Computers in Biology and Medicine 110 (2019) 218–226

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

IQ estimation by means of EEG-fNIRS recordings during a logical-mathematical intelligence test

Shabnam Firooz, Seyed Kamaledin Setarehdan*

Control and Intelligent Processing Centre of Excellence, School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran

ARTICLE INFO

Keywords:
Electroencephalogram
Functional near-infrared spectroscopy
Intelligence quotient
Linear discriminant analysis
Linear regression
Principal component analysis
Raven progressive matrix test
Support vector regression

ABSTRACT

Intelligence differences of individuals are attributed to the structural and functional differences of the brain. Neural processing operations of the human brain vary according to the difficulty level of the problem and the intelligence level of individuals. In this study, we used a bimodal system consisting of functional Near-Infrared Spectroscopy (fNIRS) and Electroencephalogram (EEG) to investigate these inter-individual differences. A continuous wave 32-channel fNIRS from OxyMonfNIRS device (Artinis) and 19-channel EEG from (g.tec's company) were utilized to study the oxygenation procedure as well as the electrical activity of the brain when doing the problems of Raven's Progressive Matrix (RPM) intelligence test. We used this information to estimate the Intelligence Quotient (IQ) of the individual without performing a complete logical-mathematical intelligence test in a long-time period and examining the answers of people to the questions. After EEG preprocessing, different features including Higuchi's fractal dimension, Shannon entropy values from wavelet transform coefficients, and average power of frequency sub-bands were extracted. Clean fNIRS signals were also used to compute features such as slope, mean, variance, kurtosis, skewness, and peak. Then dimension reduction algorithms such as Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) were applied to select an effective feature set from fNIRS and EEG in order to improve the IQ estimation process. We utilized two regression methods, i.e., Linear Regression (LR) and Support Vector Regression (SVR), to extract optimum models for the IQ determination. The best regression models based on fNIRS-EEG and fNIRS presented 3.093% and 3.690% relative error for 11 subjects, respectively.

1. Introduction

Intelligence means the ability to learn and understand concepts [1]. In this regard, there are 8 kinds of intelligence, including linguistic, logical-mathematical, spatial, musical, corporal-kinesthetic, intra-personal, interpersonal, and naturalistic [2,3]. One of the most interesting aspects of the psychophysiology is the investigation of the function of the human brain when performing logical-mathematical intelligence tests. Over the last two decades, non-conscious cognitive procedures have shown a significant effect on problem-solving when dealing with intelligence tests [4]. These processes vary according to the individual's neuro-cognitive identity, which are modulated based on the internal mechanisms achieved or improved throughout life [1,2].

In recent years, many studies have been conducted on brain activities when implementing intelligence tests. However, there are still many procedures remaining unknown in this area such as the estimation of the accurate value of the Intelligence Quotient (IQ), which requires further discovery and investigation. Having sufficient information about the brain cognitive processes while performing the logical-mathematical problems allows using this information to improve the learning and teaching methods. In addition, it provides an accurate description of the cognitive identity of individuals [2]. The functional differences of the brain of individuals in various mental activities could also be utilized for the IQ estimation to choose suitable persons for specific jobs.

To discover the nature of the cognitive procedures of the brain during problem-solving, one can use various brain mapping systems such as electroencephalogram (EEG), functional Magnetic Resonance Imaging (fMRI), and functional Near-Infrared Spectroscopy (fNIRS). Several authors have used EEG to study the effects of various factors such as aging, having different skills, and different levels of intelligence on solving different arithmetic problems [5–11]. These studies usually employ Raven's Progressive Matrix (RPM) intelligence test to discover the hidden events happening inside the brains of different people. They used a reduced 15-item version of the original Raven test [12] and recorded EEG from the scalp of the subjects. Predictability, self-

* Corresponding author.
E-mail address: ksetareh@ut.ac.ir (S.K. Setarehdan).

https://doi.org/10.1016/j.compbiomed.2019.05.017
Received 10 January 2019; Received in revised form 20 May 2019; Accepted 20 May 2019

similarity, and the stable behavior of the EEG signal grew when the brain concentrated on one cognitive problem [13]. Patterns of information processing in an individual with high and low intelligence vary depending on the type of mental activity and its degree of difficulty. Positron Emission Tomography (PET) and EEG studies have suggested that there is a negative correlation between the intelligence and brain activities [14,15], while new studies from fMRI report positive correlation in most cases [16,17].

