Classification of Frontal EEG Signals of Normal Subjects to Differentiate Gender by Using Artificial Neural Network

Salmi Abdul Ghani, Norliza Zaini, Haryanti Norhazman, Ihsan M. Yassin, Maizura M. Sani
Faculty of Electrical Engineering (Electronic Computer),
Universiti Teknologi MARA, Selangor, Malaysia,
salmiag90@gmail.com

***Abstract*—Varying mental states of an individual can influence their brainwave patterns. This is also true for individuals of different gender, where a male's EEG signal is different from a female's EEG signal. This provides a context for our research, where our main aim is to classify different patterns of EEG based on different gender. This paper presents our initial study to classify gender of normal subjects based on their frontal EEG signals. Forty normal subjects have participated in this experiment, where their EEG signals have been recorded for analysis purpose. The recorded raw EEG data is first pre-processed and filtered into 4 different frequency sub-bands. Two types of analysis were then conducted; the first analysis took into consideration all four sub-bands of frontal EEG, whereas the second analysis only considered two sub-bands namely alpha and beta bands. The features extracted from the selected sub-bands are in the form of EEG Energy Spectral Density (ESD) values, which are then fed into an Artificial Neural Network (ANN) classifier for classification purpose; i.e. to distinguish between male and female. Based on results obtained from the analysis, it is found that higher classification accuracy can be achieved from combining four sub-bands when compared to if only two sub-bands (alpha and beta) are being considered.**

***Index Terms*—Classification; EEG Signals; Frontal; ESD; ANN; Gender.**

I. Introduction

As how mental state of an individual affects their EEG brainwave frequency, different gender is also found to give effects onto brainwave signals. In this context, some researchers have investigated the effects of age and gender towards the characteristics of sleep EEG power spectral density (PSD) [1]. Through the experiments, they found out that females experience higher PSD during non-rapid eye movement (NREM) sleep in Delta, Theta and low Alpha frequency range (0.25 – 9.0 Hz) and also in high Sigma frequency compared to males. Another report also studied on EEG signals that assumed to be affected by age and gender in the development of normal child. The studies found out that gender differences exist, where males produce more alpha and less theta compared to females [2]. This shows that male and female have different characteristics in their brainwave, which also vary according to different mental states.

EEG signals can be categorized by frequency range called sub-bands. Different kinds of information are carried by each sub-band. Four sub-bands can be used to represent the EEG signals, which are Delta frequency (1-4Hz); associated with deep sleep, Theta frequency (4-8Hz); correlates with light sleep, Alpha frequency (8-12Hz); indicate calmness and peace, and Beta frequency (13-21Hz); present thinking and focusing state [3]. Normal people are likely to have a high level of EEG Alpha brainwave (that indicate calmness and peaceful state) and a low level of EEG Beta brainwave [4]. Frontal cortex is found to be highly connected with emotional processing [5]. However, until now, not many researches have explored the correlations between frontal asymmetry with gender.

There are many types of classifiers exist such as Support Vector Machine (SVM), k-Nearest Neighbour (kNN), Fuzzy C-Means (FCM) and Artificial Neural Network (ANN). All of these classifiers can be used to classify both linear and non-linear types of data set. In this preliminary study, this research will only focus on exploring the use of ANN as the classifier. Among all classifiers, ANN is one of the most popular classifiers being used by most researchers nowadays in many applications, e.g. pattern recognition, prediction, classification, modelling and control. This is due to the fact that ANN is able to give promising results, e.g. in agriculture sector [6]. It has been proven to be a good classifier based on its ability in performing accurate classification of data. Previous researchers have used ANN in classifying EEG signals of normal and epileptic subjects [7]. The results obtained have proved that ANN can successfully classify the EEG signals with high accuracy measure. The major advantage offered by ANN is its ability to learn and understand patterns based on sets of data, thus making it a smart tool for classification. The neural network is also able to process the rules of input and output relationship from the data set [8].

