Evolution in Electrical and Electronic Engineering Vol. 4 No. 1 (2023) 415-422

EEEE

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/eeee
e-ISSN : 2756-8458

Removal of ECG Artifacts from EEG Signals

Khaliesah Johari[1], Ashok Vajravelu[1]*

[1]Faculty of Electrical and Electronic Engineering,
Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, Johor, MALAYSIA

*Corresponding Author Designation

DOI: https://doi.org/10.30880/eeee.2023.04.01.049
Received 16 January 2023; Accepted 28 March 2023; Available online 30 April 2023

Abstract: Any unusual feature in brain functioning, structure or biochemical levels are referred as brain abnormalities. The brain abnormalities, deformities or dysfunction will affect the whole body. Electroencephalogram (EEG) tests are taken in order to diagnose many diseases caused by brain abnormalities such as sleep disorders, head injuries, Alzheimer's disease, Epilepsy, brain hemorrhage and etc. However, an EEG recording could have many types of noises or artifacts that came from the blinking of the eye, heartbeat, muscle movement and many other types of noises which will contaminate the EEG recording, decreasing the accuracy of the EEG recording. In this paper, machine learning algorithm was proposed and used to remove electrocardiogram (ECG) artifacts from the EEG signal. ECG artifacts are noises that came from the beating of the heart. The Independent Component Analysis (ICA) algorithm was the machine learning algorithm that was used for the artifact removal. It was implemented by using Python code and was executed in Google Colaboratory. A completely clean EEG signal free from any artifacts is impossible to obtain but by implementing this machine learning algorithm to remove ECG artifacts from EEG signal, it will produce a better EEG signal.

Keywords: Electroencephalogram (EEG), Electrocardiogram (ECG), Independent Component Analysis (ICA)

1. Introduction

An electroencephalogram (EEG) is a test that detects abnormalities in brain waves or in the electrical activity of the brain using small metal discs (electrodes) attached to the scalp. The brain communicates through electrical impulses and are active even when a person is sleeping. This activity shows up as wavy line on an EEG recording.

Epilepsy is a common condition that affects the brain and causes frequent seizures [1]. These seizures are bursts of electrical activity in the brain that temporarily affect how it works. According to World Health Organization (WHO) in ATLAS Country Resources for Neurological Disorders second edition 2017 [2], epilepsy is the most common chronic brain diseases that affects people of all ages. Approximately more than 50 million people worldwide are diagnosed with this condition with 80% of them are from low- and middle-income countries. Epilepsy was diagnosed with various methods including observing brain activity abnormalities through EEG recording [3].

*Corresponding author: ashok@uthm.edu.com

The importance of improving the EEG signals has been proved and the presence of artifacts will contaminate the EEG signals. Removing artifacts is an important process of improving the overall accuracy of EEG signals. Some of the artifacts affecting the EEG signals originated from electrooculogram (EOG), electrocardiogram (ECG) and electromyogram (EMG). EMG is essentially electrical "noise" generated by facial muscle activity near the electrode. EOG is electrical noise generated by eye movement. ECG is electrical noise generated by the heart rhythm.

This project is proposed for the analysis of the brain activities of human with EEG. Since there are different types of artifacts, this project will specifically about removing ECG artifacts from the EEG signals. Removal of ECG artifacts will give better EEG signals for diagnosis.

2. Materials and Methods

This project simulated two sets of coding into Google Collaboratory website. The first set of the coding has used ICA to blindly separates the event related potentials (ERP) data of EEG into temporally and spatially independent components (ICs) in 6 different conditions which are control condition, moving head, moving jaw, blinking, swallowing and clenching teeth.

The second set of the coding has used PCA to decompose correlated variables into uncorrelated variables. These uncorrelated variables were then decomposed into temporally and spatially ICs and next, ICs with ECG artifacts were removed. The second coding also illustrated the separation of ICA components spatially based on its power spectral density (PSD) and topology. These ICA components were marked into three clusters. This EEG data was acquired from a 60-channel electrode cap and was acquired simultaneously with the MEG. This datasets description is for the MNE datasets used in this project.