EEG is very powerful to detect transient processes in the range of 100 ms and helps understand the functional pattern of the brain; however, there are several limitations in this regard. There is no one-to-one correspondence between any specific frequency band activities (i.e., delta, theta, alpha, beta, and gamma) or Event-Related Potentials (ERP) and cognitive activity or a mental process. Further, eye artifacts affect the signals recorded in the frontal area and thus should be removed from EEG signals. Therefore, some studies have employed fNIRS to study various events of the brain during cognitive activities because of its higher temporal and spatial resolution compared to fMRI and EEG, respectively.

Studies have suggested that fNIRS is a valuable tool for illustrating the relationship between the function of the cortical regions of the brain during cognitive activities [18]. fNIRS extracts hemodynamic response function (HRF) of the brain via measuring local changes of the concentration of oxygenated (HbO2) and deoxygenated hemoglobin (Hb) [19,20]. fNIRS is integrated over a longer timeframe and provides a better localization compared to EEG [20]; it represents only the local activity of the brain directly below the probes, while, this feature is difficult for EEG because of the volume conduction effect [21]. This modality is cost-effective, portable, silent, low noise, easy to handle, and provides a better temporal resolution in comparison to fMRI [22].

Some studies have reported the use of a hybrid EEG-fNIRS method in Brain-Computer Interfaces (BCI) systems to classify the cognitive and motor tasks [23]. Hong et al. [24] combined common fNIRS features (e.g., the mean of the signal) with the power spectrum density (PSD) from EEG to improve the accuracy of classification of the mental arithmetic and word formation tasks for locked-in syndrome (LSI) patients. The EEG-fNIRS combination could also be useful to investigate the brain patterns and relationships between the hemodynamic response and electrical activity [25] during various tasks such as emotional [26] and stress processing [27].

Despite the works in the context of synchronous processing of EEG and fNIRS imaging modalities in different fields, most studies have used EEG to investigate the brain procedures when solving logical-mathematical problems. To the best of our knowledge, no research has simultaneously benefited from EEG and fNIRS for analyzing the information processing of the brain while solving the problems with different difficulty levels. These modalities are supplementary tools that can reveal uncovered phenomena of the brain. The temporal resolution provided by EEG complements the spatial resolution obtained by fNIRS. In this regard, there is no interaction between the electrical and optical activation of the brain [25,26]. Further, it is also possible to compare the efficiency of the two methods and select the optimal method.

One of the major contributions of this paper is the examination of the neuronal activation (obtained by EEG) and the local hemodynamic response of the brain (acquired by fNIRS) while performing a mathematical-logical intelligence test. A secondary contribution of the present study is the use of a combination of neuronal and hemodynamical features to estimate the IQ of subjects regardless of their answers to the intelligence test.

The purpose of this paper is to modify the paradigm of the experiment such that the weaknesses of previous researches are addressed. Previous studies have often identified low- and high-intelligence subjects based on their responses during the experiment, but they may be stressed, and thereby the function of their brain might be affected. Therefore, we do not use the answers of subjects during the experiment to categorize them as low- and high-intelligence groups. Instead, we utilize another mathematical-logical intelligence test (similar to Raven's Progressive Matrices test) to cluster low-and high–intelligence individuals before the experiment. Typically, intelligence scales use different classification labels for different IQ score ranges [28,29]. Here, we consider IQ $<$ 120 and IQ $\geq$ 120 to cluster the subjects into low- and high-intelligence groups, respectively.

We also eliminate the time limitations of each question to decrease the stress of subjects and each subject is free to answer as many questions as possible in each task block. Before initiating the experiment, the subjects are told that their focus on questions is important and it does not matter how many questions they answer. Furthermore, all questions are presented in an ascending order in terms of difficulty and thus the fatigue effect on the subjects could be observed in difficult questions. To resolve this problem, we randomly assign the blocks of questions with varying difficulties to the subjects.

The paper is organized as follows. Section 2 describes our experiment paradigm and dataset. This section also explains our methodology including feature extraction, feature selection and regression methods. The experimental results are explained in section 3. In section 4, we discuss some experimental results. Finally, we conclude the paper in section 5.

