An Investigation of Brain Signal Characteristics between Hafiz/Hafizah Subjects and Non-Hafiz/Hafizah Subjects

S F Abdul Halim1*, S A Awang1, S Mohamaddan2

1Biomedical Electronic Engineering Department, Faculty of Electronic Engineering Technology, Universiti Malaysia Perlis, Arau, Malaysia
2Department of Bioscience and Engineering, College of System Engineering and Science, Shibaura Institute of Technology, Fukasaku 307, 3378570 Saitama, Japan

*sitiabdulhalim@studentmail.unimap.edu.my

Abstract. Tahfiz education has gain its popularity among Malaysians thus expand the circle of hafiz and hafizah all over the country. This study has been done to investigate effect of memorizing Al-Quran by determining the difference between hafiz/hafizah subjects and non-hafiz/hafizah subjects in terms of their focus using brain signal characteristics. 10 subjects (5 hafiz/hafizah and 5 non-hafiz/hafizah) have been participated in this study. Database of EEG was recorded by using EegoSport (ANT Neuro, ES-230, The Netherlands) while listening no music, rock music, instrumental music and Al-Quran audio simultaneously with Continuous Performance Task (CPT). The classification has been done by using machine learning method. Decision Tree method have obtained the highest accuracy (96.63%) for PSD Burg by using beta wave. The finding shows that hafiz/hafizah group were more focus in all given tasks compared to non-hafiz/hafizah group. Statistical analysis by using Wilcoxon Signed-Ranks Test found that the designed methodology was significant with 95% confidence interval.

1. Introduction

Human brain contains billions of neurons which are the building blocks of our brain. The network of nerves provides communication between neurons produce brain signals called electroencephalogram (EEG) [1]. Each person has a distinct EEG patterns possibly measure the unique physical characteristic of a person [2]. The brain activity such as thinking or body movement stimulate the electrical activity of nerves causes voltage fluctuation in microvolts (μV). Neuroscientists today collect EEG data non-invasively from human scalp using an electroencephalography due to its simplicity and economical [3]. The collected data could discover hidden information from EEG by applying signal analysis to the retrieved data.

Different kind of activities will produce different EEG signal patterns categorized into several types according to the specific range of frequencies. There are five frequency ranges of EEG signal which are alpha, beta, gamma, theta and delta that determine the human brain states either relax, active, focus, meditation or deep sleep respectively [3],[4]. In a few decades ago, researchers started to emphasis on the study of focus and relaxation [5],[6],[7]. Various analysis methods [8],[9],[10] were performed to identify the changes of these two parameters for different individuals especially in neuroscience field.

This paper demonstrates a significant difference between two groups of people which are hafiz/hafizah and non-hafiz/hafizah based on their brain signal characteristics considering the nature of
EEG which can generate a few similar signal characteristics within the same class. This research has been done by extracting the most informative features of focus/attention followed by comparing the characteristics of EEG biomarkers features shown by both groups. The results were analysed in terms of their focus using alpha, beta and gamma frequency bands.

2. Data Acquisition
Prior to the experiment, an experimental protocol was designed as shown in Figure 1 based on objective of this research.

![Figure 1. Experimental protocol.](image)

All subjects should take 4 tasks; Task A, Task B, Task C and Task D with duration of 3 minutes each. Each tasks run simultaneously with Continuous Performance Test (CPT) [10] using random alphabets presentation. Task A, B, C and D were designed with no music, rock music, instrumental music and Al-Quran audio respectively. The EEG data collected a minute before task, three minutes during task and a minute after task [2]. Database collected based on EEG signal of 10 subjects which are all Muslim, 2 male and 8 female subjects, with age ranging between 20 to 30 years old.

EEG device which is EegoSports device (ANT Neuro, ES-230, The Netherlands) [11] has been used for the data acquisition. It contains 32 channels of electrodes attached to the scalp mapping the international standards EEG geometrical sites, the 10-20 system [12]. The sampling rate set to 512 Hz which records 512 data per second. The scale set to 70μV and frequency range defined from 1 Hz to 100 Hz to obtain all EEG frequency bands for analysis.

