UNIVERSITY OF PADOVA

DEPARTMENT OF INFORMATION ENGINEERING

Master Thesis in

ICT for Internet and Multimedia - Life and Health

"Advanced Pipelines For Artifact Removal From EEG Data"

Supervisor:
Prof. Giulia Cisotto

Master candidate:
Milad Nasiri Nejaddafchahi

Co-supervisor**:**
Prof. Viviana Betti
Dr. Ottavia Maddaluno

April,11th 2022

Academic Year:2021/2022

Contents:

List Of Images

List Of Tables

Abstract

Feature extraction and working with EEG data has become one the most challenging studies these years. The raw EEG signal has various artifacts and needs to be detected and separated from brain components.This study is part of ERC. For removing artifacts from EEG data , this procedure done by a method known as “semi-automatic ICs selection pipeline”.This method was developed and verified by Cosynclab directed by Prof. Betti in Rome (where I spent my internship). In particular, the thesis work aims to investigate another method for complementing semi-automatic ICs selection pipeline and evaluate results which conveys to increasing the accuracy of semi-automatic ICs selection pipeline.The ICA algorithm derives independent sources from highly correlated EEG signals statistically without concern for the actual location or configuration of the EEG signal source . It is used to locate concurrent signal sources that are either too close together or too broadly scattered to be separated using conventional localization techniques. The primary issue in understanding ICA output is determining the right dimension of the input channels and the physiological and/or psychophysiological relevance of the resulting ICA source channels.With semi-automatic ICs selection pipeline method more than 2600 ICs evaluated and 405 ICs labeled as brains and the rest classified as artifacts. To evaluate these 405 ICs and increase possible accuracy another method was used known as ICLabel. ICLabel projects had been proposed by EEGLAB. This is a method based on Deep Learning and provides classification based on EEG IC classifier[1]. After running and comparing the two methods pipeline, then,we designed an application for comparison and visualization output for both methods which name is IC selection.With this application we realize some modification needed for future steps for labeling with semi-automatic ICs selection pipeline method and some artifacts could change from artifacts to brain.

Acronyms

- IS:Information systems
- EEG: Electroencephalogram
- fMRI: functional magnetic resonance imaging
- EPSP: Excitatory postsynaptic potentials
- CNS: central nervous system
- AP: Action potential
- IPSP: Inhibitory postsynaptic potential
- EMG: Electromyogram
- Na+: Sodium
- K+: Potassium
- Ca: Calcium
- ERS: Event-related synchronization
- ECG: Electrocardiogram
- MRI: Magnetic resonance imaging
- PET: Positron emission tomography
- SPET: Single-photon emission tomography
- ADCs: Analog-to-digital converters
- EOG: Electrooculography
- ADCs:Analog-to-Digital Converters
- ICA: Independent Component Analysis
- ACF: The autocorrelation function
- PSD: Power Spectral Densities
- DFT: Discrete Fourier Transform
- FFT: Fast Fourier Transform
- BSS :Blind Source Separation
- SOBI: Second-order blind identification
- CBSS :Convolutive Blind Source Separation

Introduction

The neural activity of the brain can be recorded from the 23rd week of prenatal development and provide the status of the whole body[2]. This valuable information provides motivation to implement advanced technology to investigate more and more about the brain and its signals. In this regard, depth understanding of brain activities, neuronal functions, and neurophysiological properties of the brain are essential for every researcher working with this data.

An electroencephalogram (EEG) signal is a measurement of currents that flow during dendritic excitations of many pyramidal neurons in the cerebral cortex. This current creates a magnetic field that may be measured using electromyogram (EMG) machines, as well as a secondary electrical field that can be measured using electroencephalogram (EEG) systems. When EEG data is collected through recording equipment, signal artifacts are more prevalent[3] These artifacts have the potential to compromise the EEG data's quality. A thorough understanding of the many artifacts is required in this case to eliminate the artifacts or noise effectively. Abnormalities are undesired signals often caused by noise in the environment, experimental mistakes, or physiological artifacts.Additionally, extrinsic artifacts are caused by external causes such as the environment or experiment mistakes. Intrinsic artifacts are those caused by the body itself [4–6]. EEG signals are frequently affected by these three artifacts (eye blink, muscle activities, heartbeat).semi-automatic ICs selection pipeline is a binary classification method to classify each ICs. The output of this method is pure brain components or pure artifacts components.However, each ICs has a combination of information such as brain and artifacts; extracting brain components belonging to each ICs is the main motivation for implementing another method published in 2019 known as ICLabel.

This study is a step in this direction to preprocess EEG data and compare available methods for preprocessing. Thesis aims to evaluate an algorithm designed by University La Sapienza with another method known as the ICLabel project. Here, an ICLabel aims to be defined to classify ICs[1]. The method is based on machine learning . The framework would be trained with simulation and different EEG datasets recorded from several subjects[6]. The HCP dataset has been used as a worldwide known dataset and in this thesis only 31 subjects were evaluated and classified by the semi-automatic algorithm and the output of the algorithm validated and corrected by members of COSYNCLAB. These 31 subjects were first labeled by semi-automatic ICs selection pipeline and finally controlled by Professor Betti. Then this dataset runned by ICLabel and the result of these two methods compared together. Comparing results represent good agreements between ICLabel and semi-automatic ICs selection pipeline. Significantly developed initially for EMG data and adjusted by Professor Betti and CosyncLab for EEG.furthermore, additional information can be integrated from ICLabel to help the dataset. Finally, an application, namely "IC selection "was developed to compare the two methods' output simultaneously. This application help CosyncLab members investigate more brain ICs and more accurate labeling. modifying criteria implemented in semi-automatic ICs selection pipeline would increase this method accuracy as future study.

Chapter 1

1 Background

1.1 Neuro-anatomy of the brain

In this chapter I introduce some basic information of brain anatomy and brain activities.Then EEG and artifacts described with the method to detect artifacts .The human brain is the most complex object known in the world[7]. The human brain receives and processes information from the outside world and depicts the nature of the mind and psyche. Intelligence, creativity, emotions, and memory are some of the things created by the human brain. The brain receives information through the five senses: sight, smell, hearing, touch, and taste. The brain collects these messages in an understandable way and stores this information in specific and often mysterious areas.

Different parts of the brain are involved in the formation and planning of movements. For example, in the occipital lobe, there is a visual recognition system that after image recognition in this part, information is transferred to the cerebral cortex then relevant motion commands are sent to the muscles. The basal ganglia and cerebellum play a vital role in this transformation. The cerebellum is the center that receives sensory information quickly and sends information to other centers. Another function of the cerebellum is to create softness and coordination in movement. One of the essential functions of the cerebellum is to prepare internal models of the motor task. In the meantime, the task of the Basal ganglia is controlling movements and, to some extent, learning movement skills. In this section, various brain centers related to movement were examined. In the following, the four main parts of the brain and the overall function in the production of movements are described.

The human brain is made of four main lobes[8]: frontal lobe, parietal lobe, occipital lobe, and temporal lobe, which can be seen in Figure[1] shows.

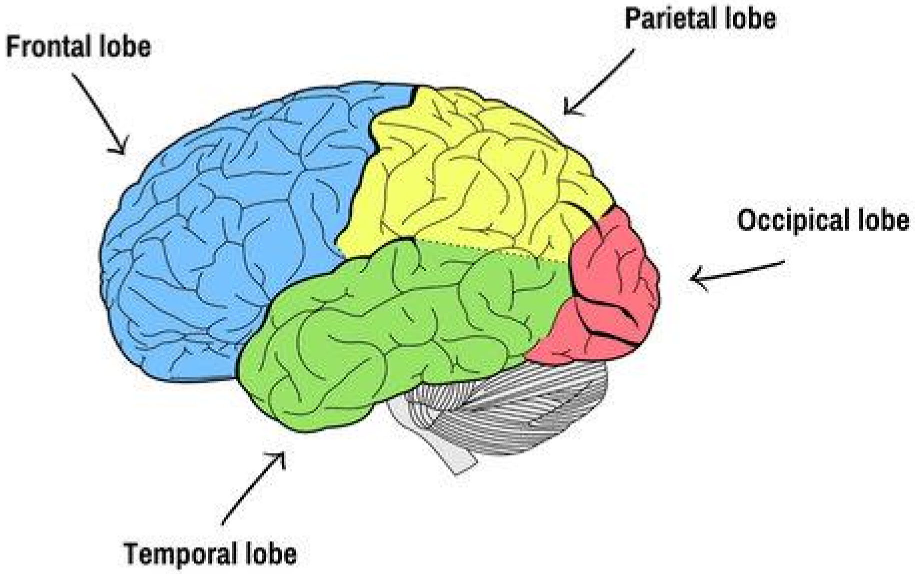


Figure 1. The brain is divided into four major lobes (Frontal lobe, Parietal lobe, Occipital lobe, Temporal lobe) [9]

The frontal lobe is located in the front of the hemispheres of the brain. This section is responsible for analytical and complex activities and plays an important role in the development of memory, feeling, and learning. The frontal lobe has the highest number of dopamine-sensitive neurons. Dopamine constitutes about 80% of the catecholamine content in the brain and plays an important role in Limiting and selecting sensory information transmitted to the frontal lobe of the brain[8,10]. The frontal lobe includes the main areas such as the Prefrontal Cortex, Broca's area, Primary motor cortex, Premotor cortex, posterior parietal cortex, and supplementary motor area. The primary motor cortex is the administrative part and controls autonomic movements. The supplementary motor area helps the premotor cortex to plan complex movements. The posterior parietal cortex declares the position of the body to the premotor cortex. Prefrontal Cortex Plays a role in planning and decision making.

In this way, it first collects and combines various information (including sensory information, memory, or emotional components) from a wide range of cortical and subcortical structures and plays a key role in creating a mapping between sensory inputs of thought and action.

The purpose of this mapping is to represent the internal goals (internal states of the system) and to create an individual rule to take action.

The parietal lobe is located above the back of the head and below the forehead, and this part plays an important role in perceiving the sense of touch and muscle position sensors, gathering sensory information from different parts of the body and the connections between them. The parietal is more concerned with understanding the senses of the body and the proprioceptive. This part is located just above the ears on both sides of the head. This area receives auditory information from the cochlea and takes part in visual processing. Auditory processing, recognition, and perceptions of objects are also combinations of visual and auditory data in memory. The occipital lobe is located at the back of the skull. Visual processing and coordinating eye movements are the duty of this part, including five sections of visualizations.

Neural Activities

The central nervous system (CNS) is made up of nerve cells and glial cells that exist between the nerves. Nerve cells react to stimuli and communicate across great distances. Each nerve cell is made up of three components[11]: axons, dendrites, and cell bodies. The brain generates electrical current primarily by pumping positive ions such as sodium, Na+, potassium, K+, and Ca calcium through the membranes of neurons. Differential electrical potentials are created by the summation of postsynaptic graded potentials from pyramidal cells, which form electrical dipoles between the soma and apical dendrites of neurons. (Figure 2) represents the structure of a neuron.

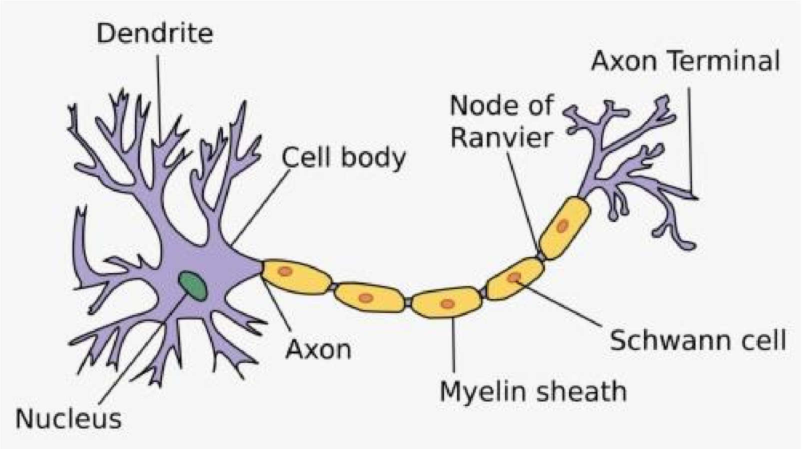


Figure 2. structure of neuron[12]

The human head is made up of several layers, including the scalp, skull, and brain, as well as numerous more thin layers in between. The skull attenuates impulses about a hundredfold more than soft tissue. Most of the noise is created inside the brain (internal noise) or on the scalp (external noise) (system noise or external noise). About One hundred thousand billion neurons are developed at birth when the central nervous system (CNS) becomes complete and functional [13]. This makes an average of 5 × 10 ^{5} neurons per cubic mm. Neurons are interconnected into neural nets through synapses. The number of synapses per neuron increases with age, whereas the number of neurons decreases with age.

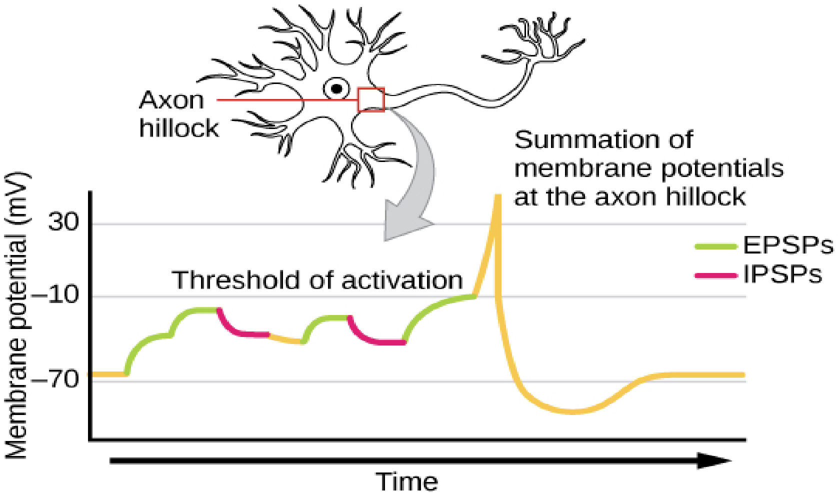


Figure 3. Neuron membrane potential changes [14]

Action potential (AP)

The action potential is the electrical signal sent by a neuron (AP). APs occur because of an ion exchange across the neuron membrane. Between 1 and 100 m/s is the conduction velocity of action potentials. A typical AP can be observed in Figure[3]. The AP surge is mostly due to the opening of Na (sodium) channels. Excitable cells have Na and K channels with gated sodium and potassium channels that open and shut in response to membrane voltage. By opening the gates of Na channels, positive charge-carrying Na may enter the cell. This results in a positive membrane potential (depolarization), resulting in a spike [15].

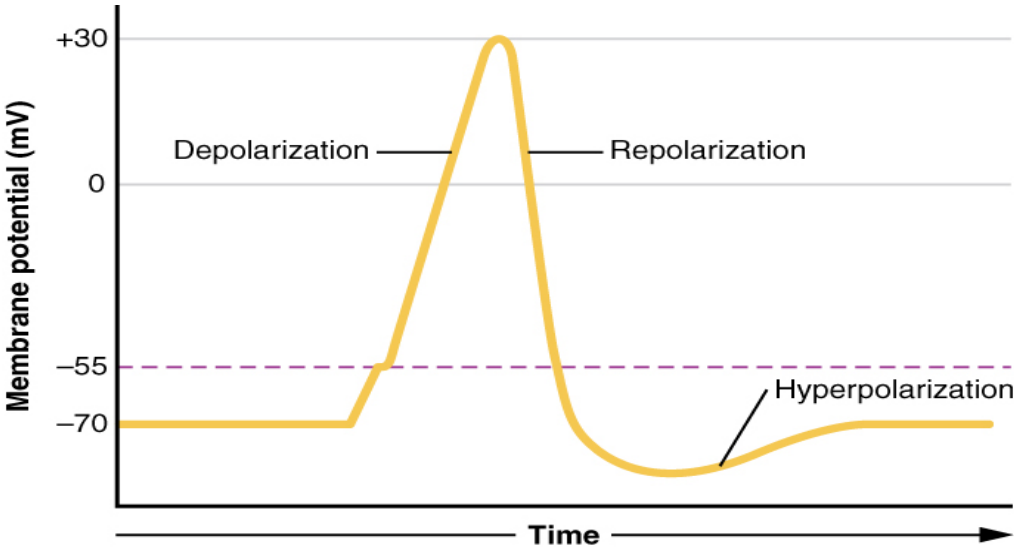


Figure 4. Typical AP
14,16

1.2 Electroencephalogram(EEG)

EEG Generation

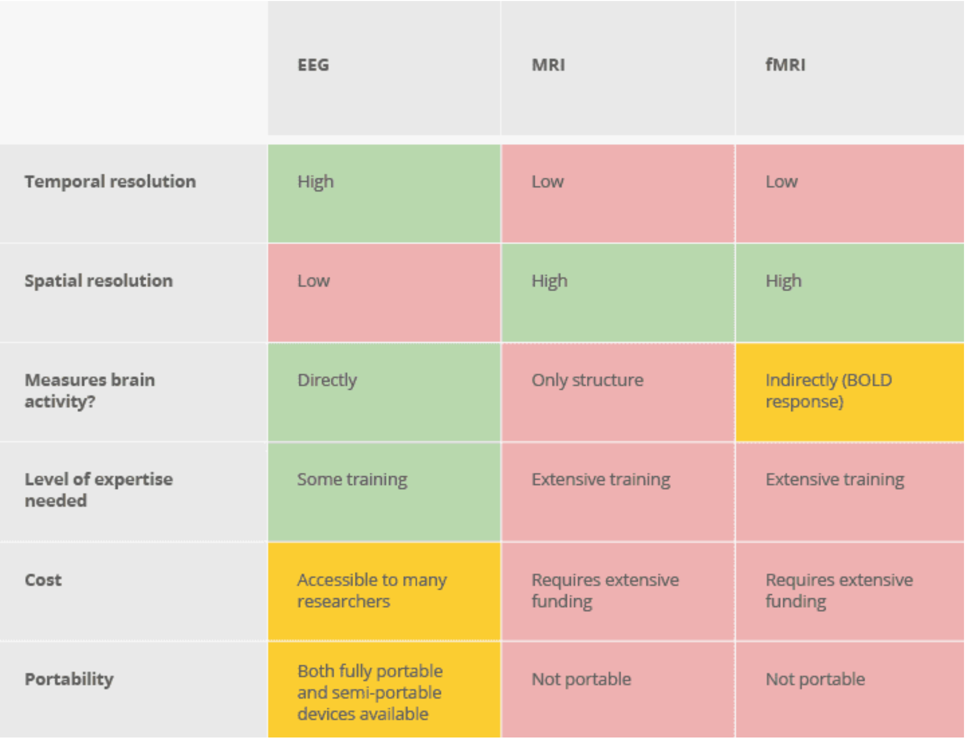
An electroencephalogram (EEG) signal is a measurement of currents that flow during dendritic excitations of many pyramidal neurons in the cerebral cortex. This current creates a magnetic field that may be measured using MEG machines, as well as a secondary electrical field that can be measured using electroencephalogram (EEG) systems.

The first to record electrical activity in the human brain was German neurophysiologist Hans Berger, who did so in the late 1920 [17]. Before this discovery, EEG [18] was mostly used for clinical purposes, such as investigating brain disorders. A typical way to find neural activity is by putting electrodes on the scalp to detect electrical signals generated by the brain. However, the activity of a single brain cell is not detectable even by highly sensitive devices. Thus, EEG measurements are associated with measurements of electrical activity generated by populations of neurons. Neurons that fire together, and that connection helps electrodes measure activity at the scalp. This electrical activity change over time is mapped as a waveform. Since EEG readings monitor an ongoing neural activity, they do excellent one-time readings. There are billions of neurons within the human brain, each in contact with the other and communicates via electrical and chemical signals. Neurons have cell bodies with extensive dendrites that carry neurotransmitters to and from receptors on the receiving dendrites. Neuron membrane potential changes in response to neurotransmitter binding to receptors, resulting in postsynaptic potentials, which are electrical signals generated at the synapse. In postsynaptic potentials, an increase or decrease in membrane potential corresponds to excitatory or inhibitory, respectively [18].

Additionally, because of the overlap in time between the spatial distribution of electrical sources and the arrival of the signals at the scalp, electrical signals become smeared and may be difficult to locate. EEG has a temporal resolution range that extends milliseconds. Because it provides a clear advantage when studying the time course of neural activity, EEG is a better method for analyzing neural activity than functional magnetic resonance imaging (fMRI), which studies the temporal activity of the brain in increments of seconds. Intracranial recordings, which are limited in practical use due to the necessity of invasive procedures, outstrip the temporal resolution of EEG by a small margin. Because EEG is a non-invasive technique.it directly records real-time neural activity without invasive methods.

The ability to acquire signals and pictures from the human body has become critical for the early detection of a wide range of disorders. EEG, MEG, or fMRI may all be used to measure functional and physiological changes in the brain. However, fMRI use is quite restricted compared to EEG or MEG for a variety of fundamental reasons. In figure[4] a comparison between EEG and MRI and fMRI is mentioned. Numerous mental activities, brain illnesses, and brain malfunctions cannot be detected with fMRI due to their little influence on the quantity of oxygenated blood. Access to fMRI (and, more recently, MEG) devices is restricted and expensive. Although EEG has a high spatial resolution, it is restricted by the number of recording electrodes (or the number of coils for MEG).

Recent EEG devices include a computerized system capable of storing and analyzing data in digital format. Variable settings, stimulations, and sampling frequency are possible with computerized systems, and some are equipped with basic or powerful processing tools [19]. The EEG is converted from analog to digital using multichannel analog-to-digital converters (ADCs). EEG signals have an effective bandwidth of roughly 100 Hz. This bandwidth was maybe even less than half of this amount in many applications. As a result, capturing EEG signals at a minimum frequency of 200 samples/s is often sufficient. The quantization of EEG data is very precise, up to 16 bits, which necessitates a large amount of memory for preserving them - particularly for sleep EEG and epileptic seizure monitoring. Various electrode types are employed, including gel-free, pre-gelled, and reusable disc electrodes, saline-based electrodes, needle electrodes, and helmet electrodes. The electrodes used for EEG recording and their appropriate operation are critical for obtaining high-quality data and a cap is often used for multichannel recordings with a large number of electrodes used.The conversion from analog to digital EEG employs multichannel analog-to-digital converters (ADCs). The raw EEG signals have amplitudes in the volt range and have frequencies up to 300 Hz[2]. To preserve useful information, signals must be amplified before the ADC and filtered, either before or after the ADC, to remove noise and prepare the signals for processing and viewing. The filters are built to introduce no alteration or distortion to the signals. Highpass filters with a cut-off frequency typically less than 0.5 Hz are used to eliminate obtrusive shallow frequency components such as those associated with breathing. On the other hand, high-frequency noise is reduced by using lowpass filters with a cut-off frequency of roughly 50–70 Hz [2].



	EEG	MRI	fMRI
Temporal resolution	High	Low	Low
Spatial resolution	Low	High	High
Measures brain activity?	Directly	Only structure	Indirectly (BOLD response)
Level of expertise needed	Some training	Extensive training	Extensive training
Cost	Accessible to many researchers	Requires extensive funding	Requires extensive funding
Portability	Both fully portable and semi-portable devices available	Not portable	Not portable

Figure 5. EEG advantage and Potentials
20

International 10-20 system

One of the common typical electrode setup (also known as 10–20) [21]. Additional electrodes are sometimes utilized to record the EOG, ECG, and EMG of the eyelid and surrounding muscles. In the figure[5] electrode location has been shown. Typically, the earlobe electrodes labeled A1 and A2 are used. For example, C3 and C4 may be utilized to capture signals associated with right and left finger movement for brain-computer interface (BCI) applications. Additionally, the ERP P300 signals may be recorded using F3, F4, P3, and P4.

Two distinct ways of recording are used, namely differential and referential. Additionally, there are reference-free recording approaches that make use of a shared average reference.

The EEG data shown in Figure [6] were collected from typical adult brain activity using a 10–20 setting scheme.

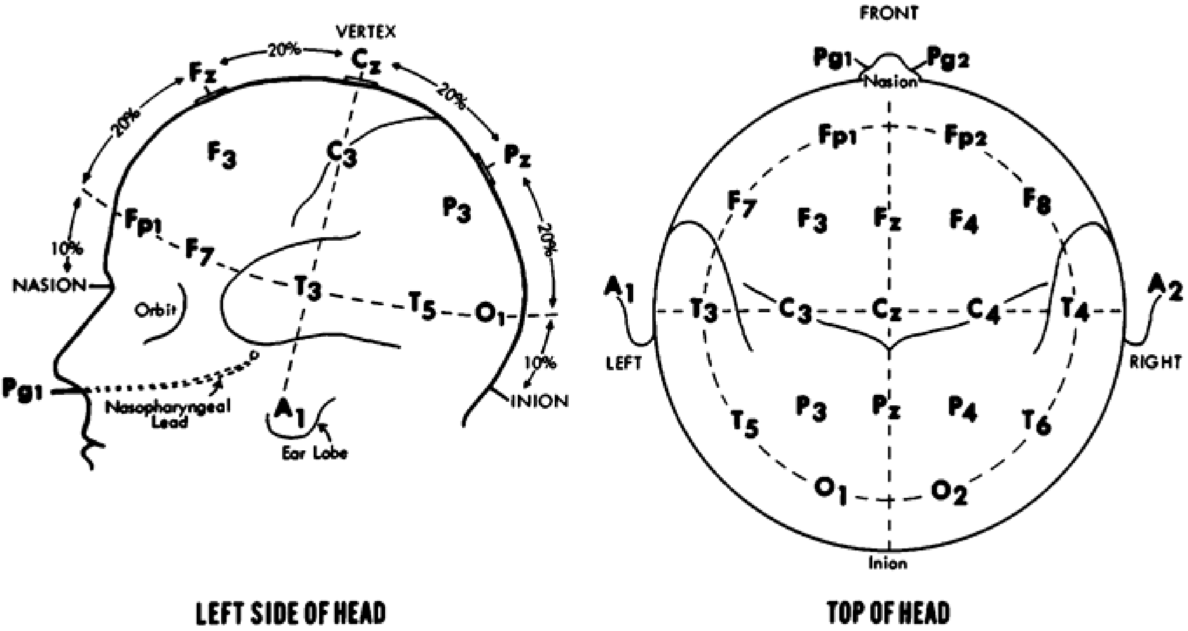


Figure 6. EEG electrode position(21 electrodes)
22

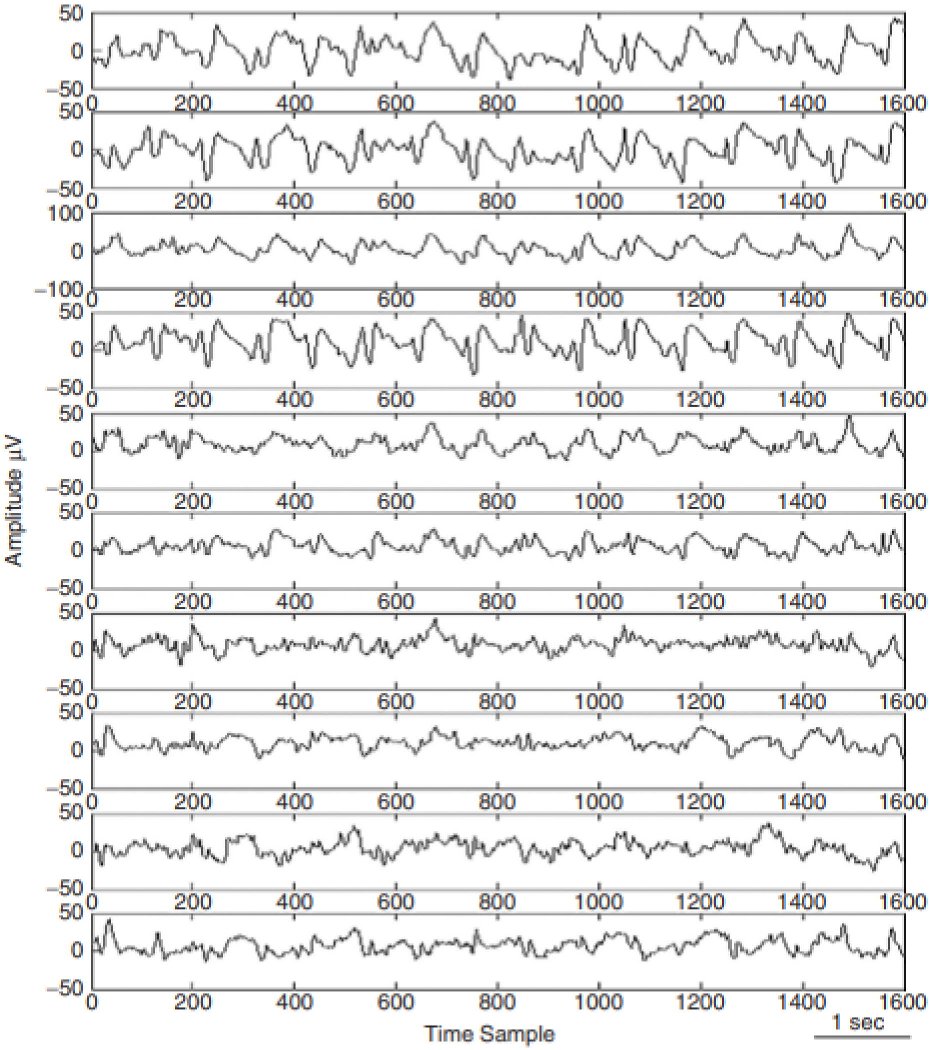


Figure 7. Typical set of EEG signal during time
2

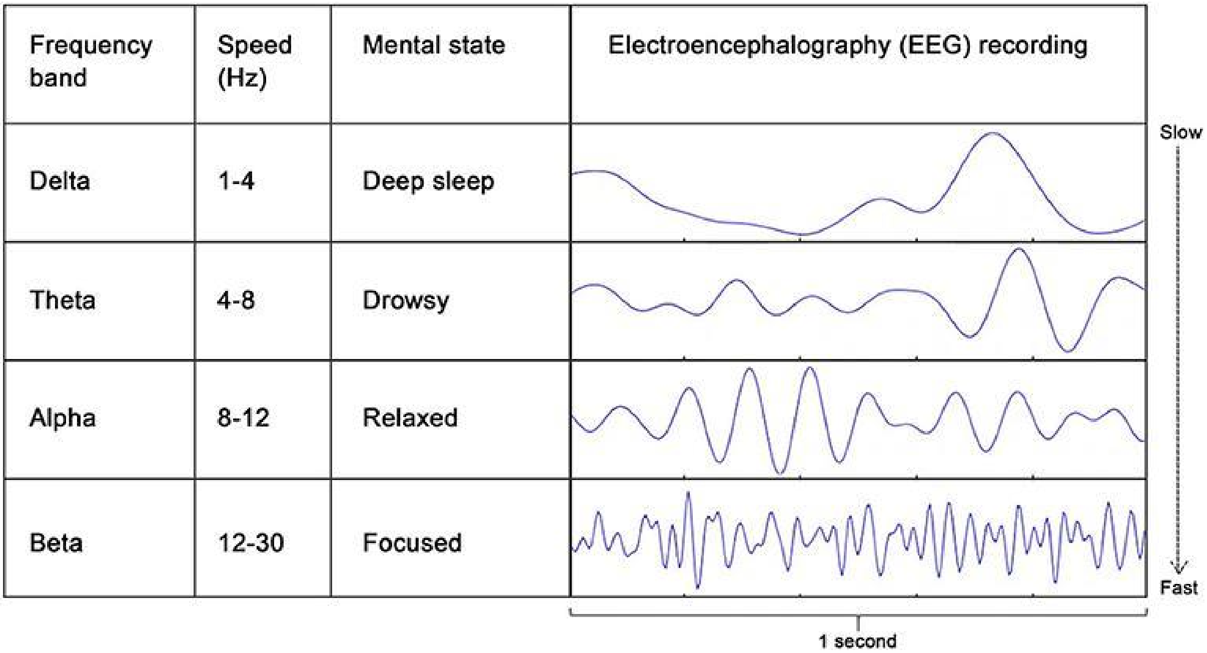
1.2.1 Brain Rhythms

The amplitudes and frequency of such signals vary across human states, such as awake and sleep, in healthy people. There are five primary brain waves.The five brain waves are Delta, Theta, Alpha, Beta, and Gamma [13] , each with a distinct frequency range. These frequency ranges are denoted by the letters alpha 8 Hz-12 Hz (α), theta 4 Hz to 8 Hz(θ), beta 12 Hz-40 Hz(β), delta 0.5Hz to 4 Hz (δ), and gamma 40 Hz-100 Hz ((γ). [23]

Delta waves occur at frequencies between 0.5 and 4 Hz and are typically connected with deep sleep. It is quite simple to mistake artifact signals produced by the neck and mouth muscles with the actual delta response. The signal of interest begins deep inside the brain and undergoes significant attenuation as it passes through the skull.

Theta waves occur between 4 and 7.5 Hz and occur as awareness begins to lapse into drowsiness. Theta waves have been linked to intuitive access, creative creativity, and profound concentration. Increased theta wave activity in the awake adult is abnormal and is produced by a variety of medical conditions [23].

Alpha waves are the most prominent rhythm in all brain activity and may span a wider range than previously believed. The frequency of alpha waves is typically between 8 and 13 Hz, and they appear as a circular or sinusoidal-shaped signal. With their eyes closed, the majority of participants generate some alpha waves, which has led to the notion that it is nothing more than a scanning pattern created by the visual areas of the brain[24]. A beta wave is the brain's typical waking rhythm linked with active thinking, active attention, external focus, and problem-solving in normal individuals. When a person is in a panicked condition, a high-level beta wave may be obtained. These rhythms have very small amplitudes. Above 30 Hz (most often 45 Hz), the gamma range, also known as the rapid beta wave, may be utilized to diagnose brain disorders. Additionally, the beta wave band has been shown to be a useful indicator of the brain's event-related synchronization (ERS) and may be used to indicate the locus for right and left index finger movement, right toe movement, and the rather large and bilateral region for tongue movement figure [8] below demonstrates the typical normal brain rhythms at their characteristic amplitude values.



Frequency band	Speed (Hz)	Mental state	Electroencephalography (EEG) recording
Delta	1-4	Deep sleep	
Theta	4-8	Drowsy	
Alpha	8-12	Relaxed	
Beta	12-30	Focused	

Figure 8. Typical rhythms of brain
2,25

1.2.2 Artifact affecting EEG

EEG Artifacts:

Following the presentation of the EEG and EEG signal concepts and various methods for signal collection in this section , certain artifacts used in this investigation were discussed. First discussed are Some common artifacts with details After that, how these artifacts affected EEG signals will be presented [3]. When EEG data is collected through recording equipment, signal artifacts are more prevalent [4] These artifacts have the potential to compromise the EEG data's quality. A thorough understanding of the many artifacts is required in this case to eliminate the artifacts or noise effectively. Abnormalities are undesired signals often caused by noise in the environment, experimental mistakes, or physiological artifacts.Additionally, extrinsic artifacts are caused by external causes such as the environment or experiment mistakes. Intrinsic artifacts are those caused by the body itself [3,4][4][5] .Although EEG signal affected by these three artifacts (eye blink, muscle activities, heartbeat), there are some more artifacts such as Line noise channel noise rarely affected EEG signals [5] . In figure number 8 , the shape of each noise is presented separately.

Eye Artifacts

Ocular artifacts are caused by eye movement and blink that may propagate over the scalp and be detected by EEG activity. The electrooculogram may be used to record such ocular impulses (EOG). EOG has a much larger amplitude than EEG and a frequency comparable to EEG impulses [26][5] It's important to note that EEG data may be affected by EOG, but EEG can also impact EOG data.

Muscle Artifacts

EMG artifacts may be created by any muscle contracting or stretching close to the signal recording locations or by the person speaking, sniffing, or swallowing. Compared to EOG and eye tracking, it is relatively challenging to derive activity from a single-channel measurement. Muscle artifacts detected by electromyogram (EMG) have an extensive frequency range of 0 Hz to >200 Hz in general [27].

Heart artifact

cardiac artifacts may be generated when electrodes are placed on or near a blood vessel[26,28],where the expansion and contraction of the heart cause movement of the electrodes.plus artifacts,with a frequency of roughly 1,2 Hz,may exist inside the EEg as an identical waveform,making them difficult to eliminate[26,28].another kind of cardiac activity called ECG records the electrical signal produced by the heart[29].In contrast to pulse artifacts ECGs may be measured with a consistent pattern and recorded independently of brain activity;hence,eliminating such artifacts may be more straightforward if a reference waveform is used.

Other artifacts

Apart from the mentioned distortions, external sources of artifacts also have a detrimental influence on EEG measurement. Instrument artifacts, a form of extrinsic artifact, are caused by electrode misalignment and cable movement. These artifacts are easily erased with the proper process and preparation. Electromagnetic interference generated by the environment is another external artifact that impairs EEG recordings. Due to the recognizable frequency range, such artifacts from ambient sources may be readily reduced using a simple filter. Despite white noise's vast frequency spectrum, a high-frequency filter can still eliminate the bulk of artifacts. Because the same brain area's action may appear in many channels, coherence across EEG channels introduces a volume conduct artifact . has pertinent literature dealing with volume conduct artifacts.