2. Materials and methods

2.1. Data acquisition

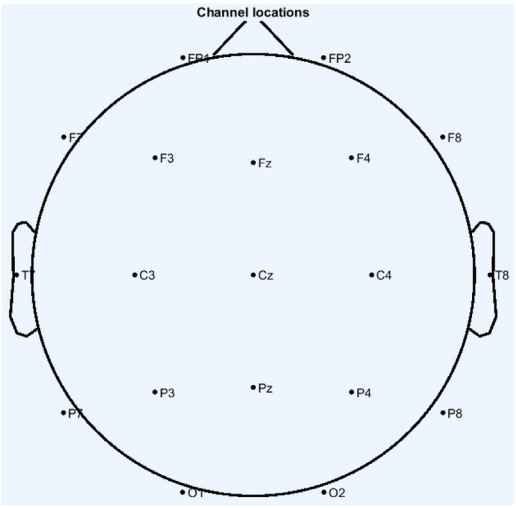
In this study, to collect our dataset, we used 11 healthy Persian-speaking graduate students (5 females and 6 males) between the ages of 24 and 30 years (27.48 $\pm$ 3 years). All subjects were right-handed, had normal or corrected vision, and had no history of neurological illness. They were asked to be comfortable and limit their head movements as much as possible while performing the experiment. Our dataset was balanced and consisted of 6 low- and 5 high-intelligence subjects. In this way, we did not need a class balancing technique in the case of unbalanced dataset [30].

Prior to the experiment, the subjects participated in the process of the standard Cattle IQ test, Scale 3 (similar to Raven's Progressive Matrices test) to calculate the reference IQ score [31–33]. We also utilized the RPM test when recording EEG and fNIRS signals, consisting of 60 questions of intelligence based on the logical-mathematical domain. These problems were divided into five sets of 12 problems each (sets A, B, C, D and E), thereby increasing the difficulty both within and across sets [34]. The RPM test [35] is a non-verbal test where each question includes a matrix of geometric figures with one entry missing. The subject should select the correct missing entry of each question from a set of choices to complete the logical series [13].

Considering the use of fNIRS and the need for a block-based design of the experiment, we utilized 60 questions of the RPM test. We put these questions into four blocks (A, B, C, and D) according to their difficulty levels. This experiment consisted of four sets randomly displayed on a computer screen. Each set was composed of 30 s of rest (R) and 60 s of task (T). The questions of each task block were randomly selected and displayed. Depending on the speed of responsiveness, each subject could answer a different number of questions in each task block. The total recording time for each subject was 390 s (considering 30 s after the four sets at the end of the experiment).

EEG was recorded through 19 scalp electrodes (g.tec's 80-channel biosignal amplifier named g.HIamp)[1] placed in accordance with the international 10/20 systems [36]. The signals were recorded from the following channels: Fp1, Fp2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, O1, and O2. Prior to the test, EEG signals were recorded with open and closed eyes. A further EEG recording was captured during the Raven's intelligence test. EEG signals were digitized at a

[1] www.gtec.at/Products/Hardware-and-Accessories/g.HIamp-Specs-Features.



(a)

(b) (c)

Fig. 1. The Optodes and electrodes placement; (a) the location of EEG electrodes and (b), (c) fNIRS Optodes placement and channel configuration for problem solving task.

sampling rate of 512 Hz. Fig. 1 (a) displays the placement of EEG electrodes on the head of the subjects.

In this study, fNIRS signals were recorded via a 48-channel MR compatible OxyMonfNIRS from Artinis[2]. This device is capable to measure oxy-, deoxy, and total hemoglobin concentration variations. MRI, EEG and ECG does not interfere with the optical signal. The sampling rate was 10 Hz. Optodes and templates were positioned over the prefrontal, frontal, and parietal lobes to record 32-channel fNIRS data at two wavelengths of 760 and 845 nm. We applied a 3 cm emitter-detector distance, as shown in Fig. 1 (b, c). We performed the feature extraction and dimensionality reduction steps on EEG and fNIRS signals, separately, and then employed the combination of efficient EEG and fNIRS features as well as various regression methods to estimate the IQ value of the subjects (Fig. 2).

2.2. Pre-processing

EEG was filtered between 0.5 and 45 Hz using a band-pass FIR filter [37,38]. After subtracting the baseline based on the mean of each data channel, common artifacts were eliminated with the help of EEGLAB and ADJUST via independent component analysis (ICA) based on runica algorithm [39] under Matlab. One example of the raw and pre-processed EEG signals is displayed in Fig. 3.

The fNIRS data from each subject were processed channel-by-channel [40]. To remove the baseline drift, we fitted one line to each rest block using linear regression analysis [41] and then subtracted the subsequent task block from this linear model along with the mean of its corresponding rest block. In fact, we considered each rest block as a baseline for its subsequent task block. A fourth-order band-pass Butterworth filter between 0.01 and 0.2 Hz was applied to remove the low frequency (< 0.01 Hz) and high-frequency (> 0.2 Hz; heartbeat, respiratory rate, instrument noise) information from the fNIRS signal [24,42,43]. We utilized the clean fNIRS data to identify significant channels. Fig. 4 demonstrates an example of an fNIRS signal in one of the channels placed in the frontal lobe, in both raw and pre-processed conditions.