The experiment has been done in a silent laboratory environment with suitable lighting. Subjects asked to sit comfortably about 70 cm in front of a laptop. EEG cap was placed on the subject’s scalp. The cap was connected to the amplifier and then to the tablet with EEG064 software. In order to increase the conductivity and provides a stable electrode attachment, electrode gel was injected in each electrodes channel allowing the transmittance of electrical signals. Then, impedance was reduced until all channels show measurement below 5 kΩ [2]. The recording started according to the experimental protocol when subject was ready in a relax condition. Subject requested to minimize body movements thus reduces artifacts in the recorded signal.

3. EEG Signal Processing
EEG signal processing was divided into three stages which are pre-processing, feature extraction and classification as shown in Figure 2.
3.1. **Pre-processing**

Pre-processing is a crucial stage for artifacts and noise removal which interfere the raw EEG signals. This stage also extracts specific frequency of brain waves thus produce database for this study.

Firstly, eye-blinking artifacts (80-200μV) were removed by using Independent Component Analysis (ICA). ICA is a powerful algorithms which can deal with all kinds of artifacts occurred in EEG recordings [13]. ICA separates a multivariate signal into additive subcomponents in case of blind source separation. Sparse representation is a powerful technique for extracting prominent features of a pattern from the ICA component [13]. It employed 20 ICA components for the measurement and detection. The accuracy of this method (automatic detection) is 89.6% which represents a better estimate than that obtained by an extreme machine learning classifier [14].

Next, bandpass filter was applied to the data. Butterworth bandpass filter used to obtain the specific frequency range of alpha, beta and gamma wave [5],[15] which is related to focus and relaxation [16]. Butterworth bandpass filter frequencies were set to 8-13Hz, 13-30Hz and 30-42Hz for alpha, beta and gamma respectively. Fast Fourier Transform (FFT) was used to check whether filtered frequency signal was exactly on specified frequency range or not thus determines the suitable filter order. Power line interference was automatically removed by bandpass filter as it lies on a higher frequency which is 50Hz.

3.2. **Feature Extraction**

Feature extraction stage was divided into two. Linear feature extraction has been done by using Power Spectral Density (PSD) [2] Welch and Burg method. Since EEG is non-linear signal in nature, thus Approximate Entropy (ApEn) and Hurst Exponent (HE) method were also applied. Thus, the performance of linear and non-linear algorithm can be compared.

PSD Welch spectrum distributes power signal with frequency which measures the strength of energy. The analysis provides mathematical equation of frequency analysis for a complex waveform. $\hat{P}_{Welch}$ defines as estimator of spectral density that average the modified periodograms, $\hat{P}_d$ on the windowed sections from possibly overlapped sections, $L$ shown in Eq(1).

$$\hat{P}_{Welch}(f) = \frac{1}{L} \sum_{t=0}^{L-1} \hat{P}_d(f)$$ (1)
PSD Burg Method estimates the power spectral density (PSD) of the input frame. This method fits an autoregressive (AR) model to the signal by minimizing (least squares) the forward and backward prediction errors. Such minimization occurs with the AR parameters constrained to satisfy the Levinson-Durbin recursion. Burg method is expressed by \( \hat{P}_{\text{Burg}} \) where \( f_s \) is the sampling frequency, \( c \) is the order of the model, \( f \) is the frequency, and \( a(k) \) are the AR model parameters.

\[
\hat{P}_{\text{Burg}}(f) = \frac{1}{f_s^c \left[ 1 + \sum_{k=1}^{c} a(k)e^{2\pi jkf / f_s} \right]}
\]

Approximate Entropy (ApEn) is a technique used to quantify the amount of regularity and the unpredictability of fluctuations over time-series data. It is expressed by \( \text{ApEn} \) where \( m \) is an integer represents the length of compared run of data and \( r \) is a positive real number specifies a filtering level.