In the scope of our research, the process of distinguishing either male or female based on the individuals' EEG signals is considered as one of the important features in EEG-based profiling mechanism. Such profiling mechanism can be incorporated into an application, e.g. biometrics identification system that requires the automatic capturing of profile data based on EEG signals e.g. gender, age-range, race and habits [9]. In such application, identifying gender of an individual is part of the demographic identification, which is a significant step that can lead to categorization of profile and personality.

In this context, this paper presents an initial study to classify EEG signals of normal subjects based on the extracted frontal Energy Spectral Density (ESD) values on selected sub-bands to differentiate male and female subjects by using ANN as the classifier.

II. METHODOLOGY

A. Subjects

All subjects involved in this experiment are students of UiTM Shah Alam. A total of 40 healthy subjects (right-handed) that comprise of 12 males and 28 females have volunteered to participate in this experiment. All subjects have reported free from having any mental disorder history and not on medication. Prior to the experiment session, each subject was given a consent form to be filled up so that they are well-informed regarding the experiment and its procedure. Subjects were also requested not to consume any caffeine or drugs prior to the session. The whole experimental procedure has been approved by the ethics committee to be conducted.

B. Apparatus

To record EEG signals of subjects during the experiment sessions, the Emotiv Epoc Neuro-headset was used (see Figure 1). This device offers a high-resolution neuro signal acquisition.

Figure 1: Emotiv Epoc Neuro-Headset

Emotiv Epoc Neuro-headset consists of 14 channels and 2 reference channels (CMS and DLR) positioned at the relative locations on the scalp, following the International 10-20 Standard as shown in Figure 2. This device was tested prior to the experiment to ensure that it has no significant risk and pain-free. The sampling rate of this device is 128Hz. In this study, EEG signals acquired from only 8 electrodes namely AF3, AF4, F3, F4, F7, F8, FC5 and FC6 will be used in the analysis. These electrodes are those positioned at the frontal region, 4 from each hemisphere (left and right). This selection of electrodes is due to the fact that the frontal asymmetry is normally associated with emotion based on previous research [10].

During EEG recording session, an isolated earphone is used in order to reduce external noise and improve the quality of EEG signals being recorded. Prior to the experiment session, to determine whether a subject is normal or stress (on a varying range of severity), self-reported questionnaire Depress Anxiety Stress Scale (DASS) questionnaire [11] is used. This questionnaire is chosen in recognizing a subject's mental state because of its excellent core stability and sequential constancy and provides a better separation of the features of given scales [12].

C. Data Collection

The experiment sessions were organized into 2 stages. First, the subject was given a DASS questionnaire to answer in order to classify them either into normal or stress (varied level of severity) group. Such process of excluding stress subjects is performed in order to make it a controlled

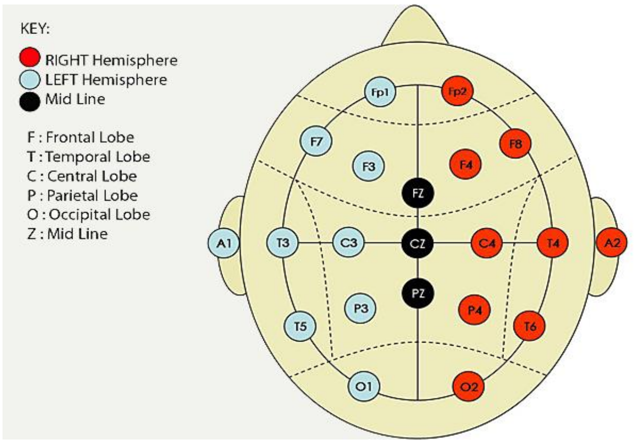

Figure 2: The International 10-20 system

experiment and to ensure that the process of distinguishing male and female EEG pattern is not being affected by EEG stress data, which may have varying patterns. A profile enquiry was also given to capture their profile details, e.g. gender and age-range. Next, the subject's EEG was recorded in the mode of "relaxing" and "doing nothing" for five minutes, where the subject was requested to close her eyes, while sitting comfortably on a chair in a dim lighted room. Once recorded, the EEG data was ready for pre-processing. The EEG data were first recorded in the 'edf' file via the Emotiv Testbench software and converted into a readable 'csv' file for offline processing purpose by using Matlab software.