The average of temporal profiles of oxyhemoglobin was computed across all subjects (N = 11). Thereafter, *t*-test [44] was applied to examine the difference of distinct task blocks (high-low, medium-low, and high-medium) of these profiles and to identify the effective fNIRS channels discriminating tasks including the problems with different difficulty levels. Afterward, the channels with p-values smaller than 0.05 were marked as the significant channels for each condition. Having reviewed overlapping significant channels in three conditions,

[2] https://www.artinis.com/oxymon.

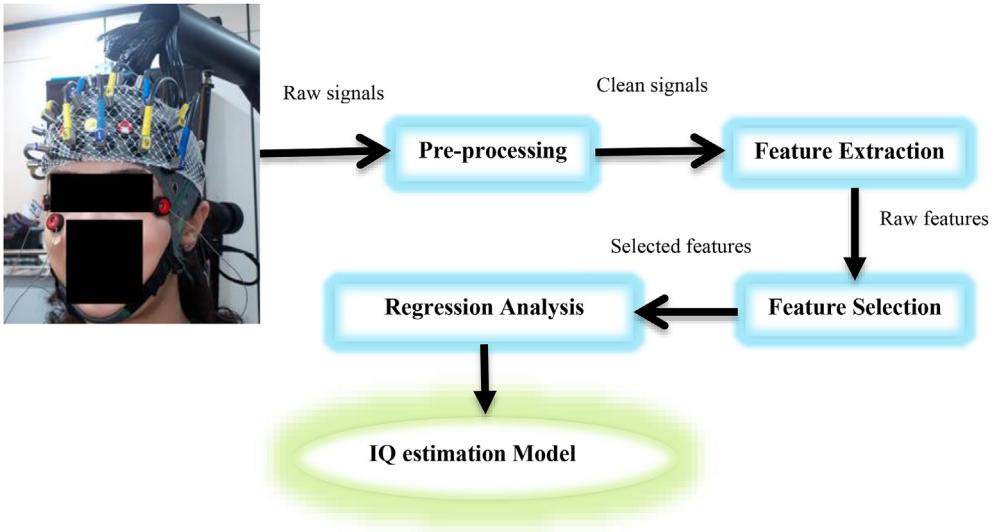


Fig. 2. The overall scheme of the IQ estimation procedure.

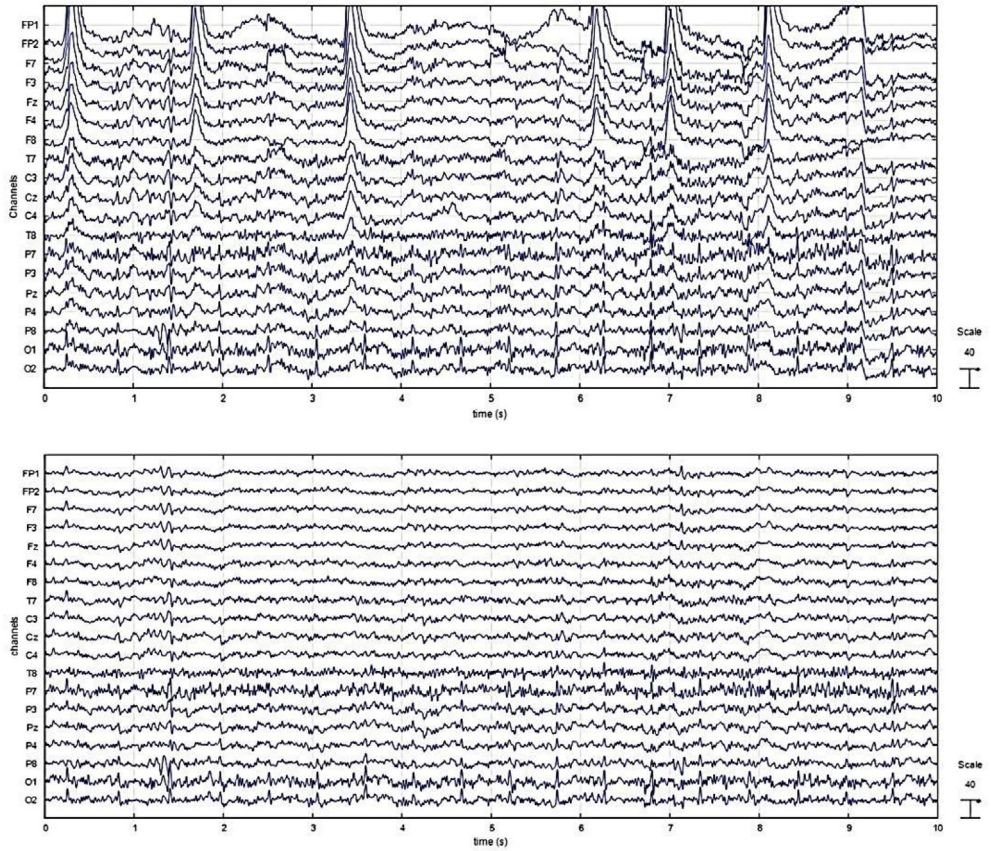
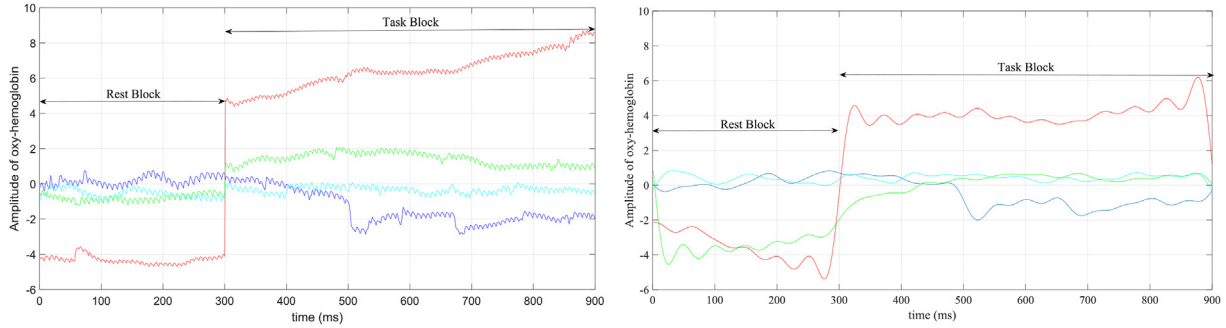