\[
\text{ApEn} = \phi^m(r) - \phi^{m+1}(r)
\]

Hurst Exponent is referred to as the index of dependence or index of long-range dependence. It quantifies the relative tendency of a time series either to regress strongly to the mean or to cluster in a direction. The generalized Hurst, \( H \) defines as \( \log \frac{R}{S} \) over period; where \( R \) is the difference between maximum and minimum from the mean, \( S \) is the standard deviation and \( T \) is the period of the sample data.

\[
H = \frac{\log \frac{R}{S}}{\log T}
\]

3.3. Classification
The classification stage tests the database using supervised methods which are Decision Tree, k-Nearest Neighbors (kNN), and Random Forest (RF).

k-Nearest Neighbors (k-NN) is used for classification or regression according to k-closest training examples in the feature space [4]. Eq(5) defines the formula for k-NN where \( d \) is Euclidean distance, \( N \) is the number of features, \( x \) is coordinate value in x-axis and \( y \) is coordinate value in y-axis.

\[
d(x, y) = \sqrt{\sum_{t=0}^{N} (x_i - y_i)}
\]

Decision Tree (DT) is a supervised classification technique based on the division of a complex problem into several subproblems, and repeating this process recursively through the generation of a tree. Decision tree defines as \( E(S) \) where \( S \) is current state and \( p_i \) is probability of an event of state \( S \) or percentage of class \( i \) in a node of state \( S \).

\[
E(S) = \sum_{i=1}^{c} -p_i \log_2 p_i
\]

Random Forest (RF) classifier uses an ensemble learning method for classification consists of large number of individual decision trees that operate as an ensemble. Eq(6) was used to determine branch of each nodes where it average the decisions from multiple trees to find an answer.

4. Results and Discussions
Most informative features of EEG signals during performing task and the characteristic of EEG features between non-hafiz/hafizah and hafiz/hafizah have been analysed and discussed in this section.

4.1. Pre-processing
In pre-processing stage, ICA method was applied to the raw signal to remove eyes blinking artifacts remains clean signals as shown in Figure 3. However, after noise removal the signal generated has altered as ICA reconstructed the signal by approximation to the other channels.
Figure 3. EEG plot shows EEG signal before and after filter using ICA. Blue = channel before rejection, red = after rejection.

Figure 4. Filtered frequency domain for alpha, beta and gamma.

Figure 5. Filtered time domain alpha, beta and gamma.

Figure 4 shows alpha beta and gamma waveforms during listening to Al-Quran plot by using the filter order 6, 8 and 9 respectively. Gamma signal obtained the highest amplitude of more than 3.5\(\mu\)V followed by beta 3.3\(\mu\)V and alpha which is 2\(\mu\)V. The complexity of the signals is clearly shown in Figure 5 where there is no specific patterns, thus it is very difficult for analysis to be done.

Figure 5 shows alpha, beta and gamma waveforms in time domain graph obtained after bandpass filter was applied. Alpha is the least frequent signal followed by beta which is more frequent and the highest frequent signal is gamma. As referred to the frequency domain graph, alpha lies at the lowest frequency range thus in time domain alpha appear only a few times, the wavelength is longer hence it takes longer time to complete one complete cycle. The cycle determines frequency \((f=1/t)\) of different types of waves; alpha (8-13Hz), beta (13-30Hz) and gamma (30-42Hz).

4.2. Feature Extraction

Feature extraction take out only the most important features thus reduced the signal’s dimension. Figure 6 shows segmentation or framing of signal for each second. Feature extraction algorithms were implemented to the framed part of the signal and moving the frame through the whole signal. This approach allows the spectral characteristics to be processed constantly for each frame. Thus, it produced only selected features from each frame to reduce the size of the original data for efficient analysis.

Figure 6. Segmentation of data without overlapping frame/window.
Feature extraction method includes linear and non-linear method due to the nature of EEG signal which is non-linear signal. Figure 7 shows the example of power spectral density plot by using Welch method for alpha frequency range of F3 during task. The bandwidth is half of sampling frequency thus the plot ends at 256Hz. Features such as mode, median and mean extracted from this plot used as inputs to the classifiers [14].