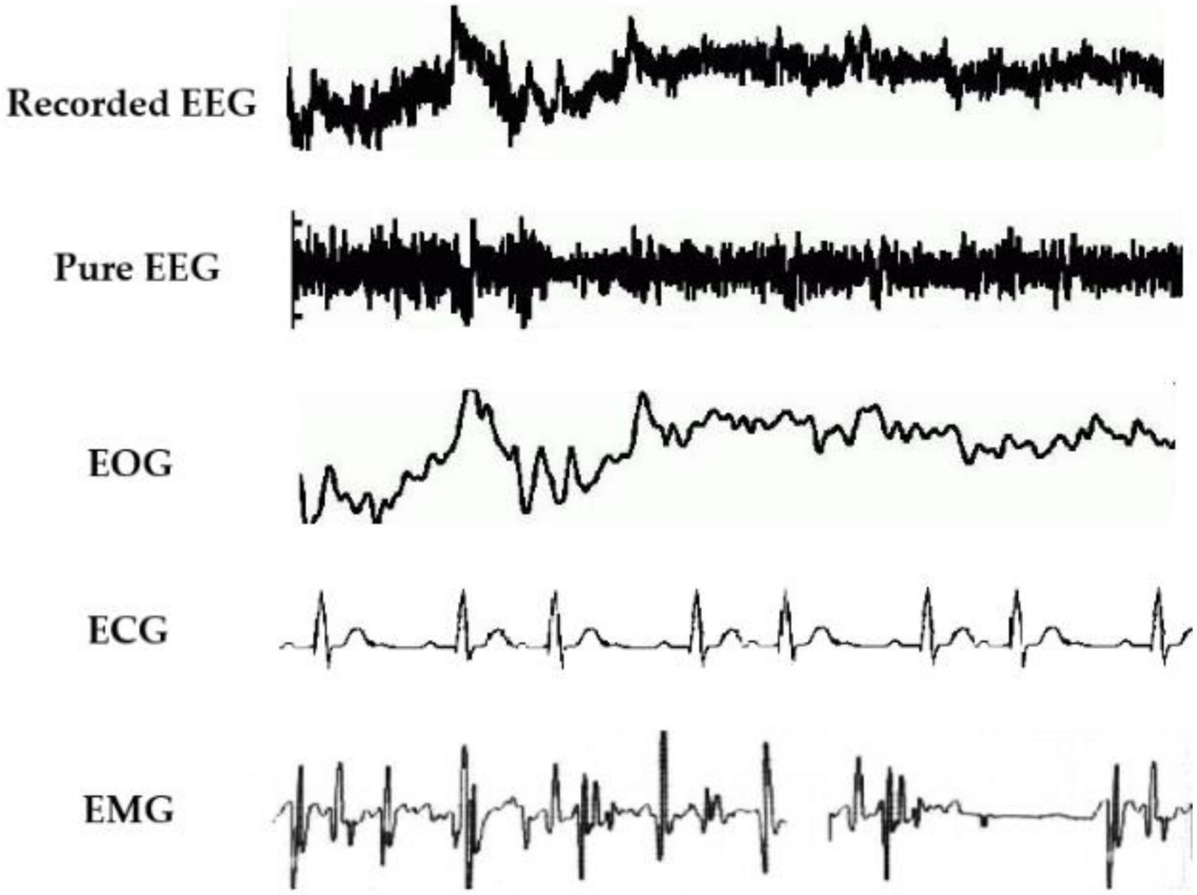


Figure 9. physiological artifact present in EEG
26

Apart from the mentioned distortions, external sources of artifacts also have a detrimental influence on EEG measurement. Instrument artifacts, a form of extrinsic artifact, are caused by electrode misalignment and cable movement. These artifacts are easily erased with the proper process and preparation. Electromagnetic interference generated by the environment is another external artifact that impairs EEG recordings. Due to the recognizable frequency range, such artifacts from ambient sources may be readily reduced using a simple filter. Despite white noise's vast frequency spectrum, a high-frequency filter can still eliminate the bulk of artifacts. Because the same brain area's action may appear in many channels, coherence across EEG channels introduces a volume conduct artifact that has pertinent literature dealing with volume conduct artifacts.

Removal of artificial segments:

This was an early strategy for dealing with artifacts. One method for reducing the impacts of artifacts is to reject or cancel the epoch or segment of EEG classified data as artifactual. The biggest drawback of this strategy is that it also eliminates critical EEG information, resulting in data loss [30][31] nowadays, thanks to recent advanced methods, this method is not very usable. However, in some situations, such as offline analysis or training a classifier, this strategy may still function pretty well[32]. There are two categories for artifact removals by maintaining the feature of signal: regression or filtering and segmentation of EEG data into other regions.

Notch filters with a 50 Hz null frequency are often required to reject the powerful 50 Hz power source.The following are patient-related or internal artifacts: body movement, EMG, ECG (including pulsation), EOG, ballistocardiogram, and perspiration. The system artifacts include interference from the 50/60 Hz power supply, impedance fluctuation, cable faults, electronic component noise, and electrodes with unbalanced impedances. Often, these artifacts are greatly minimized during the preprocessing step, and valuable data is recovered. Some methods for removing these kinds of artifacts and more detail about artifacts will be discussed in the follwoin sections .In the following figure [10] displays a set of standard EEG signals affected by the eye-blinking artifact. Also, multichannel EEG set with the ECG signals visible across the occipital electrodes shown in figure number 10 and 11].

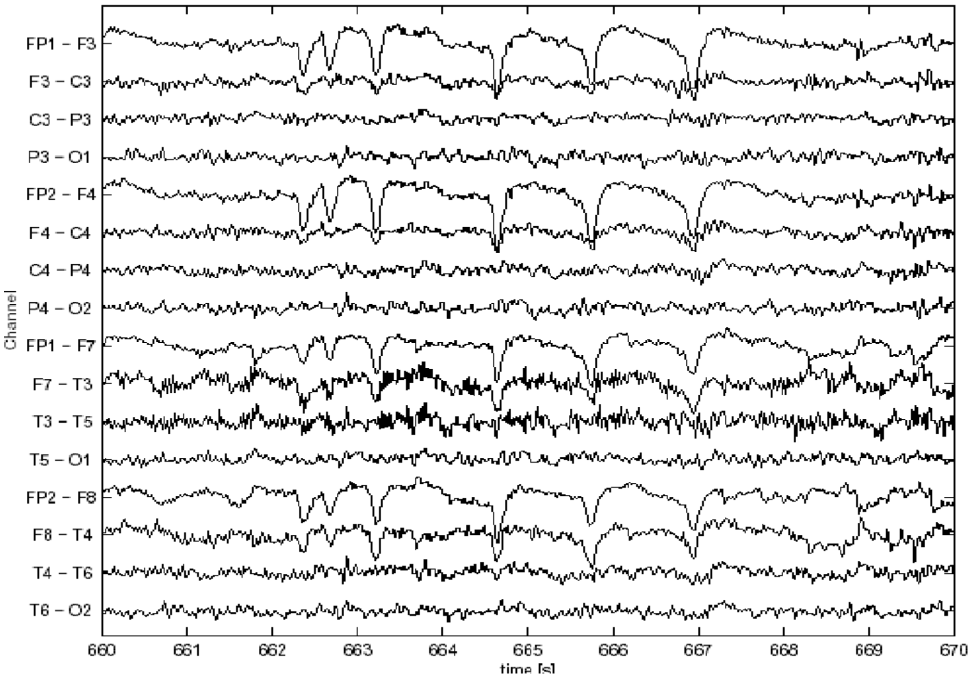


Figure 10. EEG recording contaminated by Eye Blinks
33,34

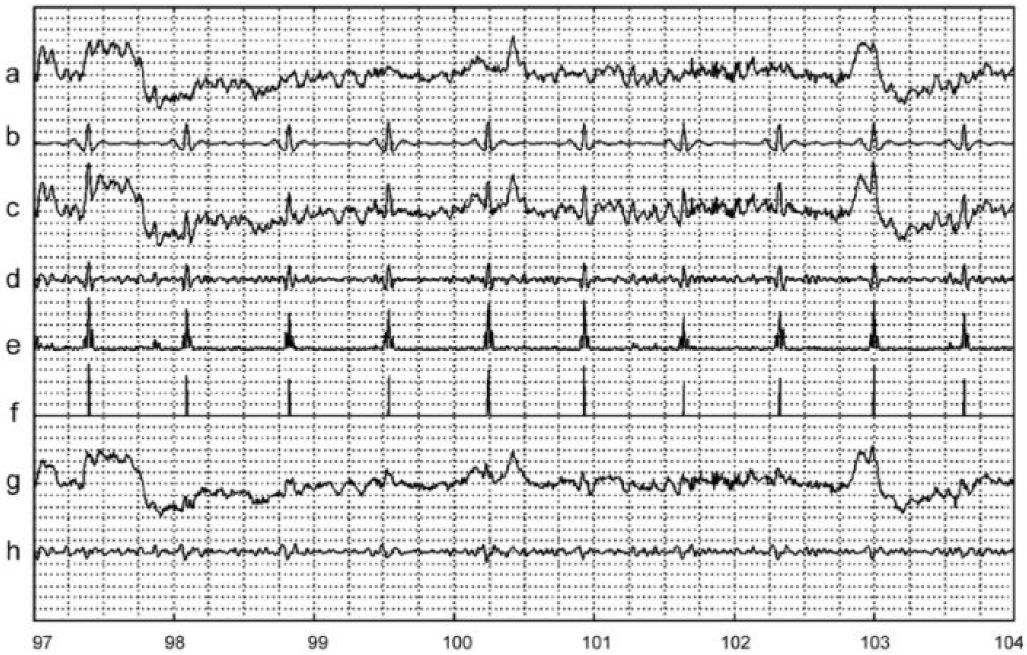


Figure 11. EEG set with the ECG signals
3

Skewness

Skewness measures the asymmetry of the probability distribution of a real-valued random variable around its mean in probability theory and statistics. The skewness value may be positive, zero, harmful, or undefined. When a distribution is unimodal, negative skew often suggests that the tail is on the left, whereas positive skew indicates that the tail is right. Skewness does not follow a simple rule in circumstances when one tail is extended while the other is fat. For instance, a zero value indicates that the tails on both sides of the mean balance out overall; this is true for asymmetric distribution, but it may also be valid for an asymmetric distribution with one tail being long and thin and the other being short and fat [35] where μ is the mean, σ is the standard deviation, E is the expectation operator [36] μ3 is the third central moment, and κt is the t-th cumulants. Figure [12] represent the situation of Skewness[[37]]

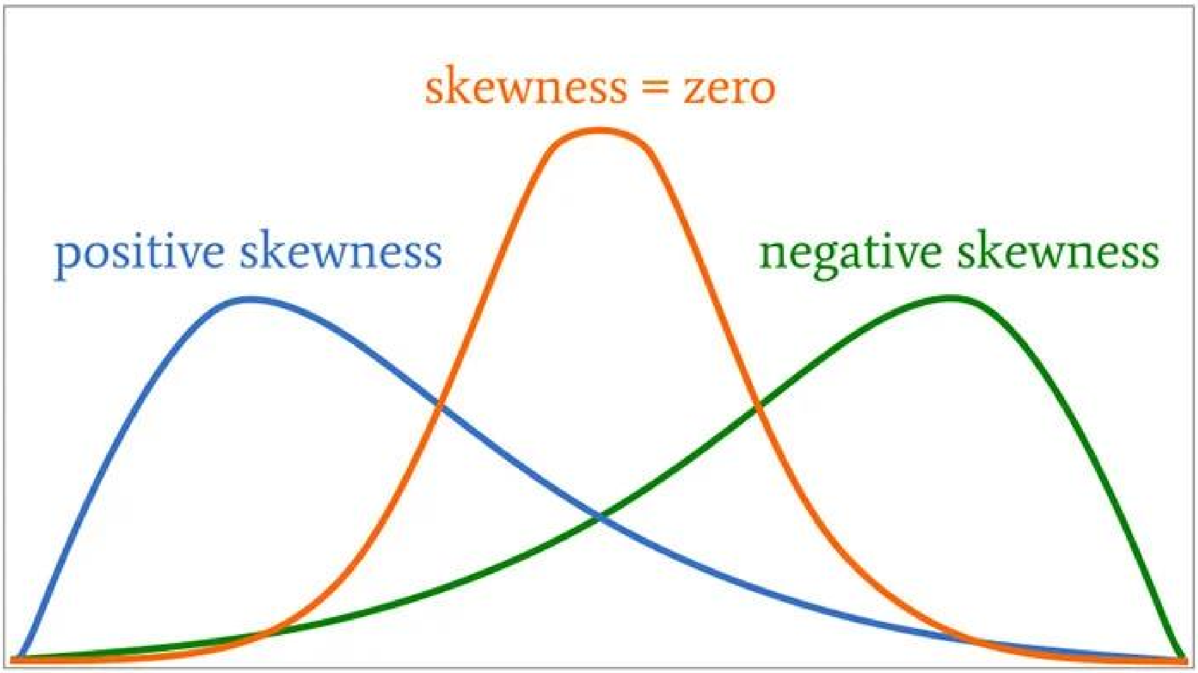


Figure 12. Skewness poison

Kurtosis

In probability and statistics, kurtosis is a measure of the "tailedness" of the probability distribution of a real-valued random variable. Increased kurtosis suggests a greater number of severe deviations (or outliers), rather than a more normal distribution of data. The uniform distribution, which is devoid of outliers, demonstrates a kurtosis. Kurtosis value distributions have longer tails than Gaussian distributions and so produce more outliers than the normal distribution.[38] distribution. The kurtosis for a real signal x(n) is defined as:
mi[x(n)] is the ith central moment of the signal x(n). Figure[13]represent distribution of Kurtosis .Therefore, an excess or normalized kurtosis is often used and defined as:

$$Ex\ kurt\ = \frac{m_4\ \ [x(n)]}{m_2^2\ \ [x(n)]} - 3$$

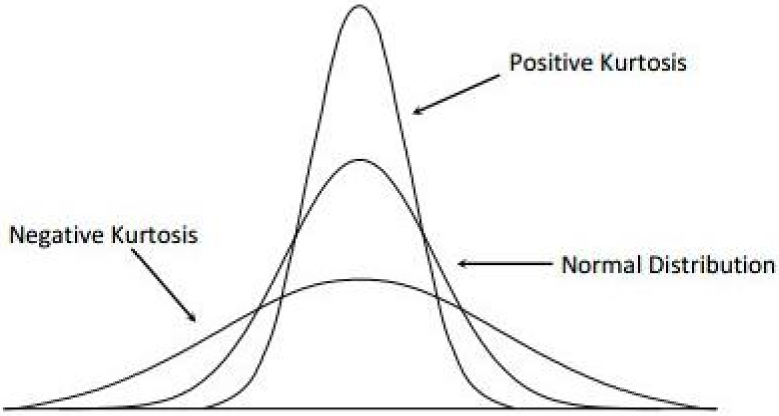


Figure 13. Distribution of Kurtosis
39

The power spectral density (power spectrum)

The power spectral density (power spectrum) reflects the 'frequency content' of the signal or the distribution of signal power over frequency. Any physical signal may be divided into a set of discrete frequencies or a spectrum of frequencies throughout a continuous range using Fourier analysis. The spectrum is the statistical average of a particular signal or kind of signal (including noise) as viewed in terms of its frequency content. When the energy of a signal is concentrated around a discrete period, the energy spectral density may be computed. The power spectral density (PSD) is then used to refer to the spectral energy distribution discovered over a unit period. A physical process's overall power (or variance) is obtained by summing or integrating the spectral components (in a statistical process). The power spectrum is critical in statistical signal processing and stochastic process analysis, and a wide variety of other fields of physics and engineering. Typically, the procedure is time-dependent[[40]][[41]].

Factors Affecting the Estimation of the Power Spectral Density in EEG

Estimating power in various frequency ranges is the most often done analysis in EEG research. The power spectral density (PSD) is approximated using a (discrete) Fourier transform, or DFT, which produces information on the power of each frequency component. PSD estimation is dependent on a variety of factors, including the length of the window, the degree of overlap between the windows, and the number of DFT points[[41]]. Programming languages such as MATLAB and Python have algorithms for computing the DFT for a given signal or time series by using the fast Fourier transform (FFT) technique. So the effect of each factor (length of windows, degree of overlap between the windows, and the number of the DFT point)will be discussed in more detail in the next section. here is mentioned some small impact of these feature on PSD[[42]][[43]]

1. A smooth PSD can be generated by the Welch method; the Welch method is an improvement over the usual method for calculating the periodogram spectrum. It minimizes noise in the estimated power spectra at the expense of frequency resolution. Due to the noise introduced by imprecise and finite data, Welch's noise reduction techniques are often needed [[42]]
2. In low-frequency analysis, small windows size decreases the PSD performance calculations

3. Large windows may increase the resolution of PDS and bring the noise. considering the frequencies of interest is important to choose windows size
4. The number of DFT points should always be greater than or equal to the length of the window
5. Due to the significant correlation between the windows, highly overlapping windows do not always result in smoother PSD estimations.[42]

1.3 Independent Component Analysis(ICA)

Signals in the mixture are linear combinations of the signals in the sources. This is in contrast, Independent Component Analysis (ICA) is a term referring to a collection of data from many random variables that are intended to be represented as a linear combination of observations from various other random variables that are independent of one another. Many artifacts make the analysis of clinical EEG signals difficult; also, rejecting damaged EEG segments results in severe data loss. Recent efforts to characterize EEG sources have mostly focused on spatial separation and localization of source activity.Independent component analysis is known as the process of optimizing the degree of statistical independence between outputs using estimated contrast functions. ICA implemented in electroencephalographic (EEG) data analysis. The primary goal of ICA on a random vector is to find a linear transformation that reduces the statistical dependency between the signal's components. In practice, the ICA algorithm has been used to investigate the problem of source identification and localization. The ICA algorithm derives independent sources from highly correlated EEG signals statistically without concern for the actual location or configuration of the EEG signal source sources[2].Recently, more emphasis has been made to ICA approaches as a remarkable and deconstructed tool[44]; ICA was more effective than artificial neural networks and adaptive schemes. Indeed, the ICA approach is well suited for signal source separation when the sources are statistically independent and meet certain additional characteristics.Additionally, the independent signal source is considered to have the same number of sensors as the N sensors. Thus, by using the ICA approach, it is possible to extract signals from several sources. If we suppose that the complexity of EEG dynamics can be described as a collection of a small number of statistically independent brain activities, then the EEG source analysis issue meets the ICA assumption. The primary issue in understanding ICA output is determining the right dimension of the input channels and the physiological and/or psychophysiological relevance of the resulting ICA source channels. The ICA model of the EEG ignores the known variable synchronization of separate EEG generators by common subcortical or cortical influences. ICA is used to locate concurrent signal sources that are either too close together or too broadly scattered to be separated using conventional localization techniques[45 46].Each estimated source is called an independent components (IC)[47]

Here is the mathematical part of the ICA[2] calculation algorithm by considering y(n) as the multichannel signal , yi(n) as constituent signal components, so yi(n) are independent if:

$$p_y\left(y(n)\right) = \prod_{i=1}^{m} \quad p_y\ (y_i\ (n)\ )\ \forall_n$$

where p(Y) is the joint probability distribution , py(yi(n)) are the marginal distributions and m is the number of independent components .

1.4 Data set of semi-automatic ICs selection pipeline

the data has been used in this research acquired form 31 subjects eating breakfast. The current data set is the resting state over 31 subjects with more than 2600 ICs.

Chapter 2

2 Materials and methods

In chapter one, some general information of the brain, its activities, the concept of EEG, and signal acquisitions has been described .Also in the first chapter, that EEG signal has a lot of various artifacts that need to be cleared for feature extraction and study. Many methods have been published for feature extraction and filtering EEG signals. In this section implication and the software that have been used in this study will be evaluated. Then more details about algorithm and methods used for filtering and feature extraction will be discussed with more details.the aim of this study is to evaluate two different methods knowns as semi-automatic ICs selection pipeline and IClabel.Firstly EEG signal preprocessing with Matlab by Fieldtrip toolbox. Then for the second time, that EEG signal tried to be preprocessed by another EEGlab Plugin called ICLabel, and finally, the result of both outputs was compared. By considering these methods first, more details about the Field trip and EEGLab will be presented.pre-processing for the method will be described in following sections.

2.1 EEG Signal Preprocessing Pipeline

This section discusses two pipelines and algorithms to select ICs.

2.1.1 Nonlinearity and nonstationary of EEG signals

An EEG signal may be regarded as the output of a nonlinear system that is deterministically described. Understanding such a system is highly complex. Several measurements derived from chaos theory and time series analysis may be used to describe the nonlinear behavior of EEG data.

Nonstationarity may be determined by observing the statistics of signals with varying time delays. If these statistics do not fluctuate over time, the signals are considered stationary. In EEG distribution, mean and covariance features vary significantly across segments. As a result, EEGs are deemed stationary only during brief periods. This is true just for regular brain functions, not mental or psychical activities. Nonstationary can be observed during eye blinks, awareness and attention fluctuations, and event-related potential (ERP) and evoked potential (EP) signals.

The variation in the distribution of the signal parts could be calculated employing both the Gaussian process's parameters and the distribution's deviation from Gaussian. By measuring some parameters, for example, Kurtosis ,Kulback-laibler(KL) distance, one over F, (power spectral density)Spectflat and skewness can check The non-Gaussianity of signals.

$$\mu_3 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} = \frac{E\left[(X-\mu)^3\right]}{\left(E\left[(X-\mu)^2\right]\right)^{\frac{3}{2}}} = \frac{\kappa_3}{\kappa_2^{\frac{3}{2}}}$$

48,49

$$kurt = \frac{m_4\,[x(n)]}{m_2^2\,[x(n)]}$$

mi[x(n)] is the ith central moment of the signal x(n). Figure[13]represent distribution of Kurtusis .Therefore, an excess or normalized kurtosis is often used and defined as:

$$Ex\ kurt = \frac{m_4\,[x(n)]}{m_2^2\,[x(n)]} - 3$$

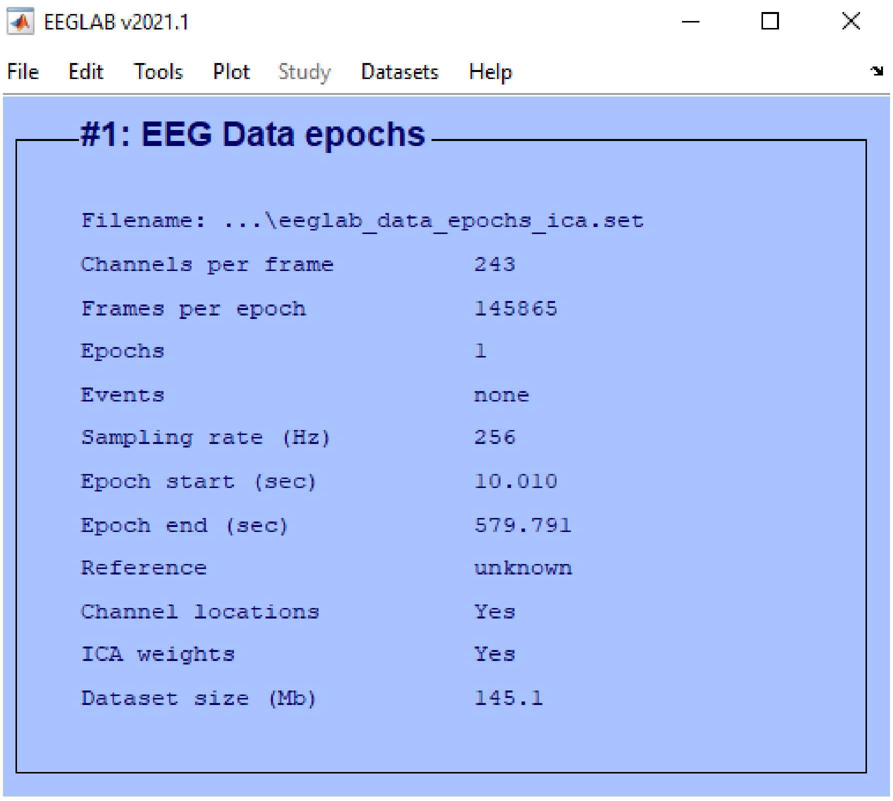
48–50

2.1.2 Implementation EEGLAB and FieldTrip

Arnaud Delorme and Scott Makeig founded EEGLAB in 2000 at The Salk Institute in La Jolla, California. EEGLAB was released in its entirety in 2002 at UCSD's Swartz Center for Computational Neuroscience. EEGLAB employs the DIPFIT plugin for essential source localization, including a FieldTrip source localization method. EEGLAB is a MATLAB toolkit for processing continuous and event-related EEG, MEG, and other electrophysiological data utilizing independent component analysis (ICA), time/frequency analysis, and other techniques, such as artifact rejection. EEGLAB combines and extends Making's ICA/EEG toolkit and offers a graphical user interface. EEGLAB is available at[51] http://www.sccn.ucsd.edu/eeglab/.

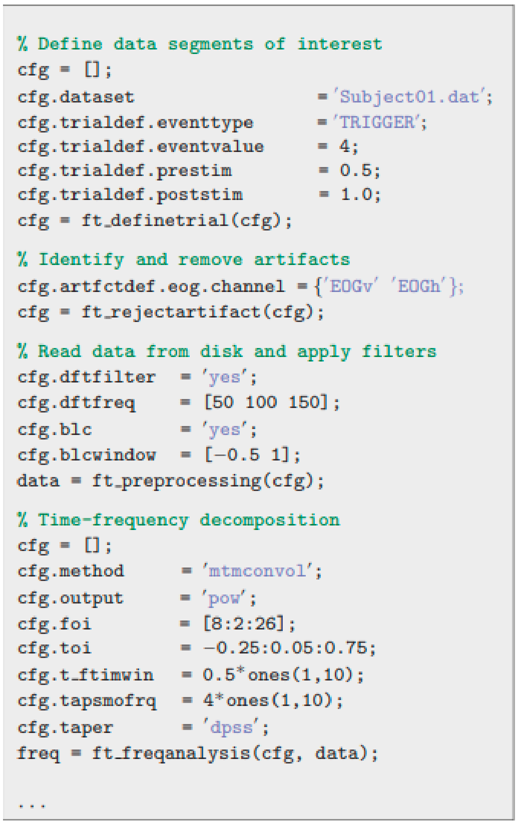
A critical objective of the FieldTrip project is to provide a centralized platform for experimental scientists and method developers.[[48]] Both programs are compatible with MATLAB (The Mathworks, Inc.). While EEGLAB enables the processing of several continuous data segments, it lacks the flexibility of FieldTrip when it comes to processing data epochs of varying duration. The third-party extension (plugin) feature in EEGLAB offers a free platform for creating and releasing additional tools by several organizations and individuals. EEGLAB and FieldTrip may be included in the MATLAB pipeline and run concurrently. Some functions are available to convert data structures between the two toolboxes, such as the EEGLAB eeglab2fieldtrip.m and fieldtrip2eeglab.m functions - note that they are intended to convert specific data structures (FieldTrip has several of them) and are not intended to be general-purpose. To complete the control over the processing pipeline, it may also use EEGLAB since it provides direct access to low-level processing capabilities and includes extensive documentation. Additionally, the EEGLAB Extension Manager enables the user to assist users in maintaining current versions of your extensions.[[52]]

In the Fieldtrip, typically, analysis scripts involve a series of FieldTrip function calls. Typically, each analysis stage is handled by a separate high-level FieldTrip function. An example of EEGLab can be seen in figure [14]. To give a better idea in the figure [15] presented analysis pipeline schematically [[48]][[52]]



```
#1: EEG Data epochs

Filename: ...\eeglab_data_epochs_ica.set
Channels per frame            243
Frames per epoch              145865
Epochs                        1
Events                        none
Sampling rate (Hz)            256
Epoch start (sec)             10.010
Epoch end (sec)               579.791
Reference                     unknown
Channel locations             Yes
ICA weights                   Yes
Dataset size (Mb)             145.1
```

Figure 14. EEGLAB User Interface



```
% Define data segments of interest
cfg = [];
cfg.dataset                   = 'Subject01.dat';
cfg.trialdef.eventtype        = 'TRIGGER';
cfg.trialdef.eventvalue       = 4;
cfg.trialdef.prestim          = 0.5;
cfg.trialdef.poststim         = 1.0;
cfg = ft_definetrial(cfg);

% Identify and remove artifacts
cfg.artfctdef.eog.channel = {'EOGv' 'EOGh'};
cfg = ft_rejectartifact(cfg);

% Read data from disk and apply filters
cfg.dftfilter   = 'yes';
cfg.dftfreq     = [50 100 150];
cfg.blc         = 'yes';
cfg.blcwindow   = [-0.5 1];
data = ft_preprocessing(cfg);

% Time-frequency decomposition
cfg = [];
cfg.method      = 'mtmconvol';
cfg.output      = 'pow';
cfg.foi         = [8:2:26];
cfg.toi         = -0.25:0.05:0.75;
cfg.t_ftimwin   = 0.5*ones(1,10);
cfg.tapsmofrq   = 4*ones(1,10);
cfg.taper       = 'dpss';
freq = ft_freqanalysis(cfg, data);

...
```

Figure 15. Sample of analysis script FieldTrip[48]

2.2 Artifact managing with FieldTrip

1: Rejection of noisy electrodes or parts of the signal in which we observe artifacts. Here we need to identify and remove the contaminated data and just analyze the intact data. In this manner, we can avoid channels with a bad connection or ruined trials. This approach is useful for the elimination of signal artifacts and also the subject's inappropriate behaviors.
2: Using bandpass signal filters and also the ICA method to eliminate artifact contribution in the data. FieldTrip[53] has provided a bunch of functions based on which we can detect artifacts. Consider that it is easier to eliminate artifacts from continuous data before segmentation; it would let us use filters without edge effects. By considering the signal's artifact as a linear summation to the pure signal, we can use filtering for some types of artifacts. For example line noise in 50-60 Hs and also high-frequency muscle noises.

2.3 semi-automatic ICs selection pipeline and ICLable projects

semi-automatic ICs selection pipeline
An algorithm which in this report we know as Rome algorithm was designed by the member of Consynlab in the University La Sapienza.

Rome criteria:

According to twelve criteria Rome algorithm classify each ICs, only all these twelve criteria match the output labeled as brain otherwise labeled as artifact. Rome criteria will be mentioned in below table,One over F is the feature of 1/f , Specflat shows property of being smooth , and Kurtosis represents Gaussian criterion. The other features in the table represent EOG,which is an artifact related to electro-oculographic. EKG related to heart artifacts and VEOG also relet to eyes artifacts (These criteria are fixed for all IC in all subjects). After running the algorithm, the output is a binary classification as Brain or artifacts corresponding to the final label for each ICs.
All the labels for each ICs controlled by the members of Consylab and some of them corrected and their final label changed.
After running all data by the algorithm and the output classification corrected and controlled in order to compare the accuracy of the result, we used another method known as the ICLabel project and we compared the result to correct possible misclassified ICs components and

increase the performance of the Rome algorithm. In this section more detail about ICLabel and how this algorithm design and works will be decrease with full details

Criteria	Thresholds
OneOverF	>0.91
Specflat	<3
Kurtosis	>15
elecSig[HEOG]	>0.1
elecPow[HEOG]	>0.25
elecSpe[HEOG]	>0.95
elecSig[EKG]	>0.1
elecPow[EKG]	>0.25
elecSpe[EKG]	>0.95
elecSig[VEOGR]	>0.15
elecSig[VEOGR]	>0.25
elecSpe[VEOGR]	>0.95

Table 1. semi-automatic criteria

2.4 ICLabel project

This is the second method we used to classify our Ics, This new approach recently published is an automatic EEG independent component classifier plugin for EEGLAB[54]. The main aim of this method is classifying ICs with different components . More details will be discussed about this method in the following sections. This is the link to download the plugin https://github.com/sccn/ICLabel.

In the output of each ICs some information can be observed are follow :

1/**Scalp topography images** These square images, 32 pixels to a side, are calculated using a slightly modified version of the topo plot function in EEGLAB

2/ **Channel-based scalp topography measures**: this feature and scalp topography images can calculated by using the function topoplotFast.m .this is a 32x32 pixel greyscale image. Each,black dot in the figure[17] illustrates the location and name of the electrode, the different colors in this map represent the power or energy in that area. For example, here, the red area shows the most activities based on that.

3/ Power spectral densities (PSD) described before its densities from 1 to 100 Hz[[1]]

4/ Autocorrelation functions: Autocorrelation functions can be calculated using file eeg_autocorr.m The autocorrelation function (ACF) gives insight into the distribution of hills and valleys over the surface.

Also, this feature used as input for defining machine learning model will be mentioned. more detail about this feature and their characteristic are as follow also, in figure [17] illustrate the graphical representation of them.

5/ECD model: this one used in ICLable ,a single and bilaterally symmetric equivalent current dipole, The ECD model is fitted using EEGLAB's DipFit plug-in, which determines the dipole locations and moments corresponding to the IC scalp topography[1].

The figure below illustrates the output for each IC number in MATLAB. The time series to the top-right shows IC activity, as does the plot to the bottom-left. Downright represent PSD or power Spectrum.

The variation in the distribution of the signal parts could be calculated employing both the Gaussian process's parameters and the distribution's deviation from Gaussian. by measuring some parameters, for example, Kurtosis ,Kulback-laibler(KL) distance, one over F, (power spectral density)Spectflat and skewness can checked The non-Gaussianity of signals.

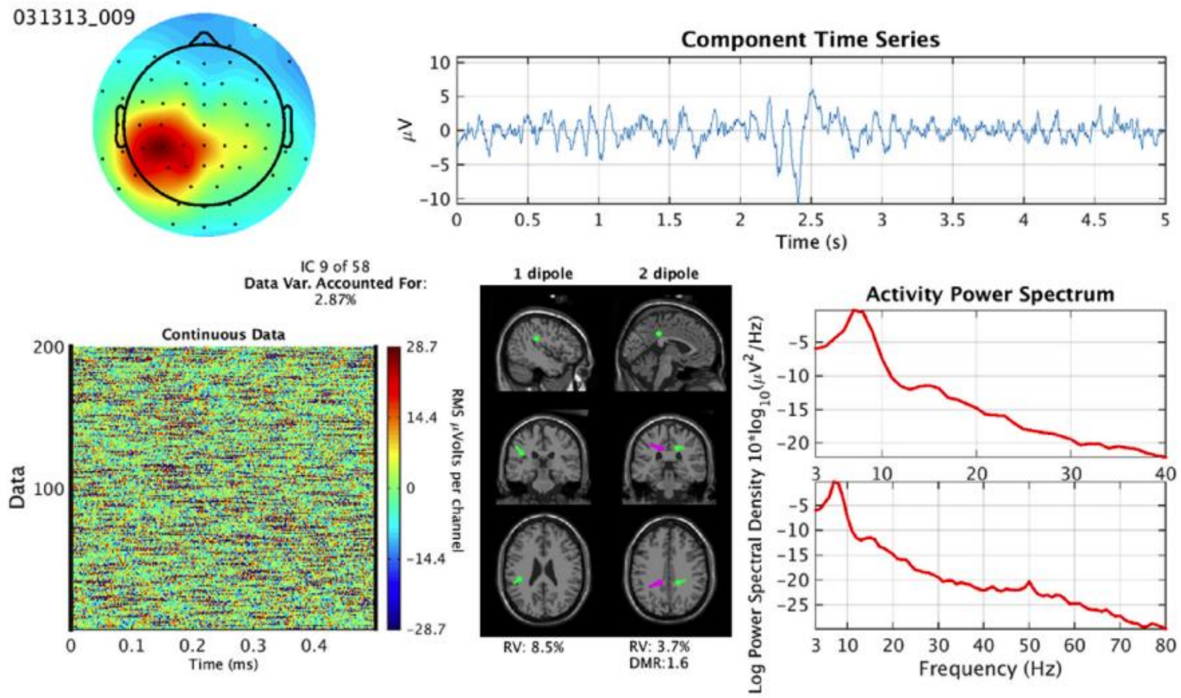


Figure 16. ICLabling features extraction 55

2.4.1 ICLable website and tutorial

ICLabel data set and categories:

Data set:

the ICLabel dataset contains spatiotemporal measures for over 200,000 [62]ICs from more than 6000 EEG recordings and matching component labels for over 6000 of those ICs. Its dataset collects extracted characteristics from EEGLAB-discovered, anonymized EEG samples in .set files format. Deconstructed by ICA and include details about the channel's location. by using this code

ICL_feature_extractor.m

This data set used to train these approaches, and the website used to gather and crowd-sourced IC labels for the dataset. The classifier is available for download through the EEGLAB extensions manager as ICLabel or directly from https://github.com/s.after

ICLabel website :

The ICLabel website (https://iclabel.ucsd.edu/tutorial) , This website was designed to label the IC without any previous labels. After logging into the website, there is some free tutorial to train the participant or voluntary person who wants to label these IC . with this training tutorial, and the labeling accuracy could be increased. the overview of this tutorial step by step can be seen in following sections.