Fig. 3. One example of EEG signals before and after pre-processing.

17 channels were identified as the significant channels among 32 existing channels. Table 1 presents the p-values of 32 channels based on the pre-defined conditions. The location of the significant channels placed on the subjects head is marked in Fig. 5. We applied these significant channels to extract useful features for estimating the IQ of the subjects.

2.3. Feature extraction

In this section, the methods for EEG and fNIRS analysis are examined to obtain further information on the human brain while performing the Raven's intelligence test. Note that the features described below have been separately extracted from each block including rest (30 s) and task (60 s) blocks. The length of the segments of EEG and fNIRS was fixed for feature extraction. Finally, we had two kinds of feature resulted from rest and task blocks for the IQ estimation.



(a) (b)

Fig. 4. One channel (Rx4b-Tx7) of fNIRS signals before (a) and after pre-processing (b).

Table 1
The p-values of 32 fNIRS channels based on three pre-defined conditions to identify the significant channels; the significant channels are denoted by red color.

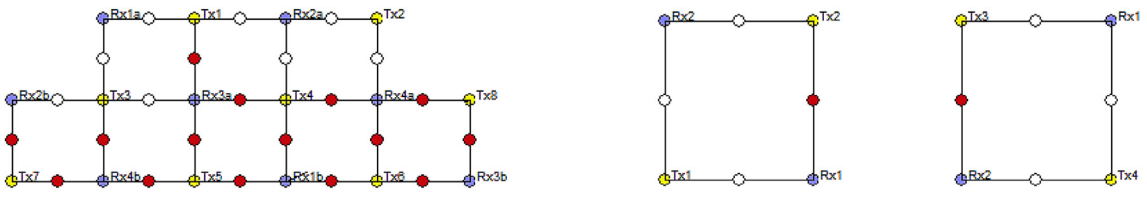
Channel number	High-Low (condition 1)	Medium-Low (Condition 2)	High-Medium (Condition 3)
1	1.0720 e-16	0.0085	7.5632 e-22
2	1.0145 e-05	0.0830	1.7940 e-04
3	0.6853	1.631 e-04	2.0698 e-04
4	0.0031	3.9197 e-16	0.0924
5	0.6819	6.9305 e-29	4.4035 e-23
6	0.0032	0.0652	0.0507
7	1.1947 e-11	0.0077	1.0079 e-07
8	8.7211 e-04	0.0104	5.1177 e-10
9	0.0043	0.6009	0.0012
10	2.1264 e-07	0.5584	2.7686 e-11
11	0.8180	1.1220 e-92	3.8986 e-24
12	0.0175	0.0037	7.0086 e-12
13	0.0012	1.0014 e-20	4.6113 e-10
14	3.9050 e-13	8.7959 e-08	2.3950 e-08
15	1.3974 e-12	4.4213 e-04	5.3398 e-10
16	2.6862 e-11	2.1345 e-07	3.9902 e-47
17	0.0143	1.2871 e-60	3.8435 e-43
18	0.7849	3.4001 e-73	9.6283 e-27
19	2.7093 e-08	2.7827 e-09	0.7539
20	8.0563 e-04	1.0353 e-07	1.6164 e-16
21	0.3424	1.1844 e-07	7.3345 e-08
22	8.3425 e-12	0.8282	3.7448 e-16
23	2.1010 e-16	1.1568 e-09	9.5778 e-47
24	0.6448	5.7897 e-71	1.8186 e-32
25	0.2228	3.0972 e-13	1.6511 e-10
26	1.1064 e-27	7.0378 e-87	1.4679 e-10
27	0.0029	1.7625 e-48	5.7413 e-10
28	2.4230 e-10	7.2133 e-95	1.0540 e-118
29	1.1570 e-22	2.3812 e-24	7.8542 e-72
30	0.0271	2.1580 e-162	2.6361 e-136
31	6.3027 e-10	1.3086 e-138	1.7040 e-142
32	0.1486	6.6437 e-94	1.9495 e-33