D. Pre-processing EEG Raw Data

The raw EEG signals captured for subjects of both genders will be analyzed but need to be pre-processed beforehand. Such signal pre-processing includes artifacts removal operation, which is crucial in order to obtain clean EEG data that is free from noise and unwanted frequencies. Some of the artifacts are in the form of signals due to eye movement and blinks that could affect the quality of the EEG data.

2 minutes of the recorded EEG signals will be cut off (the first one minute and the last one minute). Then the 3 remaining minutes of EEG data will go through the band pass filtering (Hamming window) into four different frequency bands, Delta (0-4Hz), Theta (4-8Hz), Alpha (8-13Hz) and Beta (13-30Hz). Fast Fourier Transform (FFT) is being used to extract features which are Power Spectral Density (PSD) and Energy Spectral Density (ESD). ESD value is calculated by dividing the area of PSD curve with frequency bands. ESD is chosen for feature calculation because it covers whole energy distribution for each frequency band (range of frequency) [13]. In this paper, Asymmetry value of ESD is captured, which will then be fed into the classifier. Equation (1) shows the equation for EEG sub-bands hemispheric asymmetry [14]:

$$\textit{EEG hemispheric asymmetry} = \ln(\textit{Right}) - \ln(\textit{Left}) \qquad (1)$$

E. Artificial Neural Network

ANN is one of the artificial intelligence (AI) techniques for data processing that simulates the signal processing in human brain. ANN is defined as a 'black box' that will perform some computational processing based on the given inputs to produce some outputs. Based on the biological neural network function it promotes, ANN is capable to perform complex analysis, e.g. pattern recognition, function approximation, estimation, and controlling [15].

Figure 3 shows the architecture of a multilayer perceptron (MLP) of neural network model (feed-forward network) with one hidden layer. Generally, MLP comprises of an input layer, one or more hidden layer(s) of hidden units and an output layer. The computation generated underneath the ANN modelling is based on a weighted sum and bias of its input for each of the processing elements [16]. Initializing different number of hidden units before training sets of data may influence the learning and training process and may produce better result.

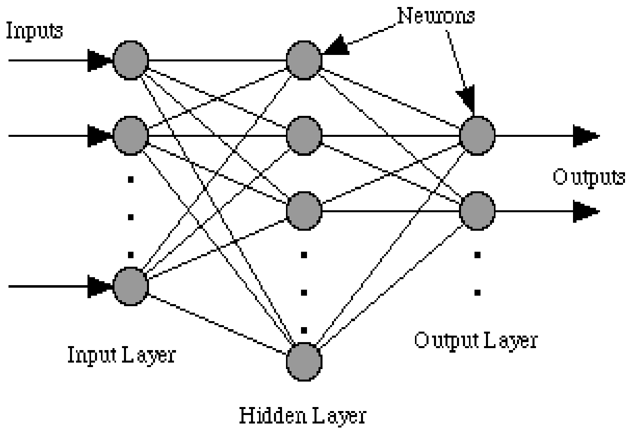

Figure 3: The MLP feedforward neural network architecture

F. ANN Classification

Figure 4 shows the overall stages of the ANN classifier using pattern recognition tool. 3 basic steps in the ANN modeling are training, validating and testing. To start with the ANN pattern recognition, the data set will be divided into 2 sets randomly. First set of 70% will be used for training process, while the remaining 30% will be used for both validating and testing process.

