

RESEARCH ARTICLE

Person identification from EEG using various machine learning techniques with inter-hemispheric amplitude ratio

Isuru Jayarathne^{1*}, Michael Cohen¹, Senaka Amarakeerthi²

1 Spatial Media Group, University of Aizu, Aizu-Wakamatsu, Fukushima, Japan, **2** Dept. of Information and Communication Technology, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka

* d8191101@u-aizu.ac.jp

Abstract

Association between electroencephalography (EEG) and individually personal information is being explored by the scientific community. Though person identification using EEG is an attraction among researchers, the complexity of sensing limits using such technologies in real-world applications. In this research, the challenge has been addressed by reducing the complexity of the brain signal acquisition and analysis processes. This was achieved by reducing the number of electrodes, simplifying the critical task without compromising accuracy. Event-related potentials (ERP), a.k.a. time-locked stimulation, was used to collect data from each subject's head. Following a relaxation period, each subject was visually presented a random four-digit number and then asked to think of it for 10 seconds. Fifteen trials were conducted with each subject with relaxation and visual stimulation phases preceding each mental recall segment. We introduce a novel derived feature, dubbed Inter-Hemispheric Amplitude Ratio (IHAR), which expresses the ratio of amplitudes of laterally corresponding electrode pairs. The feature was extracted after expanding the training set using signal augmentation techniques and tested with several machine learning (ML) algorithms, including Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and k-Nearest Neighbor (kNN). Most of the ML algorithms showed 100% accuracy with 14 electrodes, and according to our results, perfect accuracy can also be achieved using fewer electrodes. However, AF3, AF4, F7, and F8 electrode combination with kNN classifier which yielded 99.0±0.8% testing accuracy is the best for person identification to maintain both user-friendliness and performance. Surprisingly, the relaxation phase manifested the highest accuracy of the three phases.

OPEN ACCESS

Citation: Jayarathne I, Cohen M, Amarakeerthi S (2020) Person identification from EEG using various machine learning techniques with inter-hemispheric amplitude ratio. PLoS ONE 15(9): e0238872. <https://doi.org/10.1371/journal.pone.0238872>

Editor: Francesco Pappalardo, Università degli Studi di Catania, ITALY

Received: March 14, 2020

Accepted: August 25, 2020

Published: September 11, 2020

Copyright: © 2020 Jayarathne et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Introduction

A person identification system verifies the identity of a given individual from a set of people. In contrast, authentication uses different classification methods (such as one-class classification, template matching, and score level fusion) to confirm identity. Both identification and authentication have the same pre-processing and feature-extraction steps. An identification

process is often conducted in police departments using records or visual information of arrested criminals as biometrics. Even though potential applications in this domain are fewer than those for authentication, the process can be enhanced to create an authentication system. Currently, there are several strategies for both cases, including knowledge (such as of a pass-code), possession (such as of an ID card), and biometric traits. Biometric-based techniques use biological or physiological attributes such as fingerprint, palm-print, iris, and voice to identify someone. Biometrics are usually more convenient and intimately personal compared to other strategies.

A brain-computer interface (BCI) or neural-control interface (NCI) represents direct communication between an external device and an enhanced or exposed brain. Besides functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), which observe changes of blood flow, other methods can track electrical activity, such as electroencephalography (EEG) and magnetoencephalography (MEG). EEG is a non-invasive, electrophysiological monitoring method to record electrical activity of a human brain. International standards have been established regarding placement of electrodes over the human scalp with different resolutions. The 10-20 system is the most popular for low-resolution capture, and the 10-5 system, which allows more than 300 electrodes, has highest resolution. Since EEG is innocuous, medically safe, it has attracted considerable attention from the scientific community for such applications as machine learning, robotics, and health care. This study mainly focuses on person identification using EEG analysis to find unique brain signal patterns.

0.1 Related research

Relevance for individually personal information in EEG signals was revealed in the early 1930s [1]. Even though EEG has low spatial resolution compared to fMRI, several studies have shown considerable accuracy in this domain [2–4]. Pozo-Banos et al. [5] reported in a comprehensive review about EEG subject identification that EEG has subject-specific information. Jayarathne et al. [6] and Almeahmadi et al. [7] compiled comprehensive surveys of EEG-based access control systems. These surveys reported suitability, state of the art of EEG for person identification, and important parameters which must be adjusted to implement an authentication system with desired security and usability levels. To find unique features of EEG signals, researchers have used different kinds of stimulation, including visually evoked potentials (VEP) [8], auditory stimulation, motor movement, and assigned math problems [9]. Also, devices used to capture EEG signals have various attributes, such as number of electrodes, electrode type (dry or wet), and wired or wireless computer connection. Chen et al. have shown that wet electrodes allow better signal quality than dry ones [10]. Nevertheless, dry electrode headsets provide more user-friendliness. Despite perfect accuracy having been achieved in several studies, most systems are not very convenient or user-friendly.

The brain emits different EEG signal patterns according to stimuli and responses. Researchers have confirmed that as stimulus sets become increasingly complex, distinctiveness of subjects becomes higher [5]. Moreover, visual stimulation is the most promising and popular modality in this field. Poulos et al. [11, 12] tested two different approaches with the same computational geometry classification algorithm without using any specific task. They achieved 91% and 95% accuracy for Auto Regression (AR) type of alpha-band EEG signals, extracting spectral features via Fourier Transform (FT), and 84% maximum accuracy for Learning Vector Quantizer (LVQ) network with AR parameters as features [2]. Singhal et al. [13], Das et al. [8], Palaniappan et al. [14], and Yazdani et al. [15] used VEP as a stimulation when capturing EEG signals, achieving 78%, 94.08%, 94.18%, and 100% accuracy respectively. Most approaches have used temporal domain features and achieved considerable accuracy.

In our approach, a new feature called Inter-Hemispherical Amplitude Ratio (IHAR) is introduced, and features are extracted from various mental task data sets to find uniqueness. A similar technique (inter-hemispheric amplitude relationship) was used by Goldstein et al. to observe differences of REM and NREM sleep patterns of humans, cats, and rabbits [16]. Goldstein's approach emphasizes differences of REM and NREM stages of EEG signal for easy observation, although a classification technique was not used. Even though various time- and frequency-domain features have been proposed, feature extraction method with low computational complexity is better for embedded system implementations with low computational power budget. Several studies can be found about various asymmetries of the head, concerning such topics as skull asymmetry [17], structural cerebral asymmetry [18], and functional cerebral asymmetry [19]. Functional asymmetry has been analyzed using fMRI, PET [20], and EEG in different domains. Based on functional asymmetry of the brain, IHAR was calculated to find subjective features to distinguish people.

Classification is the most popular supervised learning problem domain in machine learning (ML). This process is concerned with identifying the sub-population to which each given observation belongs, based on a training set of tagged data containing observations whose sub-populations are known. There are various algorithms to solve this kind of problem, including Linear Discrimination Analysis (LDA), Support Vector Machine (SVM), k-Nearest Neighbor (kNN), and Decision Trees. Before training these algorithms, features must be extracted from the raw data. The above-described IHAR feature was extracted and tested with different ML classification algorithms, discussed in next section. Furthermore, we noted the accuracy while reducing the number of channels to increase user-friendliness of EEG capturing devices, because less calibration time is required for minimal electrode setups.

1 Methodology

1.1 Experimental setup and apparatus

The same dataset used in a previous experiment [21] was used to perform the analysis reported here. Twelve mentally healthy subjects (6 males and 6 females, ages 24 to 45) were recruited for this experiment, including students and academic staff at the Sabaragamuwa University of Sri Lanka. An equal number of males and female subjects were selected to avoid gender bias. The experiment was performed after obtaining written consent. The Sabaragamuwa University of Sri Lanka ethics committee waived the need for ethical approval. Since our focus is to implement an identification system for a small group of users, we decided to conduct this experiment with a limited number of users. However, scalability issues of the proposed approach are discussed in the results and discussion section.

Subjects were instructed to stay still while capturing EEG readings, to reduce artifacts of head motion. The EMOTIV EPOC+ headset was used to collect signals in a calm and silent room. Wang et al. [22] had compared the EMOTIV headset with a clinical grade EEG device and confirmed that it produces reliable signals. Fig 1 shows the electrode distribution of the headset. The EPOC+ headset has 14 wet electrodes, and saline solution was used to increase their conductance. It took 10 to 15 minutes to set up the headset for each subject. A software package comes with an accessory to inject time-stamp markers, which can be used to separate trials and phases. The device was precisely tested before the experiment. The effect of ambient electromagnetic waves, including power lines, was minimized by isolating the experimental location, away from EMI (electromagnetic interference) of noise-producing equipment.

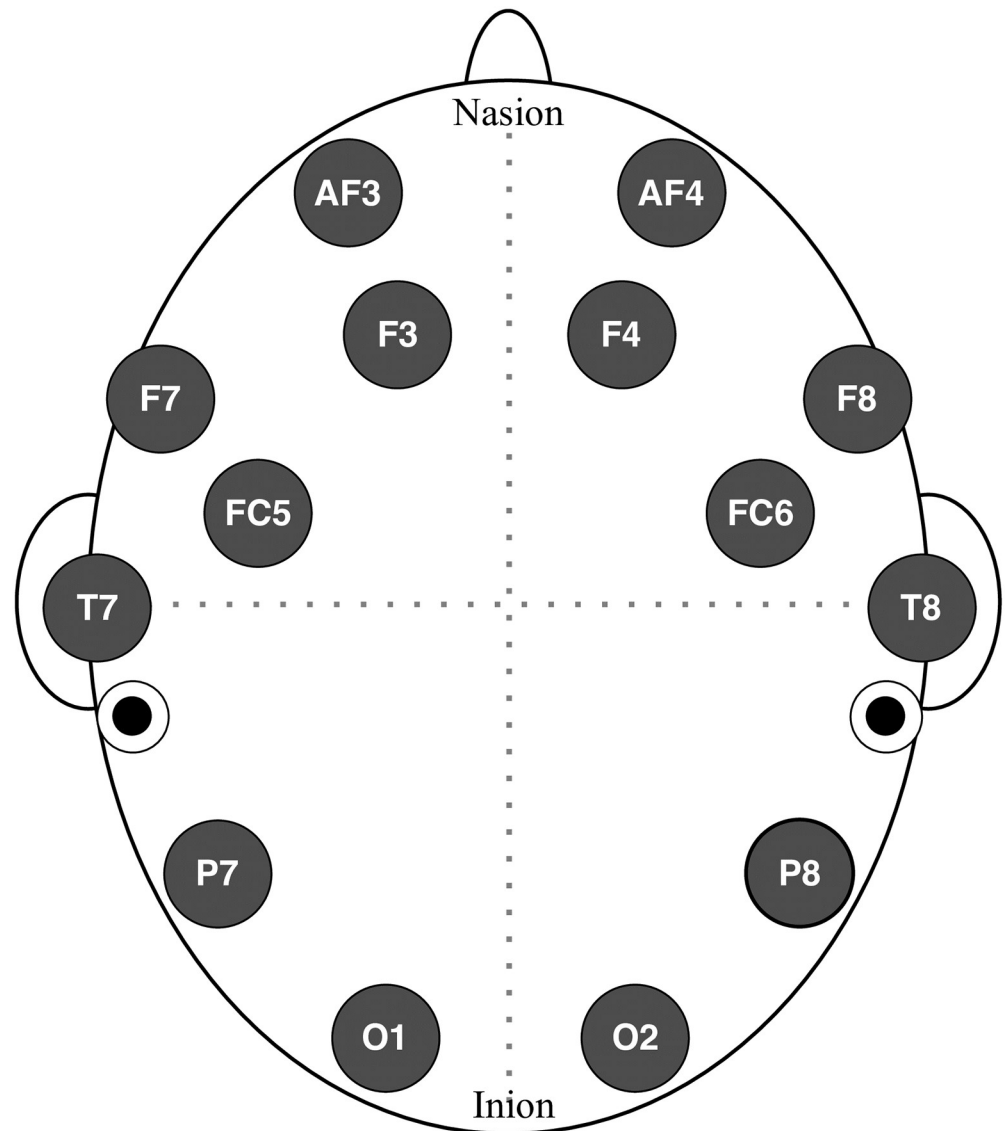


Fig 1. Electrode placement of the EMOTIV EPOC+ EEG headset (AF: Anterior-Frontal, F: Frontal, FC: Frontal-Central, T: Temporal, P: Parietal, O: Occipital), where odd numbers denote left hemisphere and even numbers represent right. Black dots indicate location of reference electrodes (“device ground” or “chassis ground”).