For EEG frequency analysis, clean EEG signals were obtained for five standard EEG bands: delta, theta, alpha, beta, and gamma [45]. Next, we extracted various features from each sub-band to estimate the intelligence level of each subject. The average power of each frequency sub-band was computed according to (1); Where $S_{xx}(\omega)$ and $P_{freq-band}$ represent the power spectral density (PSD) and power of the signal in a given frequency $[\omega_1, \omega_2]$, respectively. $E[x^*[n]x[n+m]]$ is the auto-correlation function of the discrete-time signal x[n].

$$S_{xx}(\omega) = \sum_{m=0}^{N-1} E[x^*[n]x[n+m]]e^{-\frac{j\omega m}{N}}$$
$$P_{freq-band} = \frac{1}{\pi}\int_{\omega_1}^{\omega_2} S_{xx}(\omega)d\omega \quad (1)$$

Many studies have utilized the wavelet transform as a powerful approach for feature extraction. The wavelet transform is appropriate for non-stationary signals such as EEG, so we also applied discrete wavelet transform (DWT) to the EEG signals to present a multi-resolution time-frequency description of the signal [46–48]. The original signal could be decomposed into different decomposition levels; however, high decomposition levels significantly increase the computational cost and the time required for the entire process. Therefore, in this work, we performed Daubechies 8 (db8) mother wavelet with five decomposition levels (corresponding to the number of frequency sub-bands) to calculate the Shannon entropy [49] of the detailed coefficients of each sub-band.

The application of linear approaches to study non-linear systems such as the brain does not seem to reveal the nature of dynamical events occurring in the brain. If we consider the brain as a chaotic or quasi-chaotic system, it gives us the opportunity to model the brain in a more realistic, albeit though not simpler, way [13]. Chaotic properties such as fractal dimensions (FD) are employed to describe non-linear time series such as EEG signal. FD is a non-local characteristic which measures the complexity of fundamental patterns or self-similarity features in the signal. Self-similarity means that the properties of a very small portion are the same as the larger portion of the signal. For more than



(a) (b)

Fig. 5. The schematic of fNIRS transmitters (yellow), receivers (blue), channels (white) and the significant channels (red) configuration on the (a) prefrontal, frontal, and (b) parietal lobes.