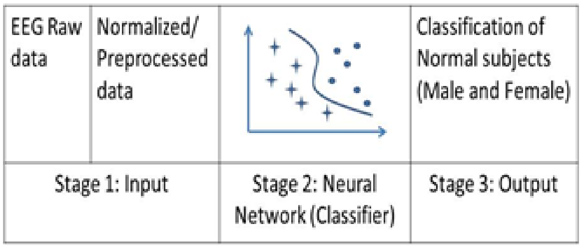

Figure 4: ANN pattern recognition stages

The inputs to the ANN classifier are in the form of frontal EEG signals from the different sub-bands, grouped into male and female data. The first analysis considers all the 4 sub-bands (Alpha, Beta, Delta and Theta) whereas the second analysis will just consider data from the 2 sub-bands (Alpha and Beta). The results from these two analysis will be compared as to identify which dataset will help to classify better, i.e. to distinguish between male and female.

Using MATLAB version R2013a, all inputs and targets need to be created and save in 'mat' file. The targets are indicated with 0 for female and 1 for male. The net function will be used; i.e.: net = patternnet (no. of hidden units). This function will execute the pattern recognition process based on the configured number of hidden units. Different number of hidden units and threshold are configured for both analysis. In this experiment, default training algorithm is used i.e. trainscg (scale conjugate gradient). Transfer function of tangent sigmoid (TANSIG) is used in both hidden and output layer since this experiment performs a pattern classification process [17]. In each analysis, the input data is basically divided into a default ratio, 70:15:15 which are for training, validating and testing respectively. Each computation of ANN may produce different results due to different initial conditions and sampling. Hence, to ensure the experiment can be repeated and re-produce the same results, the generator state is set to a fixed value. This setting will generate the same set of random numbers each time the optimization process executes and then being applied to MLP weights for processing [17].

III. EXPERIMENTAL RESULTS

This section presents the analysis and results of the classification of different gender based on frontal EEG data. Based on the two different analysis, either all 4 sub-bands or only 2 sub-bands (Alpha and Beta) are considered. For both analysis, ANN is trained with varying number of MLP structure as shown in Table 1.

Table 1
MLP structures tested.

Network	MLP Structure	
	2 Bands	4 Bands
1	2-5-1	4-5-1
2	2-10-1	4-10-1
3	2-15-1	4-15-1
4	2-20-1	4-20-1

Table 2 shows the percentage of classification accuracy based on different sub-bands for 4 different numbers of hidden units. Based on 2 sub-bands (Alpha and Beta) analysis, the result shows that the highest percentage of accuracy obtained by ANN classification is 72.5% for 5 and 15 hidden units with a threshold value of 0.5 and 10 hidden units with a threshold value of 0.6. Figure 5 shows the confusion matrix plotted for 5 hidden units. 29 data out of 40 data are found to be correctly classified according to gender targets.

In the analysis that considers all 4 sub-bands, the result shows that the highest percentage of accuracy obtained from this ANN classification model is 80% for 5 hidden units with a threshold value of 0.5. Figure 6 shows the confusion matrix plotted for 5 hidden units. 32 data out of 40 data are found to be correctly classified according to gender targets.

By comparing the values as shown in Table 2 and confusion matrix (as shown in Figure 5 and 6), higher classification accuracy in differentiating gender is obtained in the analysis based on all sub-bands features (Alpha, Beta, Delta and Theta) by using 4-5-1 structure. Low classification accuracy was obtained when the analysis is based on only 2 sub-bands (Alpha and Beta). High accuracy value could be obtained by a small number of hidden units in the ANN classifier may be due to the properties of the data, which is less complex. Whereas, larger size of hidden units that results in low classification accuracy is due to too many parameters that need to be considered.

Table 2
ANN Classification Accuracy of Different Sub-bands.