Figure 7. Welch Power Spectral Density of alpha waveform.

4.3. Classification

4.3.1. The most informative features of attention for EEG signal characteristics among hafiz/hafizah. The most informative feature was determined by using classification method. Several types of classifiers were used to compare the accuracy obtained from different feature extraction methods applied. The classification results found that beta wave obtained the highest accuracy. Thus, further analysis concentrates on beta wave only. Figure 8 and Table 1 show the general structure of classifier and classification results for beta wave with different classifiers and feature extraction method respectively.

Table 1. Classification result of beta wave.

| Feature Extraction \ Classifier       | Decision Tree (%) | KNN (%) | Random Forest (%) |
|--------------------------------------|-------------------|---------|-------------------|
| PSD Welch                            | 96.43             | 78.97   | 69.21             |
| PSD Burg                             | 96.63             | 60.90   | 74.58             |
| Approximate Entropy                  | 94.28             | 55.82   | 58.24             |
| Hurst Exponent                       | 93.44             | 36.22   | 39.44             |
The highest accuracy was obtained by DT (96.63%) implementing PSD Burg method. This shows that the combination of PSD Burg and DT are able to extract more informative features than the other methods. It can be concluded that this is most suitable to be used for this case study.

4.3.2. Characteristics of EEG biomarkers features between hafiz/hafizah subject and non-hafiz/hafizah subject. The characteristics of EEG biomarkers feature was obtained from PSD Welch method. The data was taken from Channel F3, by analysing the maximum value of alpha, beta and gamma for each task. Result was tabulated based on group of hafiz/hafizah and non-hafiz/hafizah as shown in Table 2.

| Task      | Wave | Hafiz/hafizah | Non-hafiz/hafizah |
|-----------|------|---------------|-------------------|
| No Music  | Alpha| 18.33 ±4.32   | 17.91 ±1.88       |
|           | Beta | 24.18 ±3.39   | 21.33 ±2.56       |
|           | Gamma| 17.71 ±3.41   | 16.12 ±3.06       |
| Rock      | Alpha| 19.48 ±2.94   | 18.87 ±1.59       |
|           | Beta | 24.57 ±4.30   | 21.11 ±2.93       |
|           | Gamma| 16.91 ±4.03   | 19.11 ±5.30       |
| Instrumental | Alpha| 19.43 ±4.52   | 19.27 ±2.19       |
|           | Beta | 25.30 ±4.21   | 24.53 ±3.02       |
|           | Gamma| 17.44 ±5.85   | 19.04 ±4.57       |
| Quran     | Alpha| 19.83 ±5.42   | 18.63 ±3.19       |
|           | Beta | 25.40 ±3.21   | 23.79 ±3.32       |
|           | Gamma| 18.17 ±5.68   | 18.00 ±2.02       |

Table 2 shows beta waves is dominant for both groups and all given tasks. This is due to the subjects were required to concentrate and put an attention during performing the tasks. Gamma power shows a slightly lower than alpha which concludes that the CPT task designed in this research not able to gain the highest focus for both groups.

The characteristics of EEG for both groups were determined by using maximum power of beta wave. Hafiz/hafizah group has a higher beta power in all given tasks. This shows that hafiz/hafizah were able to achieve higher focus compared to non-hafiz/hafizah subjects. Both groups had obtained the maximum power in two different tasks. Maximum power of beta wave for hafiz/hafizah group was achieved during listening to Al-Quran (25.40 ±3.21). However, non-hafiz/hafizah group shows the highest beta during listening to instrumental music (24.53 ±3.02). Thus, both Al-Quran and Instrumental music was successfully proved could be used to obtain focus with only a slight different. Results from Wilcoxon Signed-Ranks Test of Channel F3 from all subjects with α=0.05 (95% confidence level) shows that all tasks have successfully accomplished significant results of p-value less than 0.05 and z-value exceeded 1.96.