One of the contributions of this report and this thesis related to a deep understanding of how this ICLabel works. Here mentioned full detail about website, signal collecting and labeling and the user manual of the website. Then in the following sections will discuss the machine learning frame used in ICLabel and the algorithm implemented to achieve this accuracy and approach.

After registering on the website, the participants have this option also instead of setting distinct labels to each component, apply as many labels to a component as they want. These possible are as follow:

- Brain
- Eye
- Muscle
- Heart
- Line Noise
- Channel Noise
- Other

There is some valuable information for labeling and Telling Components. Apart from these 7 components will be present single by single.

Brain

Brain components are thought to be formed when patches of cortex become spatially coherent due to local field activity crossing them. Typically, this patch's electrical field may represent an "equivalent current dipole." This is comparable to a small electrical bar magnet. Because two dipoles describe some brain components more accurately, the one dipole may have a higher residual variance than planned. While brain activity occurs at frequencies greater than 200 Hz, only low frequencies are often synchronized enough to be seen in the EEG. As a result, brain components lose their efficacy at higher frequencies. Brain components display repeating patterns at specific frequencies, resulting in a power spectrum peak. These peaks are often observed between 5 and 30 Hz, with the most common being 10Hz (dubbed alpha). The most straightforward technique to determine if an ERP exists is to examine the ERP-Image.

- Here is an example of the most signification factor to detect the brain: the scalp topography shows the bipolar feature. in the second plot power spectrum decreased as frequency increased it known as one over . both these features are clearly shown in figure [17]

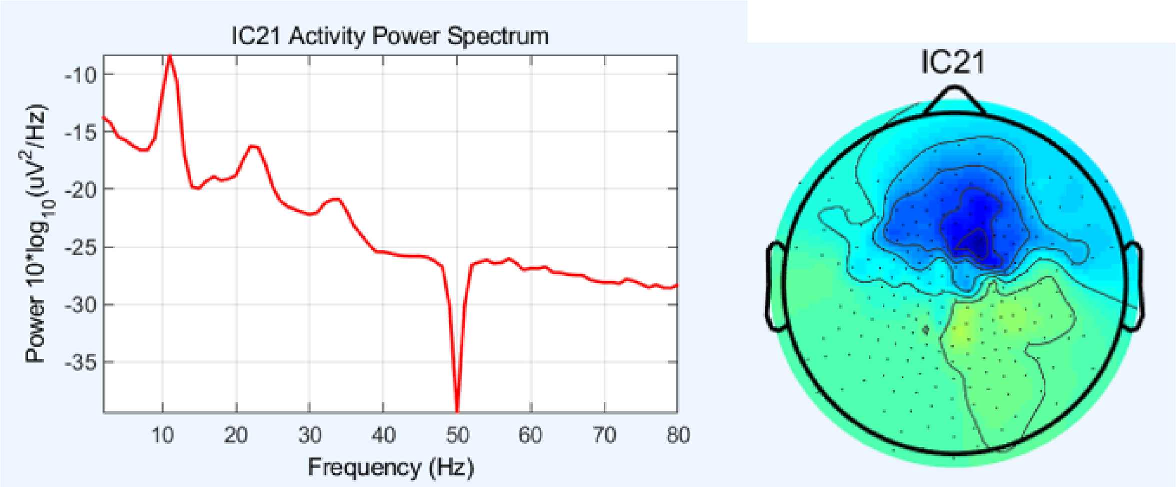


Figure 17. sample of Brain components in ICs
55,56

Eye Component:

The term "eye components" refers to the components that make up the eyes. Each retina (the portion of the eye that detects incoming light) generates an electric field that may be adequately described as an "equivalent current dipole" (ECD). Eye movement is often divided into two components: vertical movement and horizontal movement. Additional eye components, such as diagonal directions, may be identified, although uncommon and vary in each experiment. The power spectrum will vary depending on the experiment and the individual. The majority of the power will be at frequencies below 5 Hz. The bellow plot illustrates, figure [18] clearly Scalp topographies suggest ECDS near and effects of eye blink represent in time series plot.

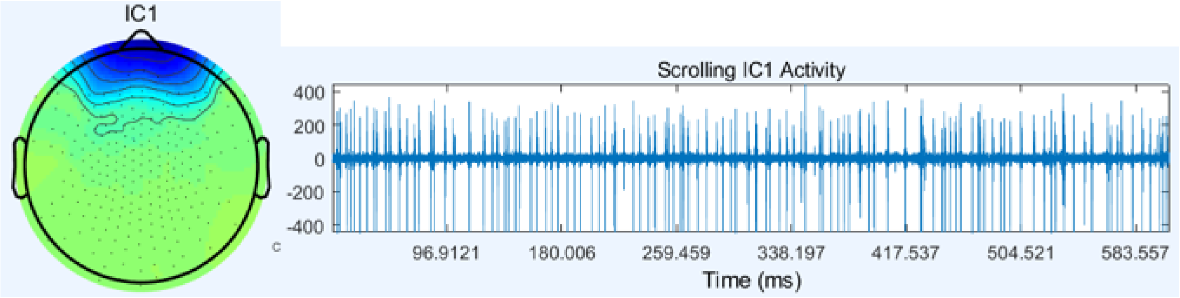


Figure 18. sample of eye components in IC

Muscle Component

Another component extracted from EEG signal and labeled by ICLabel is Muscle components.In the website, participants also have some structure to detect and label this component quickly. *Electromyography* is a term that refers to the electrical fields created by muscle activation (EMG). These components may still seem dipolar but will appear relatively shallow due to their lack of localization within the brain. A shallow dipole may be identified by the concentration of its scalp topography - the more concentrated, the shallower the dipole. The bellow plot presents some hints to detect muscle components by considering Scalp topographies. It could be disposal such as brain components but will be located outside the skull.

Heart Component

As well as muscles and eyes, the Heart is another factor effect on EEG signal, and this artifact is the same as other artifacts mentioned in previous sections. Some functional structure to detect this artifact from EEG in IC signals is a pattern formed by the Heart that is quite characteristic and is referred to as a QRS complex. These should occur about 1 HZ. Due to the distance of the Heart, the scalp map will resemble that of a very distant dipole and hence will seem practically linear. Here fig.number (19) there is a different plot, represents the topography pattern for EEG with Linear gradient scalp topography feature .

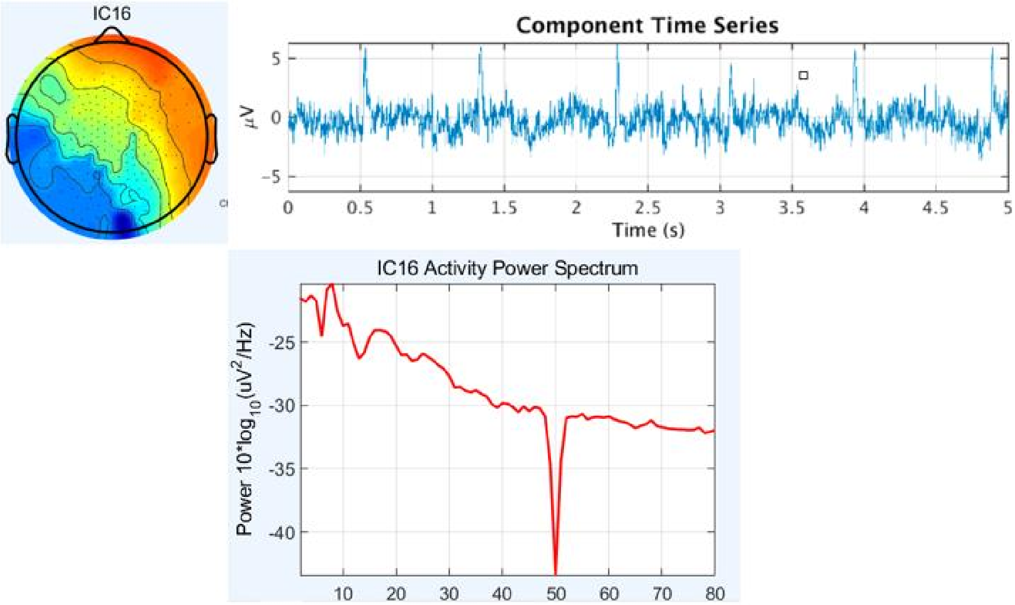


Figure 19. IC includes heart component

Line noise

This kind of noise commonly filter easily with notch filter. the main reason to create this kind of artifact is the is a component of the alternating current that powers practically all lighting fixtures and electrical devices. The frequency utilized may be 50 Hz or 60 Hz. In the following plot the clearest factor to detect this kind of noise has been mention as big peak of PSD is 50-60HZ .the the figp[20]illustrate the effects of this noice in Ics has been shown

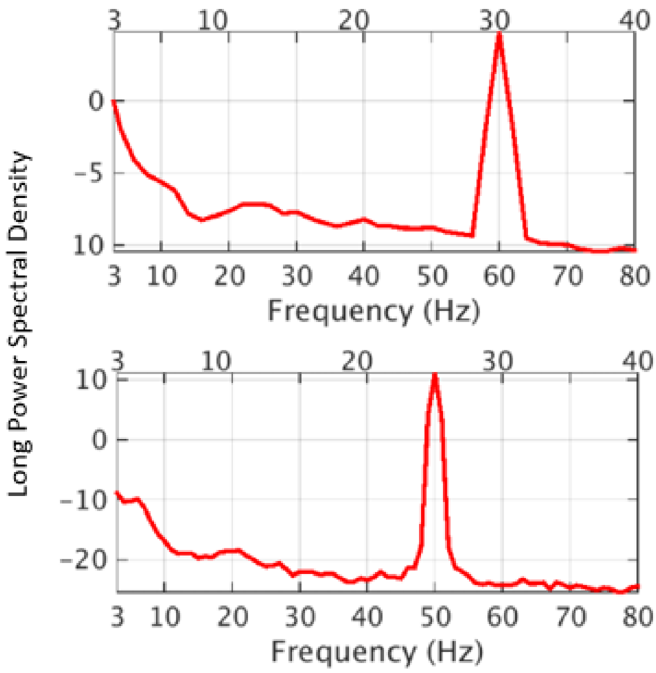


Figure 20. Effect of line noise in ICs

Channel Noise:

This artifact happens during the recording if a channel gets bumped and not affected by other channels. This artifact can see in topography as a single electrode with high energy, as mentioned in the following plot. Although they could appear the same as muscles in PSD, they have differences. As in very clear in the plot the scalp topography is only weighted on a single electrode which is the most significant pattern for this artifact figure [21] represent and toppgrapyhy map shown effect of channel noise .

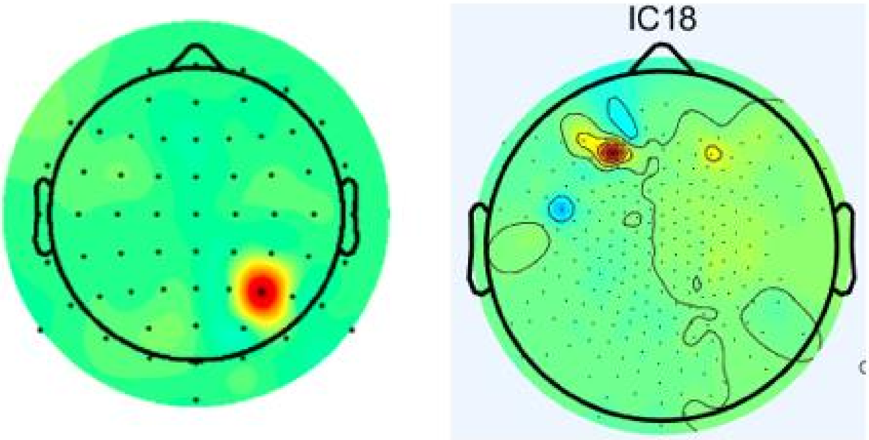


Figure 21. Effects of channel noise in topography of IC

Other noises :

Anything not matched with previous pattern mentioned so far located as OTHER category, This category include mixed signals. when participants not able to put result in any category they known it as OTHER . below Figure [22] is a sample of OTHER category.

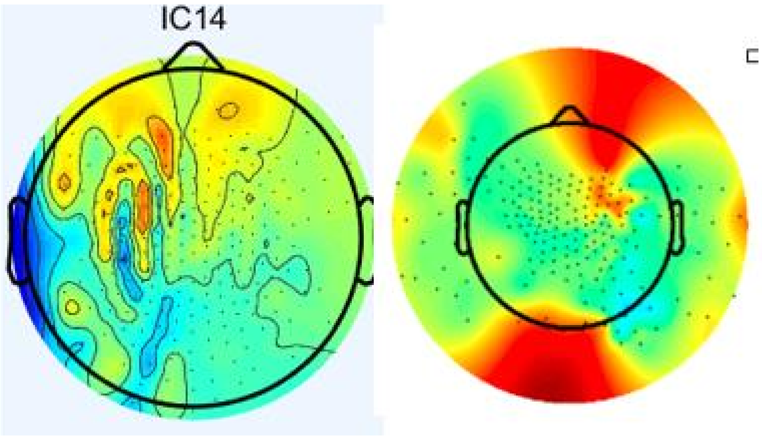


Figure 22. sample of other components in IC
57

So, according to this structure, participants train themselves to start labeling unlabeled ICs. They have three chances for labeling the ICs. first if they are sure their decision for labeling is entirely accurate, they can mark one category, the second option if they have hesitated between two or three labels for one IC, for example, participate not sure to label between muscle or brain, in this case, they can label this IC as both of them, the last option for the person does not have any idea about the component, in this case in they put this IC as? Category.

2.4.2 Model Validation

ICLabel Preprocessing:

After 250 people known as" labelers" had a contribution for labeling on 34,000 suggested labels on over 8000 ICs, as a result in final each ICs has almost four labels. It can be similar or different, and one crucial issue that needs consideration is that the knowledge of the labeler is not verified, and every person can play the role of labelers .it was one of the reasons there is the tutorial for labeling on the website. Nevertheless, this tutorial is still not trustable because participant knowledge after visiting the tutorial is unknown.

In this regard, some mathematical algorithms will be mentioned in the following parts, increasing accuracy and eliminating categories for every single label. To have unique labels for each ICs. These following methods are applied for ICs.

Crowd labeling(CL)

Crowd labeling Crowd labeling (CL)[58], also referred to as crowd consensus, is a standard method for feature extraction by visualizing and monitoring data. According to the specific features, the experts participate in labeling unable data. This manual method is known as crowd-labeling. The most straightforward crowd labeling technique gives the most frequently submitted label by majority voting for each occurrence. In the ICLabel project, crowd labeling is done by the user participating in labeling through the website and the tutorial mentioned in detail before. The output of crowd labeling is the various labeled and untreatable data. Also, doing this method is time-consuming. In this regard, this method needs more alternative ways to improve .another drawback of crowd labeling is that this method tends to put the data in one given category. Following that, classification algorithms try to estimate the class label for each occurrence in a dataset. Considering participant as a set of $\mathcal{U}$ indexed $u \in \{1, \ldots, U\}$ and their decisions(known as a vote) asset of
$\mathcal{D}$ indexed $d \in \{1, \ldots, D\}$ (D is the number of instances in this thesis D for Rome data is 2, Brain and artifacts.For the ICLabel projects set of seven categories mentioned before) producing a set of votes $\mathcal{V} = \{\upsilon \in \{1, \ldots, R\}| \, d \in \mathcal{D}, i \in \{1, \ldots, \ \}\}$[59] where U is the number of participants. The information of notation can be described in the table below.to sum up the output of the CL algorithm is an estimation of a single “true label” done by labelers.

Crowd Labeling- Latent Dirichlet Allocation (CL_LDA)

Latent Dirichlet allocation(LDA) is a generative probabilistic model for collecting discrete data such as EEG data[61].. LDA is a hierarchical Bayesian model with three levels in which each item in a collection is described as a finite mixture over an underlying set of topics. Each subject is thus represented as an endless mixture of topic probabilities derived from an underlying set of topic probabilities.

CL-LDA allows ICLabel to be[5] used in our data of CL excluding multiple-choice classification. The output of CL-LDA is the estimation of “true labels”(these reference labels are the desired output to train the ICLabel classifier) as a compositional vector with a vector of positive integers that add up to one. This vector includes the categories components as an output of ICLabel such as Brain, “Line Noise, Eye,” “Muscle,” and the other components mentioned before. CL-LDA efficiently estimates model.Parameters by taking to account the number of votes for each ICs, and the category of each ICs tend to the majority of the vote done by participants. An implementation of CL-LDA can be found at https://github.com/lucapton/crowd_labeling.

An overview of LDA and CL_LDA represent as a graphical shown in figure number 23and 24 and [6][5] :

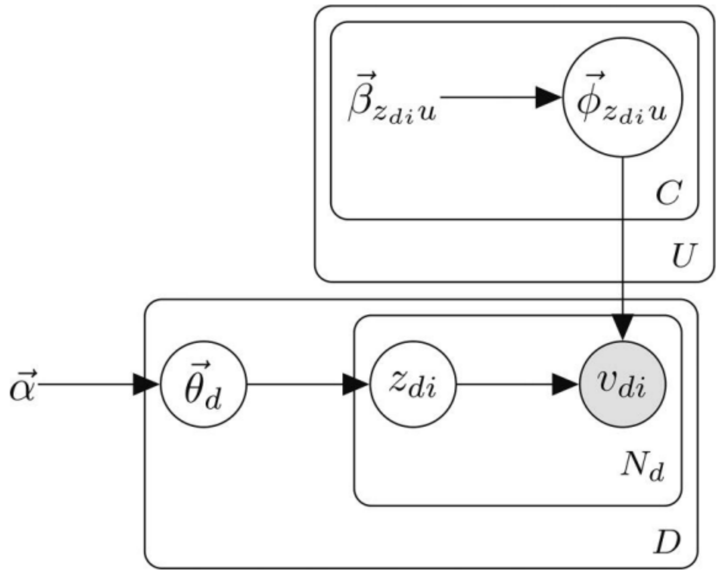


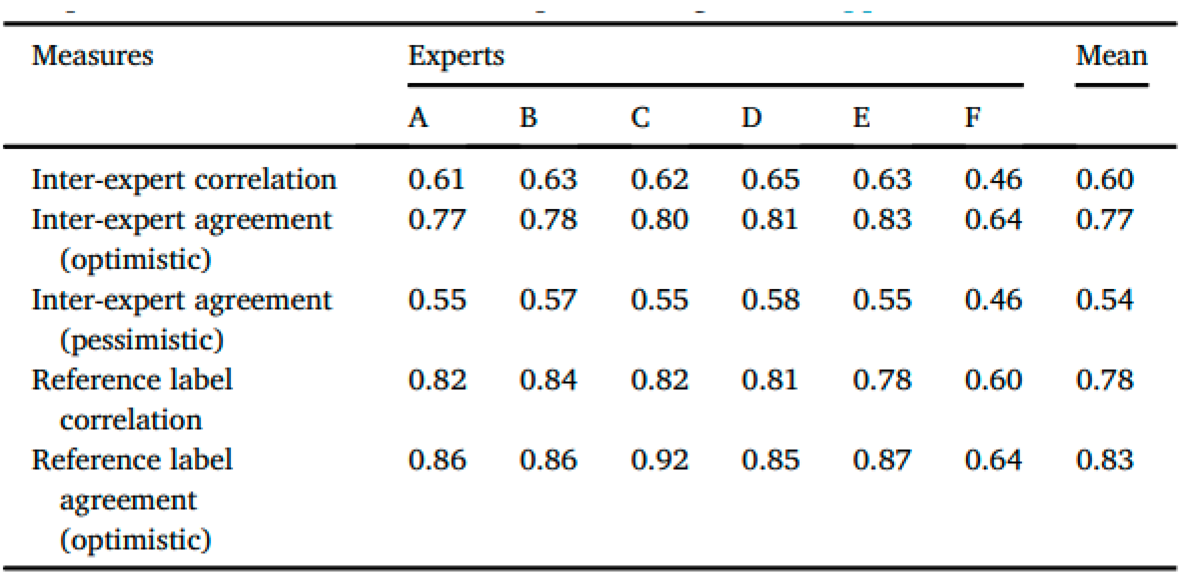
Figure 23. Graphical model for CL-LDA

Training set and ICLabel expert-labeled test set

After applying the CL_LDA algorithm in all data sets acquired by the website, the output of this algorithm is the ICs with more trustable labels. As a result, the number of ICs is less than before applying the CL_LDA algorithm and more trustable. These new ICs are known as the training set of the network. the training and validation between candidate models was done using the data labeled by everyone on the website.

Expert-labeled test set:

Because training data set several time used by participants and to avoid overlapoing other new data set not used in the training set so far is needed knonwn as expert-labeled ,ten different datasets from five different examinations with different recording environments, experimental paradigms, EEG amplifiers, electrode montages, preprocessing pipelines, and even different ICA algorithms, were used for the validation part. In contrast with the training set, these new 10 data sets were labeled by the six experts with known knowledge for labeling. These ten datasets .in this regard, exist ten data sets, from each data set extracts 13 ICs. These ICs were selected by sorting the ICs within a dataset by decreasing power and taking the union among the first five ICs, five more ICs at equally spaced intervals in descending order of source power (always including the weakest IC), and the seven ICs with highest selected class probability as per the ICLabel Beta EEGLAB plug-in for each IC category, (130 ICs in total). These 130 ICs given to experts for labeling, so these new ICs labels are trustable and accurate, known as ICLabel expert-labeled test set. The expert-labeled data was used when comparing the final version of ICLabel with other existing models like MARA and IC_MARC.In the table 3 ,first three rows illustrate how the ICs labeled by experts have similarities and correlations with other experts (experts A to F). in the last two rows present how well each expert's classifications match those of the CL-LDA-estimated reference labels. These measurements indicate that expert agreement is lower than one would assume, with an optimistic estimate of an expert agreement being just 77 percent on average. By contrast, the agreement between experts and CL-LDA-generated reference labels is always larger than or equal to the agreement between experts[1].



Measures	Experts						Mean
	A	B	C	D	E	F	
Inter-expert correlation	0.61	0.63	0.62	0.65	0.63	0.46	0.60
Inter-expert agreement (optimistic)	0.77	0.78	0.80	0.81	0.83	0.64	0.77
Inter-expert agreement (pessimistic)	0.55	0.57	0.55	0.58	0.55	0.46	0.54
Reference label correlation	0.82	0.84	0.82	0.81	0.78	0.60	0.78
Reference label agreement (optimistic)	0.86	0.86	0.92	0.85	0.87	0.64	0.83

Figure 24. Agreement between experts and between experts and CL_LDA

2.4.3 Automated electroencephalographic independent component classifier

A necessary and sometimes perplexing problem is that EEG recordings include activity from several sources other than the participant's brain. This EEG signal is not clean and does not include only the brain signals. It contains a broad range mixture of other parameters in undeniable. Task design, the environment of laboratory effect on this mixture components. It constantly monitors the difference in activity between two or more scalp electrodes.

On the other hand, as mentioned in the previous section. Electrical potentials generated by areas of locally coherent cortical field activity will reach not just the nearest EEG electrodes but practically the whole electrode montage to variable degrees. ICA, which discussed with detail such as: mathematical calculations, implementing and how to use in MATLAB with different categories has been discussed. In the Summary (ICA) may unmix and separate recorded EEG activity into maximum independent produced signals. By considering unmixed source signals have two features: spatially stationary (described before) and statistically independent of one another, and that the mixing occurs linearly, ICA estimates both a set of linear spatial filters that unmix the recorded signals and the source signals that are the products of that linear unmixing. In this research experiment, the signal was recorded by a multi-channel device, so the information was acquired by different electrolytes at a specific time. Some noises are also described with details such as Additional potentials emerge in the subject's eyes as they rotate, which project differentially to the scalp. Electromyogram graphic (EMG) activity linked with any muscle contractions that are powerful enough and close enough to the electrodes is additionally added to the EEG signals captured.

Additionally, electrocardiographic (ECG) signals from the participant's heart may be detected in scalp EEG recordings. The output of ICA includes IC, which is not labeled to include any situation of brain component. In fact, ICLabel is a new approach[60]. The ICLabel project provides improved classifications based on the qualities of an EEG IC classifier. IC classifier published in 2019 and implemented in EEG-lab. it is possible to classify these IC with ICLable(An automated solution to determining IC signal categories, referred to as IC classification).

Fig[16] illustrates the user interface of ICLabel project in MATLAB.

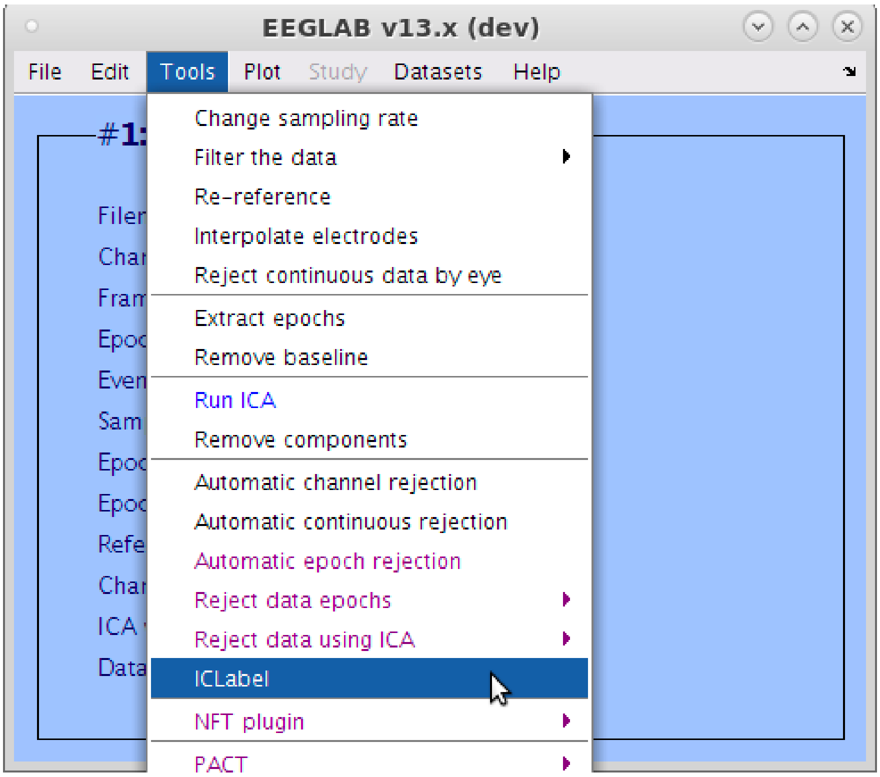


Figure 25. ICLabel project

Artificial neural network (ANN) of ICLabel:

GAN and CNN networks were used For the machine learning part of the ICLabel project. Graphically summarizes this ANN architecture of the ICLabel candidates represented in figure[27]. Three artificial neural networks (ANN) architectures were tested and compared with each other by the same CNN structure. The feature of each of these three are as follow:
1/ CNN network trained only with the labeled ICs and optimized an unweighted cross-entropy loss.
2/ CNN network trained only with the labeled ICs and optimized a weighted cross-entropy loss (WCNN) .This network acatly same as first ,only difference between cross-entropy type.
3/ The third classifier based on Semi-supervised learning adversarial networks(SSGAN. This network used noise with Zero mean, unit variance Gaussian noise characteristics because this noise worked better than constant noise . It should be mentioned in the GAN part noised used to create fake ICs feature , because it The outputs are feature vectors for each input
These three architectures were used and tested and the best performance selected as final ICLabel projects. The last version of the ICLabel projects only trained with CNN and label data without the GAN part. ICLabel, as it stands now, does not use GANS. So all that goes into ICLabel *during training* are the IC features(topo, PSD, and ACF) from labeled ICs. If we also want to consider the GAN training that was attempted, then the discriminator/classifier network (which now has an additional output class for "fake IC") is fed one of the following:
The output from the generator where ideally classifies this data as "fake."
Unlabeled data, where it ideally classifies the data as anything besides "fake."
Labeled data, where it ideally classifies the data as the correct class
All three of the above types of input have the exact feature sets/sizes to be fed into the classifier in the same way. The general structure is that each input type (topo, PSD, ACF) goes into its own initial set of layers. The results of these layers are then concatenated and go through a final layer. Here[60] contains the code to train the network.

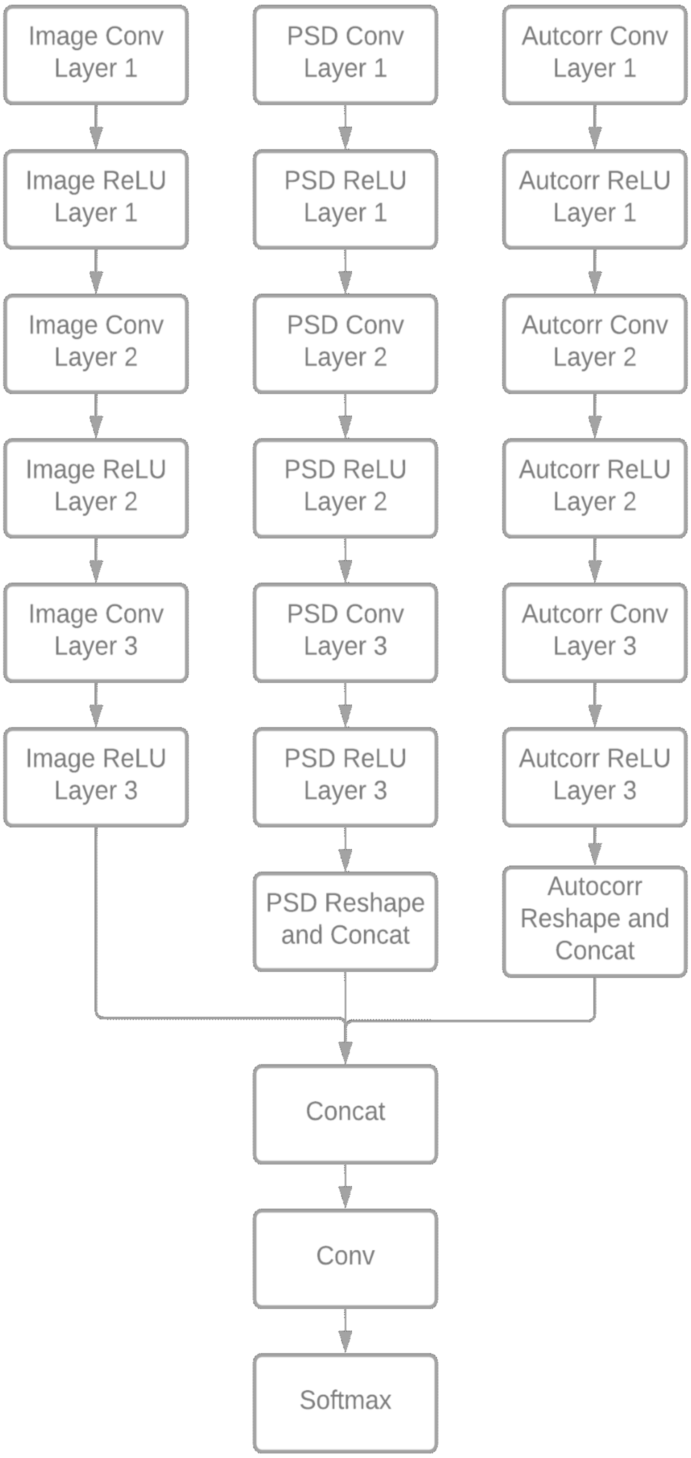


Figure 26. the graphical representation of machine learning part of ICLabel project. the input data for the first layer is the IC time course for the second layer the input data is PSD and for the last layer the input data is the IC topography

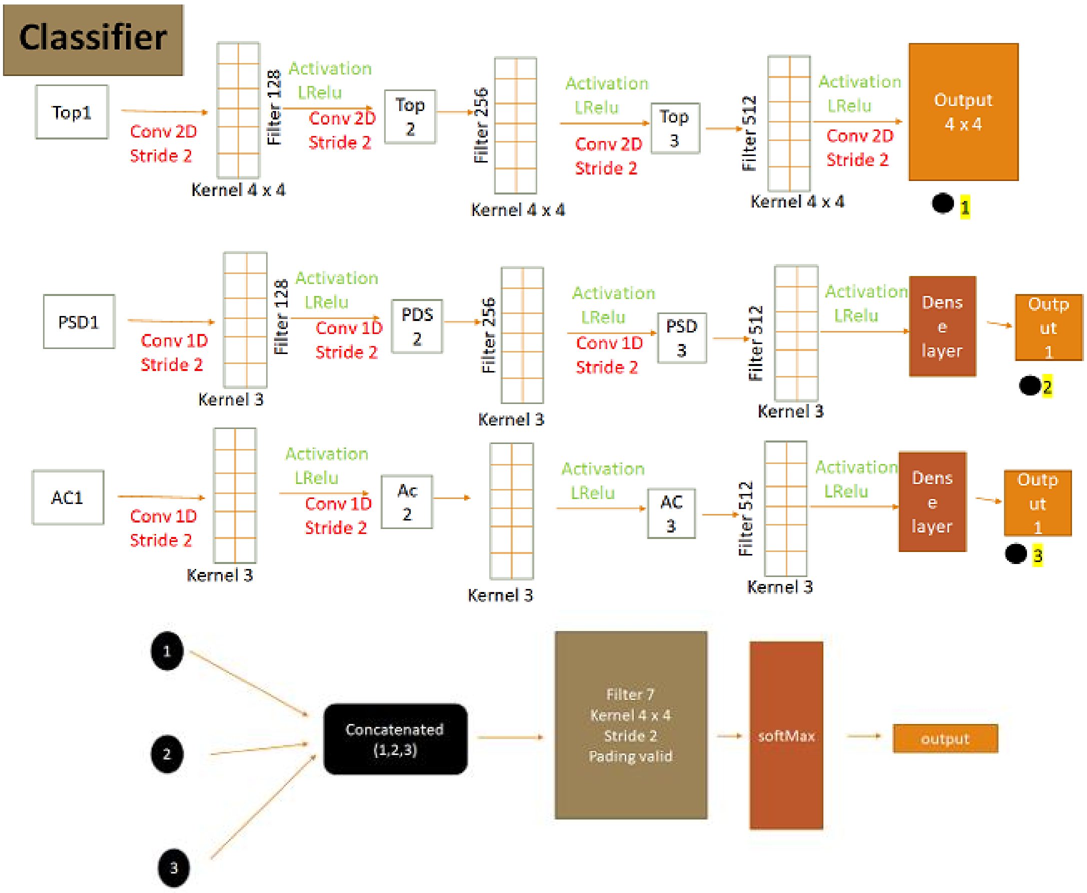


Figure 28. Graphical CNN model with the size of filter and kernels

After running the ICLabel with EEGLAB in MATLAB, the output showed as fig[29]. Each number above the topo plot represents the number of ICs, and the word below the topo plot represents the final output of ICLabel for that specific ICs for example, in the IC number 15, the final decision is Brain with more than 99% brain components. Although there are seven different components, as mentioned before in this IC, 99% of all these seven categories belong to the brain. Another example is IC number 16, and the final decision is Heart with 80% of Heart component. It means ICs number 16 maybe contain, as an example, 10% brain inside, but the majority belong to Heart artifact, so the final decision in Heart. In other words, the final decision belongs to the category having most proportion components inside the IC.By considering IC number of 17 correspond to eyes and most of the energy in the front of the head and IC number 18 channel noise exactly sam as in tutorial has been mentioned. In this regard by studying tutorial participant can make more accurate decision

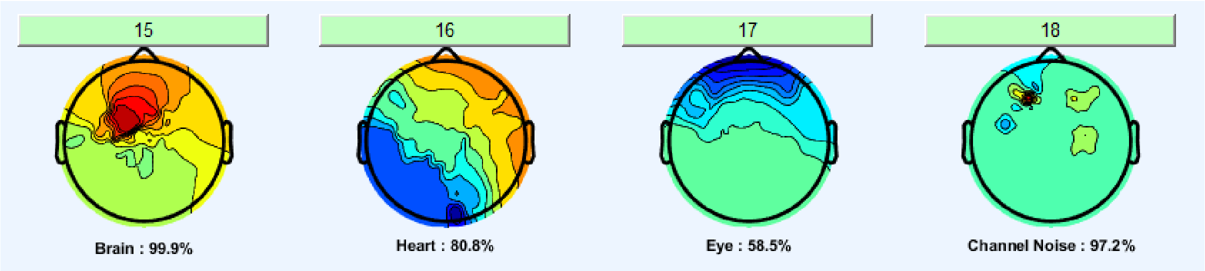


Figure 29. Example of ICLabel final out put

Following the Next step, by clicking on each button (number of IC), for example, by clicking on number 15, another window appears in figure number 31 with more information about that specific IC illustrate which included topo plot, the proportion of each seven class category (here is 99% brain) and IC activity power spectrum.