2.1 EEG preprocessing pipeline

Implementing artifacts removal from EEG signals usually consist of the combination of stages in Figure 1. Artifact removal algorithm would start with filtering the EEG signals to filter out extrinsic artifacts that are distinguishable from the signals. Re-referencing is the process of expressing the voltage at the EEG scalp channels in relation to a different, new reference.

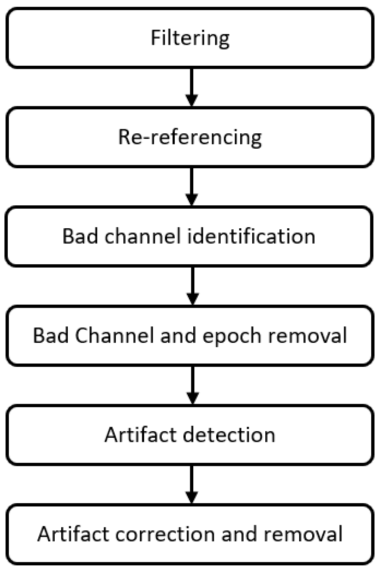


Figure 1: EEG preprocessing pipeline

The preprocessing EEG algorithm would then identify bad channels and remove it during the next stage. Artifacts are then detected by the algorithm and the artifacts are corrected and removed as the last stage in this EEG preprocessing pipeline. If the EEG preprocessing uses ICA as its algorithm, then the ICA algorithm would detect artifacts and remove it.

2.2 Simulation flowcharts

The first set of the coding started by accessing the EEG datasets stored in Google Drive. The accessed data were then loaded into Google Collaboratory. These datasets were down-sampled to 256 Hz (sampling rate) and high pass filter was applied to effectively remove extrinsic artifacts. The events related EEG data were extracted and average reference was set as reference (re-referencing).

Next, EEG channels were picked for the extracted ERP/ERF data and channel locations were set. The epochs of the ICs that have been separated temporally and spatially from the ERP/ERF data of EEG were created in the 6 different conditions. The ICs illustrated in the output include the ICs containing artifacts since this first coding only decomposed the EEG data into ICs and demonstrated it in 6 different conditions without removing ICs containing artifacts. Figure 2 shows the two simulation process.