5. Conclusion
In conclusion, for the focus/attention study, beta is the most useful frequency range where it provides a significant difference which able to differentiate between these two groups. PSD Burg was able to extract more informative features compared to the other methods. The highest accuracy (96.65%) was recorded from PSD Burg features fed to Decision Tree. It also found that hafiz/hafizah has achieved higher focus in all tasks compared to non-hafiz/hafizah subjects based on the EEG maximum power. This research should be repeated by using a larger sample about 30 subjects per group and a better design of CPT task to obtain more accurate result.
References

[1] Buckner R L and DiNicola L M 2019 The brain’s default network: updated anatomy, physiology and evolving insights, *Nature Reviews Neuroscience* 20 10 593-608.

[2] Ong Z Y, Saidatul A, and Ibrahim Z 2018 Power Spectral Density Analysis for Human EEG-based Biometric Identification, *Int. Conf. Comput. Approach Smart Syst. Des. Appl. ICASSDA 2018* 1-6.

[3] Osalusi B, Abraham A and Aborisdade D 2018 EEG Classification in Brain Computer Interface (BCI): A Pragmatic Appraisal, *American Journal of Biomedical Engineering* 81 1-11.

[4] N. Jalauadin and M. K. Mohammed Amin 2019 EEG analysis on human reflection towards relaxation of mind, *Malaysian J. Fundam. Appl. Sci.* 15 2 185-189.

[5] Ayuni S, Nasir M, Mahani W and Mahmud H W 2015 Brain Signal Analysis Using Different Types of Music, *Int. J. Integr. Eng.* 7 3 31-36.

[6] Mohamed Z, Halaby M E, Said T, Shawky D and Badawi A 2018 Characterizing focused attention and working memory using EEG, *Sensors*, 18 11 1-21.

[7] Boytsova J A, Danko S and Solovjeva M L 2018 EEG-Correlates of Mental Attention are Modified by Presence of Specifying Visual Signals, *Hum Physiol.* 44 5 600-604.

[8] Irmmischer M, Houtman S J, Mansvelder H D, Tremmel M, Ott U and Linkenkaer-Hansen K 2018 Controlling the Temporal Structure of Brain Oscillations by Focused Attention Meditation, *Hum. Brain Mapp.* 39 4 1825-1838.

[9] Liu N H, Chiang C Y and Chu H C 2013 Recognizing the degree of human attention using EEG signals from mobile sensors, *Sensors*, 13 8 10273-10286.

[10] Li J, Kronemer S I, Herman W X, Kwon H, Ryu J H, Micek C, Wu Y, Gerrard J, Spencer D D and Blumenfeld H 2019 Default mode and visual network activity in an attention task: Direct measurement with intracranial EEG, *Neuroimage,* 201 116003 1-9.

[11] eego™/mylab, ANT Neuro. [Internet] Retrieved from https://www.ant-euro.com/products/eego-ylab.

[12] Acharya U R, Sree S V, Ang P C A, Yanti R and Suri J S 2012 Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals, *Int. J. Neural Syst.* 22 2 1-9.

[13] Jiang X, Bin Bian G and Tian Z 2019 Removal of artifacts from EEG signals: A review, *Sensors*, 19 5 1-18.

[14] Zafar R, Qayyum A and Mumtaz W 2019 Automatic eye blink artifact removal for EEG based on a sparse coding technique for assessing major mental disorders, *J. Integr. Neurosci.* 18 3 217-229.

[15] Al-Galal S A Y, Alshaikhli I F T and Rahaman A W B A 2016 Automatic Emotion Recognition Based on EEG and ECG Signals While Listening to Quranic Recitation Compared with Listening to Music, 2016 6th *Int. Conf. Inf. Commun. Technol. Muslim World* 269-274.

[16] Al-Galal S A and Fakhri Taha Alshaikhli I 2017 Analyzing Brainwaves While Listening To Quranic Recitation Compared With Listening To Music Based on EEG Signals, *Int. J. Perceptive Cogn. Comput.*, 3 1 1-5.