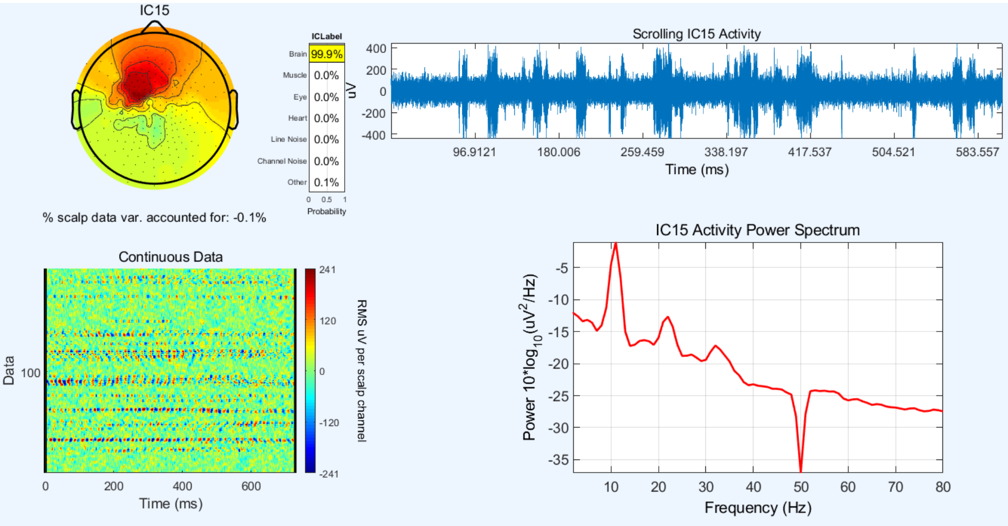


Figure 30. Information of each ICs in the output of ICLabel

2.5 Implementation of ICA and PCA in Matlab

Both PCA and ICA are implemented as functions in this package, PCA transforms multidimensional data into singular vectors corresponding to a subset of its biggest singular values. This technique essentially decomposes the input single into orthogonal components in the direction of the data's greatest variation. PCA is often used in applications involving dimension reduction.

In ICA, multidimensional data is decomposed into maximally independent components (kurtosis and negentropy, in this package). In reality, ICA often reveals disconnected underlying patterns. ICA signals do not always correlate to the directions of greatest variance; rather, they exhibit the highest degree of statistical independence.

Although the two techniques seem to be comparable, they accomplish distinct purposes. PCA is often used to compress data, i.e., dimension reduction. At the same time, ICA attempts to separate information by changing the input space into a basis that is maximally independent of the input space. Both procedures involve autoscaling the input data, i.e., subtracting each column's mean and dividing by its standard deviation. Here is the reference file for downloading these two:

https://it.mathworks.com/matlabcentral/fileexchange/38300-pca-and-ica-package

IC selection, is the application designed in this theses . The main aim of this aplication contains better illustration of all ICs in a single map and simplify the comparison between Rome and ICLable ‘s output. Also, using this app we can visualize IC’s signal and topography map fast and easy for all subjects.more details about this application will be describe in next chapter .

Chapter 3

3.Results and Discussion

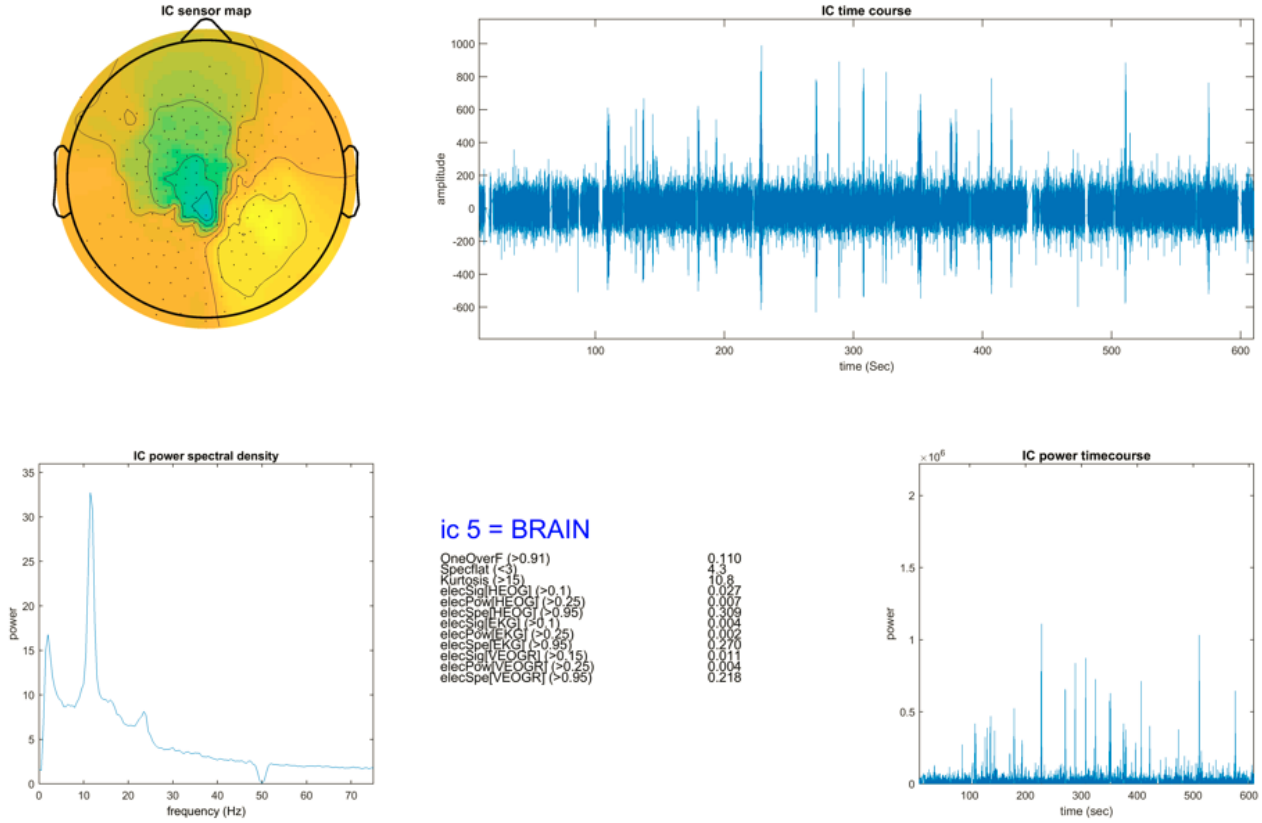
3.1 Result of semi-automatic algorithm

After running the 31 raw data with the semi-automatic algorithm, the output of each subject has several ICs with different labels, approximately 50 up to 130 ICs for each subject. The difference between the number of IC in each issue is not essential. The point is to take into consideration is the existing ICs labeled. As mentioned in chapter two, the only possible label for this method is Brain or Artifact (not Brain). All the ICs labeled controlled by the experts in Cosynclab as well they designed this algorithm. Some labeled corrected and changed the final label. This decision was made from the expert's knowledge and by observing topo plot time, criteria values, IC time course, IC power time course, IC power Spectral density. The subject's output with "ID: HM_102478" has been evaluated.

- **Brain detected as Brain**(correct decision of the semi-automatic algorithm): The figure[31] showed the ICs labeled as Brain, and experts confirm this is the right decision of the algorithm. According to the criteria, all the values match the requirements, the topo plot has bipolar representation, and a clear peak can be observed in IC power spectral density.
- **Artifact corrected as Brain** (output labeled repaired by experts): the figure[32] illustrates the algorithm that classified IC number 27 as an artifact, but experts converted this label and changed it to Brain. The algorithm ranked this IC as an artifact because one criterion is not matched (elecSig[HEOG]) as represented in figure[]. But the expert changed to Brain according to their knowledge.
- **Artifact detected as Artifact** (correct decision of the semi-automatic algorithm): figure[33] illustrate the ICs number 40 labeled as Artifact, and experts agreed with this label.

This single subject has 51 ICs, 14 of these 51 labeled as the brain. Of these 14 brain ICs, four were labeled as artifacts, but experts corrected them as an artifact, ICs numbers 27-37-45-51.

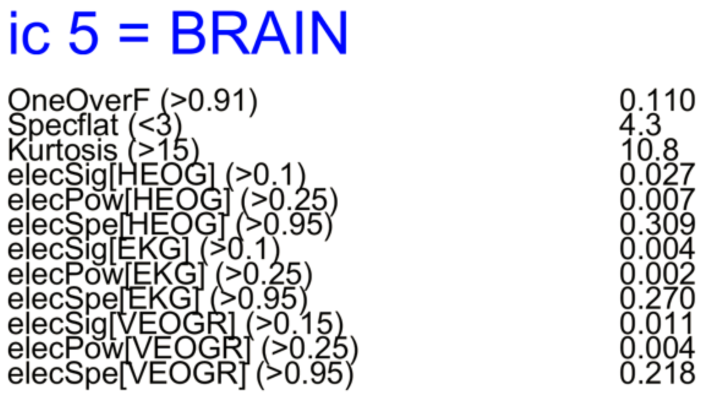
There are no ICs labeled as brain, and experts changed the label as an artifact. In the below figure, three types of labeled ICs can be see.



ic 5 = BRAIN

OneOverF (>0.91)	0.110
Specflat (<3)	4.3
Kurtosis (>15)	10.8
elecSig[HEOG] (>0.1)	0.027
elecPow[HEOG] (>0.25)	0.007
elecSpe[HEOG] (>0.95)	0.309
elecSig[EKG] (>0.1)	0.004
elecPow[EKG] (>0.25)	0.002
elecSpe[EKG] (>0.95)	0.270
elecSig[VEOGR] (>0.15)	0.011
elecPow[VEOGR] (>0.25)	0.004
elecSpe[VEOGR] (>0.95)	0.218

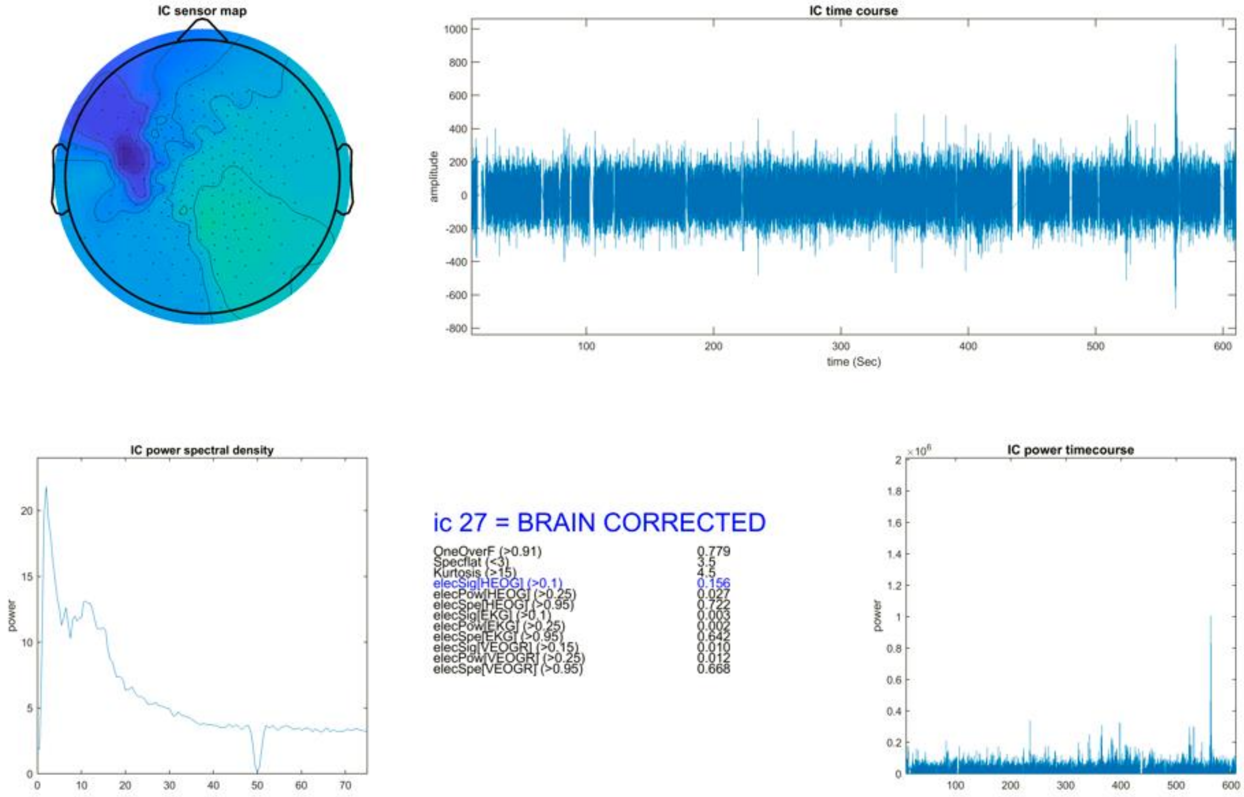
Figure 31.



ic 5 = BRAIN

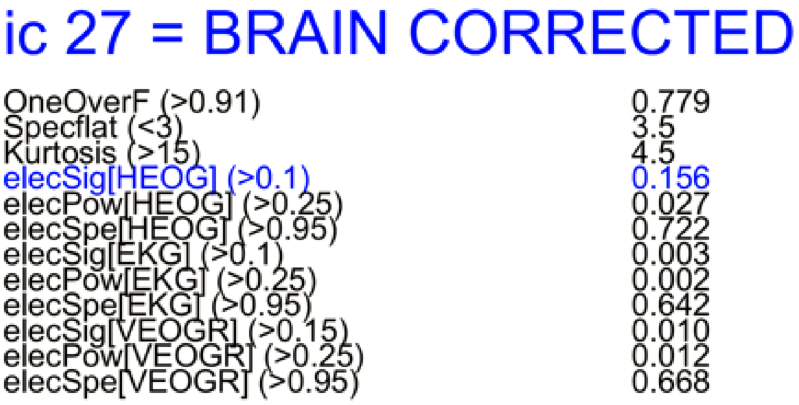
OneOverF (>0.91)	0.110
Specflat (<3)	4.3
Kurtosis (>15)	10.8
elecSig[HEOG] (>0.1)	0.027
elecPow[HEOG] (>0.25)	0.007
elecSpe[HEOG] (>0.95)	0.309
elecSig[EKG] (>0.1)	0.004
elecPow[EKG] (>0.25)	0.002
elecSpe[EKG] (>0.95)	0.270
elecSig[VEOGR] (>0.15)	0.011
elecPow[VEOGR] (>0.25)	0.004
elecSpe[VEOGR] (>0.95)	0.218

Figure 32. IC number 5 with Brain label



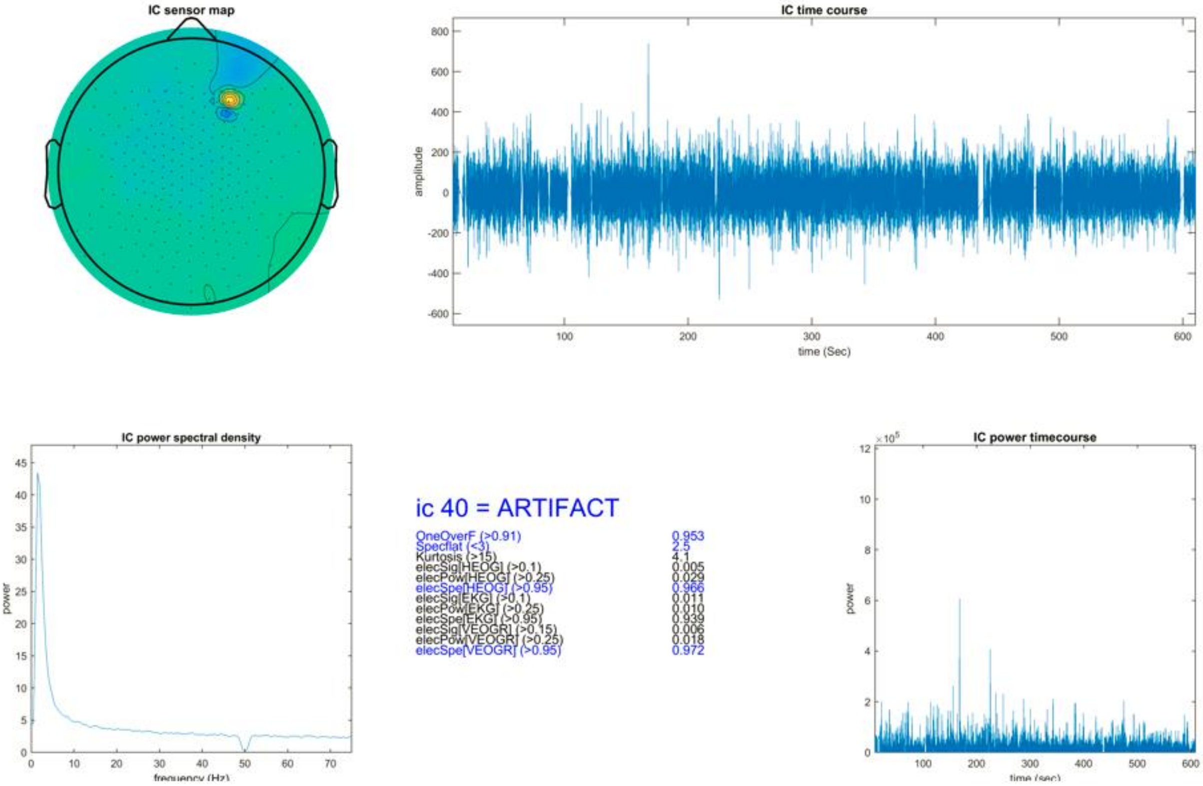
```
ic 27 = BRAIN CORRECTED

OneOverF (>0.91)          0.779
Specflat (<3)             3.5
Kurtosis (>15)            4.5
elecSig[HEOG] (>0.1)      0.156
elecPow[HEOG] (>0.25)     0.027
elecSpe[HEOG] (>0.95)     0.722
elecSig[EKG] (>0.1)       0.003
elecPow[EKG] (>0.25)      0.002
elecSpe[EKG] (>0.95)      0.642
elecSig[VEOGR] (>0.15)    0.010
elecPow[VEOGR] (>0.25)    0.012
elecSpe[VEOGR] (>0.95)    0.668
```

Figure 33.



```
ic 27 = BRAIN CORRECTED

OneOverF (>0.91)          0.779
Specflat (<3)             3.5
Kurtosis (>15)            4.5
elecSig[HEOG] (>0.1)      0.156
elecPow[HEOG] (>0.25)     0.027
elecSpe[HEOG] (>0.95)     0.722
elecSig[EKG] (>0.1)       0.003
elecPow[EKG] (>0.25)      0.002
elecSpe[EKG] (>0.95)      0.642
elecSig[VEOGR] (>0.15)    0.010
elecPow[VEOGR] (>0.25)    0.012
elecSpe[VEOGR] (>0.95)    0.668
```

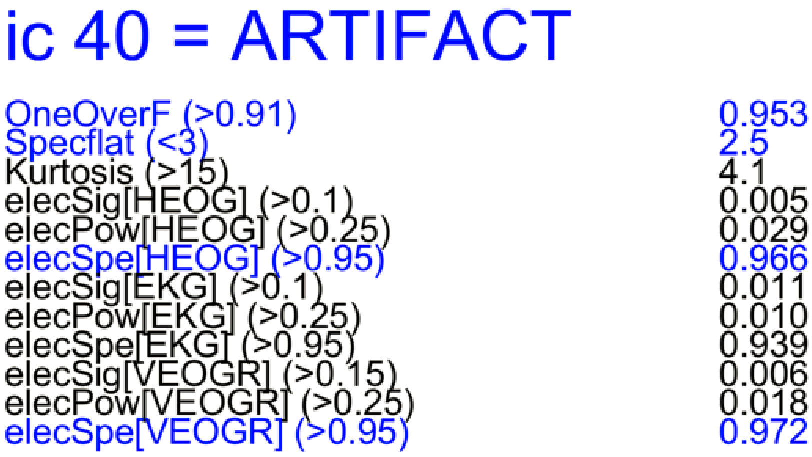
Figure 34. IC number 27 labeled artifact as output of algorithm and changed to Brain by experts



ic 40 = ARTIFACT

OneOverF (>0.91)	0.953
Specflat (<3)	2.5
Kurtosis (>15)	4.1
elecSig[HEOG] (>0.1)	0.005
elecPow[HEOG] (>0.25)	0.029
elecSpe[HEOG] (>0.95)	0.966
elecSig[EKG] (>0.1)	0.011
elecPow[EKG] (>0.25)	0.010
elecSpe[EKG] (>0.95)	0.939
elecSig[VEOGR] (>0.15)	0.006
elecPow[VEOGR] (>0.25)	0.018
elecSpe[VEOGR] (>0.95)	0.972

Figure 35.



ic 40 = ARTIFACT

OneOverF (>0.91)	0.953
Specflat (<3)	2.5
Kurtosis (>15)	4.1
elecSig[HEOG] (>0.1)	0.005
elecPow[HEOG] (>0.25)	0.029
elecSpe[HEOG] (>0.95)	0.966
elecSig[EKG] (>0.1)	0.011
elecPow[EKG] (>0.25)	0.010
elecSpe[EKG] (>0.95)	0.939
elecSig[VEOGR] (>0.15)	0.006
elecPow[VEOGR] (>0.25)	0.018
elecSpe[VEOGR] (>0.95)	0.972

Figure 36. IC number 40 labeled as artifact and experts has same idea

3.2 single-subject results from ICLabel

Once again, all the raw data were run by ICLabel in the EEGlab. In contrast with the semi-automatic algorithm, the output of this algorithm is a set of several classes, as mentioned before. Also, in ICLabel, there are two kinds of production as output

1. 1/ **class one**: a list containing the combination of 7 proportions
2. 2/ **class two:** the final output as the main result. The second class is the components having the highest proportion in class one.in figure number(37) both classes presented. the left plot illustrates the output of class two and the tight one shows the class one, for the IC number two in the subject ID: HM_102478.

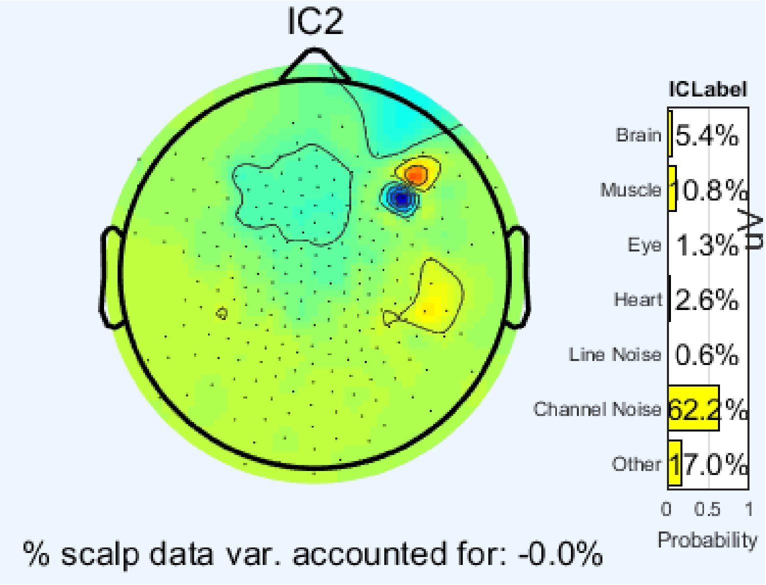


Figure 37. The output of ICLabel

According to the above figure, in the subject ID: HM_102478, the final label for IC number two is considered as Chanel noise with 62%, although there are 5.4% brain components inside these ICs simultaneously. As a result, the final label belongs to the group with more proportion. Observing other categories with different proportions could be one the advantages using the ICLabel. And we will use this proportion for more investigation in following sections.

3.3 Single subject results

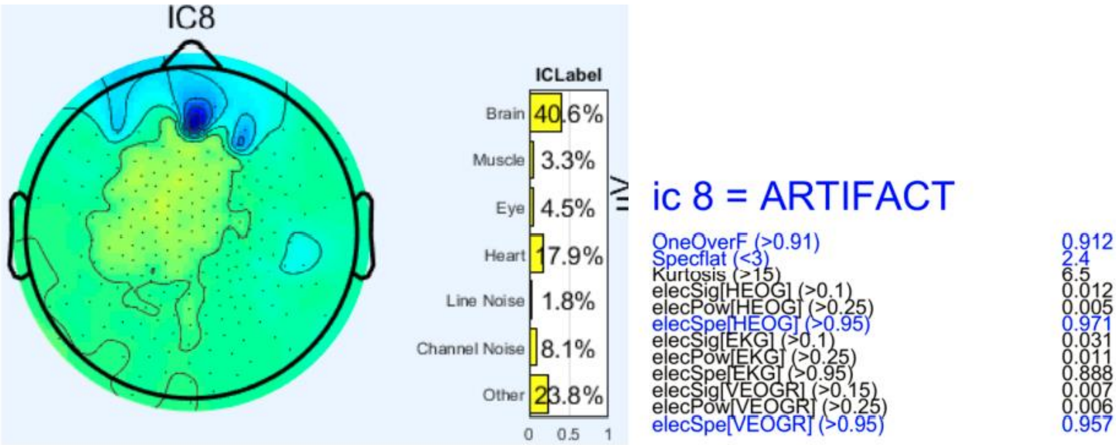
Comparing the results between ICLabel and semi-automatic algorithm for single subject ID The first aim of this thesis is to compare semi-automatic algorithm accuracy with possible and similar methods. After evaluating Mara and ICLabel we decided to compare the semi-automatic algorithm with ICLabel First, we showed all 51 ICs outputs of the subject ID: HM_102478 single by single as ICLabel output and semi-automatic algorithm output and compare the results and final label.

(The final output of ICLabel vs final output of semi-automatic algorithm) the right plot are the output of ICLabel and the left is the output of semi-automatic algorithm.

By Visual analysis of these 51 following plots, we can consider the following configurations:

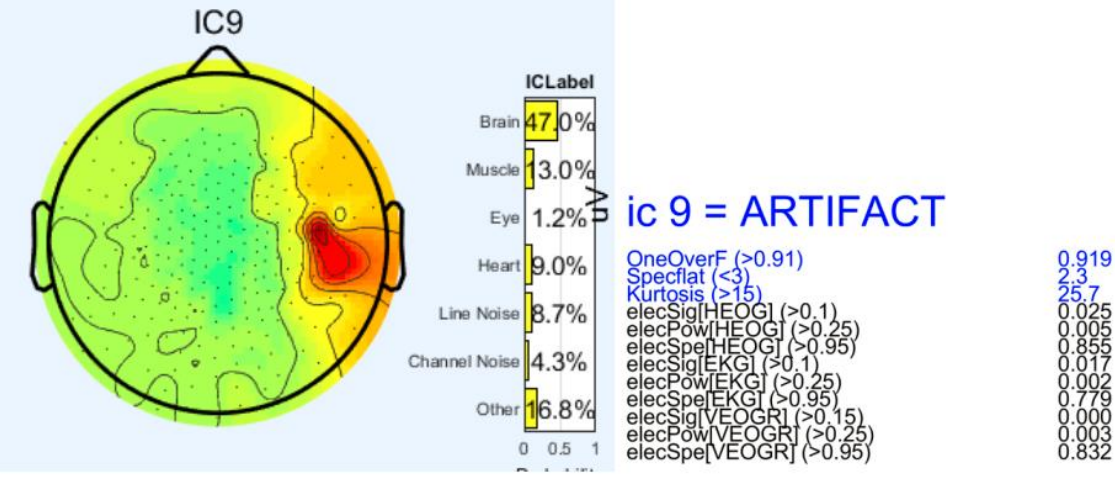
- **IC number 8**: semi-automatic algorithm + expert confirm that IC n.8 is an artifact because four criteria do not match. But ICLabel detect it as brain by 40%..figure number(38)
- **IC number 9**: ICLabel detect as a brain with 47% brain component, but semi-automatic algorithm + expert detect as an artifact because all three criteria (one over f, spectflat, and kurtosis did not match with requirement criteria).figure number(39)
- **IC number 27**: expert from semi-automatic corrected as Brain, but according to ICLabel, it has 67% brain components. The semi-automatic algorithm detects it first as an artifact because the criteria of elecSig[HEOG] is more than a defined threshold (0.156> 0.1).figure number(40)
- **IC number 33**: by 36 percent of the brain and 31 percentage of channel noise + 26 percentage of other detect as Brain in ICLabel algorithm, but semi-automatic Algorithm detects it as Artifact.figure number(41)
- **IC number 37**: in semi-automatic Expert corrected as Brain although the specflat below the criteria (2.7) ICLabel detected as brain by 40%.figure number(42)
- **IC number 45**: in ICLabel measured 40% brain and 38% other in total see as a brain. in semi-automatic, Expert corrected as a brain (Specflat 2.6 below the threshold) .figure number(43)
- **IC number 51:** ICLabel regarding as Heart 62% but in semi-automatic expert corrected as a brain. (Specflat 2.6 below the threshold).all these comparisons are available in appendix section. figure numbe(44)

By considering the output of ICLabel (automatically) which have been define by default.Some of the plots compared single by single .Here in following plots we only showed most relevant results for subject ID: HM_102478 .the left plots are the output of ICLable and on the right the output of semi-automatic ICs selection pipeline can be observed.(All 51 ICs comparions plots presented in appendix section)



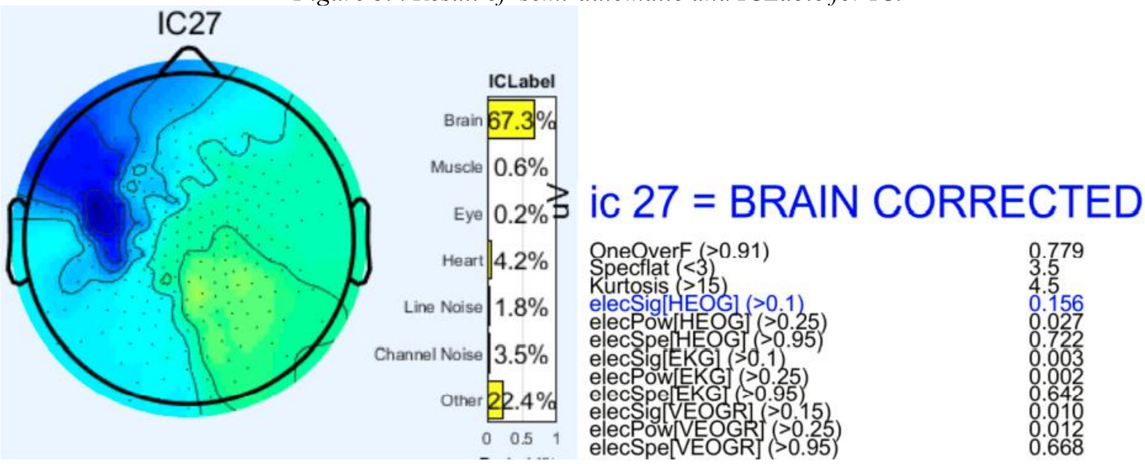
```
ic 8 = ARTIFACT
OneOverF (>0.91)          0.912
Specflat (<3)             2.4
Kurtosis (>15)            6.5
elecSig[HEOG] (>0.1)      0.012
elecPow[HEOG] (>0.25)     0.005
elecSpe[HEOG] (>0.95)     0.971
elecSig[EKG] (>0.1)       0.031
elecPow[EKG] (>0.25)      0.011
elecSpe[EKG] (>0.95)      0.888
elecSig[VEOGR] (>0.15)    0.007
elecPow[VEOGR] (>0.25)    0.006
elecSpe[VEOGR] (>0.95)    0.957
```

Figure 38. Result of semi-automatic and ICLable for IC8



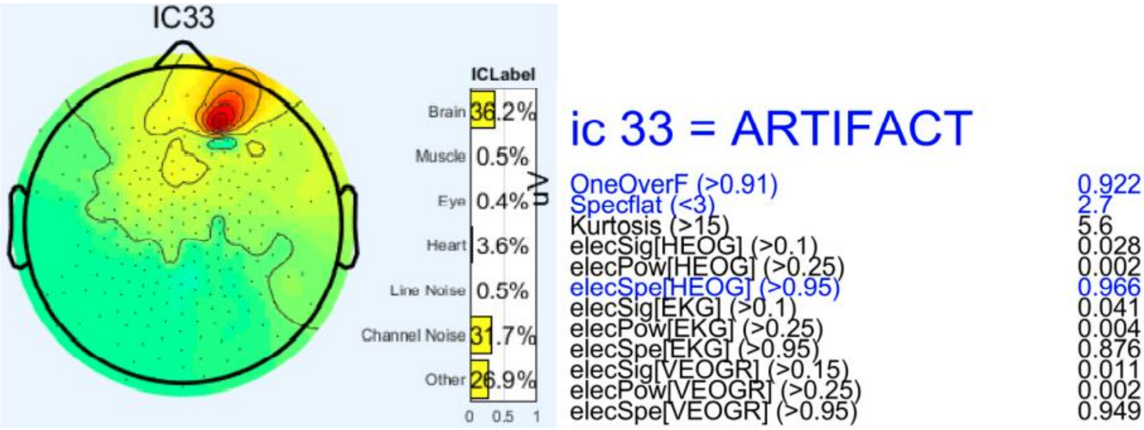
```
ic 9 = ARTIFACT
OneOverF (>0.91)          0.919
Specflat (<3)             2.3
Kurtosis (>15)            25.7
elecSig[HEOG] (>0.1)      0.025
elecPow[HEOG] (>0.25)     0.005
elecSpe[HEOG] (>0.95)     0.855
elecSig[EKG] (>0.1)       0.017
elecPow[EKG] (>0.25)      0.002
elecSpe[EKG] (>0.95)      0.779
elecSig[VEOGR] (>0.15)    0.000
elecPow[VEOGR] (>0.25)    0.003
elecSpe[VEOGR] (>0.95)    0.832
```

Figure 39. Result of semi-automatic and ICLable for IC9



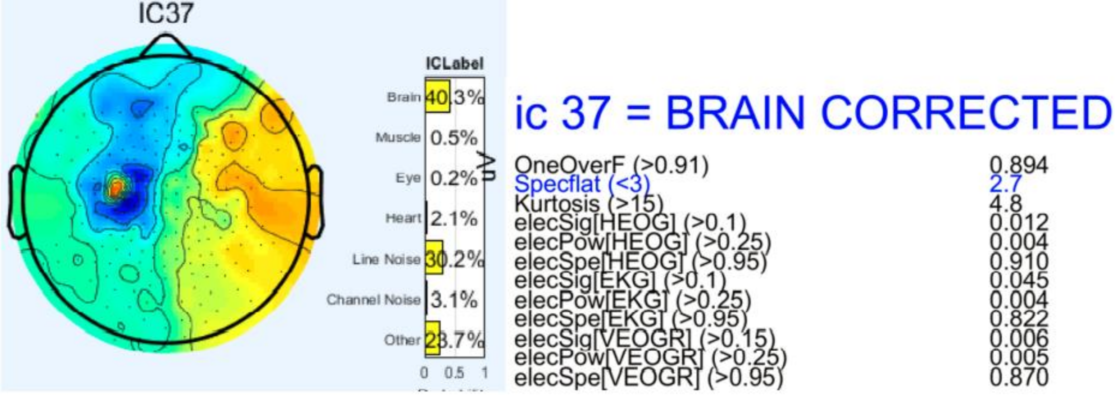
```
ic 27 = BRAIN CORRECTED
OneOverF (>0.91)          0.779
Specflat (<3)             3.5
Kurtosis (>15)            4.5
elecSig[HEOG] (>0.1)      0.156
elecPow[HEOG] (>0.25)     0.027
elecSpe[HEOG] (>0.95)     0.722
elecSig[EKG] (>0.1)       0.003
elecPow[EKG] (>0.25)      0.002
elecSpe[EKG] (>0.95)      0.642
elecSig[VEOGR] (>0.15)    0.010
elecPow[VEOGR] (>0.25)    0.012
elecSpe[VEOGR] (>0.95)    0.668
```