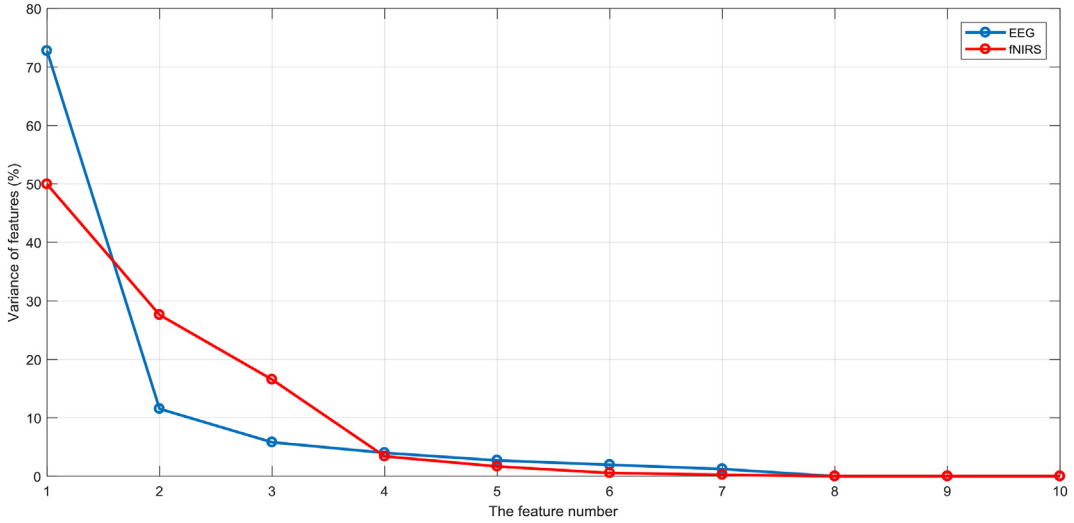


Fig. 6. The percentage of variance of first 10 EEG (blue) and fNIRS (red) principal components.

two decades, Higuchi's fractal dimension (HFD), as a nonlinear method, has claimed a special position in the analysis of biological signals including EEG [50]. HFD contains more information about the dynamics of the time series compared to other methods estimating the fractal dimension such as Hurst (H) exponent. This algorithm also uncovers the crossover behavior of the time series like EEG [51]. Accordingly, in this study, the HFD and H exponent features were computed from the rest and task blocks to investigate the predictability and chaotic status of the EEG signal.

All of the EEG feature values were scaled between 0 and 1 as shown in (2); Feat′ represents the rescaled feature values between 0 and 1.

$$Feat' = \frac{Feat - \min(Feat)}{\max(Feat) - \min(Feat)} \tag{2}$$

In the next step, the average value of fNIRS data of each subject on the significant channels was extracted to compute the key fNIRS features. Six kinds of features (mean (eqn. (3)), variance (eqn. (4)), skewness (eqn. (5)), kurtosis (eqn. (6)), slope, and peak) were calculated for different task and rest blocks [52].

$$Mean = \sum_{i=1}^{N} S_i \tag{3}$$

$$\mathrm{var}(S) = \frac{\sum_{i=1}^{N} (S_i - M)}{N} \tag{4}$$

$$skew(S) = E\left(\left(\frac{\sum_{i=1}^{N} (S_i - M)}{\mathrm{var}}\right)^3\right) \tag{5}$$

$$kurt(S) = E\left(\left(\frac{\sum_{i=1}^{N} (S_i - M)}{\mathrm{var}}\right)^4\right) \tag{6}$$

where, N is the number of observations and S_i represents the oxy-hemoglobin data. The 'peak' feature corresponds to the first maximum of the signal and its location. The 'slope' is calculated via fitting a line to the data points between the start of the block and its first peak point. Since oxy-hemoglobin is more sensitive to the cerebral activities compared to the deoxy-hemoglobin signal [18], in this study, only the concentration variations of the oxy-hemoglobin signals were considered.

2.4. Feature selection

If the above features are extracted from each subject, the feature vector dimension is high, which may result in over-fitting problem (especially in the case of small-sample size) or undermining the accuracy of regression stage. Therefore, we utilized Linear Discriminant Analysis (LDA), and Principal Component Analysis (PCA) to generate a set of EEG and fNIRS latent features according to their importance in the logical-mathematical intelligence domain [53]. LDA is a supervised algorithm, and so it uses the information of classes to find a linear transformation of features for dimensionality reduction in various BCI applications [54]. In contrast, PCA does not take information of classes into account (as an unsupervised algorithm) and uses the variance of each feature to find new features [55].

In this research, we employed the LDA as a transformation algorithm (not classification algorithm) to reduce the dimension of the input space. For dimensionality reduction, the maximum dimension of the new feature vector is the number of classes in the data minus 1 [55]. The LDA transformation is defined as

$$F_{LDA_{(s\times m)}} = F_{(s\times n)} \times W^{T}_{LDA_{(n\times m)}} \tag{7}$$

where, F denotes the original feature space, W_{LDA} represents the Eigenvector resulted from the Eigen decomposition method, and F_{LDA} shows the transformed features; (×)denotes matrix product; s and n indicate the number of observations and the original feature, respectively; and m refers to the new dimension of transformed features.Here, we have a two-class data. Note that due to the small sample size, we clustered each observation as a low-(0) or high-(1) intelligence class according to the IQ computed from the Cattle test and threshold 120. Hence, the new dimension of the feature vector is equal to 1, for both EEG and fNIRS data.