<https://doi.org/10.1371/journal.pone.0238872.g001>

1.2 Trial organization

Each experimental trial included three main phases—relaxation, visual stimulation, and mental recall—as shown in Fig 2. Calm music (“Magic Forest,” by Alexander Blu) was used to relax each subject with closed eyes for 10 s. Then a four-digit random number (white characters on black background) was shown on a screen (17” LCD monitor, 1366 × 768 pixel resolution) for 10 s. Finally, the subject was instructed to imagine the number seen on the screen for 10 s with closed eyes. The four-digit number was used assuming it helps to evoke more distinctive brain signal patterns in visual and mental recall phases. A single trial consisted of these three phases, and trials were recorded continuously with 128 Hz sampling rate. For each subject, 15 trials were conducted, separating trials using time-stamped markers. An inconvenience of the wet electrode system is that the electrodes dry out over time, lowering signal quality. Therefore, the

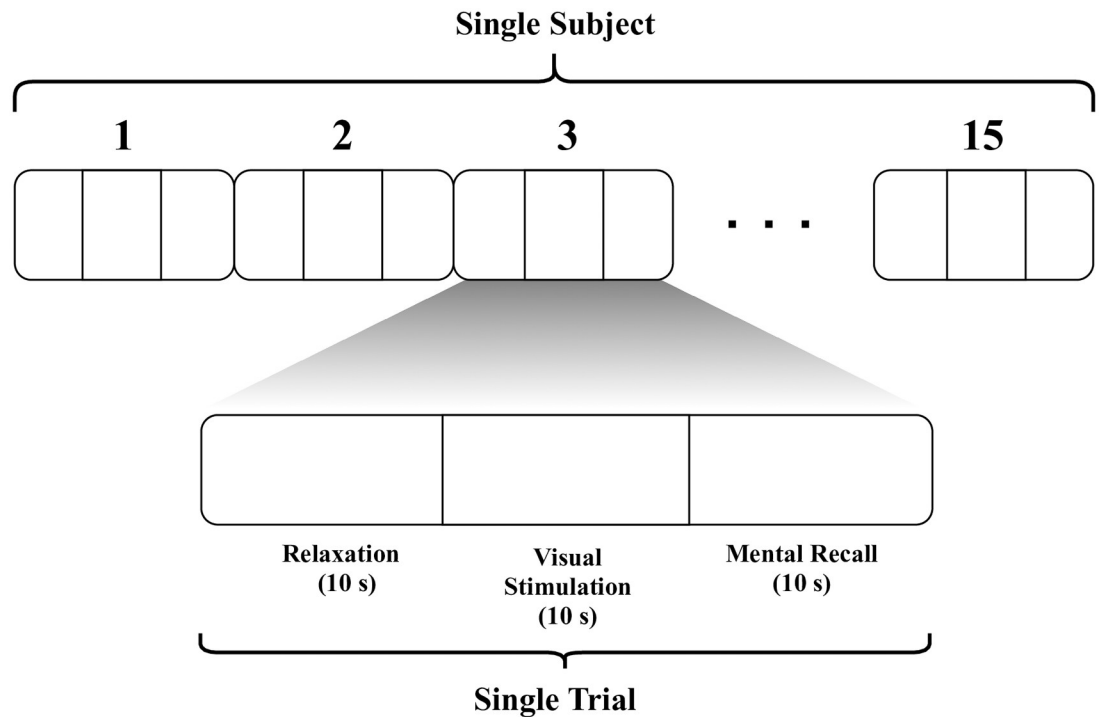


Fig 2. Structure of experiment: Each subject underwent 15 trials, and each trial comprised three phases. The stimulation phase necessarily precedes recall phase, and relaxation phase also serves as a refractory period, an implicit interval between trials, to be consistent with similar studies. (Each subject has tested for about 10 min).

<https://doi.org/10.1371/journal.pone.0238872.g002>

time allocated to each subject was determined by the interval which allowed consistent signal quality.

1.3 Pre-processing and visualization

One way of visualizing collected EEG data is by plotting a topographic distribution. Even though topographs are usually used in geomatics, here the power distribution is plotted as the third dimension (instead of elevation) as a contour map. First, the data was normalized and converted to the frequency domain, using Discrete Fourier Transform (DFT) to calculate the power of each channel:

$$\hat{x}[n] = \frac{x[n] - \min(x)}{\max(x) - \min(x)}, \tag{1}$$

where $x[n]$ is a sequence of raw samples and $\hat{x}[n]$ is the corresponding vector of normalized samples. The DFT is calculated as

$$\hat{X}[k] = \sum_{n=0}^{N-1} \hat{x}[n] e^{-j \left(\frac{2\pi}{N}\right) nk}, \tag{2}$$

where $\hat{X}[k]$ are the frequency-domain coefficients, N is the window size, and $k = 0, 1, 2, \dots, N - 1$. Finally power is calculated as

$$P = \hat{X} \cdot \bar{\hat{X}} = \sum |\hat{X}[k]|^2, \tag{3}$$

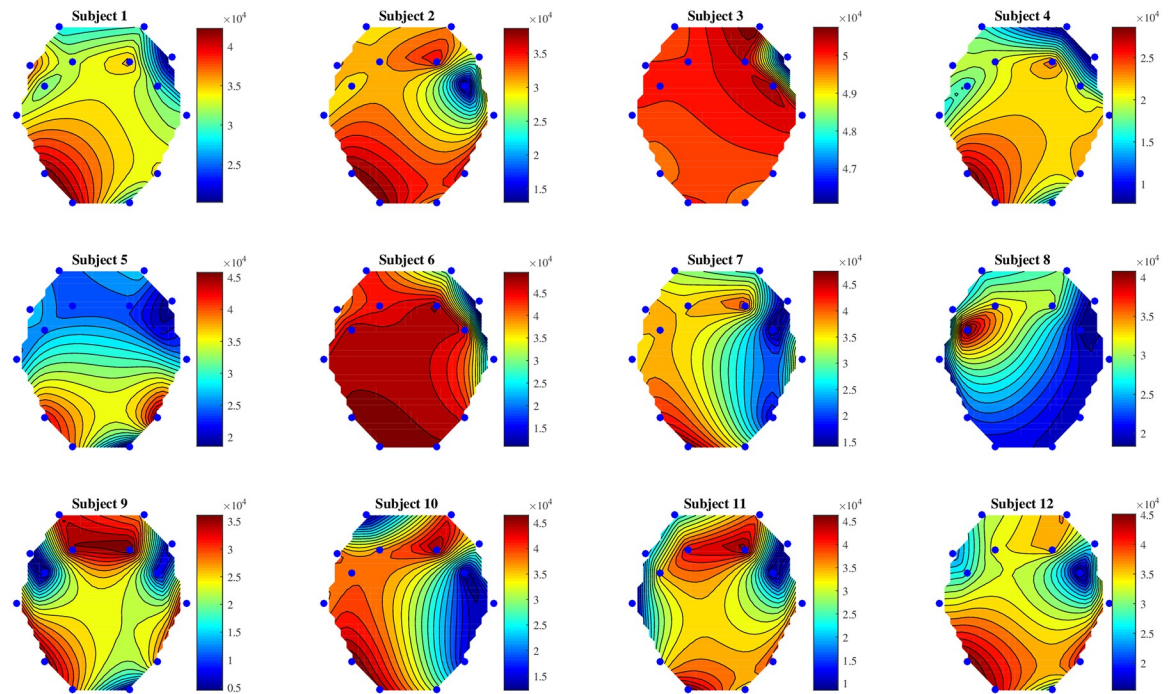


Fig 3. Topographic plots of 12 subjects for relaxation phase. (Maroon color represents high power values, and dark blue represents low).

<https://doi.org/10.1371/journal.pone.0238872.g003>

where \bar{X} is the complex conjugate of the sample sequence, and P is power (strictly speaking, energy) of the channel. After calculating the power values, Natural Neighbor Interpolation [23] is used to infer power values between the electrodes. Plotting the contours for the calculated values gives presumed topographic fields. Our MATLAB code is shared on GitHub <https://github.com/ijmax/EEG-processing/blob/master/topograph.m>.

The three trial phases (relaxation, visual stimulation, and mental recall) were separated from the continuous signals and saved as separate datasets. Each phase was set to fixed duration to maintain consistency, but the actual length differed slightly from phase to phase. Each phase for each trial for each subject had around 1280 samples because each signal was recorded at 128 Hz sampling rate for 10 seconds. The sample length of each phase was truncated to a fixed size (1280) by removing excess samples from the end. When using 14 electrode channels, each phase can be considered a 14×1280 matrix. Each phase set contains 180 trials, because 12 subjects were recruited and 15 trials were conducted for each subject. Fig 3 shows the topographic distribution plots of 12 subjects. Clear differences among these power distributions can be seen through this visualization. Every subject seems to have a uniquely personal brain wave patterns, even when performing the same task. Also, maximum and minimum power values differ from subject to subject.

1.4 Data augmentation

Data augmentation is a technique to increase the diversity of training data to improve results and avoid overfitting without collecting new data. Due to lack of trials per subject in our experiment, data augmentation techniques were used to expand the dataset. These techniques are usually domain-specific. Therefore, three signal augmentation techniques which have been used to improve classification performance of wearable sensor data was used [24].

Fig 4 shows visualization of raw and augmented signals using following techniques:

- *Jittering*: This is a process of adding random noise to the signal. Normally distributed random noise was generated with standard deviation of 0.05.
- *Random sampling*: In this process, some samples are randomly removed, interpolation of remaining samples recovering the original length of the signal.
- *Permutation*: In this process, the temporal location of within-window events are randomly shuffled. 4 randomly selected segments with random length (minimum 5 samples) were altered randomly. To allow symmetry testing such as IHAR, time-wise alignment must be preserved, so the same permutation was applied to all 14 electrode channels.

Ten trials were separated after shuffling (to avoid bias from ordering such as learning effects) from each subject as training data and each 10 second trial was used to generate 10 instances from each augmentation technique, resulting 300 new trials. Final training set contained $12 \times 310 = 3720$ trials in total. Remaining 5 trials for each subject were kept as testing set. The same process was performed for all three phases.

Scatter plots of some selected features as seen in Fig 5 were used to confirm that distribution of features of augmented signals were not changed significantly. According to the scatter plots, augmentation has widened the margins of each class while preserving the separable nature of features.