Network	No of hidden units	2 Bands (Alpha & Beta)		4 Bands (Alpha, Beta, Delta & Theta)	
		Threshold	Accuracy (%)	Threshold	Accuracy (%)
1	5	0.5	72.5	0.5	80*
2	10	0.6	72.5	0.4	70
3	15	0.5	72.5	0.6	72.5
4	20	0.3	70	0.8	72.5

* Highest accuracy.

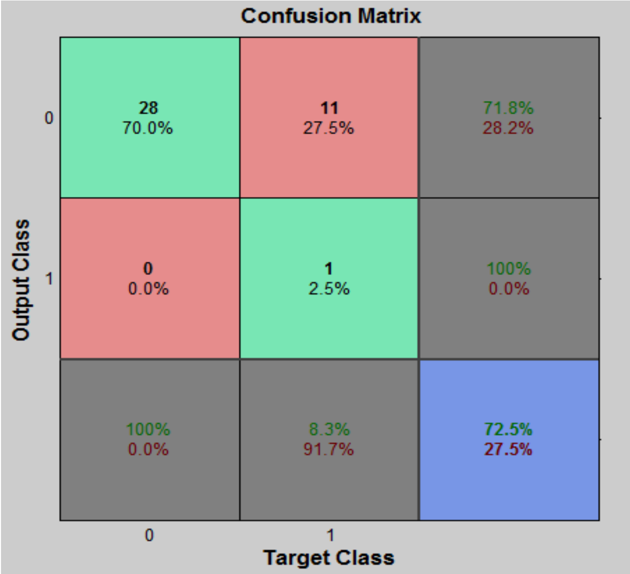

Figure 5: Confusion matrix plot based on Alpha and Beta bands (5 hidden units)

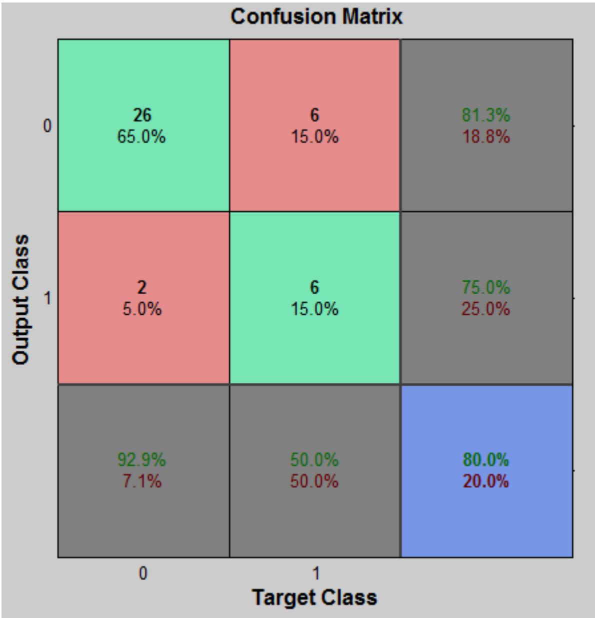

Figure 6: Confusion matrix plot based on Alpha, Beta, Delta and Theta bands (5 hidden units)

IV. CONCLUSION

The analysis presented in this paper employs ANN classification method to differentiate various EEG patterns of different gender i.e. normal male and female subjects. Based on the conducted experiment and analysis, it is shown that the ANN classifier is able to classify the different genders based on EEG signals of subjects. It is also found that this ANN classifier could perform better with inputs that are based on all EEG sub-bands; i.e. Alpha, Beta, Delta and Theta, when compared to inputs that are based on only 2 sub-bands (Alpha and Beta) from frontal EEG signals. Findings from this study may provide some initial knowledge required to develop an EEG-based profile application that can be used to capture profiling details e.g. gender, age, habits [9] etc. of a person just by analyzing their EEG signals.

ACKNOWLEDGMENT

We would like to extend our acknowledgement to Universiti Teknologi MARA (UiTM) and to those who have directly and indirectly contributed to our project. This research is funded by the Fundamental Research Grants Scheme Research, Ministry of Higher Education Malaysia (FRGS/1/2014/TK03/UITM/02/14).