Figure 40. Result of semi-automatic and ICLable for IC27



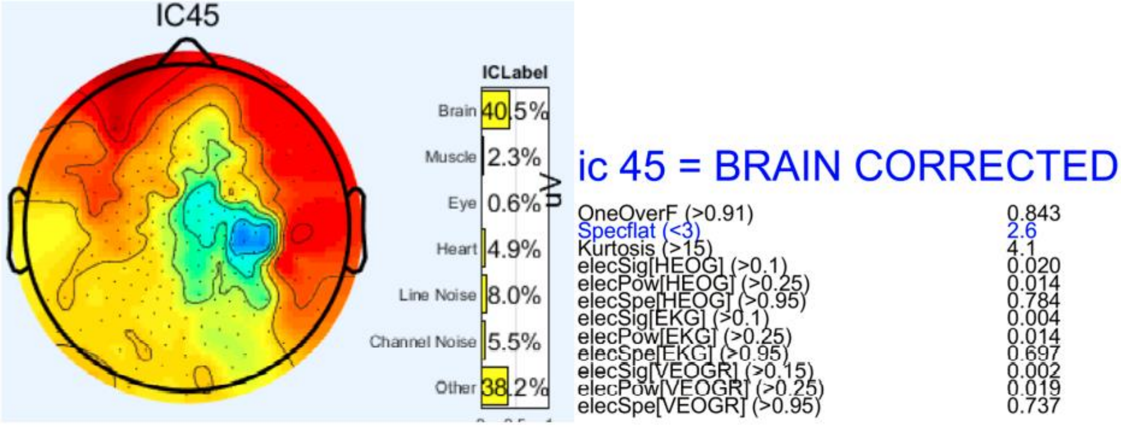
```
ic 33 = ARTIFACT
OneOverF (>0.91)             0.922
Specflat (<3)                2.7
Kurtosis (>15)               5.6
elecSig[HEOG] (>0.1)         0.028
elecPow[HEOG] (>0.25)        0.002
elecSpe[HEOG] (>0.95)        0.966
elecSig[EKG] (>0.1)          0.041
elecPow[EKG] (>0.25)         0.004
elecSpe[EKG] (>0.95)         0.876
elecSig[VEOGR] (>0.15)       0.011
elecPow[VEOGR] (>0.25)       0.002
elecSpe[VEOGR] (>0.95)       0.949
```

Figure 41. Result of semi-automatic and ICLable for IC33



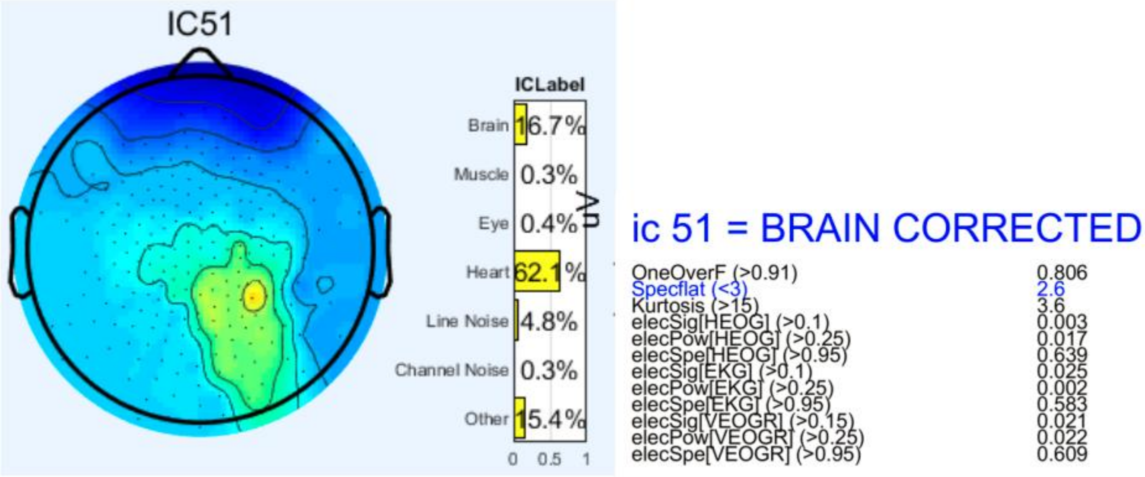
```
ic 37 = BRAIN CORRECTED
OneOverF (>0.91)             0.894
Specflat (<3)                2.7
Kurtosis (>15)               4.8
elecSig[HEOG] (>0.1)         0.012
elecPow[HEOG] (>0.25)        0.004
elecSpe[HEOG] (>0.95)        0.910
elecSig[EKG] (>0.1)          0.045
elecPow[EKG] (>0.25)         0.004
elecSpe[EKG] (>0.95)         0.822
elecSig[VEOGR] (>0.15)       0.006
elecPow[VEOGR] (>0.25)       0.005
elecSpe[VEOGR] (>0.95)       0.870
```

Figure 42. Result of semi-automatic and ICLable for IC37



```
ic 45 = BRAIN CORRECTED
OneOverF (>0.91)             0.843
Specflat (<3)                2.6
Kurtosis (>15)               4.1
elecSig[HEOG] (>0.1)         0.020
elecPow[HEOG] (>0.25)        0.014
elecSpe[HEOG] (>0.95)        0.784
elecSig[EKG] (>0.1)          0.004
elecPow[EKG] (>0.25)         0.014
elecSpe[EKG] (>0.95)         0.697
elecSig[VEOGR] (>0.15)       0.002
elecPow[VEOGR] (>0.25)       0.019
elecSpe[VEOGR] (>0.95)       0.737
```

Figure 43. Result of semi-automatic and ICLable for IC45



```
ic 51 = BRAIN CORRECTED
OneOverF (>0.91)          0.806
Specflat (<3)             2.6
Kurtosis (>15)            3.6
elecSig[HEOG] (>0.1)      0.003
elecPow[HEOG] (>0.25)     0.017
elecSpe[HEOG] (>0.95)     0.639
elecSig[EKG] (>0.1)       0.025
elecPow[EKG] (>0.25)      0.002
elecSpe[EKG] (>0.95)      0.583
elecSig[VEOGR] (>0.15)    0.021
elecPow[VEOGR] (>0.25)    0.022
elecSpe[VEOGR] (>0.95)    0.609
```

Figure 44. Result of semi-automatic and ICLable for IC51

3.4 Full dataset analysis

After the output of the two methods has been compared. We did this comparison between full datasets. 31 subjects with more than 2600 Ics. In the semi-automatic algorithm labeled 405 ICs as brain and the rest as an artifact. In figure [45], we illustrated these 405 ICs (labeled as brain in Rome)by considering the proportion of Brain in the ICLabel output The Y-axis represent number of ICs labeled as brain (in semi-automatic algorithm) and X-axes represent the proportion of brain components in the ICLabel output for each ICs. First we used an emprical thresholds of 40%. Then we investigated differents vlues when we used 80%.This reluts of this two thresholds will be presnet in following sections .

ICLabel threshold selection(40%)

As we have seen in ICLabel section, the output of ICLabel for each IC, illustrates seven IC measures in which there is a brain value existed in the IC. So, to defined an IC as a brain detected by ICLabel, we need to set a threshold value and by exceeding the brain value from this threshold, we can consider that IC as a brain reported by ICLabel. Here we have defined this threshold equal to ICP = 40 %.

For comparison between semi-automatic algorith, and ICLabel performance, we have visualized some statistics about the number of brains reported by each method. Most of the ICs labled as Brain in semi-automatic algorithm having more than 90% brain component in the output of ICLabel this a good matches between ICLabel and semi-automatic algorithm.

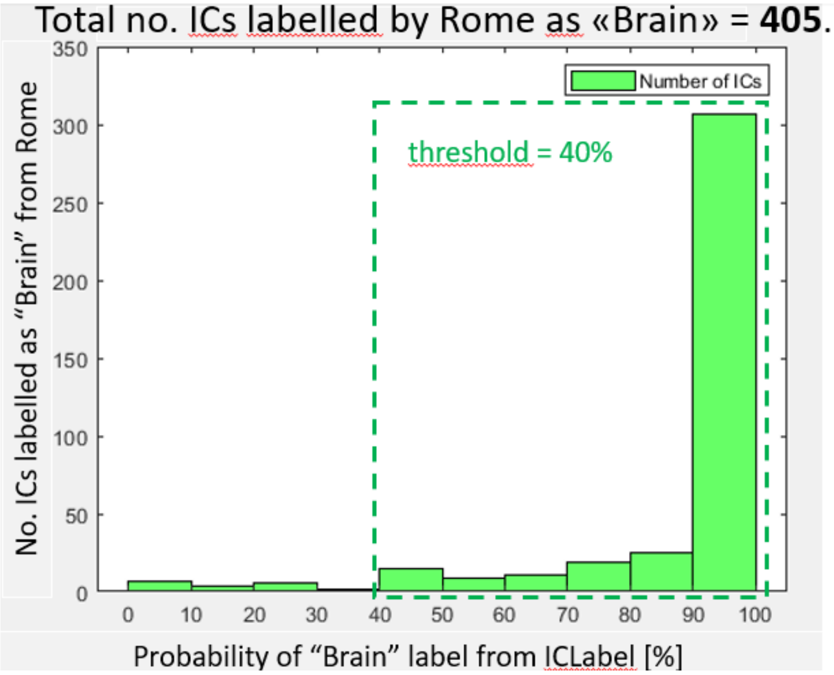


Figure 45. ICLabel threshold selection (40%)

In figure 78,we can observe that:

1. 40% is a good criteria to decide for brain IC
2. For lower this threshold, we do not gain a lot of ICs, but at the same time we probably increase the number of mis-matches
3. Setting a «brain detection threshold» (from ICLabel) to 40%(*) leads to 95% good "brain" matches this threshold has been set by visual inspection on all subjects. ICLabel uses the highest score to assign theclass

ICLabel threshold selection(80%)

1. In another comparision we changed ICLabel threshold from 40% to 80% . by observiion the result Using a threshold value for ICLabel equal to 80%, we found 81% good "brain" matches.

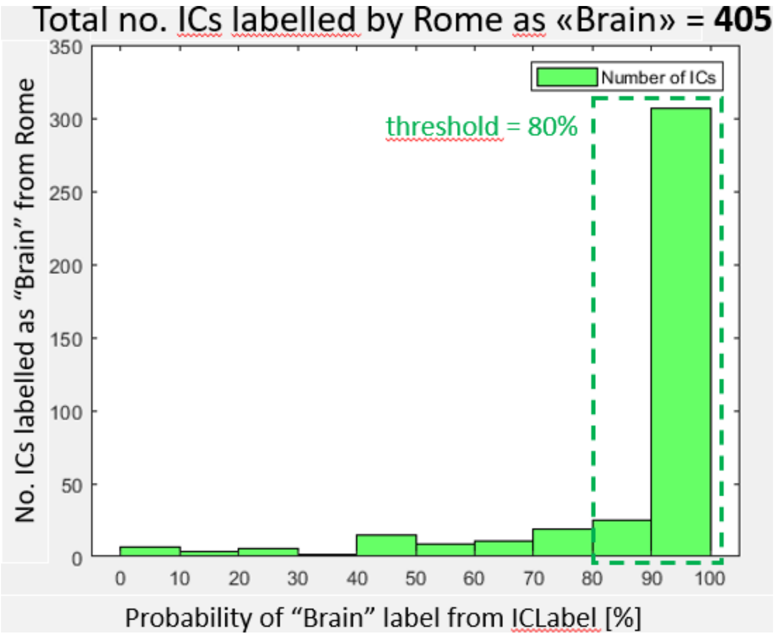


Figure 46. ICLabel threshold selection(80%)

Mis-matched impact

By using figure number (45), we decided to put the threshold as 40% for ICLabel . in this regard, with another plot, we can present how many ICs in ICLabel have more than 40% brain components in each ICs. According to the following comparison, we only consider the percentage of Brain in each ICs and the output is the ICs with more than 40% brain.

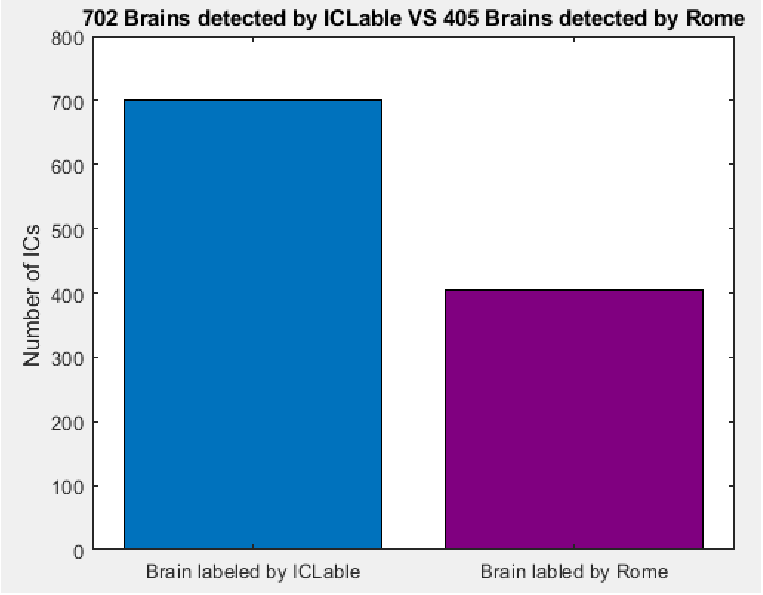


Figure 47. Mis-matches impact having 40% threshold p[38]

The Y-axes represent the number of ICs having more than 40% brain components in ICLabel and the number of ICs labeled as brain by semi-automatic algorithm. 702 ICs having more than 40% brain components. So there is big difference between the ICs labeled by semi-automatic and labeled by ICLabel can be see easily in this threshold (40%).
By changing the threshold 40% to 80% in the plot[47]it can be seen miss-matched changed sharply :

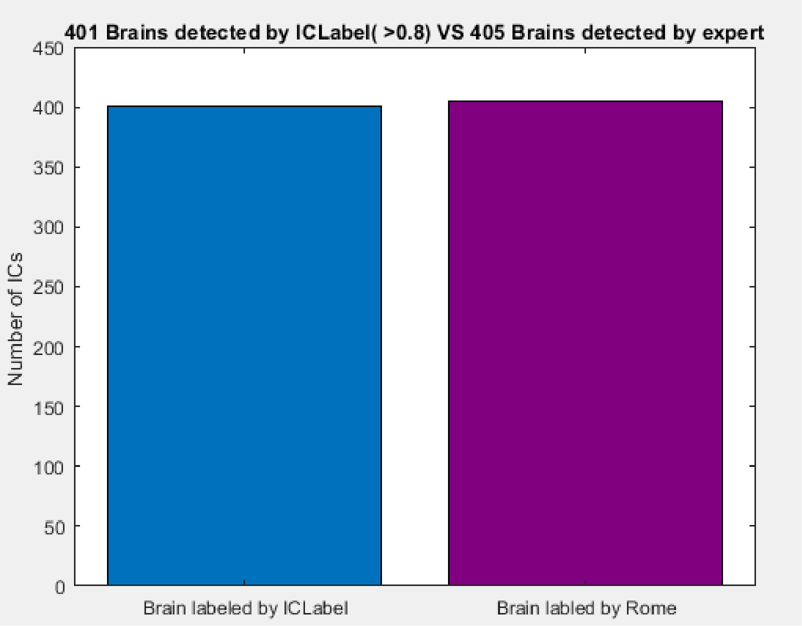


Figure 48. Mis-matches impact having 80% threshold

Comparing two methods by considering 40% thresholds

By considering 40% threshold, once again, we compared two methods single by single for each ICs in subject ID HM_102478. the results are as follow:

- **Good "brain" matches = 13 ICs** detected as Brain from both semi-automatic and ICLabel: 5-6-13-15-21-22-27-30-32-36-37-42-45
- **2 Mis-matches** detected as not brain in semi-automatic but consider as Brain in ICLable: 8-9. They were output from ICLabel as 40% and 47% "brain"
- **1 Mis-match** semi-automatic detects as Brain but ICLabel consider 16% brain components (IC number 51)

- **Good “non-brain” matches = 35 ICs** not brain from semi-automatic and have brain component below 40% in ICLabel

BRAIN ICs (semi-automatic) (good matches) with ICLabel in all 31 subjects having more than 40% brain in ICLabel

By three definitions, we designed another plot to illustrate and cover all the matches and mismatches of ICs. these three definitions are as follow:

- All colors are brain only by different Criteria (here 40%)
- Green dots: brain detect in semi-automatic and ICLabel
- Red dots: detect as <u>not Brain </u>in semi-automatic but having more than 40% components in ICLabel
- Blue dots: semi-automatic detect as <u>Brain </u>but in ICLabel having less than 40% brain components miss match blue and green.

we have illustrated each IC for all 31 subjects. Horizontal axis shows the IC number in each subject and vertical axis shows the subject number

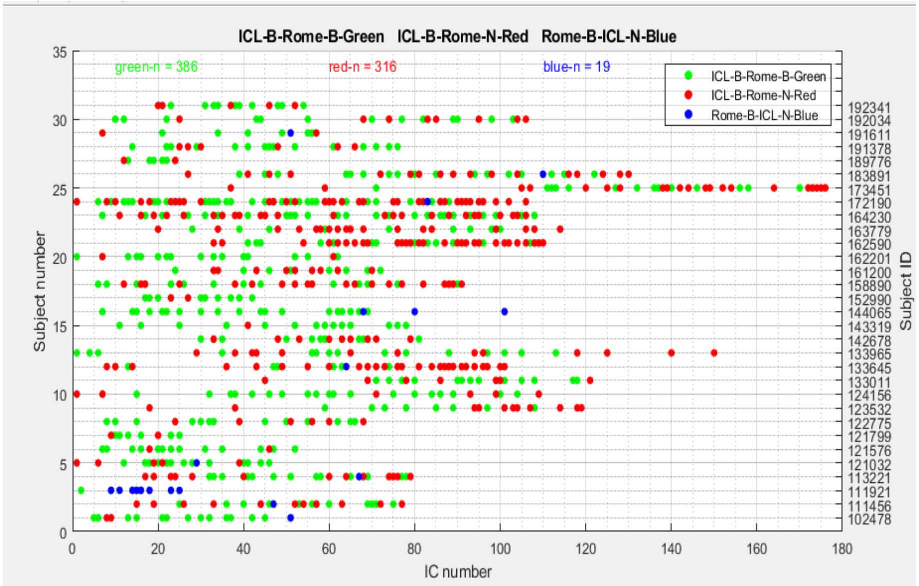


Figure 49. Joint representation of classifications from semi-automatic and ICLabel (all subjects, N=31)

Observations based on this plot:

- Good matches = 386 ICs having same label in both methods (presented by green dots)

- Brain from semi-automatic having brain components less than 40% from ICLabel includes 19 ICs
- 316 ICs,labeled as brain in ICLabel method but labeled as not brain in semi-automatic pipeline
- Most of blue dots located in subject ID 111921 some evaluations could be needed

3.5 IC selection Application

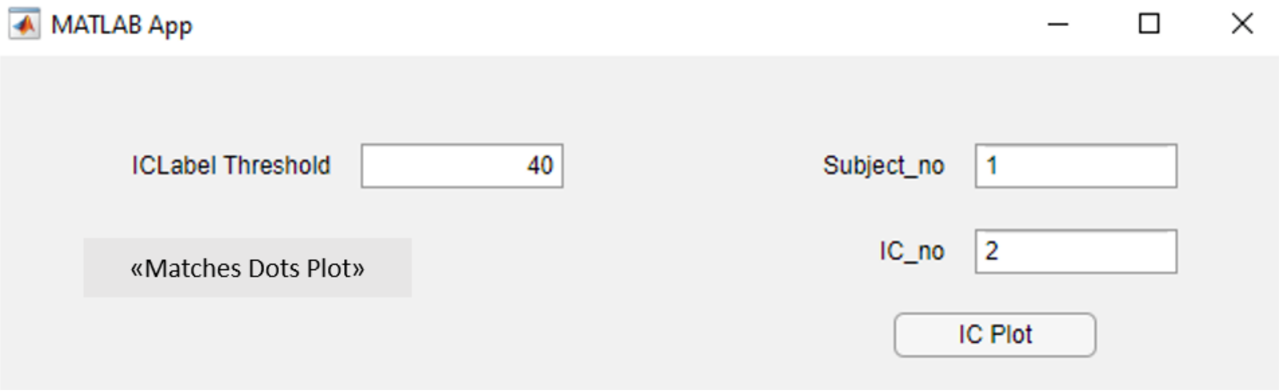
The application aims to help experts having better visualization for comparing two methods by changing thresholds.

After finding creating last following plots, we have made a GUI to give an option for user to change the thresholds and comparing the result. Then we developed an application which name is "IC_selection". First expert need to run data with two methods (with semi-automatic ICs selection pipeline and run data with ICLabel) separately. Then easily can monitor the difference number of brain ICs with different colors by using this application. In the following section the user manual and documents for installing will be describe.

There are two buttons with three options, In the left icons there is an option (ICLabel threshold) to change the thresholds and on the right, there are two options, the top which name is "subject_no" represent the subject id and the second one (right down) represent the IC number.

- By clicking the IC plot, the user can see the information related to the topographic map, PSD, fig number(52) illustrated IC plot.
- The second output is a plot with semi-automatic criteria and its threshold information. this plot plot (83) contains two dots. "Red and Black" and 12 columns. Every 12 columns represent one of the twelve criteria defined by semi-automatic. If the ICs values matched the criteria values, "red dot stays above of black dot(the except second Colum it was reversed). other views black dot located below the red dots .it can be see these twelve columns plus colored dot the ICs selected as the brain in semi-automatic and matched the criteria

In the figure [50]by clicking the IC plot the information related to topographic map, PSD, can be seen.as appears in Figure [53].The output of running this application is a combination of figures [47] and fig[46] with preferable thresholds. For example, here we set threshold to 40% and 80% and the plot appear as figure [51] and figure [52] by clicking on ICs map plot.



ICLabel Threshold 40

Subject_no 1

IC_no 2

Figure 50. Application interface

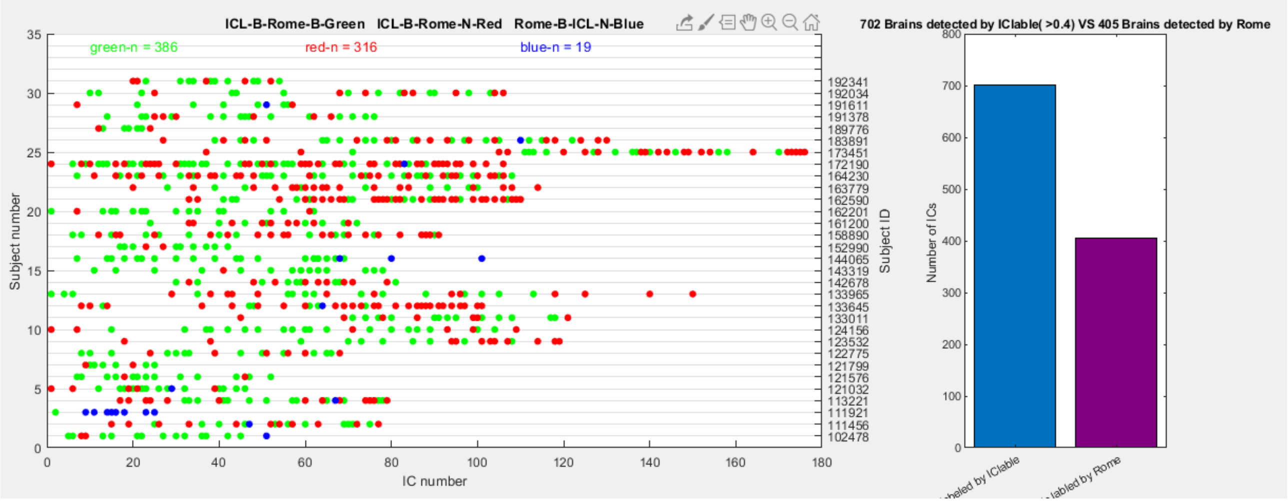


Figure 51. output of «Matches Dots Plot» button by 40% threshold

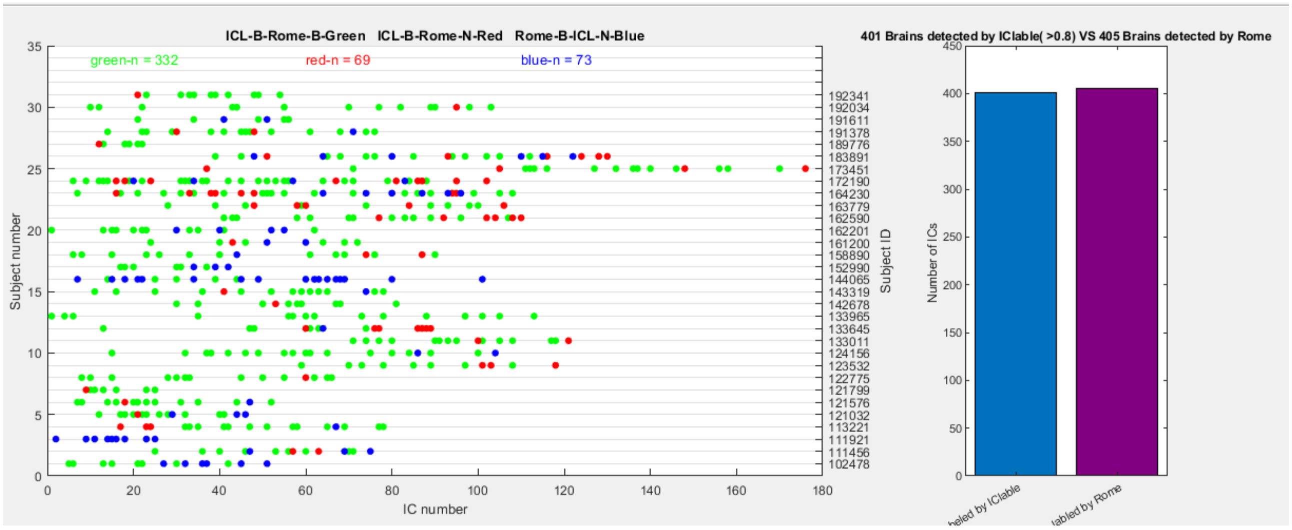


Figure 52. output of «Matches Dots Plot» button by 80% threshold

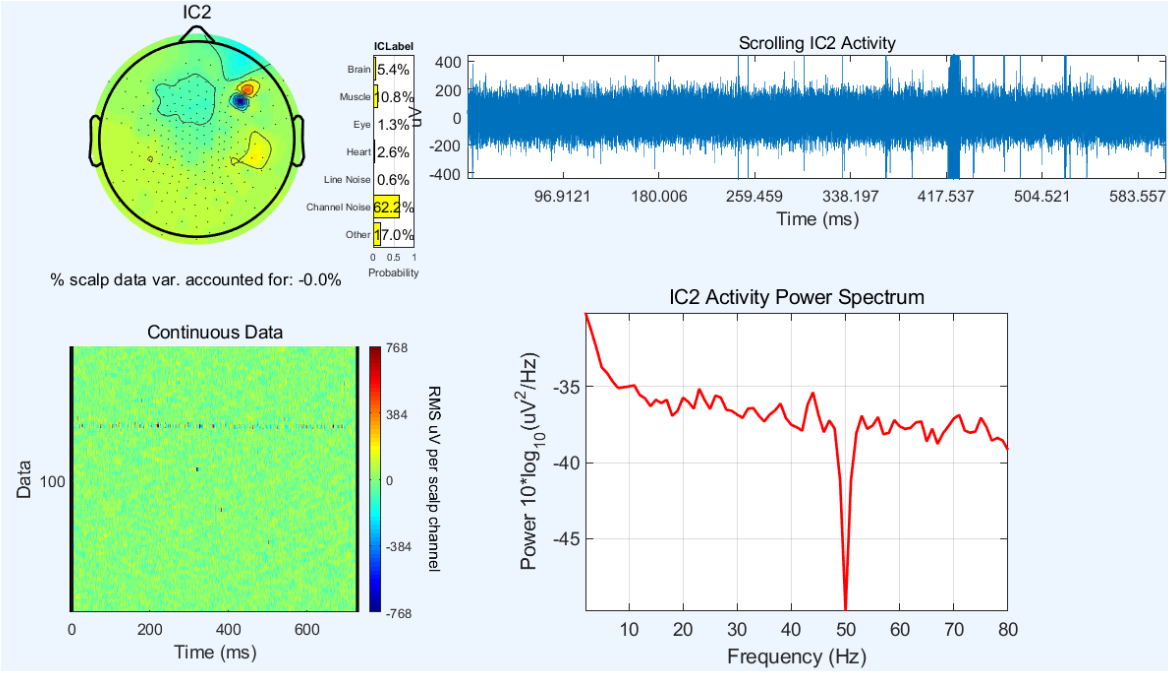
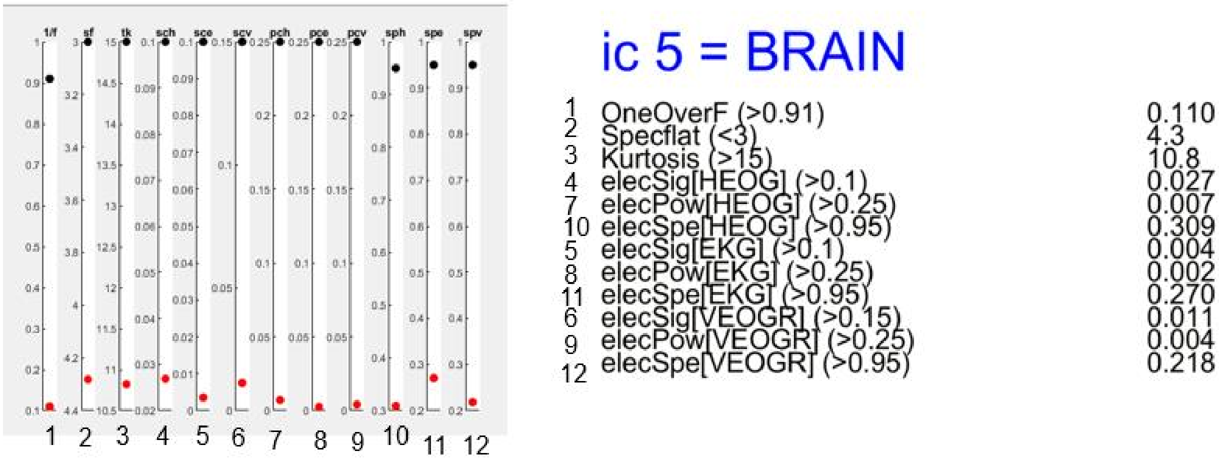


Figure 53. Information for each ICs including PSD and topography

Another plot defined to show the semi-automatic criteria and ICs values in each ICs:this plots includes 12 columns, .each 12 column corresponds to one criteria and dots with two different colors red dots and black dots.Black is the fixed values defined for semi-automatic criteria and red is Ics values for each ICs.If the black stands in up and red stand in down it mean the values matches with criteria value other views it is not match. if it matches output is a brain otherviews its a artifact. plot[54]shows a brain IC with black dot it up and red in down.



ic 5 = BRAIN

1	OneOverF (>0.91)	0.110
2	Specflat (<3)	4.3
3	Kurtosis (>15)	10.8
4	elecSig[HEOG] (>0.1)	0.027
7	elecPow[HEOG] (>0.25)	0.007
10	elecSpe[HEOG] (>0.95)	0.309
5	elecSig[EKG] (>0.1)	0.004
8	elecPow[EKG] (>0.25)	0.002
11	elecSpe[EKG] (>0.95)	0.270
6	elecSig[VEOGR] (>0.15)	0.011
9	elecPow[VEOGR] (>0.25)	0.004
12	elecSpe[VEOGR] (>0.95)	0.218

Figure 54. criteria values (matches and mis-matches

Conclusions

In this research we started to work with EEG data to find correlation between EEG data and kinematic data. In order to work with EEG at the first step we needed to have clear brain signals and Brain sources. There are several methods and approaches (semi-manual and automatic) to clear noise from the EEG data that each method has its pros and cons. Here we have compared the output of a semi-manual method (already used in Cosyncla) with ICLabel output in detection of brain courses. Semi-manual method reports a binary output for each IC as brain or noise and the ICLabel method reports a percentage of a brain for each IC. In the ICLabel method, we need to set a threshold value and by exceeding the brain value from this threshold, we can consider that IC as a brain reported by IClabel. Based on our data and analysis we have found that the best threshold here is 40%. Based on this threshold the ICLabel detects 95 percent of ICs (from ICs detected as brain by semi-manual method) as brain. It means the ICLabel is perfectly able to detect brain ICs that are already detected as brain by semi-manual method. Also, consider that ICLabel is automatic and needs less effort and time. In another point of view, ICLabel has detected about 300 brain ICs that the semi-manual method has missed. So, we propose that the combination of ICLabel with our previous semi-manual method would increase our accuracy. In this manner we have written an Matlab based application to simplify the implementation of this combination and visualize the overall result from both methods. In future we are going to improve our application to simplify and fasten working with EEG data from preprocessing to IC labeling and data manipulation based on large datasets.

References

1. ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage* **198**, 181–197 (2019).
2. Sanei, S. & Chambers, J. A. *EEG Signal Processing: Sanei/EEG signal processing*. (Wiley-Blackwell, 2007).
3. An automatic analysis method for detecting and eliminating ECG artifacts in EEG. *Comput. Biol. Med.* **37**, 1660–1671 (2007).
4. Huster, R. J. & Raud, L. A Tutorial Review on Multi-subject Decomposition of EEG. *Brain Topogr.* **31**, 3–16 (2018).
5. Minguillon, J., Lopez-Gordo, M. A. & Pelayo, F. Trends in EEG-BCI for daily-life: Requirements for artifact removal. *Biomed. Signal Process. Control* **31**, 407–418 (2017).
6. Wallstrom, G. L., Kass, R. E., Miller, A., Cohn, J. F. & Fox, N. A. Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods. *Int. J. Psychophysiol.* **53**, 105–119 (2004).
7. Ungerleider, L. G. & Haxby, J. V. 'What' and 'where' in the human brain. *Curr. Opin. Neurobiol.* **4**, 157–165 (1994).
8. The lobes of your brain - epilepsy - neurology - Epilepsy Sparks —. *Epilepsy Sparks* https://www.epilepsysparks.com/brain-lobes.
9. The lobes of your brain - epilepsy - neurology - Epilepsy Sparks —. *Epilepsy Sparks* https://www.epilepsysparks.com/brain-lobes.
10. Stuss, D. T. & Knight, R. T. *Principles of Frontal Lobe Function*. (Oxford University Press, 2013).
11. Shipton, H. W. EEG Analysis: A History and a Prospectus. *Annual Review of Biophysics and Bioengineering* vol. 4 1–13 (1975).

12. Neuron PNG Images, Transparent Neuron Image Download - PNGitem. https://www.pngitem.com/so/neuron/.

13. De, A. & Mondal, S. Yoga and brain wave coherence: A systematic review for brain function improvement. *Heart and Mind* vol. 4 33 (2020).

14. File:Figure 35 02 08.jpg. https://commons.wikimedia.org/wiki/File:Figure_35_02_08.jpg.

15. Raghavan, M., Fee, D. & Barkhaus, P. E. Generation and propagation of the action potential. *Clinical Neurophysiology: Basis and Technical Aspects* 3–22 (2019) doi:10.1016/b978-0-444-64032-1.00001-1.

16. Biga, L. M. *et al.* 12.5 The Action Potential. in *Anatomy & Physiology* (OpenStax/Oregon State University, 2019).

17. Jacks, A. S. Spontaneous retinal venous pulsation: aetiology and significance. *Journal of Neurology, Neurosurgery & Psychiatry* vol. 74 7–9 (2003).

18. Haas, L. F. Hans Berger (1873-1941), Richard Caton (1842-1926), and electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry* vol. 74 9–9 (2003).

19. Lundstrom, M. Fundamentals of Carrier Transport, 2nd edn. *Measurement Science and Technology* vol. 13 230–230 (2002).

20. Farnsworth, B. EEG vs. MRI vs. fMRI - What are the Differences? *Imotions* https://imotions.com/blog/eeg-vs-mri-vs-fmri-differences/ (2019).

21. Kim, S.-H., Lee, O.-K. & Kim, D. J. Study of Practical Method for International 10∼20 Electrode System. *Korean J Clin Lab Sci* **53**, 60–67 (2021).

22. mehta. Electrode 10-20 system. *Engineers Community* https://engineerscommunity.com/t/electrode-10-20-system/5486 (2018).

23. Machine learning with ensemble stacking model for automated sleep staging using dual-channel EEG signal. *Biomed. Signal Process. Control* **69**, 102898 (2021).

24. Zhang, H., Watrous, A. J., Patel, A. & Jacobs, J. Theta and Alpha Oscillations Are Traveling Waves in the Human Neocortex. *Neuron* vol. 98 1269–1281.e4 (2018).

25. Measuring Brain Waves in the Classroom. *Frontiers for Young Minds*

https://kids.frontiersin.org/articles/10.3389/frym.2020.00096.

26. Jiang, X., Bian, G.-B. & Tian, Z. Removal of Artifacts from EEG Signals: A Review. *Sensors* **19**, (2019).

27. Removal of ocular artifact from the EEG: a review. *Neurophysiologie Clinique/Clinical Neurophysiology* **30**, 5–19 (2000).

28. Yu, M. Removal methods of EMG Artifacts from EEG Signals. *Journal of Physics: Conference Series* vol. 1920 012076 (2021).

29. Goncharova, I. I., McFarland, D. J., Vaughan, T. M. & Wolpaw, J. R. EMG contamination of EEG: spectral and topographical characteristics. *Clinical Neurophysiology* vol. 114 1580–1593 (2003).

30. Mammone, N. & Morabito, F. Enhanced Automatic Wavelet Independent Component Analysis for Electroencephalographic Artifact Removal. *Entropy* vol. 16 6553–6572 (2014).

31. Nolan, H., Whelan, R. & Reilly, R. B. FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. *Journal of Neuroscience Methods* vol. 192 152–162 (2010).