1.5 Inter-Hemispheric Amplitude Ratio (IHAR) vs. other features (LC, WL, SSC, AR)

There is asymmetry between brain hemispheres [25]. In human neuroanatomy, brain asymmetry can be categorized into two classes: (a) neuroanatomical differences such as neuronal

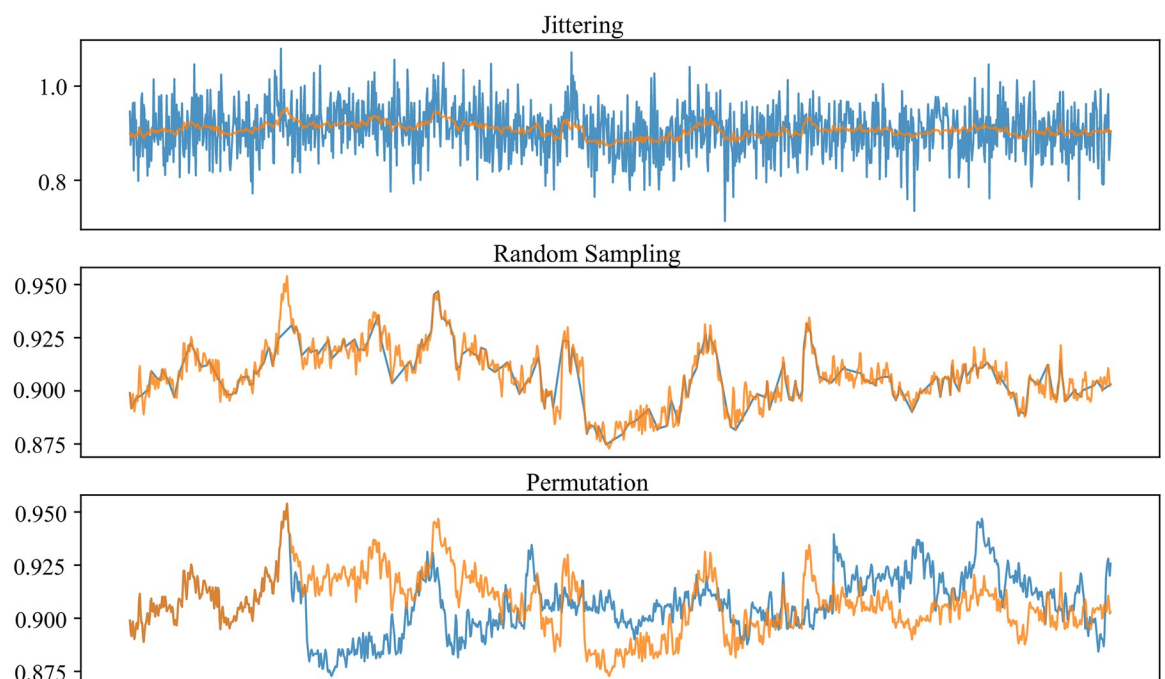


Fig 4. Visualization of raw and augmented signals subject to jittering, random sampling, and permutation techniques. Orange trace represents the original raw signal, and blue trace represents the augmented signal in each plot.

<https://doi.org/10.1371/journal.pone.0238872.g004>

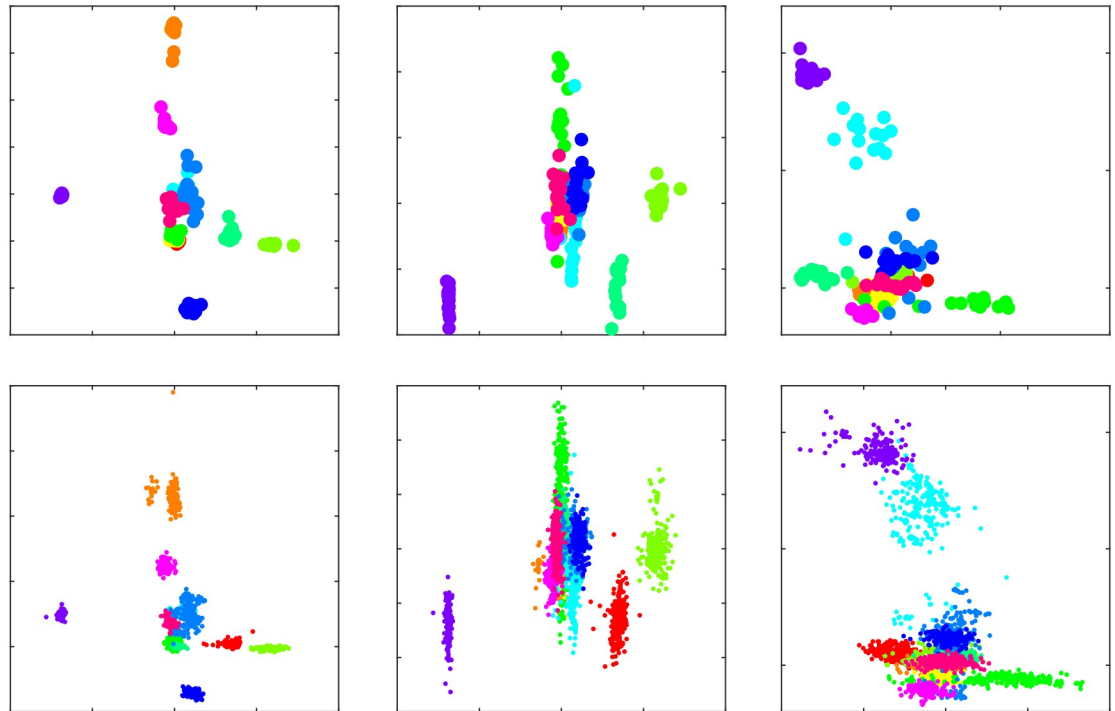


Fig 5. Distribution of features before and after the augmentation, using color to distinguish respective subjects. Top 3 plots show features of 10 training trials, while bottom 3 plots show corresponding features of augmented trials.

<https://doi.org/10.1371/journal.pone.0238872.g005>

densities, size of regions, etc., and (b) functional differences [26]. Such asymmetry can be seen in the power distributions of EEG data. All standard electrode placement systems (10-20, 10-10, 10-5) are bilaterally symmetric across theinion-nasion line, the median or sagittal plane. Also, most experimental results have shown that time-domain data contains discriminative aspects such as Autoregressive (AR) features [1], matching peaks [13, 27], and statistical features [28]. IHAR is also calculated in the time domain, and symmetric electrode pairs are needed. The proposed feature extraction method has two stages: (1) smoothing and (2) calculating ratio of corresponding channel pairs. Instead of using frequency-domain low-pass filtering for smoothing, a moving average (MA) filter [29] was used to remove high frequencies and random noise. The MA filter, applied in our analysis only for the IHAR feature, is an efficient way of extracting low-frequency signals compared to frequency-domain low-pass filters. It can also be calculated quickly, so is appropriate for modest embedded systems, such as anticipated practical deployment. Furthermore, power line noise can be easily removed from a captured raw signal [30].

A similar derived parameter, dubbed “laterality coefficient” (LC), had been introduced [31] to study motor rehabilitation after a stroke. However, LC has not been used to find subjective differences among human subjects. Comparison of feature distribution of LC and IHAR can be seen in Fig 6. The same feature pair for both IHAR and LC were plotted, and IHAR showed more discrimination compared to LC as suggested by heterogeneous distributions across varied feature pairs.

Fig 7 shows the frequency response for a MA filter with three different window sizes. Three comprehensive analyses were performed, for window sizes 4, 16, and 32. Unlike other low-pass filters, MA filter does not have a sharp cut-off. The frequency response of an MA filter

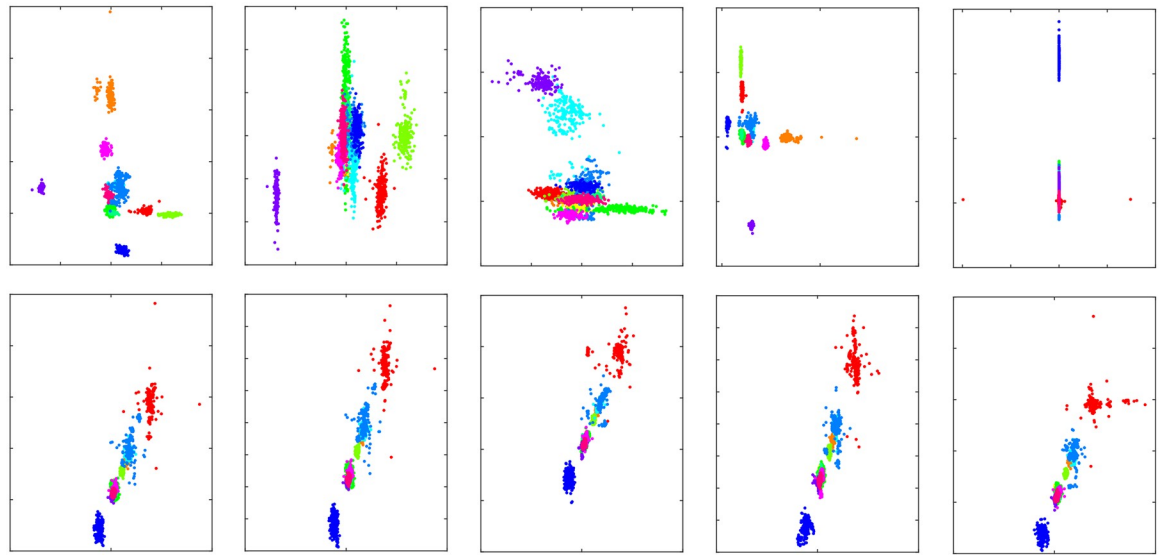


Fig 6. Comparison of selected feature distribution of IHAR and LC. Top 5 show scatter plots of IHAR, while bottom 5 plots show LC of corresponding feature pairs.

<https://doi.org/10.1371/journal.pone.0238872.g006>

can be expressed as

$$H(\omega) = \frac{1}{L} \left(\frac{1 - e^{-j\omega L}}{1 - e^{-j\omega}} \right), \tag{4}$$

where angular frequency $\omega = 2\pi f$, $e^{-j\omega}$ is complex phasor of filter transfer function, and L is window size. As seen in Fig 7, when the window size is 4, the cut-off frequency (at which

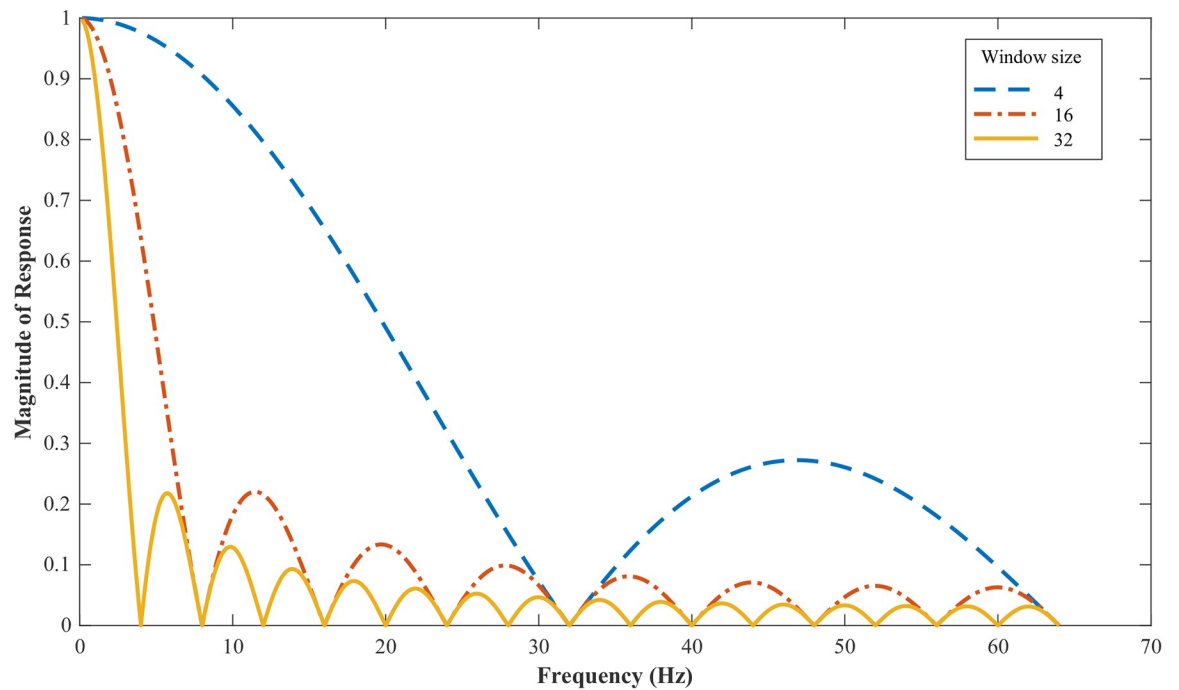


Fig 7. Frequency response of the MA filter with various window sizes.