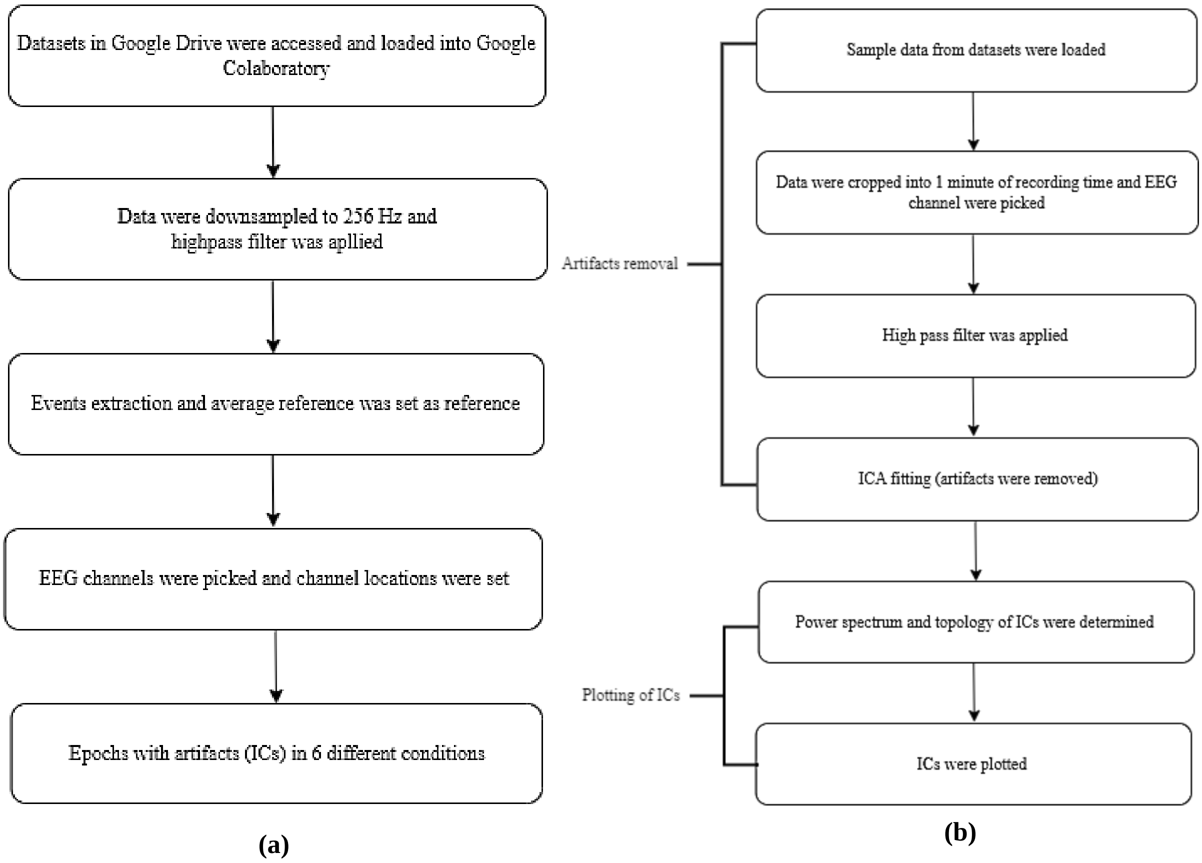


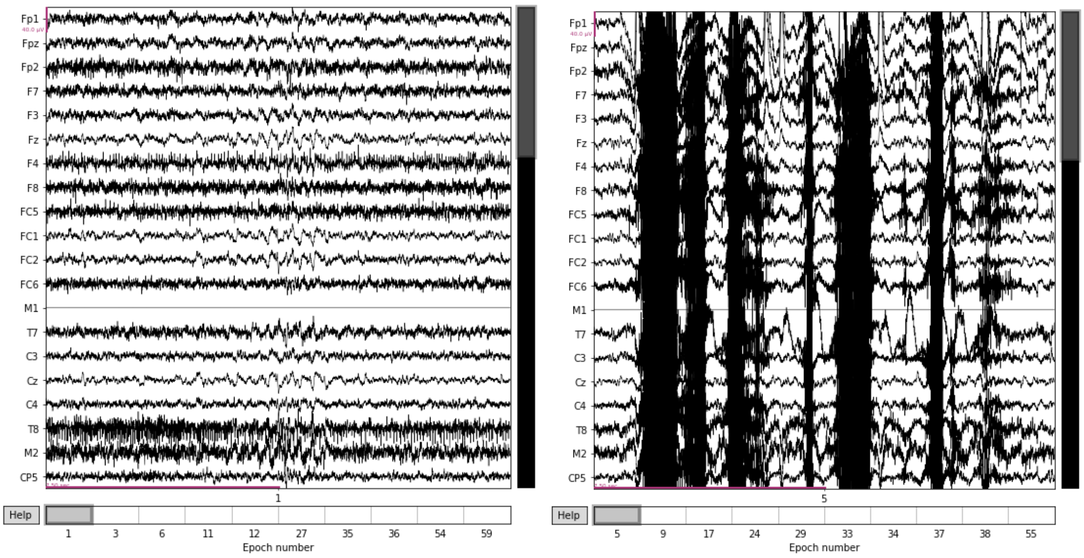
Figure 2: Simulation process (a) First coding; (b) Second coding

The second set of coding was divided into two stages which were artifacts removal and the ICs plotting. The artifacts removal stages started with loading the sample data from datasets. The data were cropped into 1 minute of recording time and EEG channels were picked. Next, high pass filter with the frequency of 2 Hz was applied. This high pass filter in ICA was applied to filter out the data prior to ICA fitting because ICA algorithm is sensitive to low frequency drifts.

After that, ICA was fitted into the processed data. The ICA removed artifacts by decomposing the principal components (PCs) that have been converted by PCA, into independent components (ICs) and removing ICs containing artifacts. The coding entered the second stage which was the plotting of ICs based on its power spectral density (PSD), topology and clusters. The ICs were plotted by determining the power spectrum and topology of each ICs before plotting. The number of clusters were stated in the coding to get the plotting of ICs marked by its clusters.

