, "Monitoring working memory load during computer-based tasks with EEG pattern recognition Methods," *Human Factors.* vol. 40, no. 1, pp. 79-91, 1998.

[5] R.N. Goodman,, J.C. Rietschel, L.C. Lo, M.E. Costanzo, B.D. Hatfield, "Stress, emotion regulation and cognitive performance: The predictive contributions of trait and state relative frontal EEG alpha asymmetry"

[6] S. Cohen, R. Kessler and L. Gordon "Measuring stress – A Guide for Health and Social Scientists," *Oxford University Press*, 1997.

[7] S. Cohen, T. Kamarck and R. Mermelstein "A global measure of perceived stress," *Journal of Health and Social Behavior*, vol. 24, no. 4, pp. 385–396, 1983.

[8] S. Cohen, T. Kamarck and R. Mermelstein "A global measure of perceived stress," *Journal of Health and Social Behavior,* vol. 24, pp. 385-396, 1983.

[9] T. T. Brunyé, S. R. Cavanagh and R. E. Propper. "Hemispheric Bases for Emotion and Memory." Frontiers in Human Neuroscience. *Frontiers Media S.A.*, 05 Dec. 2014. Web. 31 Mar. 2016.

[10] J. J. B. Allen, J. A. Coan and M. Nazarian, "Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion," *Biological Psychology*, vol. 67*,* pp. 183-218, 2004.

[11] R. J. Davidson "Anterior electrophysiological asymmetries, emotion, and depression: Conceptual and methodological conundrums," *Psychophysiology*, vol. 35, pp. 607–614, 1998.

[12] K. C. Kadosh, Q. Luo, C. Burca, M. O. Sokunbi, J. Feng, D. E. J. Linden and J. Y. F. Laub, "Using real-time fMRI to influence effective connectivity in the developing emotion regulation network," *Neuroimage*, vol. 125, pp. 616–626, 2016.

[13] B. A. Corbett, V. Carmean, S. Ravizza, C. Wendelken, M. L. Henry, C. Carter and S. M. Rivera, "A functional and structural study of emotion and face processing in children with autism," *Psychiatry Reseach*, vol. 173, no. 3, pp. 196-205.

[14] T. Ethofer, J. Bretscher, M. Gschwind, B. Kreifelts, D. Wildgruber and P. Vuilleumier, "Emotional Voice Areas: Anatomic Location, Functional Properties, and Structural Connections Revealed by Combined fMRI/DTI," Oxford press, 2011.

[15] A. Karidis, "Neurofeedback-The Scientific Evidence Grows" Perth Brain Centre. Retrieved 2016.

[16] D. E. Linden, I. Habes, S. J. Johnston, S. Linden, R. Tatineni, L. Subramanian, B. Sorger, D. Healy, R. Goebel, "Real-time self-regulation of emotion networks in patients with depression," Plos One, vol. 7, no. 6, 2012.

[17] J. A. Cameron, "Passive Infrared Hemoencephalography: Four Years and 100 Migraines," *Journal of Neurotherapy*, vol. 8, no. 3, pp. 23-51, 2005.

[18] R. Cowie and R.R. Cornelius, "Describing the emotional states that are expressed in speech," *Speech*