<https://doi.org/10.1371/journal.pone.0238872.g007>

magnitude is attenuated by 3 dB so full-scale gain is $1/\sqrt{2} \approx .7$ is ~ 15 Hz ($f_s = 128$ Hz), so the passband includes Delta (δ), Theta (θ), and Alpha (α) frequency ranges. Only δ band can be extracted when window size 16 or 32 is used. Most cognitive tasks, including relaxation and mental recall, produce low-frequency EEG signals [32]. Therefore, the above-mentioned window sizes were used to extract information in low-frequency signals. Furthermore, this filtering process eliminates power line noise, which “hum” lies between 50–60 Hz. The MA filter can be expressed as

$$x_A[n] = \frac{1}{L} \sum_{k=0}^{L-1} x[n+k], \quad (5)$$

where L is the window size, and $x[\cdot]$ are the time-domain samples. Finally, IHAR is calculated by dividing samples of a left channel by samples of its corresponding right channel:

$$x_{\text{IHAR}}[i, \tilde{n}] = \frac{x[\text{left}[i], \tilde{n}]}{x[\text{right}[i], \tilde{n}]}, \quad (6)$$

where $i = 1, 2, \dots, 7$ or total number of channel pairs, $\tilde{n} = 1, 2, \dots$, total samples in the phase sequence. $\text{left}[\cdot]$ & $\text{right}[\cdot]$ are left and right channels lists arranged to align corresponding bilateral pairs, as in

$$\begin{aligned} \text{left} &= \text{AF3, F7, F3, FC5, T7, P7, O1;} \\ \text{right} &= \text{AF4, F8, F4, FC6, T8, P8, O2.} \end{aligned}$$

For simplicity, henceforth the “IHAR” subscript is elided, and x is understood to refer to x_{IHAR} .

To validate performance of the proposed feature, several time-domain features were calculated for comparison:

- *Laterality Coefficient (LC)*: ratio of difference and sum of a bilateral EEG pair, $(L - R)/(L + R)$.
- *Waveform Length (WL)*: cumulative length of a particular segment of a signal.
- *Slope Sign Change (SSC)*: number of times the slope of a signal changes its sign.
- *Auto-regressive coefficients (AR)*: AR modeling is a process of predicting future values based on past values of time series data. Scalar values which model the prediction.

Except for LC, these features have especially been used in EMG analysis studies, including limb movement classification. Geethanjali et al. [33] conducted a performance comparison of some of these features using LDA classifier and obtained 67–100% accuracy range for pairwise mental tasks classification. Each feature was extracted using a sliding window similar to MA filter process applied in our calculation of IHAR, so the extraction process was consistent for all features types. To calculate these four comparing features, some functions in “Myoelectric control development toolbox” developed by Chan et al. [34] were used. No smoothing, low-pass filtering, or averaging was applied to these compared features (besides IHAR).

When calculating each of these four features, the kernel window convolves through rows of the trial phase matrix (14×1280). The size of the output matrix after feature calculation depends upon the window size. Then the output matrix is flattened in row-major order to make a feature vector and fed to a classifier.

1.6 Classification

The extracted feature set was tested with several ML techniques to determine which algorithms show high accuracy for person identification. Unlike authentication, person identification is a straightforward classification problem, so traditional ML algorithms can be applied. Not knowing a priori which are best suited for this problem domain, we compared four well-known techniques.

1.6.1 Linear Discriminant Analysis (LDA). LDA [35, 36] is most commonly used as a dimensionality reduction method, similar to Principal Component Analysis (PCA). Assuming equal covariance for each class, and that conditional classes are multivariate, the LDA discriminant function $\delta_k(x)$ (which separates inferred classes) can be defined as in Eq 7 and the classification rule shown in Eq 8:

$$\delta_k(x) = \log\pi_k + x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k, \quad (7)$$

where π_k denotes the prior probability that an observation belongs to the k th class, Σ is the common covariance matrix (not to be confused with the Σ operator for summation), and μ_k is the mean of class k . The classification rule can be defined as

$$\hat{G}(x) = \arg \max_k (\delta_k(x)). \quad (8)$$

1.6.2 Quadratic Discriminant Analysis (QDA). The main difference of QDA from LDA is that QDA relaxes the assumption that the inputs of every class have the same covariance. Also, class decision boundaries are not linear but quadratic [37]. Eq 9 shows the QDA discriminant function. The classification rule is the same as that for LDA.

$$\delta_k(x) = \log\pi_k - \frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) - \frac{1}{2} \log|\Sigma_k|, \quad (9)$$

where Σ_k is the covariance matrix of the k th class.

1.6.3 Support Vector Machine (SVM). SVM, a.k.a. Support Vector Network, is also a classification technique, finding the hyperplane which maximizes the margin between two classes [38]. So-called support vectors define the hyperplane that makes the separation. Linear, polynomial, Radial Basis Function (RBF), and sigmoid are the basic types of SVM kernels [39]. The SVM decision function can be expressed as

$$f(x) = \text{sign}\left(\sum_{i=1}^N \alpha_i y_i K(x_i, x) + b\right), \quad (10)$$

where N is the size of training data, K is the kernel function that measures similarity between x_i (support vector) and x (feature values), α_i is Lagrange multiplier, y_i represents the membership class of each datum (± 1), and b is a numeric constant. We used linear, polynomial, and RBF kernels to determine accuracy of the extracted features, respectively described following.

Linear SVM: a linear kernel was used, and its kernel function can be expressed as inner (“dot”) product of support vector and feature values:

$$K(x_i, x) = x_i^T \cdot x. \quad (11)$$

Quadratic and cubic ($d = 2$ and 3) SVM polynomial kernels were used,

$$K(x_i, x) = (x_i^T x + c)^d, \quad (12)$$

with $c = 1$ for both cases.

Medium and Coarse Gaussian SVM: An RBF kernel function was used with $\sigma = 16$ and 64 :

$$K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right), \quad (13)$$

where $\|x_i - x\|$ is Euclidean distance, and standard deviation σ determines the width of the Gaussian kernel.

1.6.4 k-Nearest Neighbor (kNN). kNN is the simplest machine learning algorithm, measuring distance to a given data point. k stands for the number of neighbors which should be taken into account, and accuracy can be varied by adjusting its value. There are several distance measurements, including Euclidean distance (Eq 14):

$$d(x, x') = \|x - x'\| = \sqrt{\sum_{i=1}^k w_i (x_i - x'_i)^2}, \quad (14)$$

where $d(x, x')$ is the distance between points x and x' , and w_i is edge weight.

kNN ($k = 1$) with Euclidean distance with equal weights were used in this analysis.

After extracting the IHAR feature, accuracy was tested in several ways. First we tested three phases with all 14 electrodes (7 electrode pairs). Further, accuracy was checked for three trial phases while reducing the number of electrodes. An augmented training dataset (3720 trials) was used to train all the above-discussed ML models, and remaining 60 trials (5 trials per each subject) with original signal data was used to test the models. K-fold cross-validation with 10 folds was used to validate the trained model. Accuracy was calculated as the ratio of correctly classified trials to all testing trials.

2 Results and discussion

After separating three trial phases from each single EEG signal, each phase was visualized using a custom MATLAB script. Fig 3 depicts the distribution of power of the relaxation phase, over crania of 12 subjects, showing some differences after channel-wise normalization (linear scaling with bias to remap to 0–1). Original unipolar (unbalanced) unsigned electrode data remain so after normalization. Clear asymmetry can be observed across all the nasion-inion lines. The EMOTIV Epoc+ headset uses mastoid bones as reference electrodes which helps to maintain a fair alignment of electrodes. Moreover, symmetry of electrodes was roughly checked by measuring distances from the sagittal line before recording. However, phenology discredited measurement of head shapes, and the helmet, cap, or headset were sometimes not perfectly aligned with the subject's sagittal plane. Perfectly consistent electrode alignment is practically impossible (without extreme invasive techniques such as Neural Lace [40]).

Data augmentation is used in machine learning when the training dataset is significantly small. Even if large number of trials provide more reliable classification results, such collection reduces the practicality in this domain. Therefore, data augmentation provides extra convenience to users when they record their brain signals for training. In this experiment, augmentation techniques and parameters were selected heuristically. However, drastic changes of parameters could affect classification results.

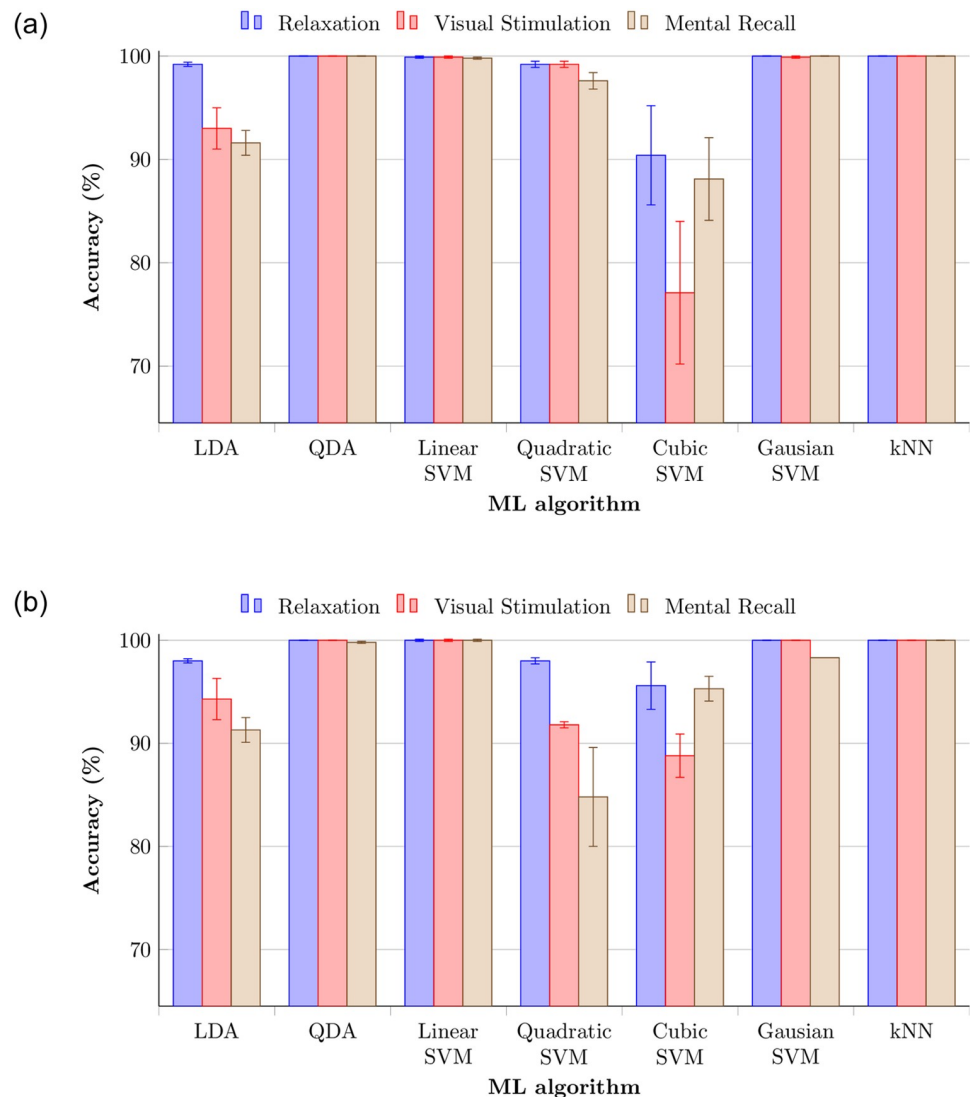


Fig 8. Accuracy of three phases with 14 electrodes for various discriminant analysis, support vector machine, and k-nearest neighbor techniques with OVO method. (a) Validation accuracies. (b) Testing accuracies.

<https://doi.org/10.1371/journal.pone.0238872.g008>

To check whether the extracted features are specific to individual subjects, statistical hypothesis testing was performed. Expressing “extracted IHAR features do not have subjective discrimination” as the null hypothesis, MANOVA (Multivariate Analysis of Variance) was used for this test. The null hypothesis was rejected for all relaxation, visual, and mental recall phases with the p -values 0.0043, 0.031, and 0.0037, respectively, which are all less than conventional threshold 0.05.