3. Results and Discussion

In the first simulation, the raw EEG datasets accessed for this project were illustrated after being filtered by high pass filter with 2 Hz. From Figure 3, the raw EEG signals from the EEG channels can be observed and analyzed. ECG artifacts are difficult to distinguish and are inconsistent. It appeared as false sharp spikes in an EEG recording and can be confused as neurological rhythms. The ECG artifacts from the EEG signals can be seen at almost every channel at different epoch numbers. Channel TB and M2 were clean EEG signals at epoch number 1 to 6. There were several other channels at different epoch numbers with clean and contaminated EEG signals. All channels have the sharp spikes of ECG artifacts at almost the same epoch numbers which were 9, approximately 16 and 20, 33, 37 and 38.



(a) **(b)**

Figure 3: (a) and (b) Raw EEG signals from datasets

The EEG signals were decomposed by ICA algorithm into independent components. There were six conditions of the subjects while EEG signals were recorded which were control condition, moving your head, moving your jaw, blink, swallow and clench your teeth. The first five decomposed ICs for the 6 different noise conditions were illustrated in the output. Figure 4 (a)-(f) are the first decomposed ICA components of each noise conditions.

Control condition is the restricted environment condition where subjects needed to stay still while EEG recording was recorded. The event related potentials or fields (ERP/ERF) are the averages of EEG trials time locked to a set of experimental events. The brain image illustrated with the graph of ERP/ERF displayed the brain activities' locations and the independent components distributions while a certain condition was taking place. For instance, during moving your head condition the brain activities were evenly distributed all over the head except for the eye location. In moving your jaw and blink condition the brain activities were focused at the front of the head except that in blink condition the ERP/ERF were more concentrated given the interval between each blink. ICA decomposed the datasets into a number of independent components each either containing artifacts or not to observe the decomposition and the distribution of ICs in different conditions.