32. Islam, M. K., Rastegarnia, A. & Yang, Z. Methods for artifact detection and removal from scalp EEG: A review. *Neurophysiologie Clinique/Clinical Neurophysiology* vol. 46 287–305 (2016).

33. Chaturvedi, P. & Gupta, L. Study and detection of eye blink artifacts in EEG signals. in *2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)* (IEEE, 2018). doi:10.1109/sceecs.2018.8546907.

34. Krishnaveni, V., Jayaraman, S., Aravind, S., Hariharasudhan, V. & Ramadoss, K. Automatic Identification and Removal of Ocular Artifacts from EEG using Wavelet Transform. (2006).

35. Skewness. https://en.wikipedia.org/wiki/Skewness.

36. Expected value. https://en.wikipedia.org/wiki/Expected_value.

37. Shoka, A., Dessouky, M., El-Sherbeny, A. & El-Sayed, A. Literature Review on EEG Preprocessing, Feature Extraction, and Classifications Techniques. *Menoufia Journal of Electronic Engineering Research* vol. 28 292–299 (2019).

38. Table 2: Skewness, kurtosis and tests of normal distribution of male height in individual counties

(županije). doi:10.7717/peerj.6598/table-2.

39. Kurtosis. https://www.daytrading.com/.

40. Abhang, P. A., Gawali, B. W. & Mehrotra, S. C. Technological Basics of EEG Recording and Operation of Apparatus. *Introduction to EEG- and Speech-Based Emotion Recognition* 19–50 (2016) doi:10.1016/b978-0-12-804490-2.00002-6.

41. Alsolamy, M. & Fattouh, A. Emotion estimation from EEG signals during listening to Quran using PSD features. *2016 7th International Conference on Computer Science and Information Technology (CSIT)* (2016) doi:10.1109/csit.2016.7549457.

42. Labs, S. Factors that Impact Power Spectral Density Estimation - Sapien Labs. *Sapien Labs | Neuroscience | Human Brain Diversity Project* https://sapienlabs.org/factors-that-impact-power-spectrum-density-estimation/ (2018).

43. Welch's method. https://en.wikipedia.org/wiki/Welch%27s_method.

44. Samadzadehaghdam, N. *et al.* Developing a Multi-channel Beamformer by Enhancing Spatially Constrained ICA for Recovery of Correlated EEG Sources. *Journal of Biomedical Physics and Engineering* (2018) doi:10.31661/jbpe.v0i0.801.

45. Sun, L., Liu, Y. & Beadle, P. J. Independent component analysis of EEG signals. *Proceedings of 2005 IEEE International Workshop on VLSI Design and Video Technology, 2005.* doi:10.1109/iwvdvt.2005.1504590.

46. Samadzadehaghdam, N., MakkiAbadi, B. & Masjoodi, S. Evaluating the Impact of White Matter Conductivity Anisotropy on Reconstructing EEG Sources by Linearly Constrained Minimum Variance Beamformer. *Advanced Biomedical Engineering* vol. 9 53–61 (2020).

47. Ventouras, E. M. *et al.* Independent component analysis for source localization of EEG sleep spindle components. *Comput. Intell. Neurosci.* 329436 (2010).

48. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 156869 (2011).

49. Understanding Parametric Tests, Skewness, and Kurtosis.

https://www.allaboutcircuits.com/technical-articles/understanding-the-normal-distribution-parametric-tests-skewness-and-kurtosis/.

50. Kurtosis. https://www.daytrading.com/.

51. EEGLAB. http://www.sccn.ucsd.edu/eeglab/.

52. Website. https://www.fieldtriptoolbox.org/getting_started/eeglab/,.

53. Introduction on dealing with artifacts. https://www.fieldtriptoolbox.org/tutorial/artifacts/.

54. sccn. ICLabel. *GitHub* https://github.com/sccn/ICLabel.

55. SCCN: Independent Component Labeling. https://labeling.ucsd.edu/tutorial.

56. SCCN: Independent Component Labeling. https://labeling.ucsd.edu/tutorial.

57. SCCN: Independent Component Labeling. https://iclabel.ucsd.edu/tutorial.

58. Pion-Tonachini, L., Makeig, S. & Kreutz-Delgado, K. Crowd labeling latent Dirichlet allocation. *Knowledge and Information Systems* vol. 53 749–765 (2017).

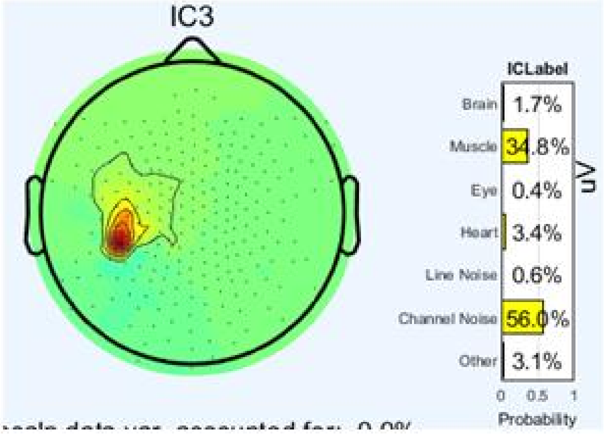
59. Yu, H. & Yang, J. A direct LDA algorithm for high-dimensional data — with application to face recognition. *Pattern Recognition* vol. 34 2067–2070 (2001).

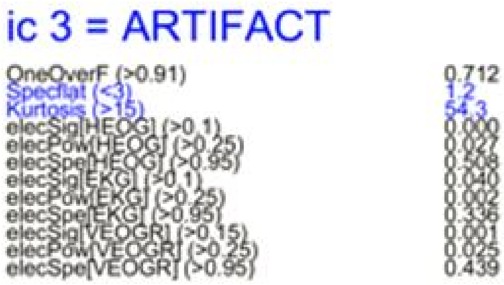
60. lucapton. ICLabel-Train/tfGAN_indvBN.py at 46b33f0e3f27782b4ed2868a189e9305cdbf4fee · lucapton/ICLabel-Train. *GitHub* https://github.com/lucapton/ICLabel-Train.

Appendix

Comparing all 51 ICs for subjects ID(HM_102478) the left plots are the output of ICLable and on the right the output of semi-automatic ICs selection pipeline can be observed.

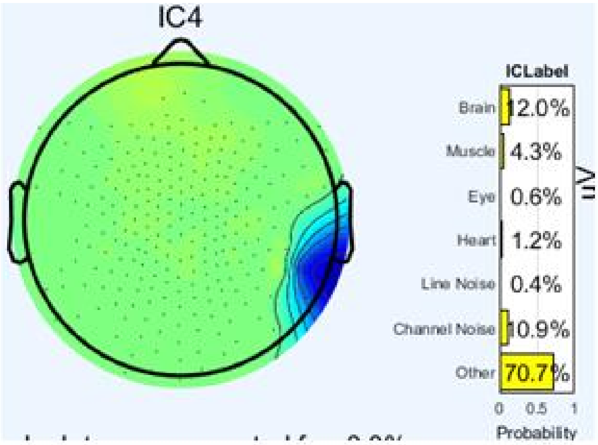
Figure 1.

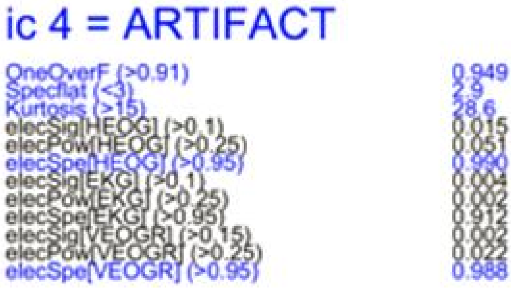




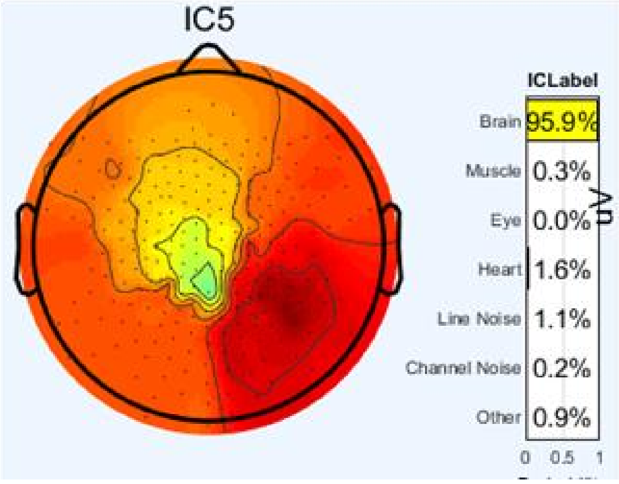
```
ic 3 = ARTIFACT
OneOverF (>0.91)            0.712
Specflat (<3)               1.2
Kurtosis (>15)              54.3
elecSig[HEOG] (>0.1)        0.000
elecPow[HEOG] (>0.25)       0.027
elecSpe[HEOG] (>0.95)       0.508
elecSig[EKG] (>0.1)         0.040
elecPow[EKG] (>0.25)        0.002
elecSpe[EKG] (>0.95)        0.336
elecSig[VEOGR] (>0.15)      0.001
elecPow[VEOGR] (>0.25)      0.025
elecSpe[VEOGR] (>0.95)      0.439
```

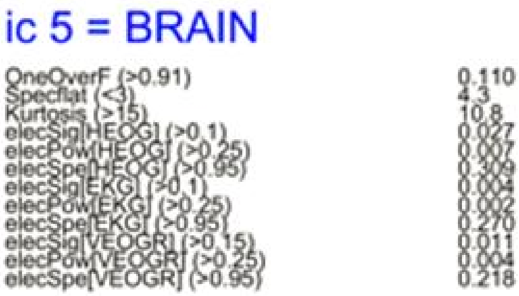
Figure 2.





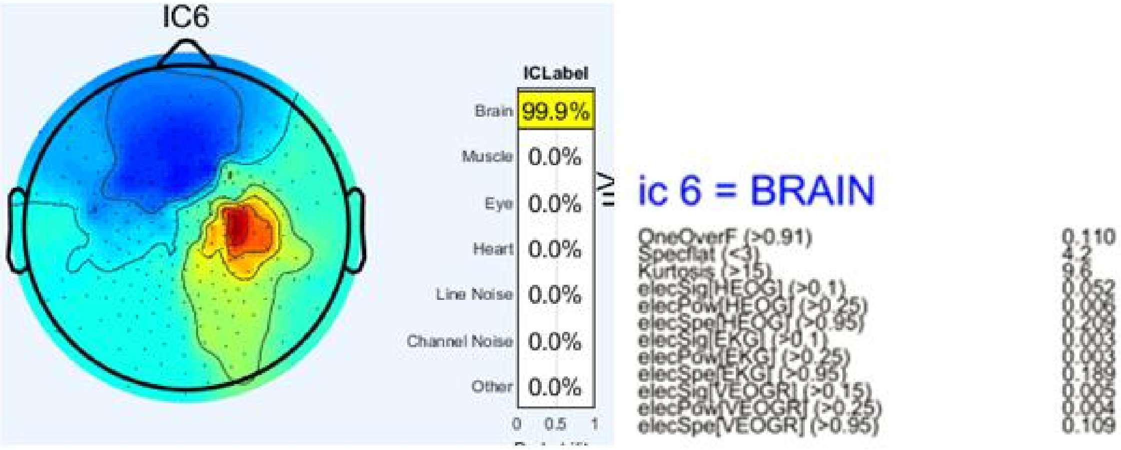
```
ic 4 = ARTIFACT
OneOverF (>0.91)            0.949
Specflat (<3)               2.9
Kurtosis (>15)              28.6
elecSig[HEOG] (>0.1)        0.015
elecPow[HEOG] (>0.25)       0.051
elecSpe[HEOG] (>0.95)       0.990
elecSig[EKG] (>0.1)         0.004
elecPow[EKG] (>0.25)        0.002
elecSpe[EKG] (>0.95)        0.912
elecSig[VEOGR] (>0.15)      0.002
elecPow[VEOGR] (>0.25)      0.022
elecSpe[VEOGR] (>0.95)      0.988
```





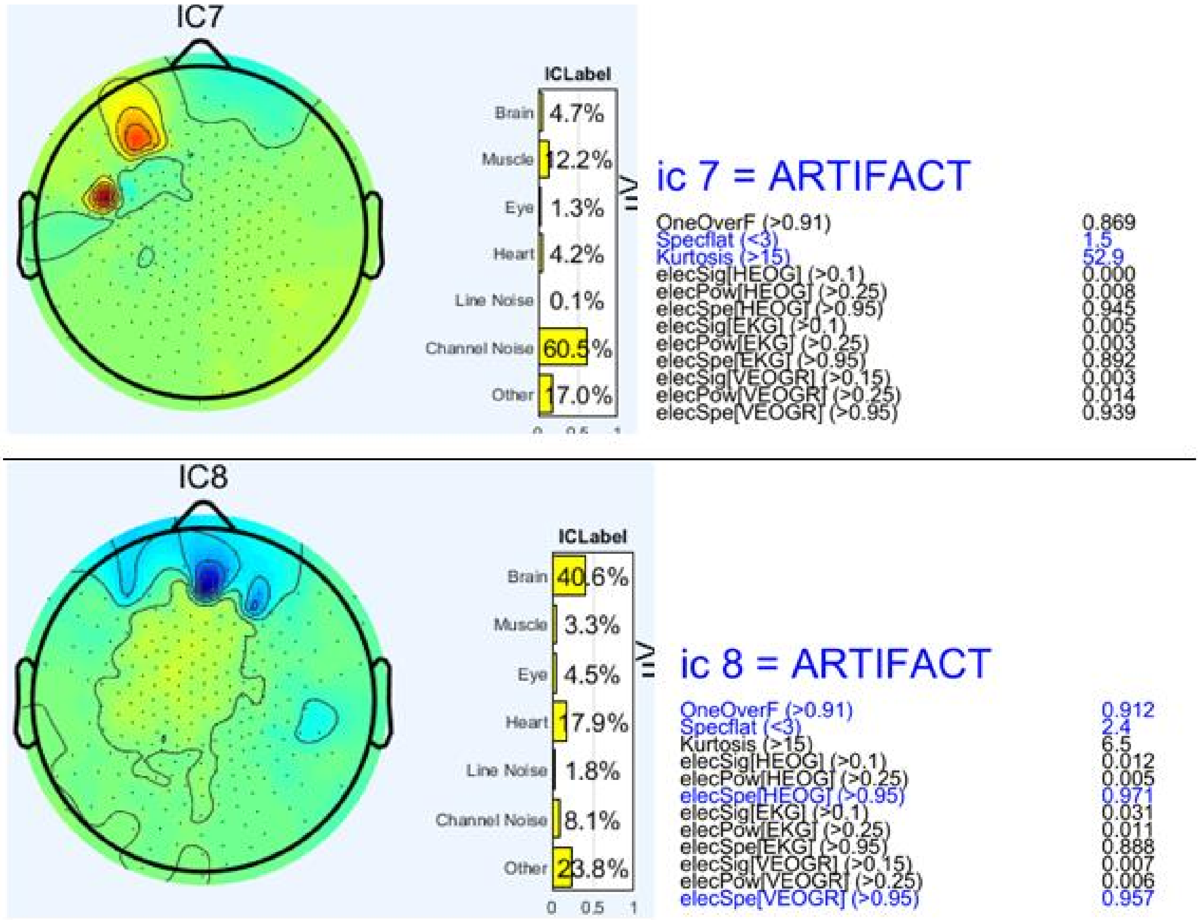
```
ic 5 = BRAIN
OneOverF (>0.91)            0.110
Specflat (<3)               4.3
Kurtosis (>15)              10.8
elecSig[HEOG] (>0.1)        0.027
elecPow[HEOG] (>0.25)       0.007
elecSpe[HEOG] (>0.95)       0.309
elecSig[EKG] (>0.1)         0.004
elecPow[EKG] (>0.25)        0.002
elecSpe[EKG] (>0.95)        0.270
elecSig[VEOGR] (>0.15)      0.011
elecPow[VEOGR] (>0.25)      0.004
elecSpe[VEOGR] (>0.95)      0.218
```

Figure 3.



```
ic 6 = BRAIN
OneOverF (>0.91)            0.110
Specflat (<3)               4.2
Kurtosis (>15)              9.6
elecSig[HEOG] (>0.1)        0.052
elecPow[HEOG] (>0.25)       0.006
elecSpe[HEOG] (>0.95)       0.209
elecSig[EKG] (>0.1)         0.003
elecPow[EKG] (>0.25)        0.003
elecSpe[EKG] (>0.95)        0.189
elecSig[VEOGR] (>0.15)      0.005
elecPow[VEOGR] (>0.25)      0.004
elecSpe[VEOGR] (>0.95)      0.109
```

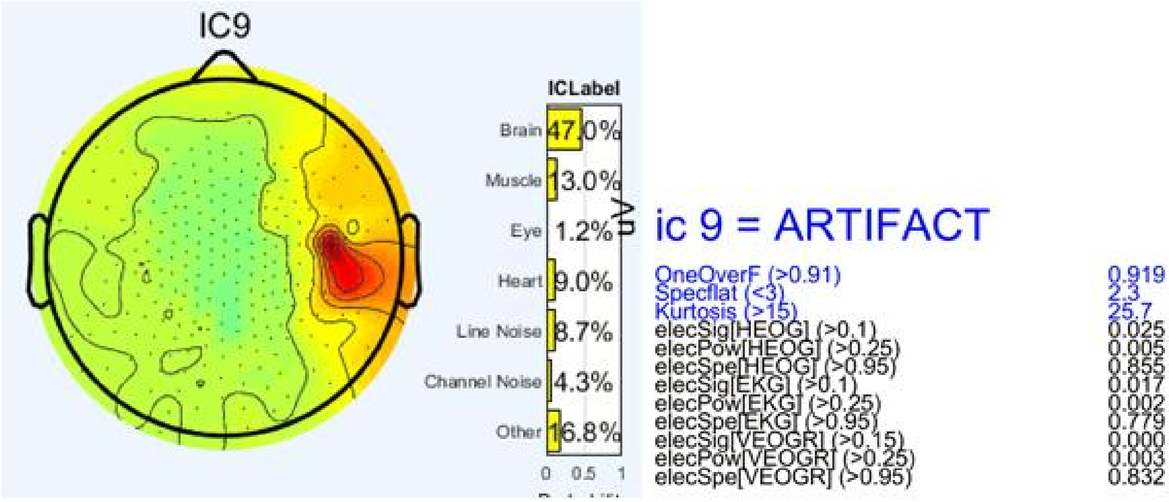
Figure 4.



```
ic 7 = ARTIFACT
OneOverF (>0.91)            0.869
Specflat (<3)               1.5
Kurtosis (>15)              52.9
elecSig[HEOG] (>0.1)        0.000
elecPow[HEOG] (>0.25)       0.008
elecSpe[HEOG] (>0.95)       0.945
elecSig[EKG] (>0.1)         0.005
elecPow[EKG] (>0.25)        0.003
elecSpe[EKG] (>0.95)        0.892
elecSig[VEOGR] (>0.15)      0.003
elecPow[VEOGR] (>0.25)      0.014
elecSpe[VEOGR] (>0.95)      0.939
```

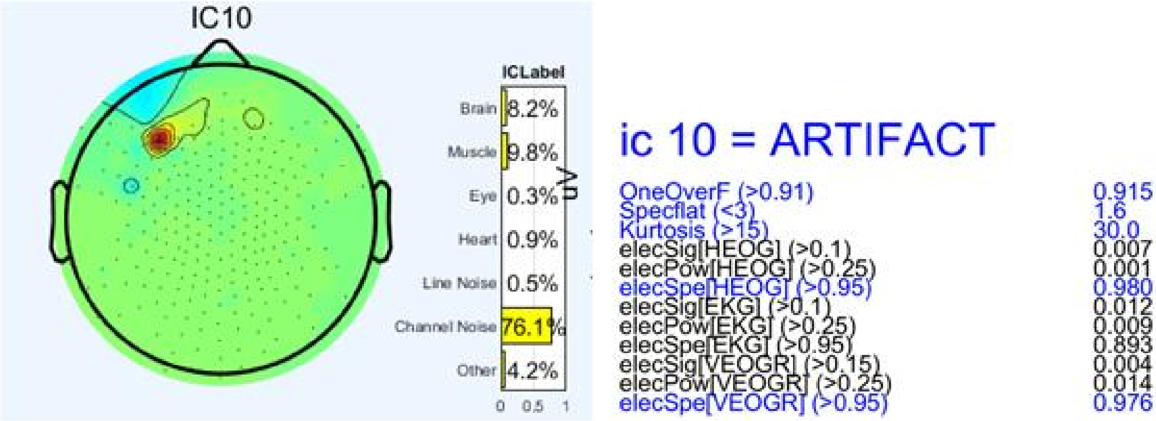
```
ic 8 = ARTIFACT
OneOverF (>0.91)            0.912
Specflat (<3)               2.4
Kurtosis (>15)              6.5
elecSig[HEOG] (>0.1)        0.012
elecPow[HEOG] (>0.25)       0.005
elecSpe[HEOG] (>0.95)       0.971
elecSig[EKG] (>0.1)         0.031
elecPow[EKG] (>0.25)        0.011
elecSpe[EKG] (>0.95)        0.888
elecSig[VEOGR] (>0.15)      0.007
elecPow[VEOGR] (>0.25)      0.006
elecSpe[VEOGR] (>0.95)      0.957
```

Figure 5.



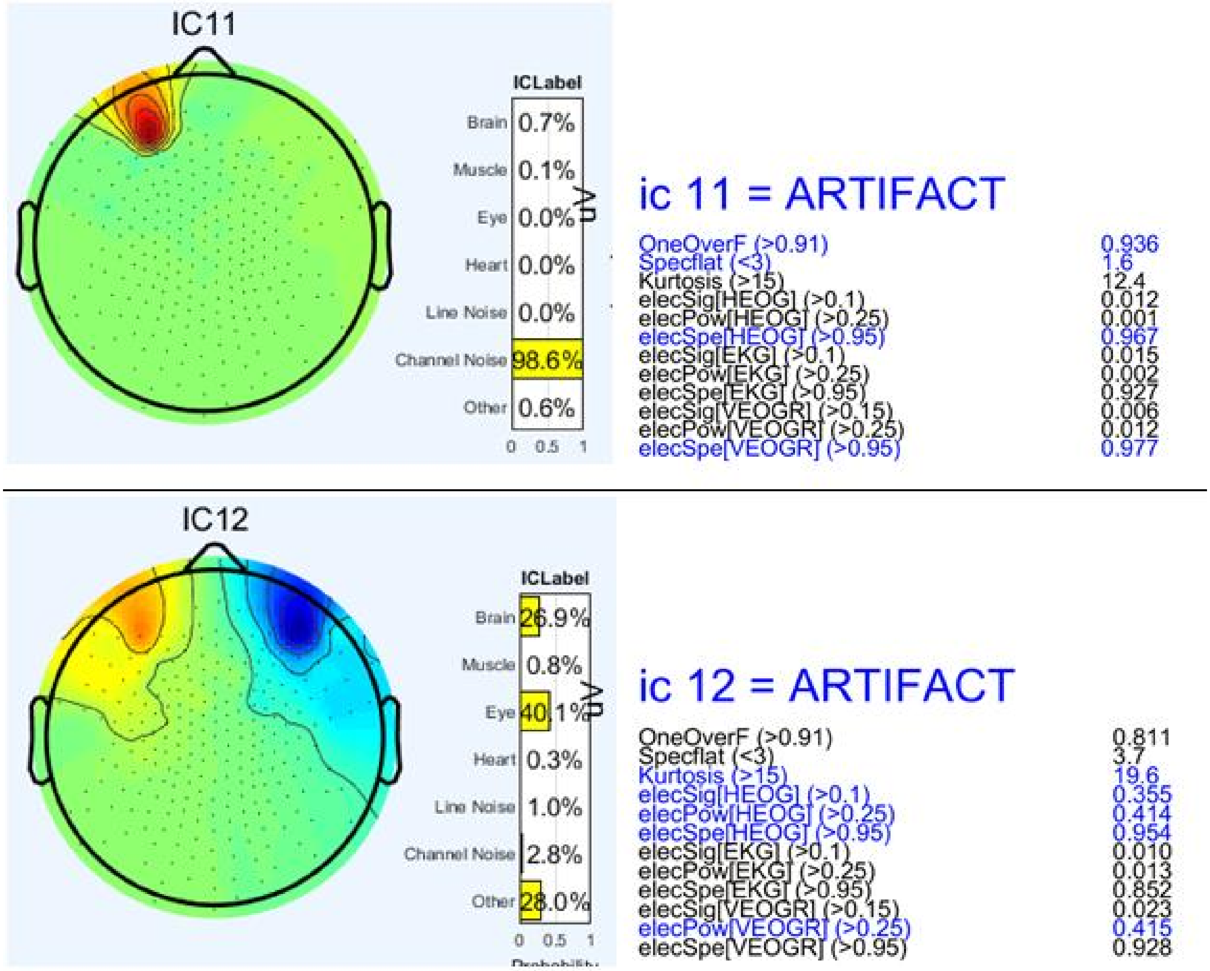
```
ic 9 = ARTIFACT
OneOverF (>0.91)        0.919
Specflat (<3)           2.3
Kurtosis (>15)          25.7
elecSig[HEOG] (>0.1)    0.025
elecPow[HEOG] (>0.25)   0.005
elecSpe[HEOG] (>0.95)   0.855
elecSig[EKG] (>0.1)     0.017
elecPow[EKG] (>0.25)    0.002
elecSpe[EKG] (>0.95)    0.779
elecSig[VEOGR] (>0.15)  0.000
elecPow[VEOGR] (>0.25)  0.003
elecSpe[VEOGR] (>0.95)  0.832
```

Figure 6.



```
ic 10 = ARTIFACT
OneOverF (>0.91)        0.915
Specflat (<3)           1.6
Kurtosis (>15)          30.0
elecSig[HEOG] (>0.1)    0.007
elecPow[HEOG] (>0.25)   0.001
elecSpe[HEOG] (>0.95)   0.980
elecSig[EKG] (>0.1)     0.012
elecPow[EKG] (>0.25)    0.009
elecSpe[EKG] (>0.95)    0.893
elecSig[VEOGR] (>0.15)  0.004
elecPow[VEOGR] (>0.25)  0.014
elecSpe[VEOGR] (>0.95)  0.976
```

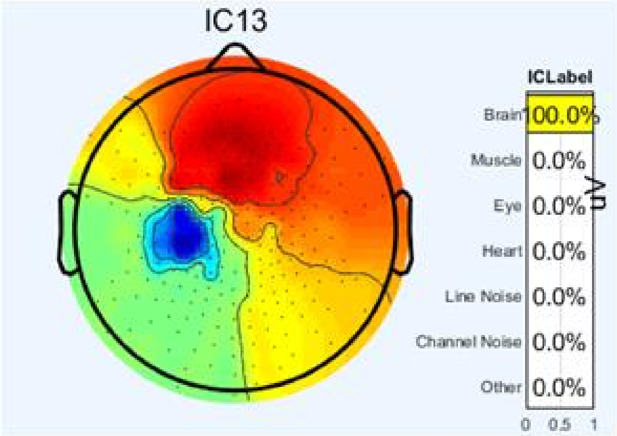
Figure 7.



```
ic 11 = ARTIFACT

OneOverF (>0.91)          0.936
Specflat (<3)             1.6
Kurtosis (>15)            12.4
elecSig[HEOG] (>0.1)      0.012
elecPow[HEOG] (>0.25)     0.001
elecSpe[HEOG] (>0.95)     0.967
elecSig[EKG] (>0.1)       0.015
elecPow[EKG] (>0.25)      0.002
elecSpe[EKG] (>0.95)      0.927
elecSig[VEOGR] (>0.15)    0.006
elecPow[VEOGR] (>0.25)    0.012
elecSpe[VEOGR] (>0.95)    0.977
```

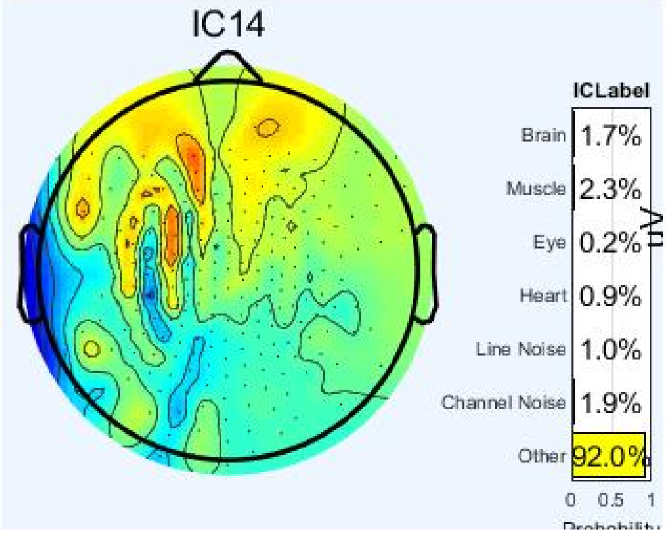
```
ic 12 = ARTIFACT

OneOverF (>0.91)          0.811
Specflat (<3)             3.7
Kurtosis (>15)            19.6
elecSig[HEOG] (>0.1)      0.355
elecPow[HEOG] (>0.25)     0.414
elecSpe[HEOG] (>0.95)     0.954
elecSig[EKG] (>0.1)       0.010
elecPow[EKG] (>0.25)      0.013
elecSpe[EKG] (>0.95)      0.852
elecSig[VEOGR] (>0.15)    0.023
elecPow[VEOGR] (>0.25)    0.415
elecSpe[VEOGR] (>0.95)    0.928
```

Figure 8.



ic 13 = BRAIN

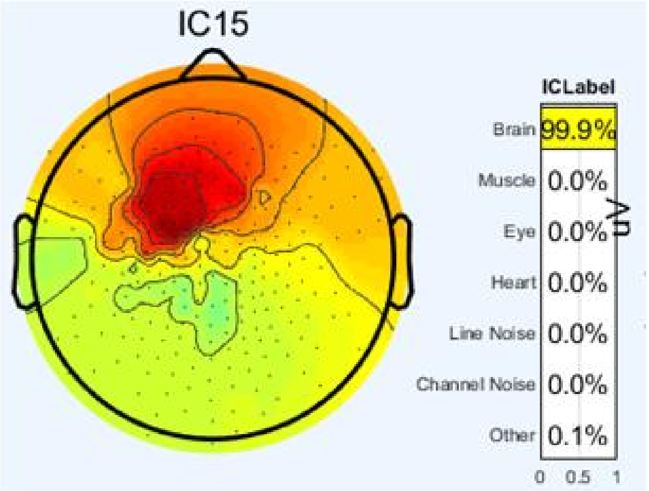
OneOverF (>0.91)	0.110
Specflat (<3)	5.1
Kurtosis (>15)	7.2
elecSig[HEOG] (>0.1)	0.047
elecPow[HEOG] (>0.25)	0.000
elecSpe[HEOG] (>0.95)	0.157
elecSig[EKG] (>0.1)	0.009
elecPow[EKG] (>0.25)	0.002
elecSpe[EKG] (>0.95)	0.147
elecSig[VEOGR] (>0.15)	0.009
elecPow[VEOGR] (>0.25)	0.006
elecSpe[VEOGR] (>0.95)	0.058



ic 14 = ARTIFACT

OneOverF (>0.91)	0.703
Specflat (<3)	1.3
Kurtosis (>15)	5.5
elecSig[HEOG] (>0.1)	0.003
elecPow[HEOG] (>0.25)	0.002
elecSpe[HEOG] (>0.95)	0.916
elecSig[EKG] (>0.1)	0.003
elecPow[EKG] (>0.25)	0.000
elecSpe[EKG] (>0.95)	0.882
elecSig[VEOGR] (>0.15)	0.001
elecPow[VEOGR] (>0.25)	0.004
elecSpe[VEOGR] (>0.95)	0.907

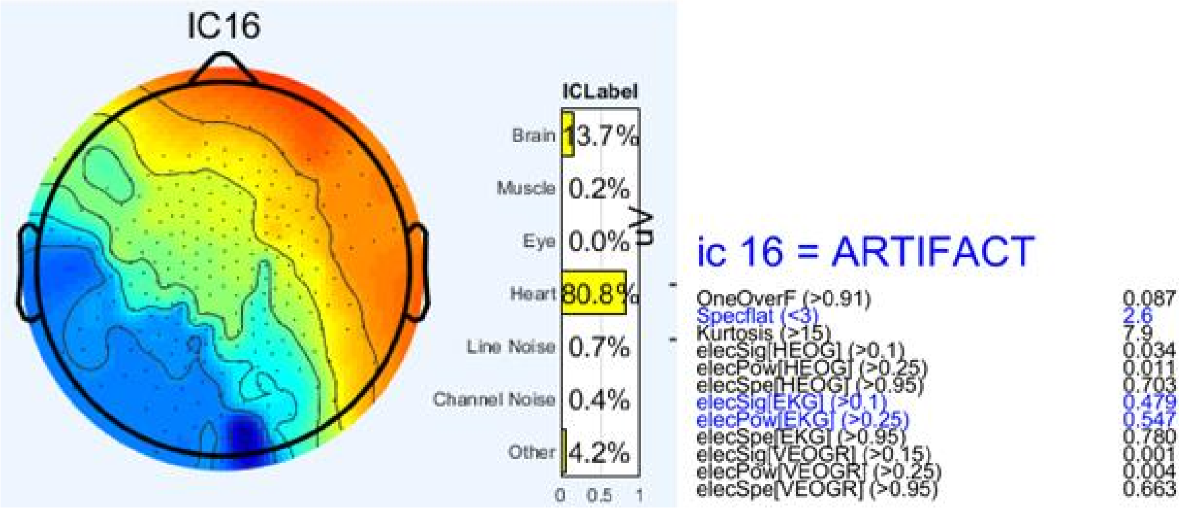
Figure 9.



ic 15 = BRAIN

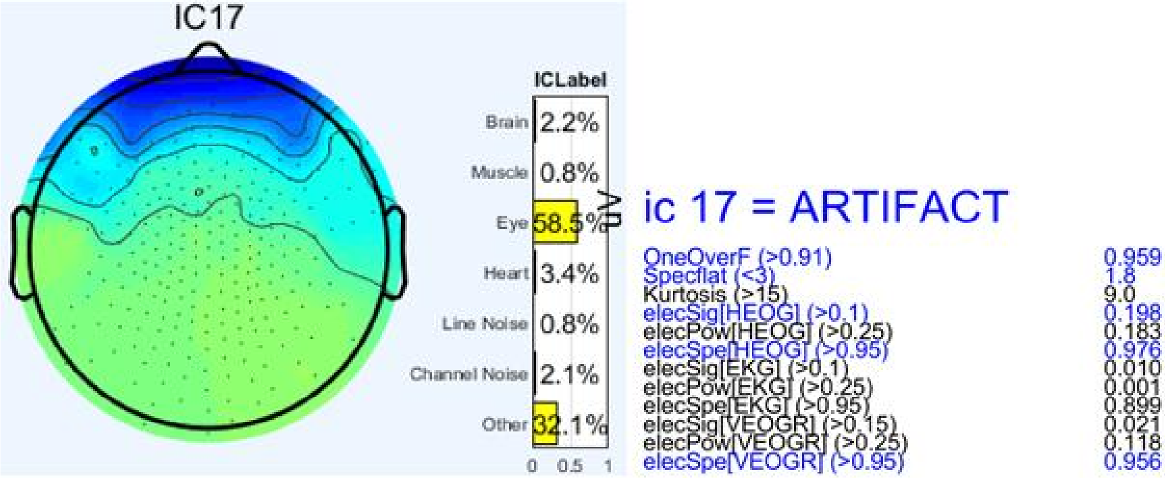
OneOverF (>0.91)	0.110
Specflat (<3)	3.5
Kurtosis (>15)	6.2
elecSig[HEOG] (>0.1)	0.013
elecPow[HEOG] (>0.25)	0.005
elecSpe[HEOG] (>0.95)	0.252
elecSig[EKG] (>0.1)	0.011
elecPow[EKG] (>0.25)	0.005
elecSpe[EKG] (>0.95)	0.228
elecSig[VEOGR] (>0.15)	0.008
elecPow[VEOGR] (>0.25)	0.008
elecSpe[VEOGR] (>0.95)	0.156

Figure 10.