Validation and testing accuracies of the three phases with 14 electrodes for the above-discussed machine learning techniques are shown in Fig 8. The data of each subject was trained against that of every other subject, which creates $N(N - 1)/2 = 66$ classifiers, the so-called One-vs.-One (OVO) multiclass classification method. This method shows even better accuracy than One-vs.-All (OVA), albeit only slightly. Validation and testing accuracies for the same data set with OVA method can be seen in Fig 9. OVO is usually better than OVA because class imbalance occurs using OVA method even if OVA is faster in training. QDA, linear SVM,

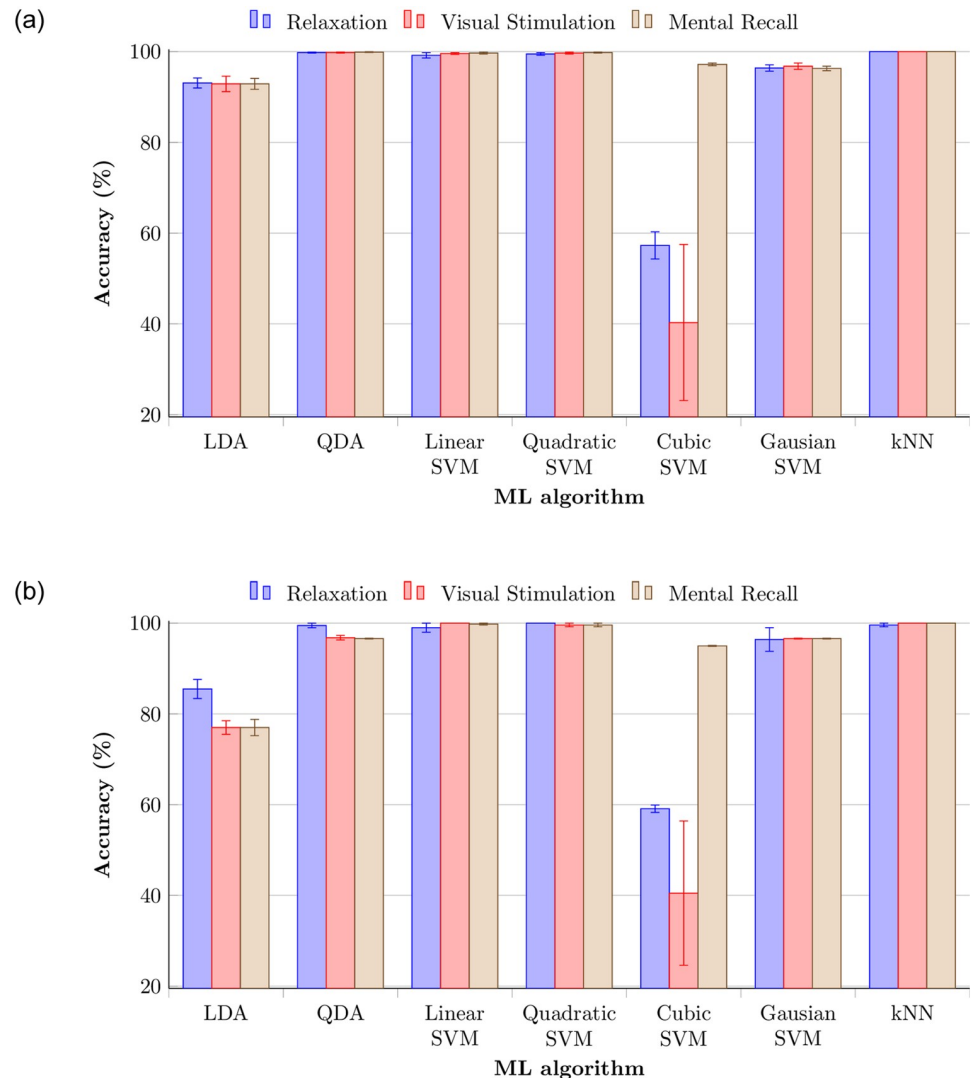


Fig 9. Accuracy of three phases with 14 electrodes with OVA method. (a) Validation accuracies. (b) Testing accuracies.

<https://doi.org/10.1371/journal.pone.0238872.g009>

Gaussian SVM, and kNN showed the highest validation accuracy for all three phases. The testing results also were close to 100% for the same ML techniques. Furthermore, relaxation phase showed higher validation and testing accuracy than other phases. Signal patterns evoked in the relaxation phase seem more stable compared to those from the other phases. Lateral asymmetry or energy difference of alpha activity between two hemispheres during the relaxation explains the high performance of relaxation phase [41]. Quadratic SVM showed better validation accuracy, but testing accuracy was slightly lower. Even though cubic SVM achieved lowest validation accuracy among all techniques, testing accuracy was slightly increased. The relaxation phase generally yielded highest accuracy with OVO method.

Comparison of performances with other time-domain features for same conditions is shown in Fig 10. All 14 channels were used to test each feature extraction method. While IHAR achieved highest accuracy, LC was the next best-performing feature, achieving

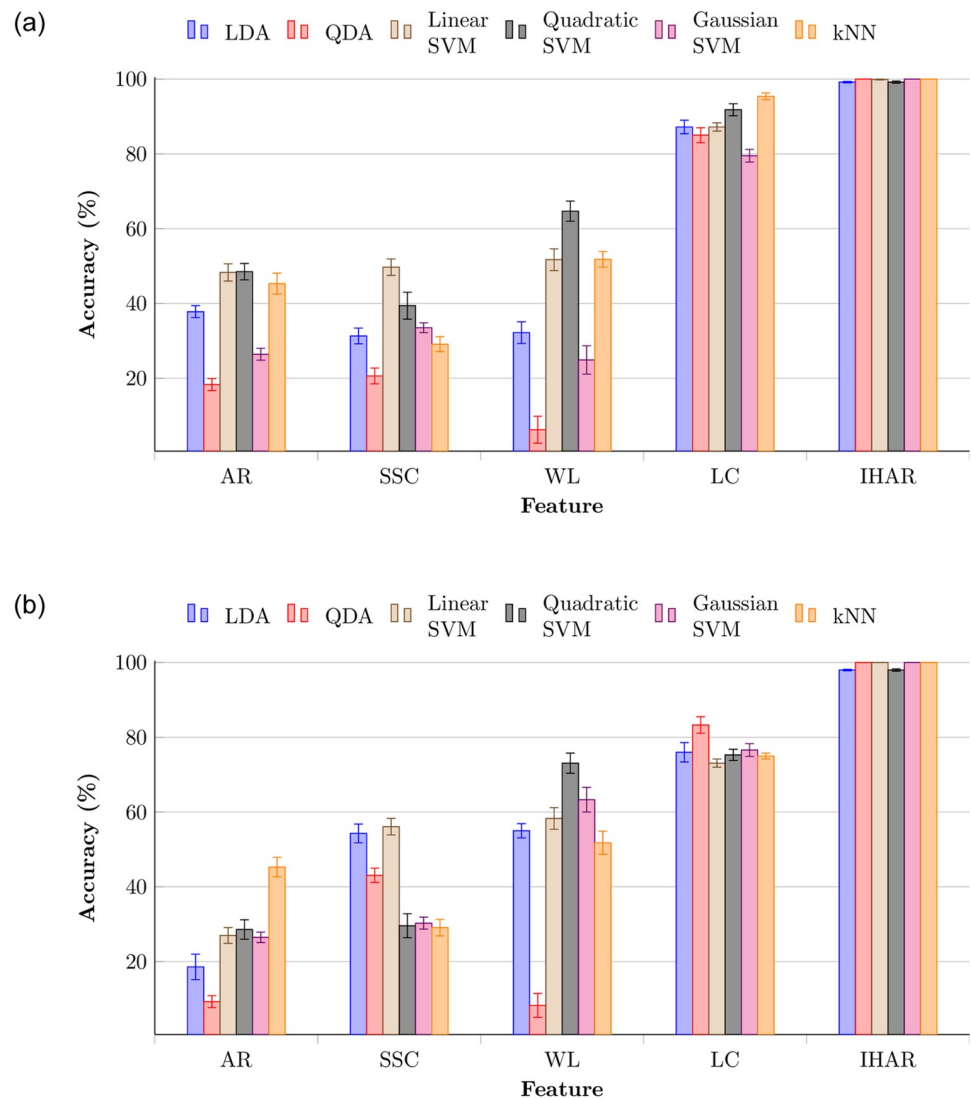


Fig 10. Comparison of accuracy across features (AR: Auto-regressive coefficients, SSC: Slope sign change, WL: Waveform length, LC: Laterality coefficient, and IHAR: Inter-hemispheric amplitude ratio) extracted during relaxation phase using all 14 electrodes. (a) Validation accuracies. (b) Testing accuracies.

<https://doi.org/10.1371/journal.pone.0238872.g010>

95.4±0.9% validation accuracy for kNN classifier, and 83.3±0.1% maximum testing accuracy for QDA. The conceptual similarity of the LC feature to proposed IHAR explains why it achieved comparable results, especially for validation (but not testing). AR, SSC, and WL showed less than 50% validation accuracies for most of the ML techniques. Even though the AR feature has been used in many EEG-related studies, it did not show considerable results for any classifier.

After selecting the best ML algorithms among tested candidates, accuracy was checked for each electrode pair. Results can be seen in Table 1. The highest testing accuracy, 94.7±1.1%, was from the FC5-FC6 electrode pair for QDA with relaxation phase. According to the results, FC5-FC6 electrode pair yielded over 90% accuracy for most of the ML algorithms. Other frontal electrode pairs F7-F8 and AF3-AF4 were the next highest. Occipital electrodes (O1-O2) showed the worst accuracy, 43.5±2.8% for the linear SVM with relaxation task. In addition,

Table 1. Validation and testing accuracies of each electrode pair for three task phases for various analysis techniques using IHAR.