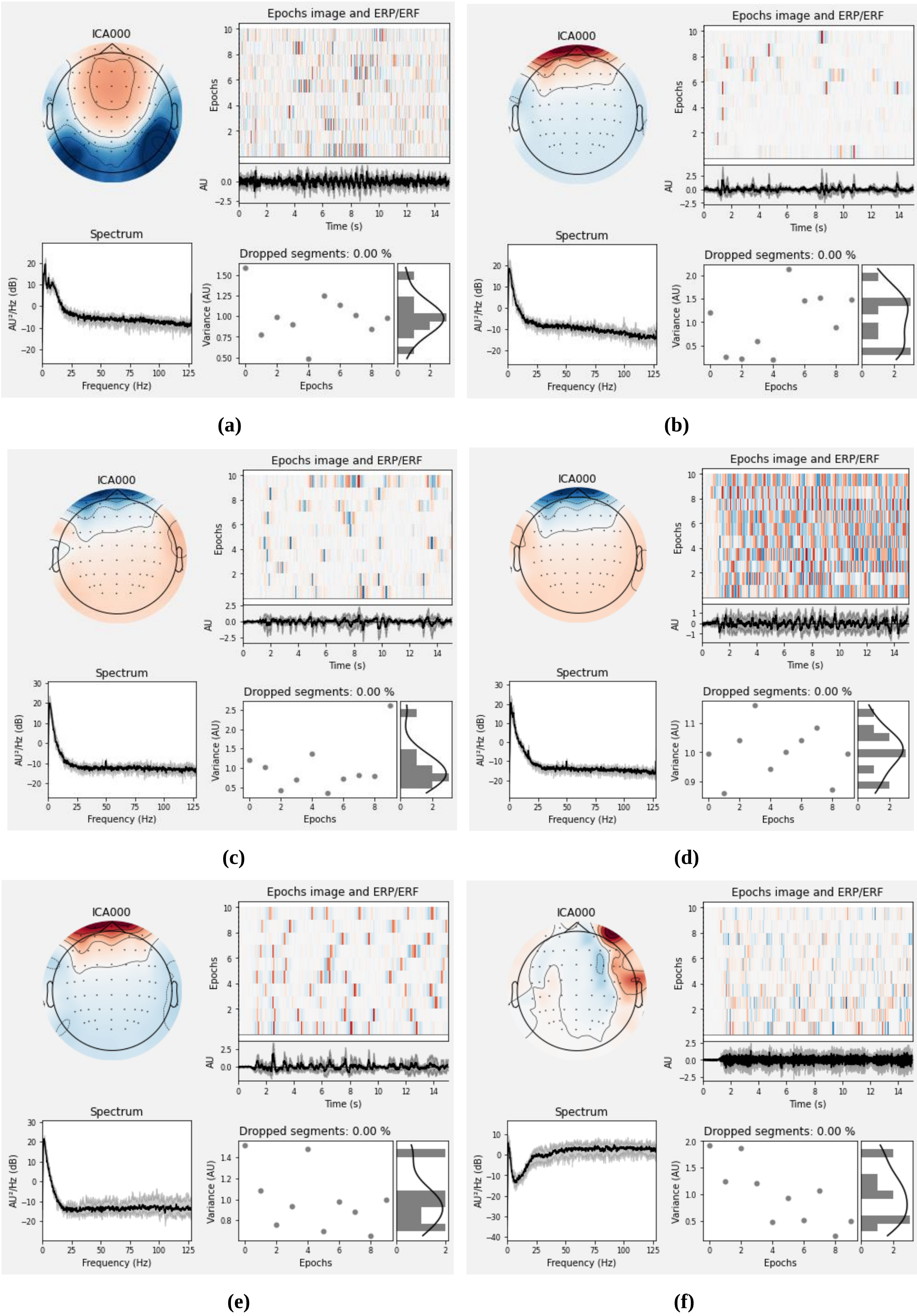
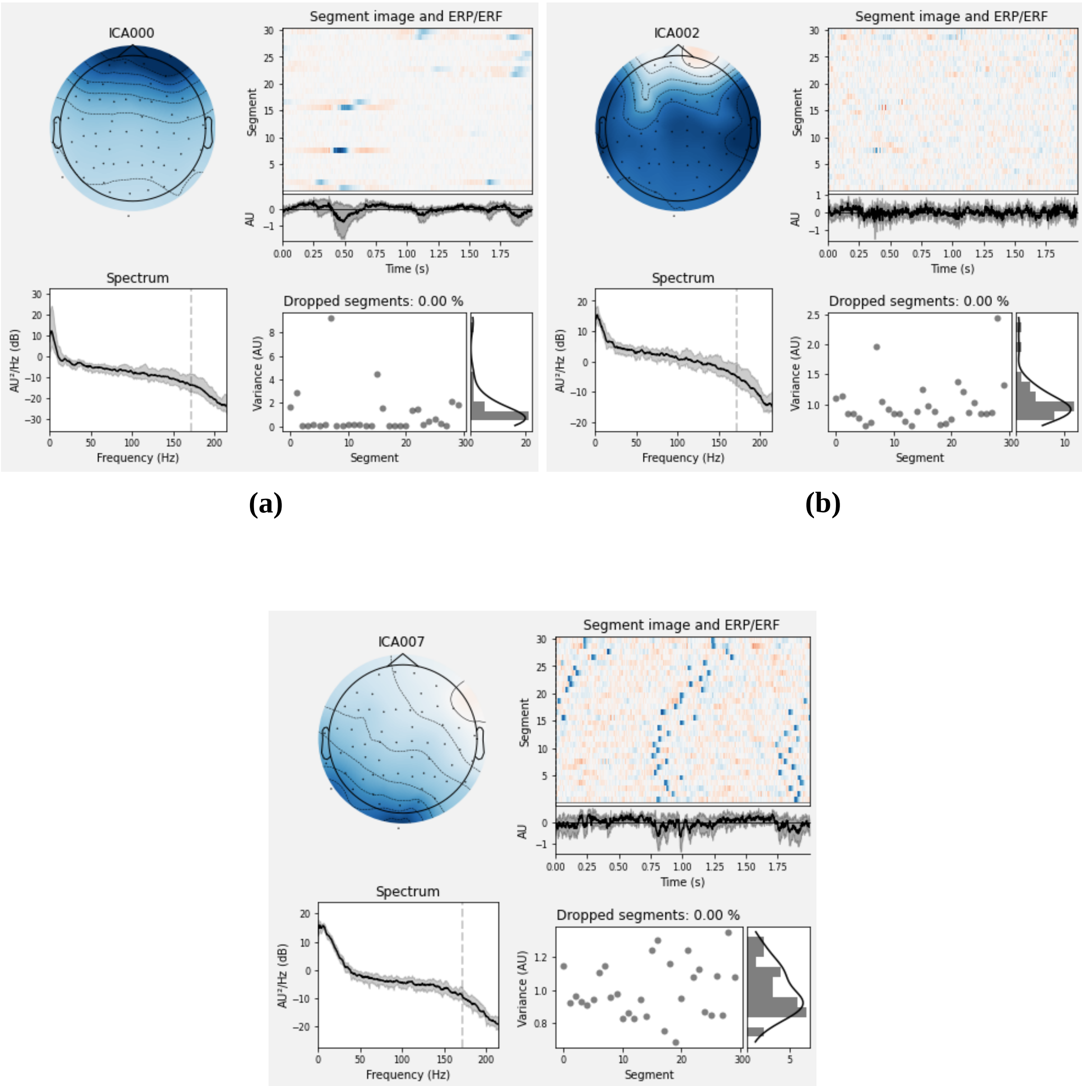