Communication, vol. 40, pp. 5-32, 2003.
[19] G. Chanel, J. J. M. Kierkels, M. Soleymani and T. Pun "Short-term emotion assessment in a recall paradigm," *Int. J. Hum.-Comput. Stud.*, vol. 67, no. 8, pp. 607-627, 2009.
[20] *R Khosrowabadi, C Quek, KK Ang, A Wahab, "ERNN: A Biologically Inspired Feedforward Neural Network to Discriminate Emotion from EEG Signal," IEEE Transactions on Neural Networks and Learning Systems, vol. 25 issue. 3, pp. 609 – 620, 2014.*
[21] B. B. Biswal, "Resting state fMRI: A personal history," *Neuroimage*, vol. 62, no. 2, pp. 938-944, 2012.
[22] R Khosrowabadi, C Quek, KK Ang, SW Tung, M Heijnen, "A Brain-Computer Interface for classifying EEG correlates of chronic mental stress," The International Joint Conference on Neural Networks (IJCNN), 2011.
[23] P. Fusar-Poli, A. Placentino, F. Carletti, P. Allen, P. Landi, M. Abbamonte "Laterality effect on emotional faces processing: ALE meta-analysis of evidence," *Neuroscience Letters*, vol. 452, no.3, pp.262-267, 2009.
[24] R.W. Thatcher, C.J. Biver and D. North, "EEG coherence and phase delays: Comparisons between single reference, average reference and current source density," *Tampa, Florida: University of South Florida College of Medicine*, 2004.
[25] N.M. Kleinhans, T. Richards, L. Sterling, K.C. Stegbauer, R. Mahurin, L.C. Johnson, "Abnormal functional connectivity in autism spectrum disorders during face processing," *Brain and Cognition*, vol. 131, no. 4, pp. 1000-1012, 2008.
[26] S. Palva and J.M. Palva, "New vistas for a-frequency band oscillations," *Trends Neurosci*. 2007.
[27] R. Khosrowabadi, Q.Hiok Chai, A. Wahab and A. Kai Keng, "*EEG-based Emotion Recognition Using Self-Organizing Map for Boundary Detection,"20th International Conference on Pattern Recognition (ICPR)*, pp. 23-26, 2010.
[28] R. Khosrowabadi, A.Wahab, K. K. Ang, and M.H. Baniasad, "affective computation on EEG correlates of emotion from musical and vocal stimuli,"*Paper presented at the international joint conference on Neural Networks, Atlanta, Georgia, USA*, 2009.
[29] J. Kim and E. Andre, "Emotion recognition based on physiological changes in music listening,"*IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 12, pp. 2067–2083,2008.
[30] M.Murugappan, N. Ramachandran, and Y. Sazali, "Classification of human emotion from EEG using discrete wavelet transform,"*Journal of Biomedical Science and Engineering*, vol. 3, no. 4, pp. 390-396, 2010.
[31] P. C. Petrantonakisand L. J. Hadjileontiadis, "Emotion recognition from EEG using higher order crossings,"*IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp.186-197, 2010.
[32] L.Yuan-Pin, W.Chi-Hong, J.Tzyy-Ping,W.Tien-Lin, J. Shyh-Kang,D. Jeng-Ren, "EEG-based emotion recognition in music listening,"*IEEE Transactions on Biomedical Engineering*, vol. 57, no.7, pp.1798-1806, 2010.
[33] L.I.Aftanas,A.A. Varlamov, S.V. Pavlov, V.P. Makhnev andN.V. Reva, "Affective picture processing: event-related synchronization within individually defined human theta band is modulated by valence dimension," *Neuroscience Letters*, vol. 303, no. 2, pp. 115-118, 2001.
[34] J.A.Coan andJ.J.B. Allen "Frontal EEG asymmetry as a moderator and mediator of emotion,"*Biological Psychology*, vol. 67(1-2), pp. 7-50, 2004.
[35] A.H.Kemp, M.A.Gray, P.Eide, R.B.Silberstein and P.J. Nathan, "Steady-State visually evoked potential topography during processing of emotional valence in healthy subjects,"*NeuroImage*, vol. 17, no. 4, pp.1684-1692, 2002.
[36] J.K.Olofsson, S. Nordin, H. Sequeira, and J. Polich, "Affective picture processing: An integrative review of ERP findings,"*Biological Psychology*, vol. 77, no. 3, pp. 247-265, 2008.

[37] A.Brovelli, M. Ding, A. Ledberg, Y. Chen, R. NakamuraandS.L. Bressler, "Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality,"*Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 26, pp. 9849-9854, 2004.

[38] C.J.Honey,R. Kötter, M.Breakspear and O.Sporns, "Network structure of cerebral cortex shapes functional connectivity on multiple time scales,"*Proceedings of the National Academy of Sciences*, vol. 104, no. 24, pp. 10240-10245, 2007.

[39] C.C.Hilgetag, G.A.P.C. Burns, M.A.O'Neill, J.W. Scannell and M.P. Young,"Anatomical connectivity defines the organization of clusters of cortical areas in the macaque and the cat,"*Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 355, no. 1393, pp. 91-110, 2000.

[40] K.J. Friston, "Functional and effective connectivity in neuroimaging: A synthesis,"*Human Brain Mapping*, vol.2(1-2), pp. 56-78,1994.

[41] S. M.Kay, "Modern Spectral Estimation: Theory and Application,"*Englewood Cliffs, NJ: Prentice Hall*, 1988.

[42] H. Wolfgang, S. Léopold, "Canonical Correlation Analysis," Applied Multivariate Statistical Analysis. pp. 321–330, ISBN 978-3-540-72243-4, 2007.

[43] K. Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space" *Philosophical Magazine,* vol. 2, no. 11, pp. 559–572, 1901.

[44] M. Kaminiski and K. Blinowska, "A new method of description of the information flow in the brain structure," Biological Cybernetics, vol. 65, no. 3, pp. 203-210, 1991.

[45] John A. Rice, "Mathematical Statistics and Data Analysis," Third Edition, Duxbury Advanced, 2006.

[46] H. Parvin, H. Alizadeh and B. Minati, "A modification on K-nearest neighbor classifier," *Global Journal of Computer Science and Technology*, vol. 10, no. 14, pp. 37-41, 2010.

[47] S. Theodoridis and K. Koutroumbas, "Pattern Recognition" *Academic Press*, 2006.

[48] A.Venkataraman, T.J. Whitford, C.F. Westin, P. Golland and M. Kubicki, "Whole brain resting state functional connectivity abnormalitiesin schizophrenia," Schizophr .Res. vol. 139, pp. 7–12, 2012.

[49] V. Tsiaras, P.G. Simos, R. Rezaie, B.R. Sheth, E. Garyfallidis, E.M.Castillo, A.C.Papanicolaou, "Extracting biomarkers of autism from MEG resting-state functional connectivity network," computers in biology and medicine, vol.41, issue. 11, pp. 1166-1177, 2011.

[50] T. Hirvikoski, M. Blomqvist, "High self-perceived stress and poor coping in intellectually able adults with autism spectrum disorder," Autism, 2014.