ML algorithm	AF3-AF4		F7-F8		F3-F4		FC5-FC6		T7-T8		P7-P8		O1-O2	
	VA	TA	VA	TA	VA	TA	VA	TA	VA	TA	VA	TA	VA	TA
LDA	86.7 ± 1.9	70.5 ± 2.2	91.8 ± 1.4	83.5 ± 1.4	67.6 ± 1.5	50.1 ± 2.4	94.3 ± 1.4	87.4 ± 1.2	82.0 ± 1.6	69.1 ± 2.2	80.2 ± 1.3	65.0 ± 1.5	65.2 ± 2.2	47.1 ± 2.4
	85.0 ± 1.4	76.1 ± 1.3	91.2 ± 1.3	85.3 ± 1.8	62.9 ± 1.8	50.8 ± 2.5	90.2 ± 1.5	78.0 ± 1.7	77.2 ± 2.3	55.0 ± 2.2	79.3 ± 2.2	75.8 ± 3.2	62.3 ± 2.4	51.1 ± 1.7
	86.9 ± 1.6	73.6 ± 1.8	92.1 ± 1.1	74.6 ± 1.7	67.3 ± 1.6	60.0 ± 1.3	93.4 ± 0.8	89.1 ± 0.8	78.9 ± 1.9	59.8 ± 1.4	79.8 ± 1.9	77.8 ± 2.0	68.6 ± 2.6	59.3 ± 2.2
QDA	86.0 ± 1.2	73.3 ± 1.5	94.9 ± 1.0	88.3 ± 2.1	60.2 ± 2.7	57.1 ± 2.3	91.8 ± 0.8	94.7 ± 1.1	79.0 ± 1.6	80.6 ± 2.1	78.8 ± 1.7	73.3 ± 2.6	59.8 ± 1.3	55.3 ± 3.2
	81.4 ± 2.1	90.1 ± 0.5	86.1 ± 1.7	83.5 ± 1.2	50.8 ± 3.1	62.8 ± 0.8	88.5 ± 1.7	82.3 ± 0.8	75.8 ± 1.5	76.0 ± 0.8	75.4 ± 1.7	72.5 ± 1.1	55.8 ± 3.1	51.0 ± 1.4
	85.1 ± 1.8	84.0 ± 1.4	94.3 ± 1.3	81.5 ± 0.5	55.0 ± 2.7	55.3 ± 2.1	94.9 ± 1.1	92.0 ± 0.1	76.5 ± 2.0	68.5 ± 0.5	78.1 ± 1.7	76.8 ± 0.9	63.1 ± 1.7	59.3 ± 0.8
Linear SVM	89.1 ± 1.4	74.6 ± 1.8	93.1 ± 1.1	89.1 ± 2.3	70.2 ± 2.3	47.8 ± 2.3	95.1 ± 1.1	86.3 ± 2.2	84.0 ± 1.1	68.8 ± 2.6	83.1 ± 2.3	61.3 ± 2.3	70.6 ± 2.5	43.5 ± 2.8
	86.5 ± 1.2	73.6 ± 1.5	92.7 ± 1.2	83.1 ± 1.6	65.8 ± 2.1	52.5 ± 2.5	90.9 ± 0.5	79.3 ± 1.6	80.6 ± 1.1	57.1 ± 1.7	82.0 ± 1.8	73.3 ± 2.6	64.6 ± 2.1	47.5 ± 2.7
	88.7 ± 1.3	75.0 ± 2.3	93.8 ± 1.4	82.8 ± 1.9	69.0 ± 3.1	61.3 ± 1.3	94.1 ± 1.0	90.0 ± 0.1	80.6 ± 1.3	61.6 ± 1.3	81.6 ± 2.1	76.1 ± 2.3	70.9 ± 2.4	57.3 ± 1.7
Quadratic SVM	93.1 ± 1.3	80.5 ± 1.2	95.3 ± 1.5	85.3 ± 2.6	79.2 ± 1.5	57.0 ± 2.3	96.6 ± 0.1	83.1 ± 1.1	89.1 ± 0.1	77.3 ± 2.1	89.7 ± 1.2	74.1 ± 1.7	81.9 ± 2.3	60.0 ± 2.2
	89.1 ± 1.4	76.1 ± 1.5	96.0 ± 1.1	85.8 ± 0.8	76.3 ± 1.1	62.5 ± 1.4	95.3 ± 0.7	84.5 ± 1.1	87.4 ± 1.2	68.3 ± 1.9	86.8 ± 1.5	74.8 ± 2.2	76.8 ± 2.2	65.8 ± 3.4
	92.7 ± 1.1	82.5 ± 2.5	96.5 ± 0.6	78.5 ± 1.2	78.5 ± 2.5	67.1 ± 1.1	96.9 ± 0.8	94.3 ± 0.8	88.1 ± 1.6	73.3 ± 3.0	87.6 ± 1.9	74.8 ± 2.7	78.2 ± 2.8	53.3 ± 2.4
Gaussian SVM	79.1 ± 1.6	69.1 ± 1.2	86.9 ± 1.9	82.8 ± 2.6	54.0 ± 2.9	56.7 ± 2.7	90.1 ± 3.5	87.2 ± 2.2	76.1 ± 1.6	77.4 ± 2.8	73.0 ± 1.4	73.1 ± 2.6	53.4 ± 1.6	71.0 ± 1.4
	78.6 ± 1.3	77.6 ± 0.8	84.3 ± 1.5	83.3 ± 0.1	46.3 ± 1.7	49.1 ± 1.1	81.2 ± 1.2	82.8 ± 1.1	66.4 ± 1.8	53.5 ± 0.9	68.7 ± 1.8	68.3 ± 0.1	49.1 ± 1.9	45.8 ± 1.4
	76.9 ± 0.9	83.1 ± 0.5	91.6 ± 1.3	85.0 ± 0.1	48.4 ± 2.2	50.3 ± 1.0	90.5 ± 1.6	93.3 ± 0.1	71.2 ± 0.8	52.3 ± 1.1	73.1 ± 1.5	63.3 ± 1.1	54.1 ± 1.9	47.6 ± 1.1
kNN	95.6 ± 1.2	75.0 ± 1.2	98.1 ± 0.5	88.0 ± 2.1	86.1 ± 2.1	51.1 ± 3.6	97.2 ± 0.9	85.6 ± 1.3	93.7 ± 1.1	72.0 ± 1.9	93.6 ± 0.4	67.9 ± 3.1	89.4 ± 1.2	59.1 ± 2.7
	93.5 ± 0.8	82.3 ± 1.1	97.6 ± 0.5	88.3 ± 0.7	83.9 ± 2.4	58.1 ± 3.0	96.7 ± 0.9	83.5 ± 0.5	92.2 ± 1.4	75.0 ± 2.3	93.8 ± 1.5	72.5 ± 1.1	85.0 ± 1.3	49.6 ± 2.3
	94.1 ± 0.5	87.8 ± 1.3	98.1 ± 0.5	88.1 ± 0.9	85.8 ± 1.8	62.8 ± 1.7	97.2 ± 0.4	94.6 ± 0.5	91.4 ± 0.8	65.5 ± 1.7	93.4 ± 1.4	65.5 ± 1.7	86.6 ± 1.3	60.5 ± 2.6

VA: validation accuracy, TA: training accuracy

White (top row in each cell): relaxation; light gray (middle cellular row): visual task; dark gray (bottom row): mental recall

Testing accuracies ≥ 90% are emboldened.

<https://doi.org/10.1371/journal.pone.0238872.t001>

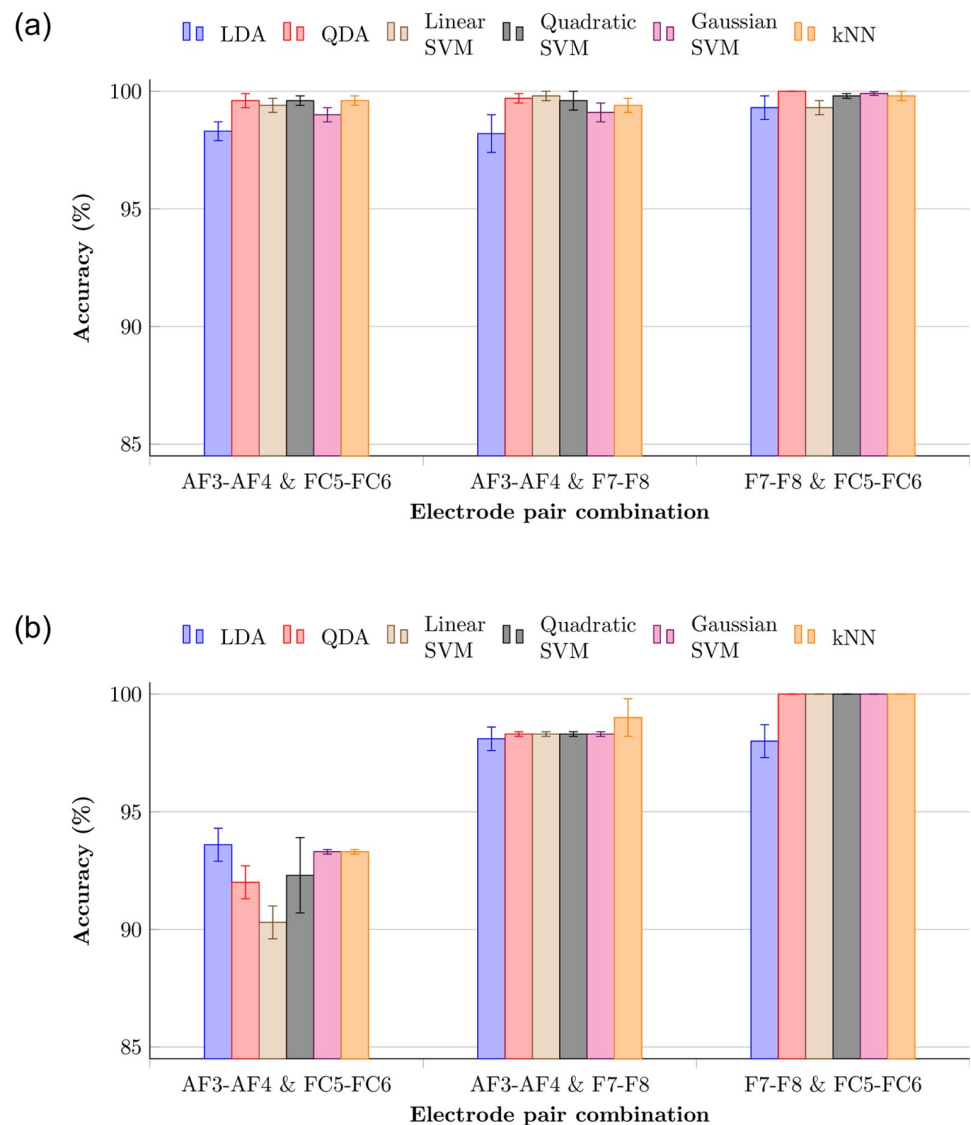


Fig 11. Accuracies of combination of 2 frontal electrode pairs which yielded higher accuracies. (a) Validation accuracies. (b) Testing accuracies.

<https://doi.org/10.1371/journal.pone.0238872.g011>

O1-O2 showed the worst testing accuracy even for the visual stimulation phase in contrast to most research findings [42, 43]. In most of the studies, so-called oddball stimulus has been used to get the P300 (positive peak at 300 ms after the stimulation) spike from the occipital area. In our study, the image (four-digit number) was shown without making any changes (blinking or switching with another image) as visual stimulation. This explains the poor performance of occipital electrodes for visual stimulation. Several studies have considered frontal asymmetry in cognitive processes of the brain [44, 45]. Our results also showed considerable accuracy for the frontal electrodes for relaxation and mental recall phases.

In order to increase accuracy further with minimum electrode setup, the model was trained again by considering sets of electrode pairs. 100% testing accuracy was shown by the four frontal electrodes (F7-F8 and FC5-FC6) for all classifiers except LDA. As seen in Fig 11, accuracy increases when combining other frontal electrode pairs. A Receiver Operating Characteristic (ROC) curve, created by plotting the true positive rate (TPR) against the false positive rate

(FPR), was plotted for all subjects with F7-F8 and FC5-FC6 electrode combination. The AUC (area under the ROC curve) value of all classifiers was unity, which means the trained model was a perfect classifier. Considering user-friendliness of specific hardware design for EEG biometrics, the anterior frontal electrodes AF3-AF4 pair is better than FC5-FC6, which is nearer to the coronal center. Therefore, the AF3-AF4 and F7-F8 electrode setup also was tested, and 99.0±0.8% (kNN), 98.3±0.1% (QDA), 98.3±0.1% (linear SVM), 98.3±0.1% (quadratic SVM), and 98.1±0.5% (LDA) testing accuracies were achieved for relaxation phase. The average AUC value for these cases lies between 0.98 and 1. These results seem to accommodate lighter devices which require less calibration time, such as Muse [46]. Even though entire hemispheres manifest asymmetry, the “frontal asymmetry” feature has been introduced in EEG analysis, calculated using frontal channels. Specifically, frontal asymmetry has been used in emotion and attention-related analysis [47, 48]. Since the frontal lobe is responsible for executive tasks [49], IHAR-based analysis for relaxation and mental recall phases shows high accuracy in frontal electrodes.

Specificity (true positive rate) and sensitivity (true negative rate) were calculated for electrode combinations which showed higher accuracies. Specificity is the ratio of number of trials that correctly classified as negatives to number of all negative testing trials, while sensitivity is the ratio of number of trials that correctly classified as positives to number of all positive testing trials. Models with high specificity values are better in the authentication domain because they decrease risk of granting access to imposters. As shown in Table 2, all electrode combinations achieved highest specificity value, which confirms suitability of IHAR method for this problem (better to tolerate occasional false rejection rather than allow false accepts).

More information can be captured using high-resolution EEG devices. Compared with consumer-grade devices, clinical grade devices have more electrodes. Even though high accuracy can be achieved with the help of many electrodes, lighter electrode setup provides more user-friendliness. As shown in Table 3, other approaches achieved high accuracy using more channels, for compound or complex tasks. Our approach achieved the same accuracy with only four frontal electrodes for a simple task. However, accuracy of the models cannot be guaranteed when increasing the number of subjects. Fig 12 shows the general degradation of accuracy while increasing the number of subjects. Average accuracy was calculated for 10 random combinations of subjects with FC5-FC6 electrode pair, which showed highest accuracy among other pairs. QDA classifier did not show rapid change of testing accuracy over increase of number of subjects. Therefore, it might be usable with more subjects without compromising accuracy significantly. In this study, 15 trials were conducted, but when increasing the number of subjects, number of trials also should be increased to maintain consistent accuracy.

Table 2. Comparison of performances for different number of electrodes.