Figure 4: ICA decomposition of ICA000 in (a) control condition, (b) moving your head condition, (c) moving your jaw condition, (d) blink condition, (e) swallow condition, (f) clench your teeth condition

In the second simulation the raw EEG data from MNE datasets were loaded and processed for ECG artifact removal with ICA. The datasets were decomposed by PCA before ICA was implemented for artifact removal. The first part of ICA decomposed the principal components (uncorrelated variables) into a number of independent components. Independent components containing ECG artifacts were removed and a clean EEG signal's ICs were reconstructed. The supposedly clean of ECG artifacts EEG signals ICs were randomly chosen and illustrated in the simulation output. Figure 5 (a), (b) and (c) are three of the remaining reconstructed independent components of EEG signals.

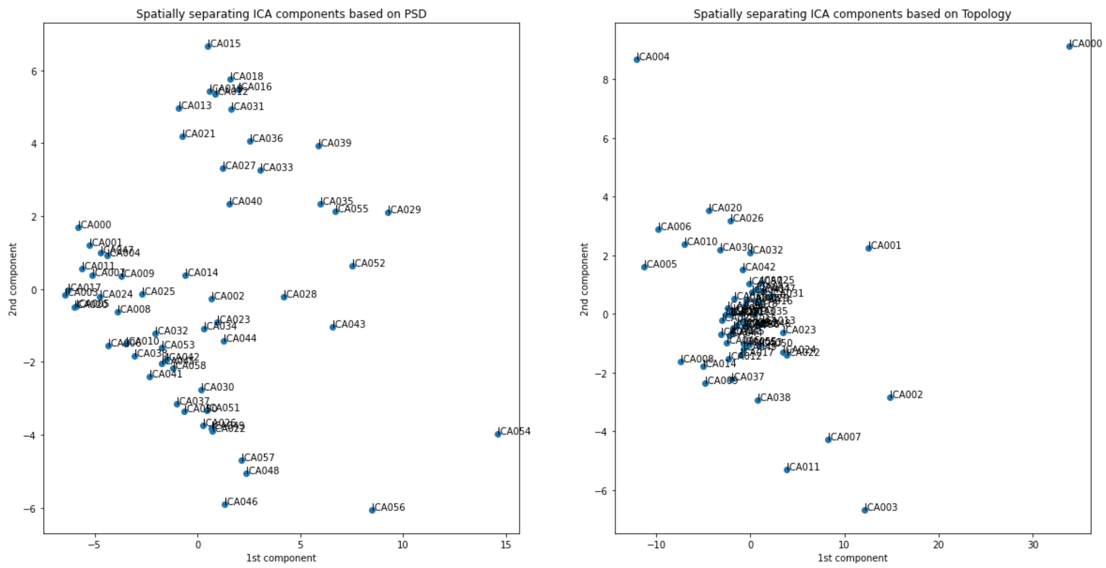


(a) **(b)**

(c)