REFERENCES

[1] J. Carrier, S. Land, D. J. Buysse, D. J. Kupfer, and T. H. Monk, 2001. The effects of age and gender on sleep EEG power spectral density in the middle years of life (ages 20–60 years old)," Psychophysiology. 38: 232-242.

[2] A. R. Clarke, R. J. Barry, R. McCarthy, and M. Selikowitz, 2001. Age and sex effects in the EEG: development of the normal child. Clinical Neurophysiology. 112: 806-814.

[3] N. M. Puzi, R. Jailani, H. Norhazman, and N. M. Zaini, 2013. Alpha and Beta brainwave characteristics to binaural beat treatment. in Signal Processing and its Applications (CSPA), 2013 IEEE 9th International Colloquium on. 344-348.

[4] T. L. Huang and C. Charyton, 2008. A comprehensive review of the psychological effects of brainwave entrainment. Altern Ther Health Med. 14: 38-50,

[5] H. Norhazman, N. M. Zaini, M. Taib, R. Jailani, and H. Omar, 2014. The investigation of alpha frontal energy asymmetry on normal and stress subjects after listening to the binaural beats 10 Hz," in Signal Processing & its Applications (CSPA). 2014 IEEE 10th International Colloquium on. 246-250.

[6] M. S. Najib, M. N. Taib, N. A. M. Ali, M. N. M. Arip, and A. M. Jalil, 2011. Classification of Agarwood grades using ANN. in Electrical, Control and Computer Engineering (INECCE), 2011 International Conference on. 367-372.

[7] L. Guo, D. Rivero, J. A. Seoane, and A. Pazos, 2009. Classification of EEG signals using relative wavelet energy and artificial neural networks. in Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation. 177-184.

[8] A. K. Jain, J. Mao, and K. Mohiuddin, 1996. Artificial neural networks: A tutorial. Computer. 31-44.

[9] Z. M. Hanafiah, M. N. Taib, and N. H. A. Hamid, 2010. EEG pattern of smokers for Theta, Alpha and Beta band frequencies. in Research and Development (SCOReD), 2010 IEEE Student Conference on. 320-323.

[10] K. Luangrat, Y. Punsawad, and Y. Wongsawat, 2012. On the development of EEG based emotion classification. in Biomedical Engineering International Conference (BMEiCON). 2012. 1-4.

[11] J. R. Crawford and J. D. Henry, 2003. The Depression Anxiety Stress Scales (DASS): Normative data and latent structure in a large non-clinical sample. British Journal of Clinical Psychology. 42:111-131,

[12] M. M. Antony, P. J. Bieling, B. J. Cox, M. W. Enns, and R. P. Swinson, 1998. Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. Psychological assessment. 10:176.

[13] N. Sulaiman, M. N. Taib, S. Lias, Z. H. Murat, S. A. Aris, and N. H. A. Hamid, 2011. Novel methods for stress features identification using

EEG signals. International Journal of Simulation: Systems. Science and Technology. 12:27-33.

[14] J. A. Coan and J. J. Allen, 2003. Frontal EEG asymmetry and the behavioral activation and inhibition systems. Psychophysiology. 40: 106-114.

[15] G. Carlos, 2011. Artificial neural networks for beginners. Artificial Neural Networks–Methodology Advances and Biomedical Applications. InTech.

[16] K. Abhishek, A. Kumar, R. Ranjan, and S. Kumar, 2012. A rainfall prediction model using artificial neural network. in Control and System Graduate Research Colloquium (ICSGRC). 2012 IEEE. 82-87.

[17] N. Mohamad, F. Zaini, A. Johari, I. Yassin, and A. Zabidi, 2010. Comparison between Levenberg-Marquardt and scaled conjugate gradient training algorithms for breast cancer diagnosis using MLP. in Signal Processing and Its Applications (CSPA). 2010 6th International Colloquium on. 1-7.