ML algorithm	FC5-FC6			AF3-AF4, FC5-FC6			AF3-AF4, F7-F8			F7-F8, FC5-FC6		
	ACC	SP	SE	ACC	SP	SE	ACC	SP	SE	ACC	SP	SE
LDA	87.4	100	78	93.3	100	98	98.1	100	98	98.1	100	98
QDA	94.7	100	98	92.0	100	96	98.3	100	100	100	100	100
Linear SVM	86.3	100	76	90.8	100	95	98.3	100	100	100	100	100
Quadratic SVM	83.1	100	96	91.5	100	98	98.3	100	99	100	100	100
Gaussian SVM	87.2	100	80	93.3	100	97	98.3	100	98	100	100	100
KNN	85.6	100	84	93.3	100	100	99.3	100	100	100	100	100

ACC: Accuracy, SP: Specificity, and SE: Sensitivity

AF: Anterior Frontal, F: Frontal, FC: Fronto-Central

<https://doi.org/10.1371/journal.pone.0238872.t002>

Table 3. Comparison with other studies.

Author	Electrodes	Task	Classified feature(s)	Classifier	Accuracy (%)
Ong et al. [50]	32	VEP	Power spectral density (PSD)	kNN	83
Alyasseri et al. [51]	6	Mental task (counting)	Multi-objective flower pollination algorithm with wavelet transform (MOFPA-WT)	ANN	85.12
Falzon et al. [42]	32	VEP	Frequency components up to 5th harmonic	kNN	91.7
Koutras et al. [52]	56	Sleep	Time-Domain Descriptors (10), Frequency-Domain Descriptors (17), Wavelet Domain Descriptors (4)	kNN	95
Fukami et al. [43]	4	VEP	5 frequency components	Mahalanobis distance	95
Yang et al. [53]	4	Motor movement, imagery	Wavelet log-DCT (WLD)	Fisher's Linear Discriminant	98.5
Kaewwit et al. [54]	4	Resting state	Combined ICA and AR	kNN	98.51
Arnau et al. [55]	32	Video stimulation	PSD	ANN	99
Thomas et al. [56]	19	Resting state	PSD	Mean correlation coefficients	99
Schetinin et al. [57]	64	Motor movement, imagery	Group Method of Data Handling (GMDH)	SVM	100
Proposed approach	4	Relaxation before stimulation	IHAR	QDA, SVM, kNN	100
	4	Visual stimulation	IHAR	kNN	98.8 ± 0.9
	4	Mental recall	IHAR	QDA, SVM, kNN	100

<https://doi.org/10.1371/journal.pone.0238872.t003>

Comparing the trial phases considered here, relaxation and mental recall showed higher accuracy. Selecting the simplest task among these three phases is complicated because there are merits and demerits for each. It is harder to concentrate or be relaxed in noisy environments, but visual stimulation would be easy. To stimulate the brain externally, some physical resource is needed, but other tasks can be performed spontaneously. In our approach, the relaxation phase showed highest accuracy for the minimal (four) electrode setup.

Mental recall of a number can be realized in many ways, including visualization (imagining an image of the number), audiation or silent verbalization (thinking of a sound, such as a voice reciting the number), synesthetic association (such as association of digits or numbers with colors, flavors, musical notes, etc.) [58] or cross-modal correspondences, cardinal correspondences (such as thinking of sets enumerated by a number's digits), and spatial analogy (such as clock hour directions).

For an n -digit integer, $n \geq 2$, the likelihood of a repeated digit is

$$p(n) = 1 - \prod_{i=1}^{n-1} \frac{(10-i)}{10}. \quad (15)$$

For a random four-digit integer, the probability is almost half that at least one digit is repeated,

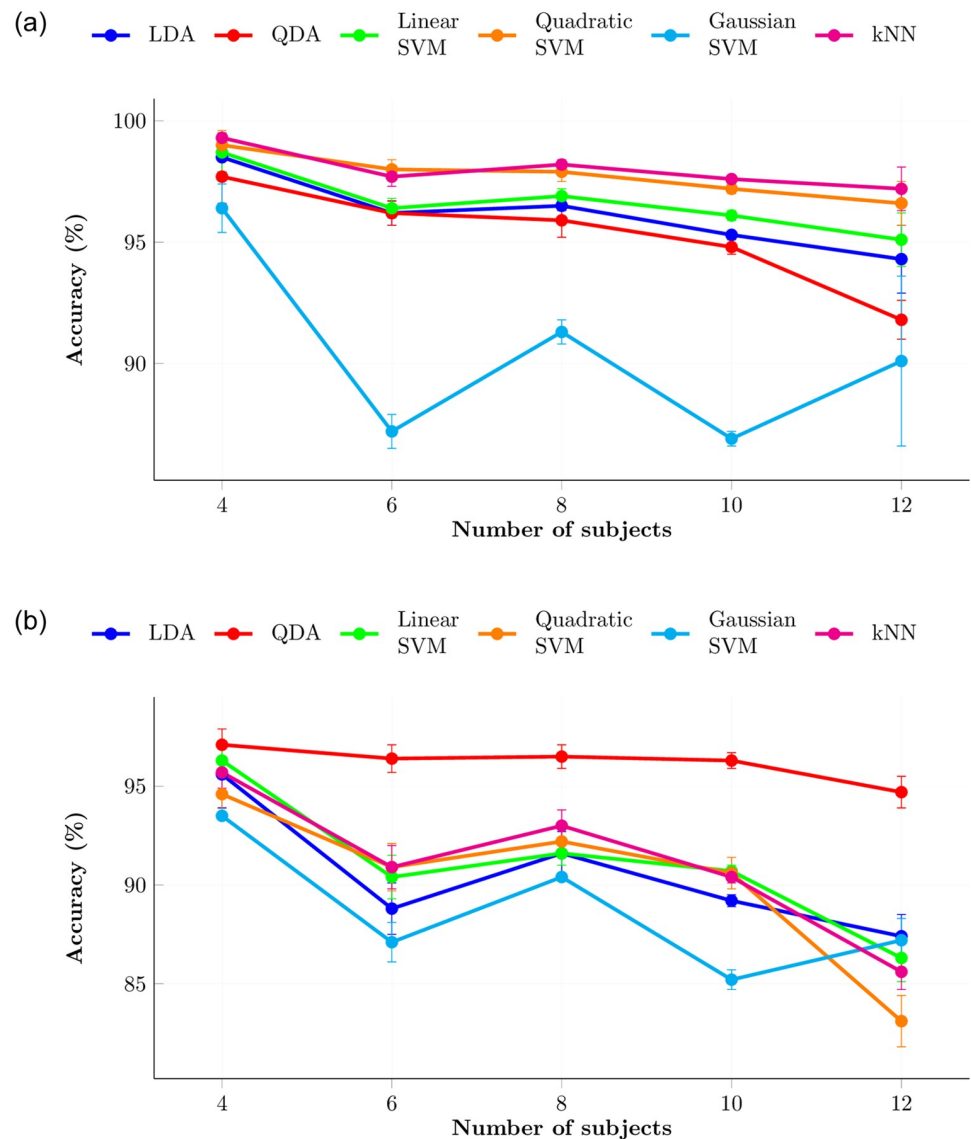


Fig 12. Variation of validation and testing accuracy for FC5-FC6 electrode pair of relaxation phase when increasing number of subjects using IHAR. (a) Validation accuracies. (b) Testing accuracies.

<https://doi.org/10.1371/journal.pone.0238872.g012>

and such a pattern could anchor its recall. All of these recall styles are possible practices, and the actual technique was left up to each subject. This vagueness could explain apparent uselessness of data from the mental recall phase.

A major problem in EEG-based biometric systems is instability of signal patterns over time. EEG patterns can change with environment, maturity, and psychological disorders [59]. This issue has not been addressed in the proposed approach. Therefore, accuracy can be degraded over considerable time period due to changing daily mental states. Even though relaxation phase has been used as refractory period in this study, merits and demerits of taking off headset cannot be addressed with this experimental setup. In addition, results can be changed with different headsets because of hardware characteristics (resolution, sampling rate, common-mode rejection ratio) and type of electrodes.

IHAR was tested with different MA windows sizes (L). No substantial differences were found for different parameter values, but comparatively better results were obtained for $L = 32$. As L decreases, the size of ripples in the frequency response increases. As shown in Fig 7, even when $L = 32$, a series of ripples is produced, although the height of the ripples is not considerable.

3 Conclusion

In this research, we successfully deployed a novel feature, Inter-Hemispheric Amplitude Ratio (IHAR), to reveal personally unique information and distinguish people using EEG signals. With the proposed feature, highest accuracy could be achieved with a lower number of electrodes for a relatively simple task. Furthermore, this approach outperforms similar approaches with less computational power, suggesting deployment in portable devices with embedded low-power microprocessors. These promising results show that this approach has practical real-world applications. Also, this approach provides more convenience when training the system because there is no need of collecting large number of training trials. In the proposed approach, highest accuracy was achieved for the two frontal pairs of electrodes—FC5-FC6 and F7-F8. As a task phase to collect EEG data, relaxation is the best according to the results. Analysis showed the most suitable ML algorithms for classification are QDA, linear SVM, quadratic SVM, Gaussian SVM, and kNN. Moreover, AF3, AF4, F7, and F8 selective electrode arrangement can be used to develop tighter hardware design with high performance. However, accuracy would be affected if number of subjects increases. Also, brain signal patterns can significantly change over time, which would affect performance. For continued study, this approach can be further examined with a clinical-grade EEG system, and should be investigated to check whether the number of electrodes can be further reduced.

Supporting information

S1 File.

(PDF)

S1 Dataset.

(ZIP)

Acknowledgments

We thank anonymous reviewers for their generous and intelligent suggestions, and also Dr. Ying Chen at the University of Aizu for her expert encouragement.

Author Contributions

Conceptualization: Isuru Jayarathne, Senaka Amarakeerthi.

Data curation: Senaka Amarakeerthi.

Formal analysis: Isuru Jayarathne.

Investigation: Isuru Jayarathne.

Methodology: Isuru Jayarathne.

Resources: Senaka Amarakeerthi.

Software: Isuru Jayarathne.

Supervision: Michael Cohen, Senaka Amarakeerthi.

Validation: Isuru Jayarathne.

Visualization: Isuru Jayarathne.

Writing – original draft: Isuru Jayarathne, Michael Cohen.

Writing – review & editing: Michael Cohen, Senaka Amarakeerthi.