Figure 5: ICA decomposition of components (a) ICA000; (b) ICA002; (c) ICA007

It can be observed that these three demonstrated ICs have regular distribution of ERP/ERF compared to the ICs from the first simulation. Hence, it can be said that the artifacts free ICs combined into a new EEG signal are cleaner than before ICA implementation. Figure 6 (a) and (b) show the decomposition of ICs plotted spatially based on its power spectral density (PSD) and topology. Power spectral density (PSD) of the signal describes the power per unit frequency of the signal in W/Hz or dBm/Hz unit. The relative magnitudes of the frequency components that make up a signal are shown in its power spectrum. The topology of the ICs plotted gave a topographic map where the distance between the ICs on the plotted graph, function as the dependencies of the components with each other. In terms of higher-order correlations or mutual information, topographically close components are relatively strongly dependent on one another.



(a) (b)

Figure 6: Spatially separated ICA components based on (a) PSD; (b) Topology

Figure 7 shows the ICA components that were marked into three clusters.

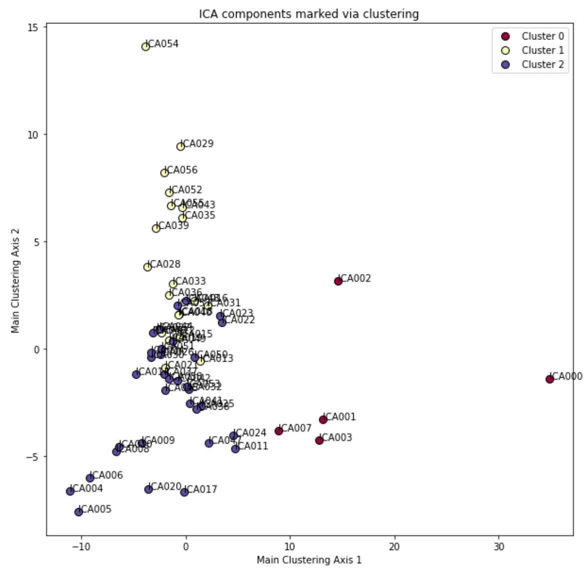


Figure 7: ICA components marked via clusters

The ICs were also marked into three clusters, cluster 0, 1 and 2. The three clusters of the ICs were the cluster of ICs that were close to each other. For instance, cluster 0 was the cluster of ICs that have the closest distance to each other. This describes each cluster to share mutual information between the components in the said cluster.

4. Conclusion

By completing this project, a simulated system to remove ECG artifacts was able to be created with the implementation of PCA and ICA algorithms. The system was created and implemented with the EEG datasets acquired from the database. The decomposition of ICA components of the EEG data in six different conditions were able to be observed closely. ECG artifacts from the datasets were able to be removed and the remaining clean EEG data components decompositions were able to be observed and compared with the components containing artifacts. It can also be concluded that blind source separation (BSS) artifacts removal technique efficiently removed artifacts. Furthermore, different artifacts removal technique can be implemented into this project which might give better results compared to the implementation of artifacts removal technique in this project. The existing artifacts removal technique can also be improved and developed further to overcome its limitations.

Acknowledgement

The authors would like to thank the Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia for its support.

References

[1] Issa, M. F., Tuboly, G., Kozmann, G., & Juhasz, Z. (2019). Automatic ECG Artefact Removal from EEG Signals. Measurement Science Review, 19(3), 101–108. https://doi.org/10.2478/msr-2019-0016.

[2] Organisation Mondiale De La Santé. Département De La Santé Mentale Et Des Toxicomanies, & Fédération Mondiale De Neurologie. (n.d.). Atlas : country resources for neurological disorders.

[3] Epilepsy - Diagnosis and treatment - Mayo Clinic. (n.d.). Www.mayoclinic.org. Retrieved April 12, 2022, from https://www.mayoclinic.org/diseases-conditions/epilepsy/diagnosis-treatment/drc-20350098#:~:text=Electroencephalogram%20(EEG)