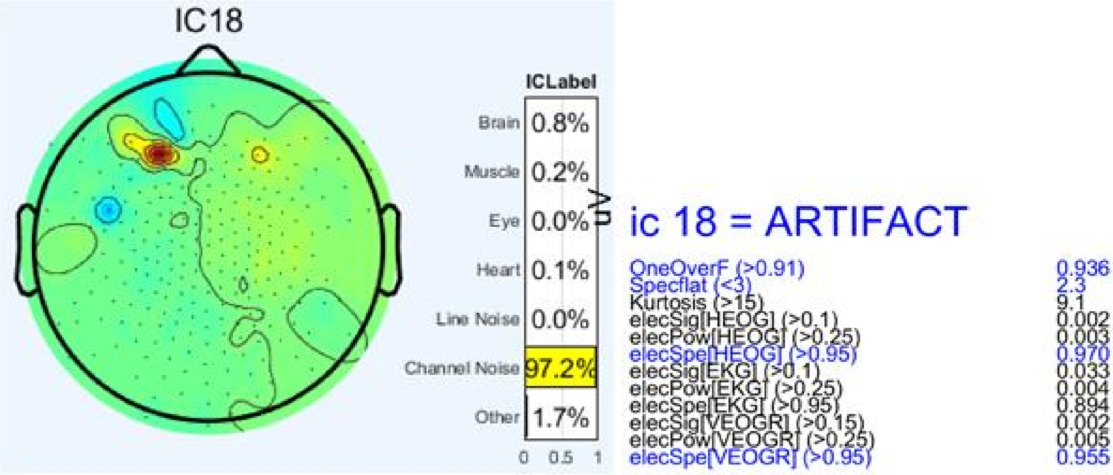
```
ic 16 = ARTIFACT
OneOverF (>0.91)                0.087
Specflat (<3)                   2.6
Kurtosis (>15)                  7.9
elecSig[HEOG] (>0.1)            0.034
elecPow[HEOG] (>0.25)           0.011
elecSpe[HEOG] (>0.95)           0.703
elecSig[EKG] (>0.1)             0.479
elecPow[EKG] (>0.25)            0.547
elecSpe[EKG] (>0.95)            0.780
elecSig[VEOGR] (>0.15)          0.001
elecPow[VEOGR] (>0.25)          0.004
elecSpe[VEOGR] (>0.95)          0.663
```

Figure 11.



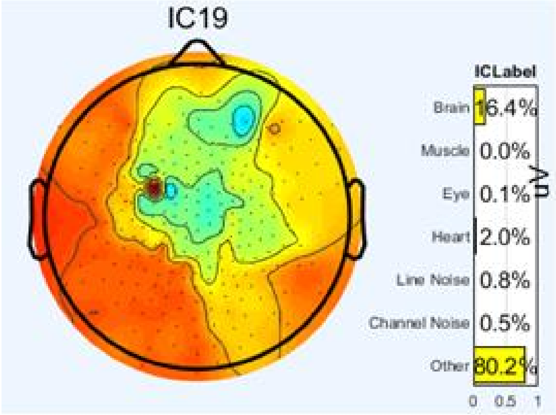
```
ic 17 = ARTIFACT
OneOverF (>0.91)                0.959
Specflat (<3)                   1.8
Kurtosis (>15)                  9.0
elecSig[HEOG] (>0.1)            0.198
elecPow[HEOG] (>0.25)           0.183
elecSpe[HEOG] (>0.95)           0.976
elecSig[EKG] (>0.1)             0.010
elecPow[EKG] (>0.25)            0.001
elecSpe[EKG] (>0.95)            0.899
elecSig[VEOGR] (>0.15)          0.021
elecPow[VEOGR] (>0.25)          0.118
elecSpe[VEOGR] (>0.95)          0.956
```

Figure 12.



```
ic 18 = ARTIFACT
OneOverF (>0.91)                0.936
Specflat (<3)                   2.3
Kurtosis (>15)                  9.1
elecSig[HEOG] (>0.1)            0.002
elecPow[HEOG] (>0.25)           0.003
elecSpe[HEOG] (>0.95)           0.970
elecSig[EKG] (>0.1)             0.033
elecPow[EKG] (>0.25)            0.004
elecSpe[EKG] (>0.95)            0.894
elecSig[VEOGR] (>0.15)          0.002
elecPow[VEOGR] (>0.25)          0.005
elecSpe[VEOGR] (>0.95)          0.955
```

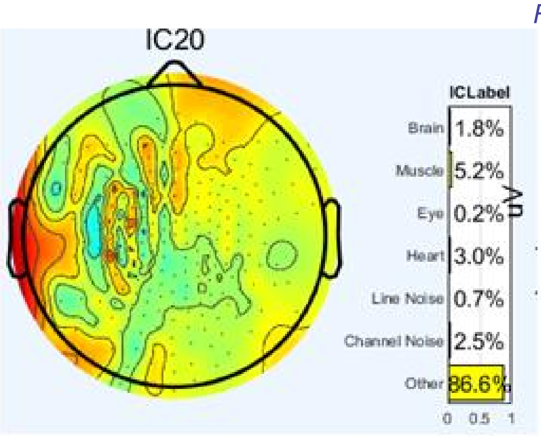
Figure 13.



ic 19 = ARTIFACT

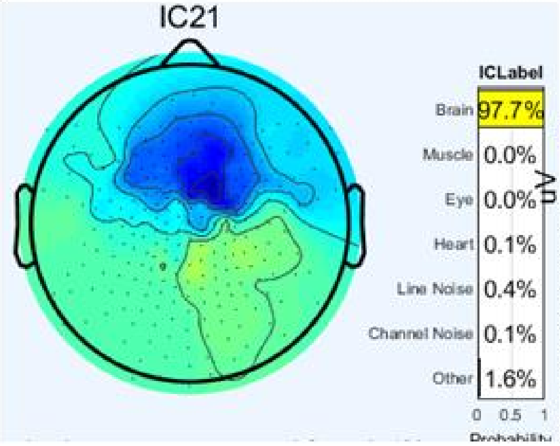
OneOverF (>0.91)	0.945
Specflat (<3)	2.2
Kurtosis (>15)	7.8
elecSig[HEOG] (>0.1)	0.008
elecPow[HEOG] (>0.25)	0.002
elecSpe[HEOG] (>0.95)	0.949
elecSig[EKG] (>0.1)	0.004
elecPow[EKG] (>0.25)	0.010
elecSpe[EKG] (>0.95)	0.867
elecSig[VEOGR] (>0.15)	0.004
elecPow[VEOGR] (>0.25)	0.005
elecSpe[VEOGR] (>0.95)	0.928

Figure 14.



ic 20 = ARTIFACT

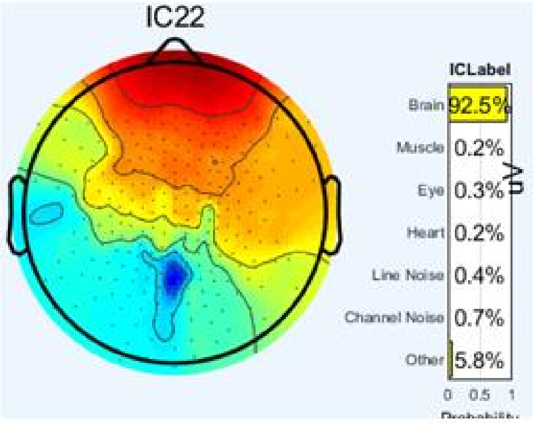
OneOverF (>0.91)	0.612
Specflat (<3)	1.3
Kurtosis (>15)	5.8
elecSig[HEOG] (>0.1)	0.005
elecPow[HEOG] (>0.25)	0.011
elecSpe[HEOG] (>0.95)	0.899
elecSig[EKG] (>0.1)	0.009
elecPow[EKG] (>0.25)	0.008
elecSpe[EKG] (>0.95)	0.833
elecSig[VEOGR] (>0.15)	0.001
elecPow[VEOGR] (>0.25)	0.011
elecSpe[VEOGR] (>0.95)	0.880



ic 21 = BRAIN

OneOverF (>0.91)	0.110
Specflat (<3)	3.4
Kurtosis (>15)	5.6
elecSig[HEOG] (>0.1)	0.014
elecPow[HEOG] (>0.25)	0.013
elecSpe[HEOG] (>0.95)	0.414
elecSig[EKG] (>0.1)	0.039
elecPow[EKG] (>0.25)	0.005
elecSpe[EKG] (>0.95)	0.377
elecSig[VEOGR] (>0.15)	0.006
elecPow[VEOGR] (>0.25)	0.006
elecSpe[VEOGR] (>0.95)	0.317

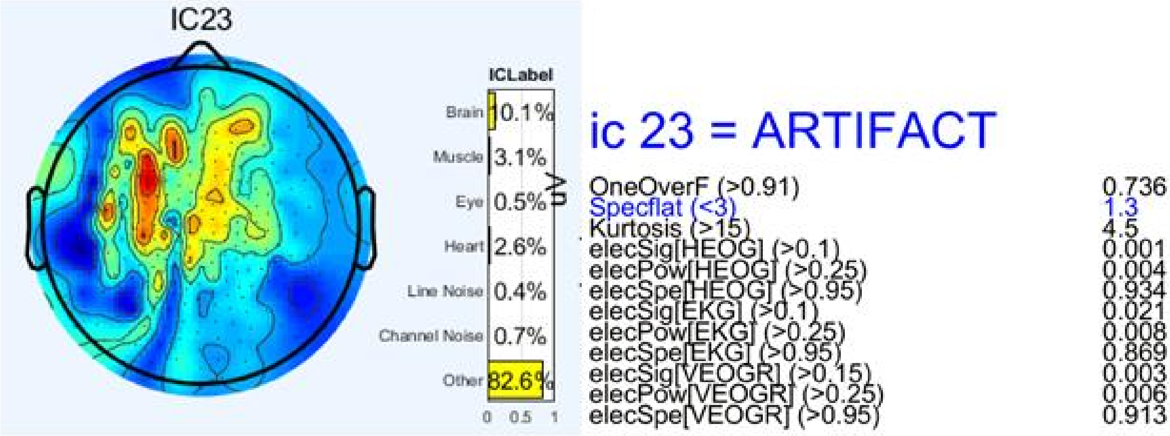
Figure 15.



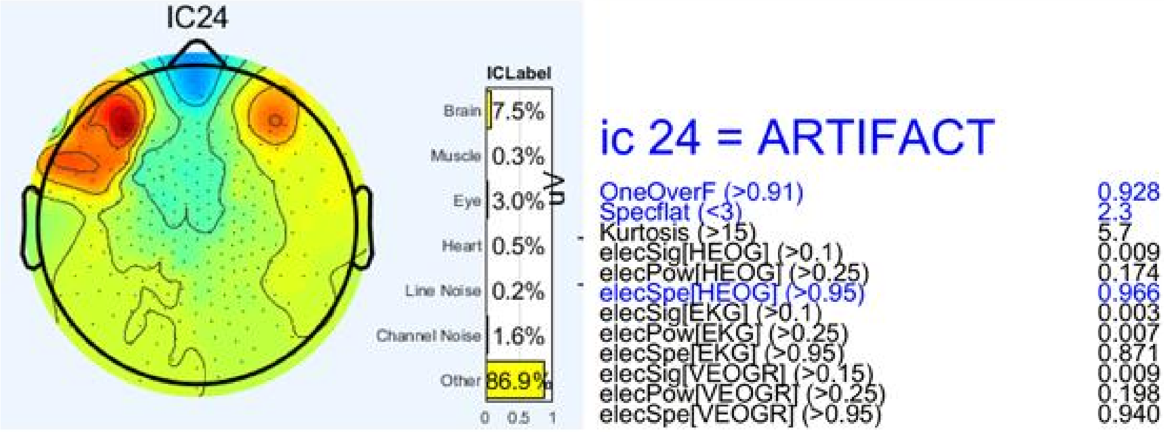
ic 22 = BRAIN

OneOverF (>0.91)	0.692
Specflat (<3)	4.2
Kurtosis (>15)	5.2
elecSig[HEOG] (>0.1)	0.045
elecPow[HEOG] (>0.25)	0.078
elecSpe[HEOG] (>0.95)	0.854
elecSig[EKG] (>0.1)	0.005
elecPow[EKG] (>0.25)	0.002
elecSpe[EKG] (>0.95)	0.748
elecSig[VEOGR] (>0.15)	0.032
elecPow[VEOGR] (>0.25)	0.101
elecSpe[VEOGR] (>0.95)	0.814

Figure 16.

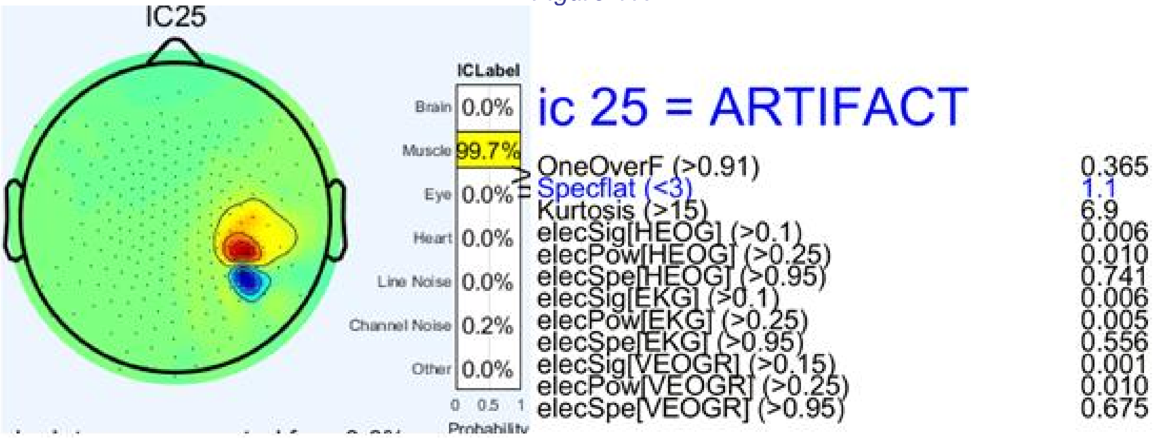


```
ic 23 = ARTIFACT
OneOverF (>0.91)           0.736
Specflat (<3)              1.3
Kurtosis (>15)             4.5
elecSig[HEOG] (>0.1)       0.001
elecPow[HEOG] (>0.25)      0.004
elecSpe[HEOG] (>0.95)      0.934
elecSig[EKG] (>0.1)        0.021
elecPow[EKG] (>0.25)       0.008
elecSpe[EKG] (>0.95)       0.869
elecSig[VEOGR] (>0.15)     0.003
elecPow[VEOGR] (>0.25)     0.006
elecSpe[VEOGR] (>0.95)     0.913
```



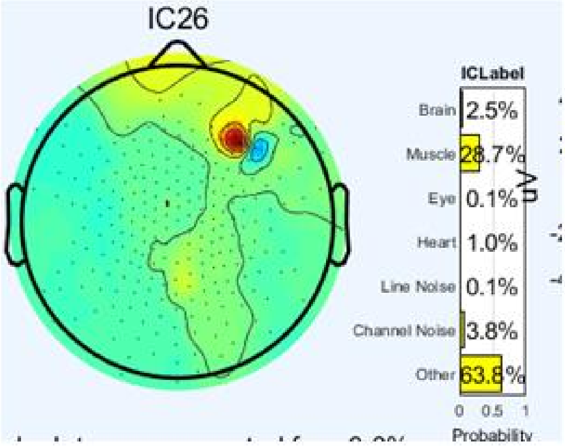
```
ic 24 = ARTIFACT
OneOverF (>0.91)           0.928
Specflat (<3)              2.3
Kurtosis (>15)             5.7
elecSig[HEOG] (>0.1)       0.009
elecPow[HEOG] (>0.25)      0.174
elecSpe[HEOG] (>0.95)      0.966
elecSig[EKG] (>0.1)        0.003
elecPow[EKG] (>0.25)       0.007
elecSpe[EKG] (>0.95)       0.871
elecSig[VEOGR] (>0.15)     0.009
elecPow[VEOGR] (>0.25)     0.198
elecSpe[VEOGR] (>0.95)     0.940
```

Figure 17.

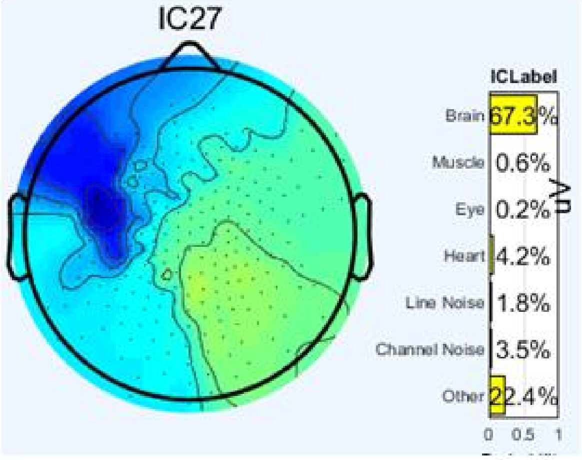


```
ic 25 = ARTIFACT
OneOverF (>0.91)           0.365
Specflat (<3)              1.1
Kurtosis (>15)             6.9
elecSig[HEOG] (>0.1)       0.006
elecPow[HEOG] (>0.25)      0.010
elecSpe[HEOG] (>0.95)      0.741
elecSig[EKG] (>0.1)        0.006
elecPow[EKG] (>0.25)       0.005
elecSpe[EKG] (>0.95)       0.556
elecSig[VEOGR] (>0.15)     0.001
elecPow[VEOGR] (>0.25)     0.010
elecSpe[VEOGR] (>0.95)     0.675
```

Figure 18.

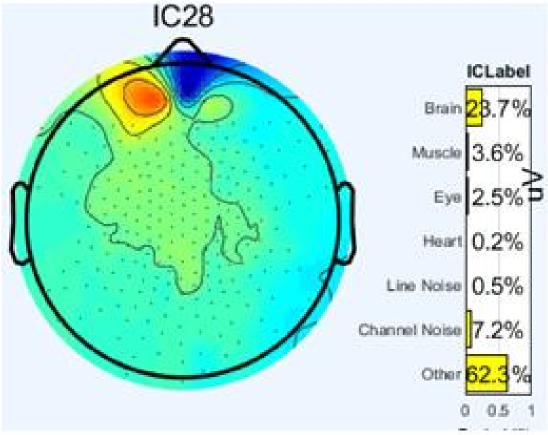


```
ic 26 = ARTIFACT
OneOverF (>0.91)          0.901
Specflat (<3)             1.4
Kurtosis (>15)            7.1
elecSig[HEOG] (>0.1)      0.014
elecPow[HEOG] (>0.25)     0.003
elecSpe[HEOG] (>0.95)     0.869
elecSig[EKG] (>0.1)       0.002
elecPow[EKG] (>0.25)      0.009
elecSpe[EKG] (>0.95)      0.802
elecSig[VEOGR] (>0.15)    0.006
elecPow[VEOGR] (>0.25)    0.004
elecSpe[VEOGR] (>0.95)    0.870
```



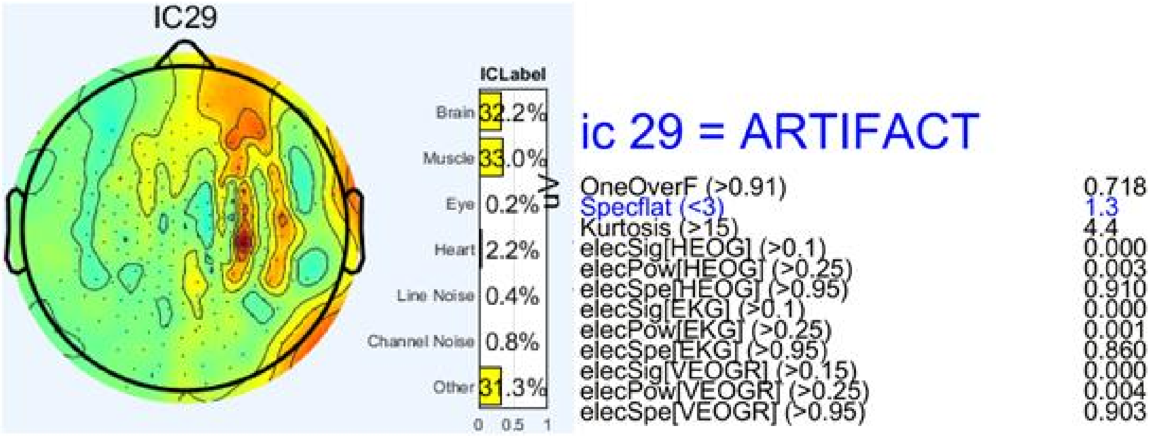
```
ic 27 = BRAIN CORRECTED
OneOverF (>0.91)          0.779
Specflat (<3)             3.5
Kurtosis (>15)            4.5
elecSig[HEOG] (>0.1)      0.156
elecPow[HEOG] (>0.25)     0.027
elecSpe[HEOG] (>0.95)     0.722
elecSig[EKG] (>0.1)       0.003
elecPow[EKG] (>0.25)      0.002
elecSpe[EKG] (>0.95)      0.642
elecSig[VEOGR] (>0.15)    0.010
elecPow[VEOGR] (>0.25)    0.012
elecSpe[VEOGR] (>0.95)    0.668
```

Figure 19.

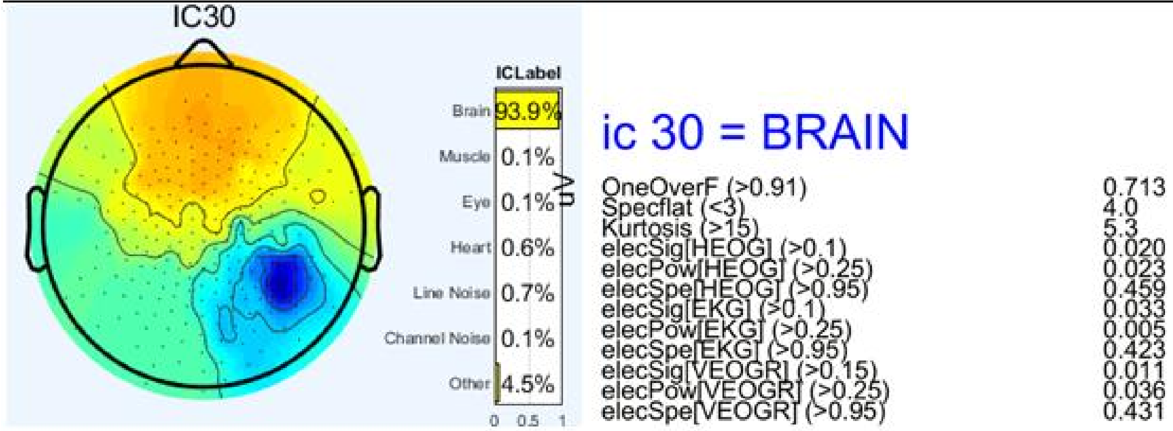


```
ic 28 = ARTIFACT
OneOverF (>0.91)          0.947
Specflat (<3)             1.9
Kurtosis (>15)            6.1
elecSig[HEOG] (>0.1)      0.029
elecPow[HEOG] (>0.25)     0.056
elecSpe[HEOG] (>0.95)     0.933
elecSig[EKG] (>0.1)       0.015
elecPow[EKG] (>0.25)      0.005
elecSpe[EKG] (>0.95)      0.853
elecSig[VEOGR] (>0.15)    0.004
elecPow[VEOGR] (>0.25)    0.065
elecSpe[VEOGR] (>0.95)    0.943
```

Figure 20.

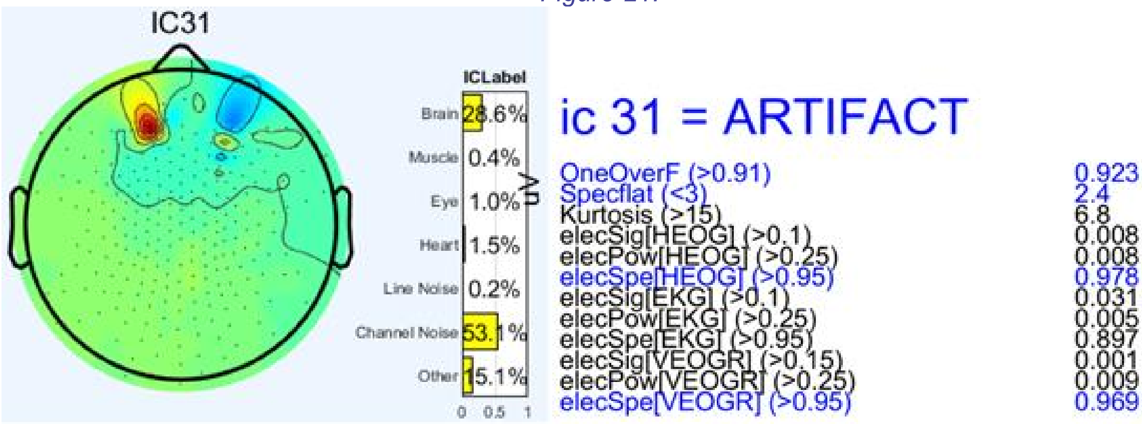


```
ic 29 = ARTIFACT
OneOverF (>0.91)           0.718
Specflat (<3)              1.3
Kurtosis (>15)             4.4
elecSig[HEOG] (>0.1)       0.000
elecPow[HEOG] (>0.25)      0.003
elecSpe[HEOG] (>0.95)      0.910
elecSig[EKG] (>0.1)        0.000
elecPow[EKG] (>0.25)       0.001
elecSpe[EKG] (>0.95)       0.860
elecSig[VEOGR] (>0.15)     0.000
elecPow[VEOGR] (>0.25)     0.004
elecSpe[VEOGR] (>0.95)     0.903
```



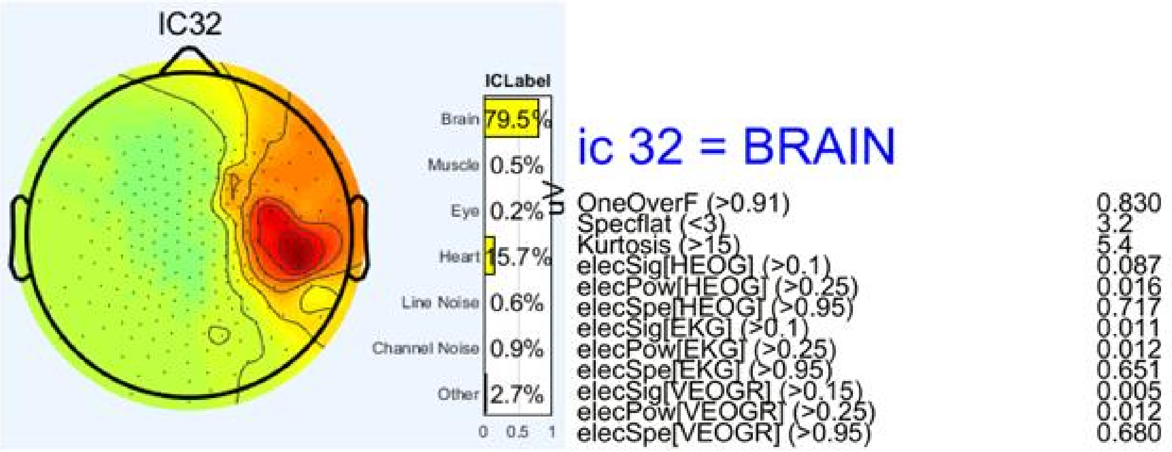
```
ic 30 = BRAIN
OneOverF (>0.91)           0.713
Specflat (<3)              4.0
Kurtosis (>15)             5.3
elecSig[HEOG] (>0.1)       0.020
elecPow[HEOG] (>0.25)      0.023
elecSpe[HEOG] (>0.95)      0.459
elecSig[EKG] (>0.1)        0.033
elecPow[EKG] (>0.25)       0.005
elecSpe[EKG] (>0.95)       0.423
elecSig[VEOGR] (>0.15)     0.011
elecPow[VEOGR] (>0.25)     0.036
elecSpe[VEOGR] (>0.95)     0.431
```

Figure 21.

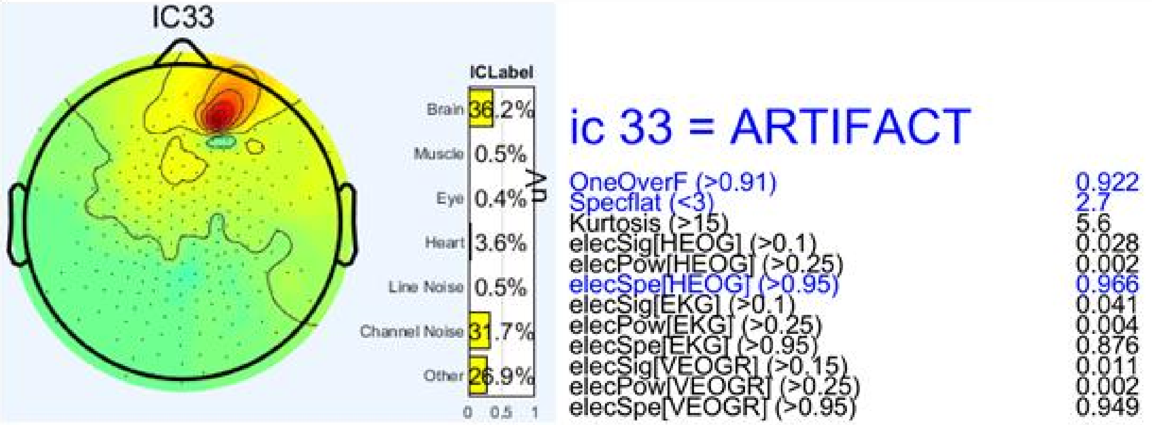


```
ic 31 = ARTIFACT
OneOverF (>0.91)           0.923
Specflat (<3)              2.4
Kurtosis (>15)             6.8
elecSig[HEOG] (>0.1)       0.008
elecPow[HEOG] (>0.25)      0.008
elecSpe[HEOG] (>0.95)      0.978
elecSig[EKG] (>0.1)        0.031
elecPow[EKG] (>0.25)       0.005
elecSpe[EKG] (>0.95)       0.897
elecSig[VEOGR] (>0.15)     0.001
elecPow[VEOGR] (>0.25)     0.009
elecSpe[VEOGR] (>0.95)     0.969
```

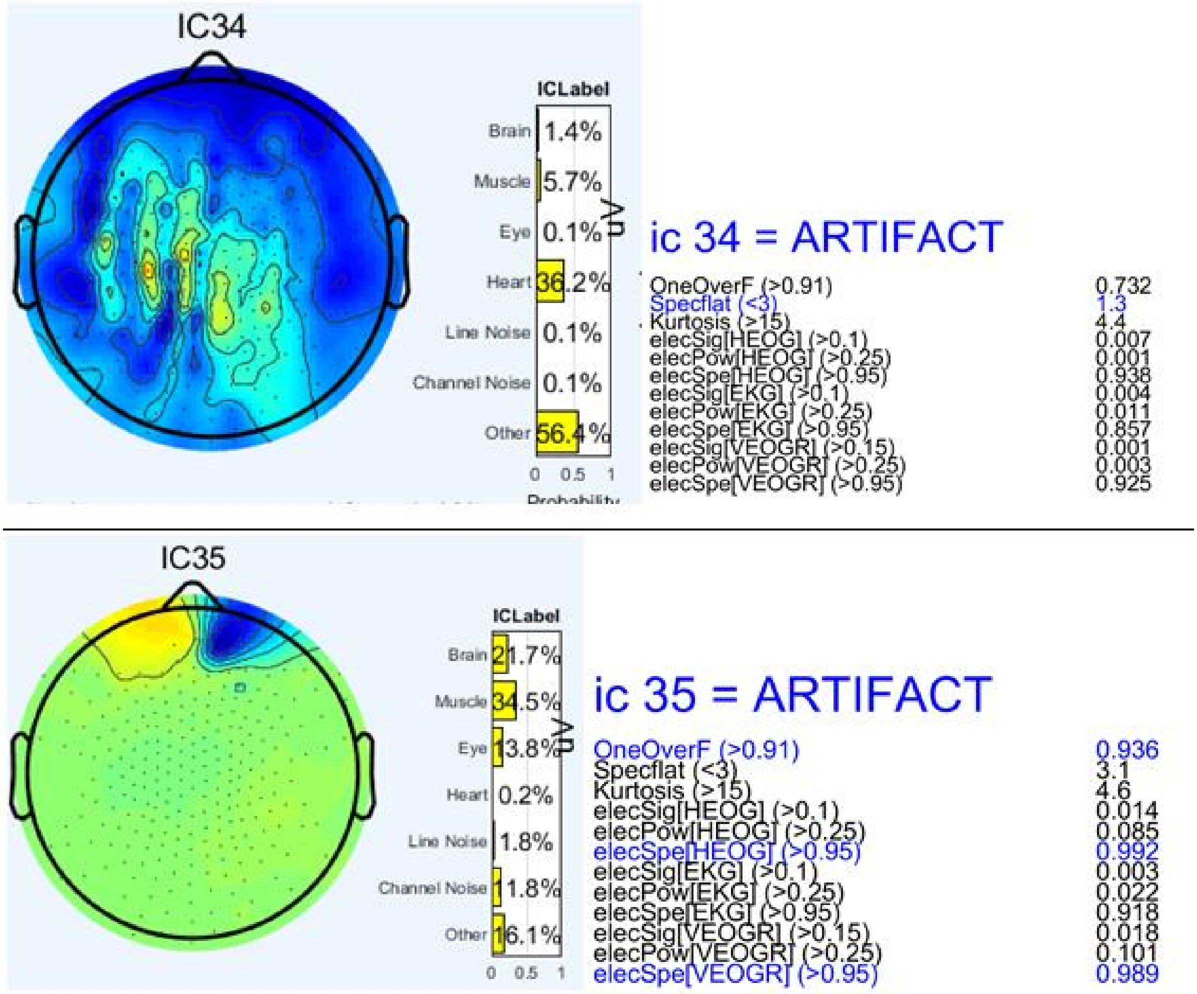
Figure 22.



```
ic 32 = BRAIN
OneOverF (>0.91)           0.830
Specflat (<3)              3.2
Kurtosis (>15)             5.4
elecSig[HEOG] (>0.1)       0.087
elecPow[HEOG] (>0.25)      0.016
elecSpe[HEOG] (>0.95)      0.717
elecSig[EKG] (>0.1)        0.011
elecPow[EKG] (>0.25)       0.012
elecSpe[EKG] (>0.95)       0.651
elecSig[VEOGR] (>0.15)     0.005
elecPow[VEOGR] (>0.25)     0.012
elecSpe[VEOGR] (>0.95)     0.680
```



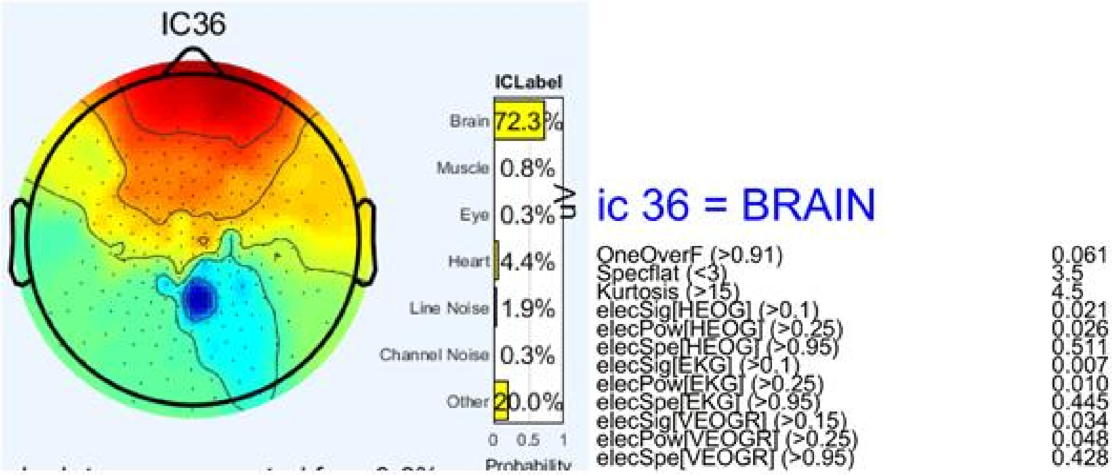
```
ic 33 = ARTIFACT
OneOverF (>0.91)           0.922
Specflat (<3)              2.7
Kurtosis (>15)             5.6
elecSig[HEOG] (>0.1)       0.028
elecPow[HEOG] (>0.25)      0.002
elecSpe[HEOG] (>0.95)      0.966
elecSig[EKG] (>0.1)        0.041
elecPow[EKG] (>0.25)       0.004
elecSpe[EKG] (>0.95)       0.876
elecSig[VEOGR] (>0.15)     0.011
elecPow[VEOGR] (>0.25)     0.002
elecSpe[VEOGR] (>0.95)     0.949
```



```
ic 34 = ARTIFACT
OneOverF (>0.91)          0.732
Specflat (<3)             1.3
Kurtosis (>15)            4.4
elecSig[HEOG] (>0.1)      0.007
elecPow[HEOG] (>0.25)     0.001
elecSpe[HEOG] (>0.95)     0.938
elecSig[EKG] (>0.1)       0.004
elecPow[EKG] (>0.25)      0.011
elecSpe[EKG] (>0.95)      0.857
elecSig[VEOGR] (>0.15)    0.001
elecPow[VEOGR] (>0.25)    0.003
elecSpe[VEOGR] (>0.95)    0.925
```

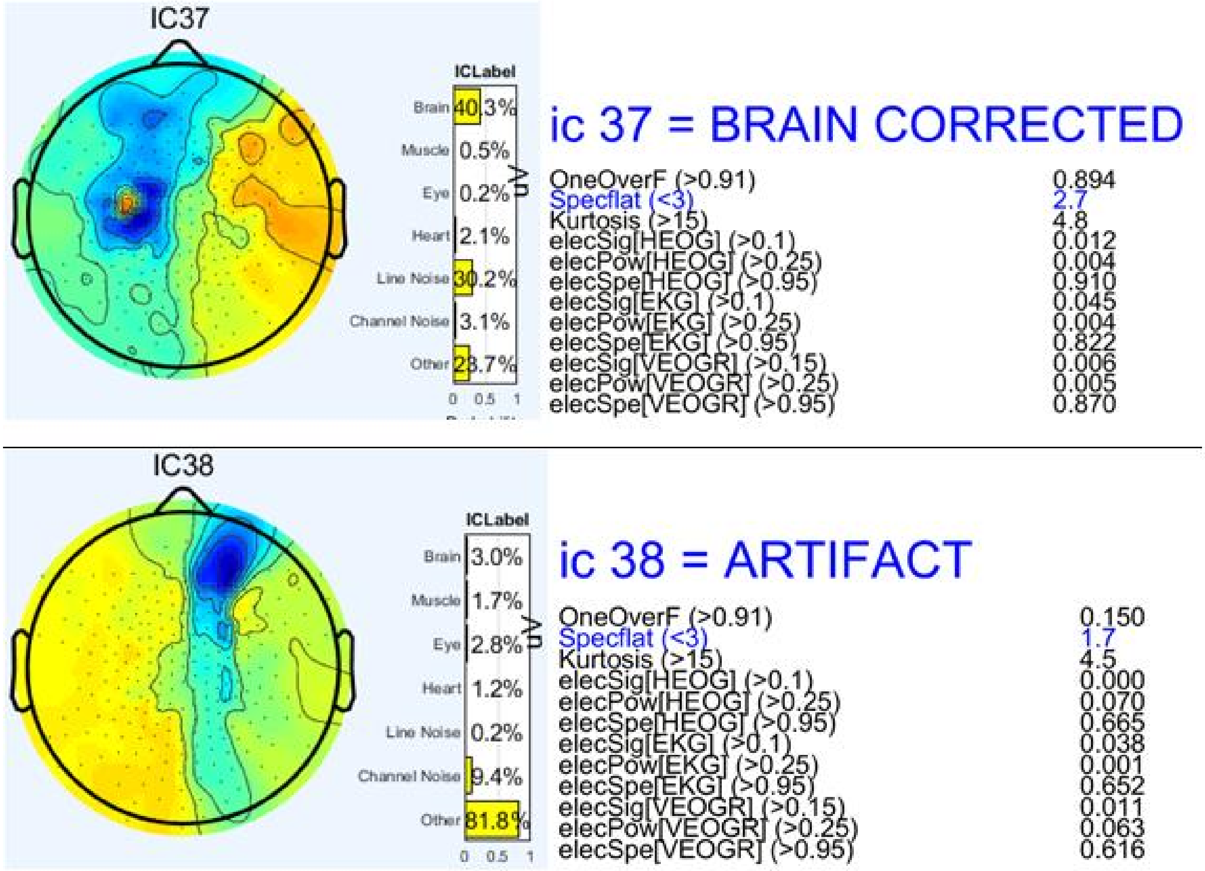
```
ic 35 = ARTIFACT
OneOverF (>0.91)          0.936
Specflat (<3)             3.1
Kurtosis (>15)            4.6
elecSig[HEOG] (>0.1)      0.014
elecPow[HEOG] (>0.25)     0.085
elecSpe[HEOG] (>0.95)     0.992
elecSig[EKG] (>0.1)       0.003
elecPow[EKG] (>0.25)      0.022
elecSpe[EKG] (>0.95)      0.918
elecSig[VEOGR] (>0.15)    0.018
elecPow[VEOGR] (>0.25)    0.101
elecSpe[VEOGR] (>0.95)    0.989
```

Figure 24.



```
ic 36 = BRAIN
OneOverF (>0.91)          0.061
Specflat (<3)             3.5
Kurtosis (>15)            4.5
elecSig[HEOG] (>0.1)      0.021
elecPow[HEOG] (>0.25)     0.026
elecSpe[HEOG] (>0.95)     0.511
elecSig[EKG] (>0.1)       0.007
elecPow[EKG] (>0.25)      0.010
elecSpe[EKG] (>0.95)      0.445
elecSig[VEOGR] (>0.15)    0.034
elecPow[VEOGR] (>0.25)    0.048
elecSpe[VEOGR] (>0.95)    0.428
```

Figure 25.