References

1. Mohammadi G, Shoushtari P, Molaee Ardekani B, Shamsollahi MB. Person identification by using AR model for EEG signals. In: Proc. World Academy of Science, Engineering and Technology. vol. 11; 2006. p. 281–285.
2. Poulos M, Rangoussi M, Chrissikopoulos V, Evangelou A. Person identification based on parametric processing of the EEG. In: Proc. Int. Conf. on Electronics, Circuits and Systems (ICECS). vol. 1. IEEE; 1999. p. 283–286.
3. Poulos M, Rangoussi M, Alexandris N. Neural network based person identification using EEG features. In: Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP). vol. 2. IEEE; 1999. p. 1117–1120.
4. Poulos M, Rangoussi M, Alexandris N, Evangelou A, et al. Person identification from the EEG using nonlinear signal classification. *Methods of Information in Medicine*. 2002; 41(1):64–75. <https://doi.org/10.1055/s-0038-1634316>
5. Del Pozo-Banos M, Alonso JB, Ticay-Rivas JR, Travieso CM. Electroencephalogram subject identification: A review. *Expert Systems with Applications*. 2014; 41(15):6537–6554. <https://doi.org/10.1016/j.eswa.2014.05.013>
6. Jayarathne I, Cohen M, Amarakeerthi S. Survey of EEG-based Biometric Authentication. In: Proc. Int. Conf. on Awareness Science and Technology (iCAST). IEEE; 2017. p. 324–329.
7. Almehmadi A, El-Khatib K. The state of the art in electroencephalogram and access control. In: Proc. Int. Conf. on Communications and Information Technology (ICCIT). IEEE; 2013. p. 49–54.
8. Das K, Zhang S, Giesbrecht B, Eckstein MP. Using Rapid Visually Evoked EEG Activity for Person Identification. In: Proc. Annual Int. Conf. of the Engineering in Medicine and Biology Society (EMBC); 2009. p. 2490–2493.
9. Chuang J, Nguyen H, Wang C, Johnson B. I think, therefore I am: Usability and security of authentication using brainwaves. In: Proc. Int. Conf. on Financial Cryptography and Data Security. Springer; 2013. p. 1–16.
10. Chen Y, Atnafu AD, Schlattner I, Weldtsadik WT, Roh MC, Kim HJ, et al. A high-security EEG-based login system with RSVP stimuli and dry electrodes. *IEEE Trans. on Information Forensics and Security*. 2016; 11(12):2635–2647. <https://doi.org/10.1109/TIFS.2016.2577551>
11. Poulos M, Rangoussi M, Kafetzopoulos E. Person identification via the EEG using computational geometry algorithms. In: Proc. 9th European Signal Processing Conf. (EUSIPCO). IEEE; 1998. p. 1–4.
12. Poulos M, Rangoussi M, Chrissikopoulos V, Evangelou A. Parametric person identification from the EEG using computational geometry. In: Proc. Int. Conf. on Electronics, Circuits and Systems (ICECS). vol. 2. IEEE; 1999. p. 1005–1008.
13. Singhal GK, RamKumar P. Person identification using evoked potentials and peak matching. In: Proc. Biometrics Symp. IEEE; 2007. p. 1–6.
14. Palaniappan R, Ravi K. A new method to identify individuals using signals from the brain. In: Proc. Information, Communications and Signal Processing, and Fourth Pacific Rim Conf. on Multimedia. vol. 3. IEEE; 2003. p. 1442–1445.
15. Yazdani A, Roodaki A, Rezatofighi S, Misaghian K, Setarehdan SK. Fisher Linear Discriminant Based Person Identification using Visual Evoked Potentials. In: Proc. 9th Int. Conf. on Signal Processing (ICSP). IEEE; 2008. p. 1677–1680.
16. Goldstein L, Stoltzfus NW, Gardocki JF. Changes in interhemispheric amplitude relationships in the EEG during sleep. *Physiology & Behavior*. 1972; 8(5):811–815. [https://doi.org/10.1016/0031-9384\(72\)90289-2](https://doi.org/10.1016/0031-9384(72)90289-2)
17. LeMay M. Asymmetries of the skull and handedness: Phrenology revisited. *J. of the Neurological Sciences*. 1977; 32(2):243–253. [https://doi.org/10.1016/0022-510X\(77\)90239-8](https://doi.org/10.1016/0022-510X(77)90239-8)
18. Watkins KE, Paus T, Lerch JP, Zijdenbos A, Collins DL, Neelin P, et al. Structural asymmetries in the human brain: a voxel-based statistical analysis of 142 MRI scans. *Cerebral Cortex*. 2001; 11(9):868–877. <https://doi.org/10.1093/cercor/11.9.868>

19. Wada JA, Clarke R, Hamm A. Cerebral hemispheric asymmetry in humans: Cortical speech zones in 100 adult and 100 infant brains. *Archives of Neurology*. 1975; 32(4):239–246. <https://doi.org/10.1001/archneur.1975.00490460055007>
20. Savic I, Lindström P. PET and MRI show differences in cerebral asymmetry and functional connectivity between homo- and heterosexual subjects. *Proc. National Academy of Sciences*. 2008; 105(27):9403–9408. <https://doi.org/10.1073/pnas.0801566105>
21. Jayarathne I, Cohen M, Amarakeerthi S. BrainID: Development of an EEG-based Biometric Authentication System. In: *Proc. Annual Conf. on Information Technology, Electronics and Mobile Communication (IEMCON)*. IEEE; 2016. p. 1–6.
22. Wang Y, Wang Z, Clifford W, Markham C, Ward TE, Deegan C. Validation of low-cost wireless EEG system for measuring event-related potentials. In: *Proc. Irish Signals and Systems Conf. (ISSC)*. IEEE; 2018. p. 1–6.
23. Ledoux H, Gold C. An efficient natural neighbour interpolation algorithm for geoscientific modelling. In: *Proc. Developments in Spatial Data Handling*. Springer; 2005. p. 97–108. https://doi.org/10.1007/3-540-26772-7_8
24. Um TT, Pfister FMJ, Pichler D, Endo S, Lang M, Hirche S, et al. Data Augmentation of Wearable Sensor Data for Parkinson's Disease Monitoring Using Convolutional Neural Networks. In: *Proc. 19th Int. Conf. on Multimodal Interaction. (ICMI)*. New York, NY, USA: ACM; 2017. p. 216–220.
25. Davidson RJ, Hugdahl K. *Brain Asymmetry*. MIT Press; 1996.
26. Toga AW, Thompson PM. Mapping brain asymmetry. *Nature Reviews Neuroscience*. 2003; 4(1):37. <https://doi.org/10.1038/nrn1009>
27. Ruiz-Blondet MV, Jin Z, Laszlo S. CEREBRE: A novel method for very high accuracy event-related potential biometric identification. *IEEE Trans. on Information Forensics and Security*. 2016; 11(7):1618–1629. <https://doi.org/10.1109/TIFS.2016.2543524>
28. Huang X, Altahat S, Tran D, Sharma D. Human identification with electroencephalogram (EEG) signal processing. In: *Proc. Int. Symp. on Communications and Information Technologies (ISCIT)*. IEEE; 2012. p. 1021–1026.
29. Smith SW. *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Pub. San Diego; 1997.
30. LeLeski JM, Henzel N. ECG baseline wander and powerline interference reduction using nonlinear filter bank. *Signal Processing*. 2005; 85(4):781–793. <https://doi.org/10.1016/j.sigpro.2004.12.001>
31. Kaiser V, Daly I, Pichiorri F, Mattia D, Müller-Putz GR, Neuper C. Relationship between electrical brain responses to motor imagery and motor impairment in stroke. *Stroke*. 2012; 43(10):2735–2740. <https://doi.org/10.1161/STROKEAHA.112.665489>
32. Birbaumer N, Elbert T, Canavan AG, Rockstroh B. Slow potentials of the cerebral cortex and behavior. *Physiological Reviews*. 1990; 70(1):1–41. <https://doi.org/10.1152/physrev.1990.70.1.1>
33. Geethanjali P, Mohan YK, Sen J. Time domain feature extraction and classification of EEG data for brain computer interface. In: *9th Int. Conf. on Fuzzy Systems and Knowledge Discovery*. IEEE; 2012. p. 1136–1139.
34. Chan AD, Green GC. Myoelectric control development toolbox. *Proc. Canadian Medical and Biological Engineering Society (CMBES)*. 2007; 30(1).
35. James G, Witten D, Hastie T, Tibshirani R. *An introduction to statistical learning*. vol. 112. Springer; 2013.
36. Raffenburgh RH. *Linear Discriminant Analysis*. Virginia Polytechnic Institute; 1957.
37. Mika S, Ratsch G, Weston J, Scholkopf B, Mullers KR. Fisher discriminant analysis with kernels. In: *Neural Networks for Signal Processing IX, Proc. Signal Processing Society Workshop*. IEEE; 1999. p. 41–48.
38. Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. *IEEE Intelligent Systems and Their Applications*. 1998; 13(4):18–28. <https://doi.org/10.1109/5254.708428>
39. Amari S, Wu S. Improving support vector machine classifiers by modifying kernel functions. *Neural Networks*. 1999; 12(6):783–789. [https://doi.org/10.1016/S0893-6080\(99\)00032-5](https://doi.org/10.1016/S0893-6080(99)00032-5)
40. Strickland E. Neuroscience Experts Weigh in on Elon Musk's Mysterious 'Neural Lace' Company. *IEEE Spectrum*.
41. Bennett JE, Trinder J. Hemispheric laterality and cognitive style associated with transcendental meditation. *Psychophysiology*. 1977; 14(3):293–296. <https://doi.org/10.1111/j.1469-8986.1977.tb01178.x>
42. Falzon O, Zerafa R, Camilleri T, Camilleri KP. EEG-based biometry using steady state visual evoked potentials. In: *Proc. 39th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*; 2017. p. 4159–4162.

43. Fukami T, Abe Y, Shimada T, Ishikawa B. Authentication System Preventing Unauthorized Access of a Third Person Based on Steady State Visual Evoked Potentials. *Int. J. of Innovative Computing Information and Control*. 2018; 14(6):2091–2100.
44. Davidson RJ. What does the prefrontal cortex “do” in affect: perspectives on frontal EEG asymmetry research. *Biological Psychology*. 2004; 67(1-2):219–234. <https://doi.org/10.1016/j.biopsycho.2004.03.008>
45. Coan JA, Allen JJ. Frontal EEG asymmetry as a moderator and mediator of emotion. *Biological Psychology*. 2004; 67(1-2):7–50. <https://doi.org/10.1016/j.biopsycho.2004.03.002>
46. InteraXon Inc. Muse— Meditation Made Easy;. <https://choosemuse.com>.
47. Wheeler RE, Davidson RJ, Tomarken AJ. Frontal brain asymmetry and emotional reactivity: A biological substrate of affective style. *Psychophysiology*. 1993; 30(1):82–89. <https://doi.org/10.1111/j.1469-8986.1993.tb03207.x>
48. Harmon-Jones E, Allen JJ. Behavioral activation sensitivity and resting frontal EEG asymmetry: Covariation of putative indicators related to risk for mood disorders. *J. Abnormal Psychology*. 1997; 106(1):159. <https://doi.org/10.1037/0021-843X.106.1.159>
49. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*. 2000; 41(1):49–100. <https://doi.org/10.1006/cogp.1999.0734>
50. Ong ZY, Saidatul A, Ibrahim Z. Power Spectral Density Analysis for Human EEG-based Biometric Identification. In: *Proc. Int. Conf. on Computational Approach in Smart Systems Design and Applications (ICASSDA)*. IEEE; 2018. p. 1–6.
51. Alyasseri ZAA, Khader AT, Al-Betar MA, Papa JP, Alomari OA, Makhadmeh JSN. Classification of EEG mental tasks using multi-objective flower pollination algorithm for person identification. *Int. J. of Integrated Engineering*. 2018; 10(7). <https://doi.org/10.30880/ijie.2018.10.07.010>
52. Koutras A, Kostopoulos GK. EEG-Based Person Identification Using Rhythmic Brain Activity During Sleep. In: *Int. Conf. on Artificial Neural Networks*. Springer; 2018. p. 682–692.
53. Yang S, Hoque S, Deravi F. Improved time-frequency features and electrode placement for EEG-based biometric person recognition. *IEEE Access*. 2019; 7:49604–49613. <https://doi.org/10.1109/ACCESS.2019.2910752>
54. Kaewwit C, Lursinsap C, Sophatsathit P, et al. High accuracy EEG biometrics identification using ICA And AR Model. *J. of Information and Communication Technology (JICT)*. 2017; 16(2):354–373.
55. Arnau-González P, Arevalillo-Herráez M, Katsigiannis S, Ramzan N. On the influence of affect in EEG-based subject identification. *IEEE Trans. on Affective Computing*. 2018; 37(1):277–289.
56. Thomas KP, Vinod AP. EEG-based biometric authentication using gamma band power during rest state. *Circuits, Systems, and Signal Processing*. 2018; 37(1):277–289. <https://doi.org/10.1007/s00034-017-0551-4>
57. Schetinina V, Jakaite L, Nyah N, Novakovic D, Krzanowski W. Feature extraction with GMDH-type neural networks for EEG-based person identification. *Int. J. of Neural Systems*. 2018; 28(06):1750064. <https://doi.org/10.1142/S0129065717500642>
58. Chun CA, Hupé JM. Are synesthetes exceptional beyond their synesthetic associations? A systematic comparison of creativity, personality, cognition, and mental imagery in synesthetes and controls. *British J. of Psychology*. 2016; 107(3):397–418. <https://doi.org/10.1111/bjop.12146>
59. Roohi-Azizi M, Azimi L, Heysieattalab S, Aamidfar M. Changes of the brain’s bioelectrical activity in cognition, consciousness, and some mental disorders. *Medical J. of the Islamic Republic of Iran*. 2017; 31:53. <https://doi.org/10.14196/mjiri.31.53>