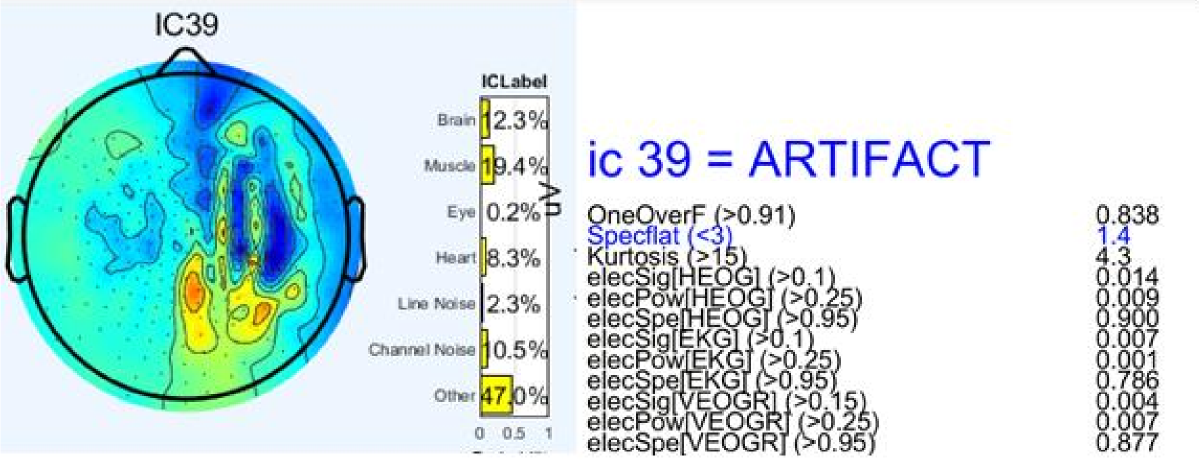
ic 37 = BRAIN CORRECTED

OneOverF (>0.91)	0.894
Specflat (<3)	2.7
Kurtosis (>15)	4.8
elecSig[HEOG] (>0.1)	0.012
elecPow[HEOG] (>0.25)	0.004
elecSpe[HEOG] (>0.95)	0.910
elecSig[EKG] (>0.1)	0.045
elecPow[EKG] (>0.25)	0.004
elecSpe[EKG] (>0.95)	0.822
elecSig[VEOGR] (>0.15)	0.006
elecPow[VEOGR] (>0.25)	0.005
elecSpe[VEOGR] (>0.95)	0.870

ic 38 = ARTIFACT

OneOverF (>0.91)	0.150
Specflat (<3)	1.7
Kurtosis (>15)	4.5
elecSig[HEOG] (>0.1)	0.000
elecPow[HEOG] (>0.25)	0.070
elecSpe[HEOG] (>0.95)	0.665
elecSig[EKG] (>0.1)	0.038
elecPow[EKG] (>0.25)	0.001
elecSpe[EKG] (>0.95)	0.652
elecSig[VEOGR] (>0.15)	0.011
elecPow[VEOGR] (>0.25)	0.063
elecSpe[VEOGR] (>0.95)	0.616

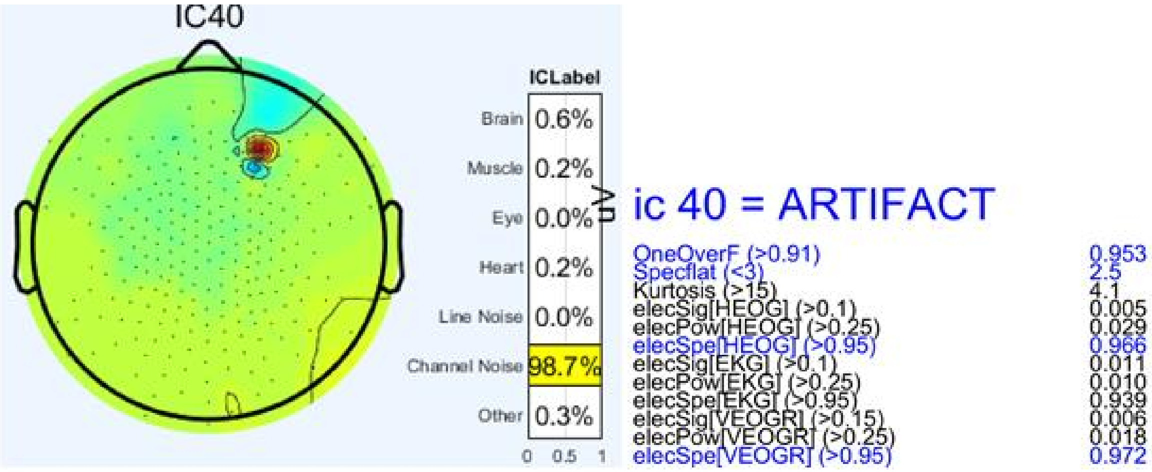
Figure 26.



ic 39 = ARTIFACT

OneOverF (>0.91)	0.838
Specflat (<3)	1.4
Kurtosis (>15)	4.3
elecSig[HEOG] (>0.1)	0.014
elecPow[HEOG] (>0.25)	0.009
elecSpe[HEOG] (>0.95)	0.900
elecSig[EKG] (>0.1)	0.007
elecPow[EKG] (>0.25)	0.001
elecSpe[EKG] (>0.95)	0.786
elecSig[VEOGR] (>0.15)	0.004
elecPow[VEOGR] (>0.25)	0.007
elecSpe[VEOGR] (>0.95)	0.877

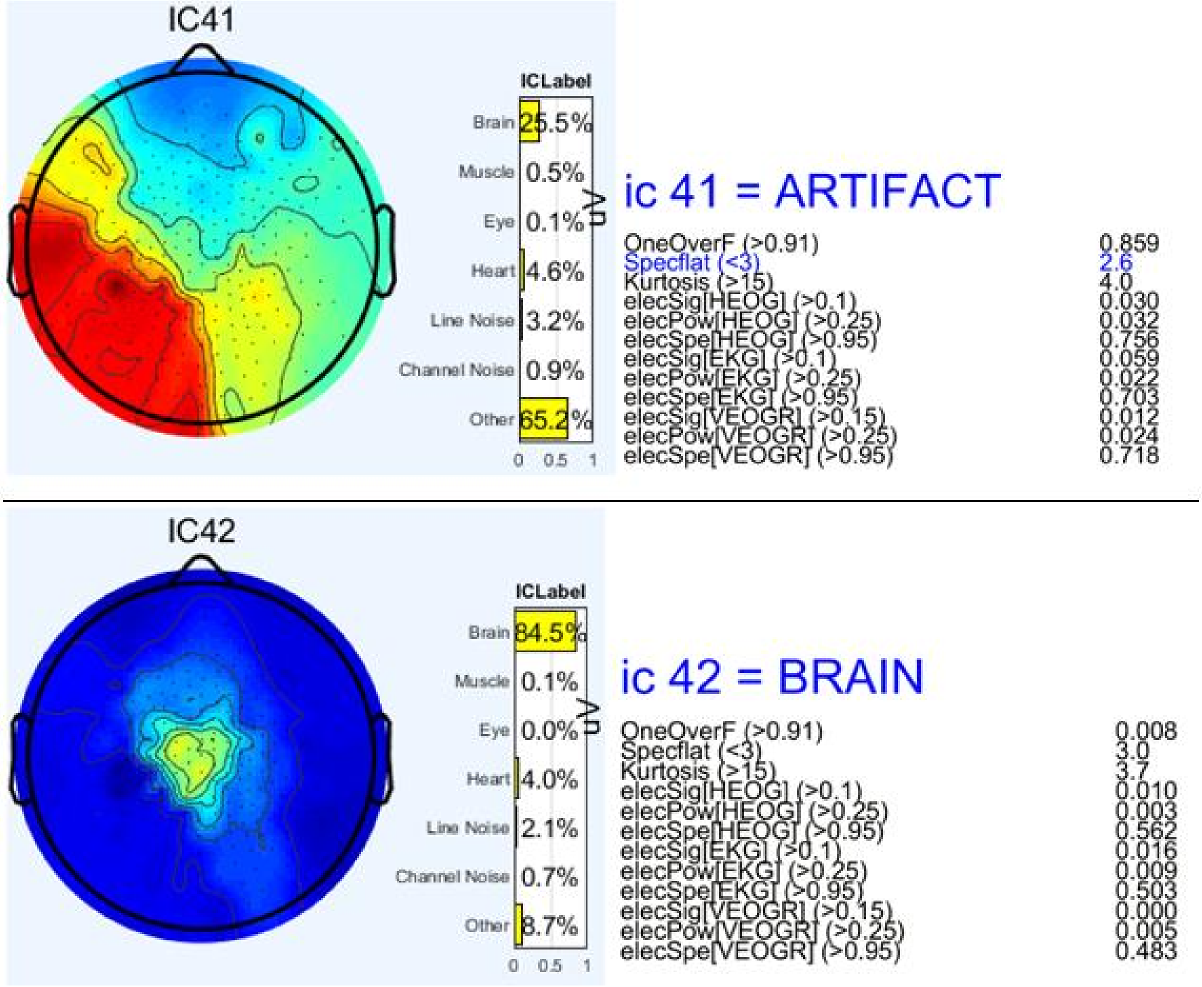
Figure 27.



ic 40 = ARTIFACT

OneOverF (>0.91)	0.953
Specflat (<3)	2.5
Kurtosis (>15)	4.1
elecSig[HEOG] (>0.1)	0.005
elecPow[HEOG] (>0.25)	0.029
elecSpe[HEOG] (>0.95)	0.966
elecSig[EKG] (>0.1)	0.011
elecPow[EKG] (>0.25)	0.010
elecSpe[EKG] (>0.95)	0.939
elecSig[VEOGR] (>0.15)	0.006
elecPow[VEOGR] (>0.25)	0.018
elecSpe[VEOGR] (>0.95)	0.972

Figure 28.



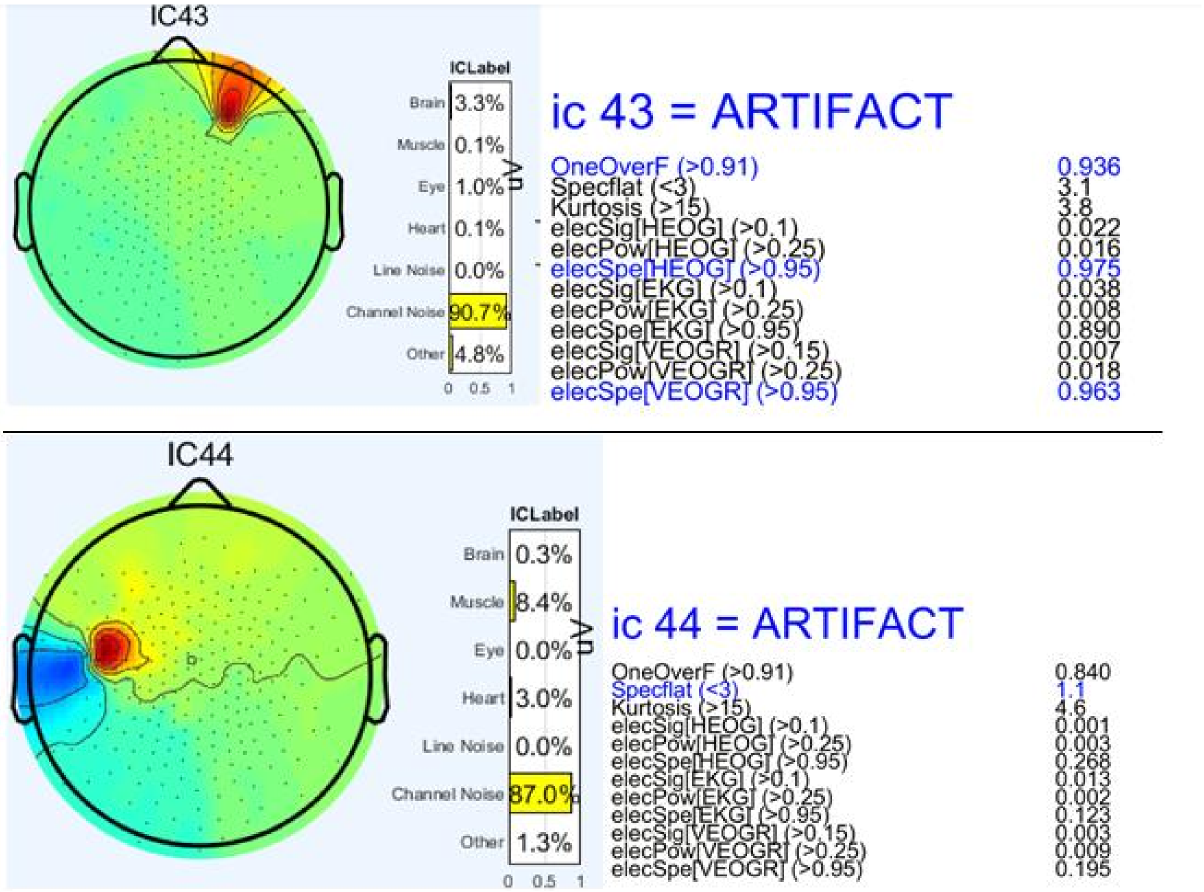
ic 41 = ARTIFACT

OneOverF (>0.91)	0.859
Specflat (<3)	2.6
Kurtosis (>15)	4.0
elecSig[HEOG] (>0.1)	0.030
elecPow[HEOG] (>0.25)	0.032
elecSpe[HEOG] (>0.95)	0.756
elecSig[EKG] (>0.1)	0.059
elecPow[EKG] (>0.25)	0.022
elecSpe[EKG] (>0.95)	0.703
elecSig[VEOGR] (>0.15)	0.012
elecPow[VEOGR] (>0.25)	0.024
elecSpe[VEOGR] (>0.95)	0.718

ic 42 = BRAIN

OneOverF (>0.91)	0.008
Specflat (<3)	3.0
Kurtosis (>15)	3.7
elecSig[HEOG] (>0.1)	0.010
elecPow[HEOG] (>0.25)	0.003
elecSpe[HEOG] (>0.95)	0.562
elecSig[EKG] (>0.1)	0.016
elecPow[EKG] (>0.25)	0.009
elecSpe[EKG] (>0.95)	0.503
elecSig[VEOGR] (>0.15)	0.000
elecPow[VEOGR] (>0.25)	0.005
elecSpe[VEOGR] (>0.95)	0.483

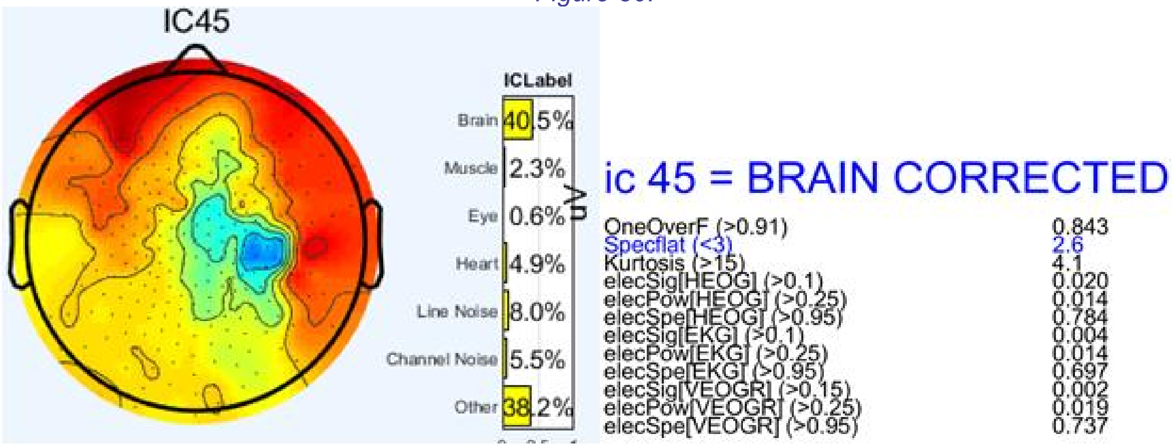
Figure 29.



```
ic 43 = ARTIFACT

OneOverF (>0.91)           0.936
Specflat (<3)              3.1
Kurtosis (>15)             3.8
elecSig[HEOG] (>0.1)       0.022
elecPow[HEOG] (>0.25)      0.016
elecSpe[HEOG] (>0.95)      0.975
elecSig[EKG] (>0.1)        0.038
elecPow[EKG] (>0.25)       0.008
elecSpe[EKG] (>0.95)       0.890
elecSig[VEOGR] (>0.15)     0.007
elecPow[VEOGR] (>0.25)     0.018
elecSpe[VEOGR] (>0.95)     0.963
```

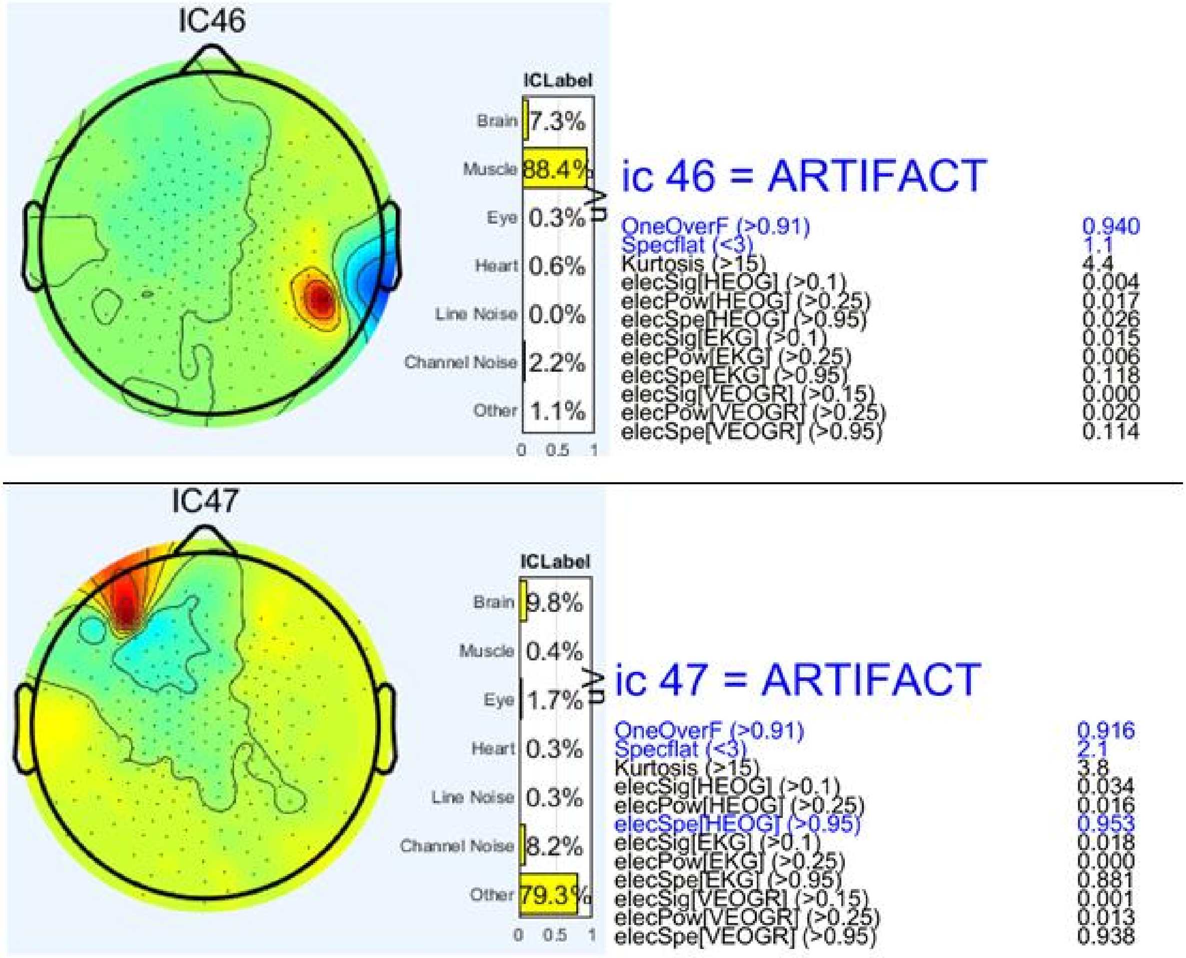
```
ic 44 = ARTIFACT

OneOverF (>0.91)           0.840
Specflat (<3)              1.1
Kurtosis (>15)             4.6
elecSig[HEOG] (>0.1)       0.001
elecPow[HEOG] (>0.25)      0.003
elecSpe[HEOG] (>0.95)      0.268
elecSig[EKG] (>0.1)        0.013
elecPow[EKG] (>0.25)       0.002
elecSpe[EKG] (>0.95)       0.123
elecSig[VEOGR] (>0.15)     0.003
elecPow[VEOGR] (>0.25)     0.009
elecSpe[VEOGR] (>0.95)     0.195
```

Figure 30.



```
ic 45 = BRAIN CORRECTED

OneOverF (>0.91)           0.843
Specflat (<3)              2.6
Kurtosis (>15)             4.1
elecSig[HEOG] (>0.1)       0.020
elecPow[HEOG] (>0.25)      0.014
elecSpe[HEOG] (>0.95)      0.784
elecSig[EKG] (>0.1)        0.004
elecPow[EKG] (>0.25)       0.014
elecSpe[EKG] (>0.95)       0.697
elecSig[VEOGR] (>0.15)     0.002
elecPow[VEOGR] (>0.25)     0.019
elecSpe[VEOGR] (>0.95)     0.737
```

Figure 31.



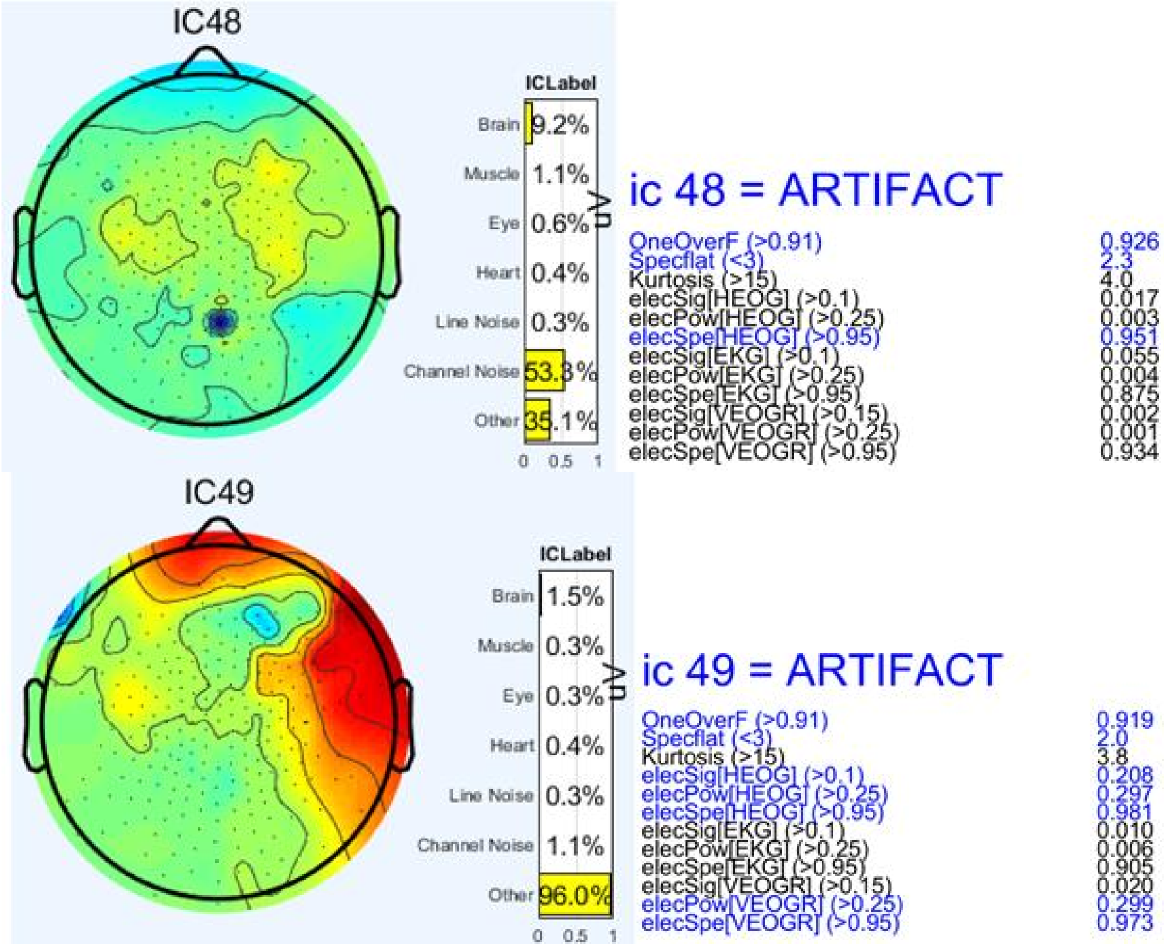
ic 46 = ARTIFACT

```
OneOverF (>0.91)          0.940
Specflat (<3)             1.1
Kurtosis (>15)            4.4
elecSig[HEOG] (>0.1)      0.004
elecPow[HEOG] (>0.25)     0.017
elecSpe[HEOG] (>0.95)     0.026
elecSig[EKG] (>0.1)       0.015
elecPow[EKG] (>0.25)      0.006
elecSpe[EKG] (>0.95)      0.118
elecSig[VEOGR] (>0.15)    0.000
elecPow[VEOGR] (>0.25)    0.020
elecSpe[VEOGR] (>0.95)    0.114
```

ic 47 = ARTIFACT

```
OneOverF (>0.91)          0.916
Specflat (<3)             2.1
Kurtosis (>15)            3.8
elecSig[HEOG] (>0.1)      0.034
elecPow[HEOG] (>0.25)     0.016
elecSpe[HEOG] (>0.95)     0.953
elecSig[EKG] (>0.1)       0.018
elecPow[EKG] (>0.25)      0.000
elecSpe[EKG] (>0.95)      0.881
elecSig[VEOGR] (>0.15)    0.001
elecPow[VEOGR] (>0.25)    0.013
elecSpe[VEOGR] (>0.95)    0.938
```

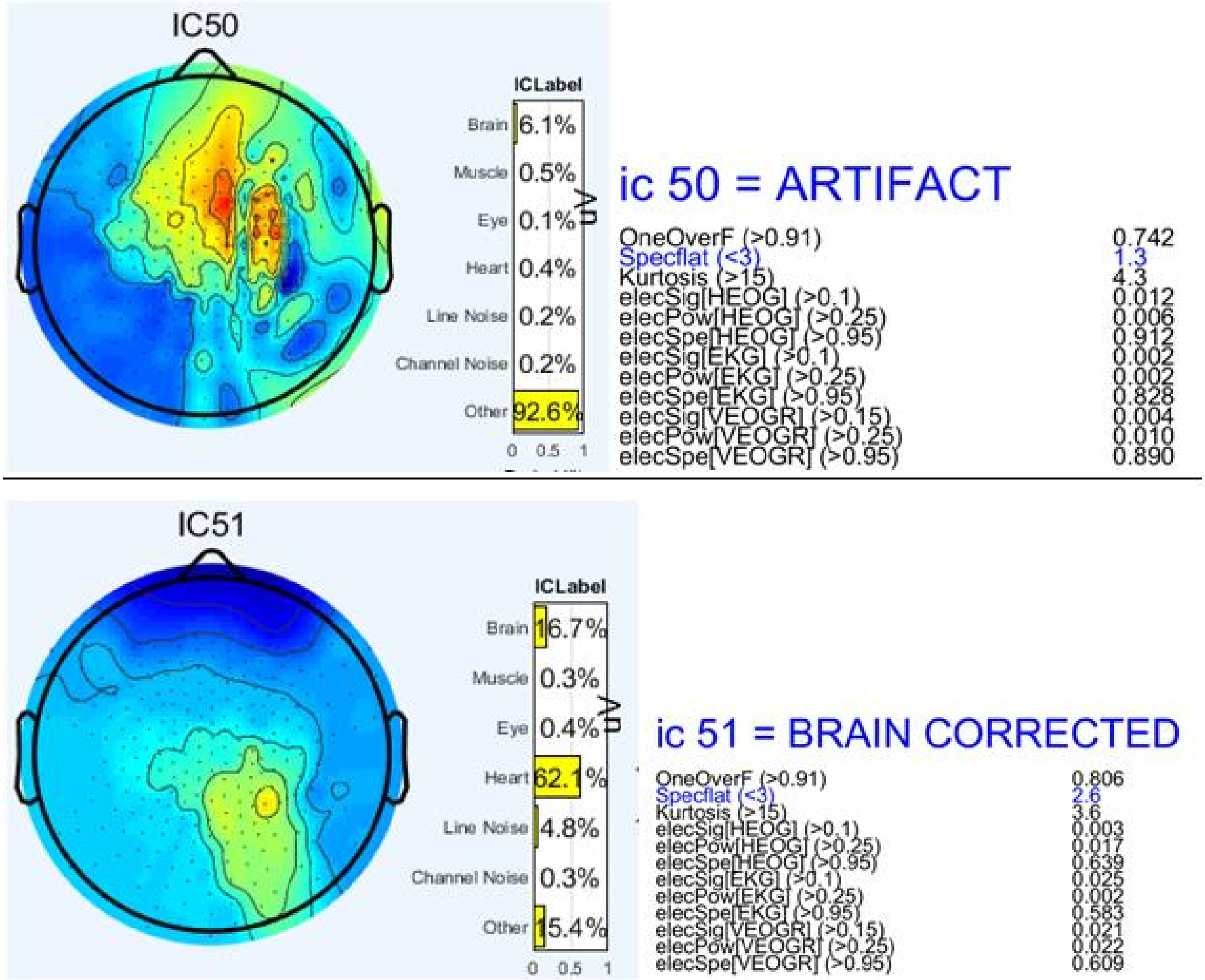
Figure 32.



```
ic 48 = ARTIFACT
OneOverF (>0.91)              0.926
Specflat (<3)                 2.3
Kurtosis (>15)                4.0
elecSig[HEOG] (>0.1)          0.017
elecPow[HEOG] (>0.25)         0.003
elecSpe[HEOG] (>0.95)         0.951
elecSig[EKG] (>0.1)           0.055
elecPow[EKG] (>0.25)          0.004
elecSpe[EKG] (>0.95)          0.875
elecSig[VEOGR] (>0.15)        0.002
elecPow[VEOGR] (>0.25)        0.001
elecSpe[VEOGR] (>0.95)        0.934
```

```
ic 49 = ARTIFACT
OneOverF (>0.91)              0.919
Specflat (<3)                 2.0
Kurtosis (>15)                3.8
elecSig[HEOG] (>0.1)          0.208
elecPow[HEOG] (>0.25)         0.297
elecSpe[HEOG] (>0.95)         0.981
elecSig[EKG] (>0.1)           0.010
elecPow[EKG] (>0.25)          0.006
elecSpe[EKG] (>0.95)          0.905
elecSig[VEOGR] (>0.15)        0.020
elecPow[VEOGR] (>0.25)        0.299
elecSpe[VEOGR] (>0.95)        0.973
```

Figure 34.



```
ic 50 = ARTIFACT
OneOverF (>0.91)              0.742
Specflat (<3)                 1.3
Kurtosis (>15)                4.3
elecSig[HEOG] (>0.1)          0.012
elecPow[HEOG] (>0.25)         0.006
elecSpe[HEOG] (>0.95)         0.912
elecSig[EKG] (>0.1)           0.002
elecPow[EKG] (>0.25)          0.002
elecSpe[EKG] (>0.95)          0.828
elecSig[VEOGR] (>0.15)        0.004
elecPow[VEOGR] (>0.25)        0.010
elecSpe[VEOGR] (>0.95)        0.890
```

```
ic 51 = BRAIN CORRECTED
OneOverF (>0.91)              0.806
Specflat (<3)                 2.6
Kurtosis (>15)                3.6
elecSig[HEOG] (>0.1)          0.003
elecPow[HEOG] (>0.25)         0.017
elecSpe[HEOG] (>0.95)         0.639
elecSig[EKG] (>0.1)           0.025
elecPow[EKG] (>0.25)          0.002
elecSpe[EKG] (>0.95)          0.583
elecSig[VEOGR] (>0.15)        0.021
elecPow[VEOGR] (>0.25)        0.022
elecSpe[VEOGR] (>0.95)        0.609
```

Figure 35. Comparing ICLabel output with semi-automatic single by single in first subject[35]