The nonlinear methods proposed for the dimensionality reduction are often not capable of outperforming traditional linear techniques such as PCA [55]. Thus, we utilized PCA method to generate the efficient artificial features for enhancing our IQ estimation model. It was also possible to compare the results of these two dimensionality reduction algorithms, LDA and PCA, and choose the best method for our study.

The idea behind PCA is that features presenting high variance are more likely to have a good split between classes. This method performs a linear transformation to generate a new set of features (artificial variables) based on a combination of the old features [56]. These new features are not of any value for us as they are simply algebraic. Fig. 6 presents the percentage of variance of the first 10 EEG and fNIRS principal components (PCs). The first five and four PCs from EEG and fNIRS accounted for approximately 96.80% and 97.51 of the entire variance, respectively. Therefore, the information loss was less than

4%. We utilized these new EEG and fNIRS features resulted from LDA and PCA algorithms to train the regression models for IQ determination.

2.5. Regression analysis for IQ estimation

We used the EEG and fNIRS features resulted from LDA and PCA algorithms to calculate the regression models for estimating the IQ of individuals. Regression analysis models the relationship between a response variable (here, the value of IQ) and a set of quantitative explanatory variables (here, a subset of fNIRS and EEG features) to determine which explanatory variables influence the response variable [57]. To find a robust model, we utilized two regression methods including linear regression (LR) and support vector regression (SVR) to predict the subjects' IQ.

Linear regression model, called general linear model, has the following form

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \quad (8)$$

where, x_k and β_k represent the quantitative explanatory variables and coefficients assigned to each explanatory variable, respectively. Also, y shows the optimal output, which is the value of IQ in this study. β_0 is a constant term as intercept, and intercept predicts the value of y when $x = 0$.

Conventional regression methods such as artificial neural networks (ANNs) require a large-sample size to guarantee the prediction precision. In comparison, in real prediction cases (such as our dataset to estimate the IQ), the sample sizes are limited due to the difficulty and cost involved in the data recording. In this regard, Support Vector Machines (SVMs) do not suffer from the problems of ANNs like under-learning, over-learning, and determination of the network structure. This method determines the decision function using a few samples called support vectors. Therefore, SVM is an optimal method for small-sample regression [58]. In this study, we exploited epsilon-SVR (or ε-SVR) [59] with three different kernel functions (i.e., Linear, Polynomial, and Radial Basis Function (RBF)) with $C = 1.0$ and epsilon = 0.1 [60] to perform the regression analysis. We compared the efficiency of three kernel functions and chose the best kernel type for the IQ estimation task.

3. Experimental results

In this section, the results of the experiments performed in this study are presented. The IQ value of the subjects was determined in two modes; using a combination of selected EEG and fNIRS features and only selected fNIRS attributes. To find the best IQ estimation model, we employed various combinations of the artificial features generated by LDA and PCA algorithms and two regression methods (i.e., LR and ε-SVR). Due to the small-sample size, we utilized a leave-one-out test; one subject was considered as a test data, whose IQ was estimated based on the model of the remaining subjects. Next, we trained our final models on all data ($N = 11$) to present an accurate model to estimate the IQ of new observations.

In linear regression models, we identified significant explanatory variables based on their p-value (the variables having p-values smaller than 0.05 are significant) and R-squared (as R-squared grows, the response variable becomes more explained by the explanatory variables) in comparison to the full model. In ε-SVR models, we evaluated the error of the models with different kernel functions and chose the best model. Based on our dataset, the optimum model was constructed via the polynomial kernel. Therefore, we used this kind of kernel function to compare the relative error of the two regression methods.

Fig. 7 shows the percentage of the relative error for the eleven subjects using different IQ estimation models and the leave-one-out evaluation.

Considering the results displayed in Fig. 7, the IQ estimation of all models was performed with an acceptable error. The PCA-based dimension reduction method presented a better performance compared to the LDA for our study. The best linear regression model based on the information of fNIRS signal and PCA algorithm is according to (9), which was trained on 11 subjects.

$$0.39 \times PC_2 + 0.55 \times PC_3 + 113.28 \quad (9)$$

where, PC_2, and PC_3 represent the second and third principal components resulted from performing PCA algorithm on the original fNIRS features.

The results of the IQ estimation procedure for PCA-LR and PCA-ε-SVR models (with leave-one-out approach) are shown in Table 2 and Table 3, respectively.

The simultaneous use of EEG and fNIRS features resulted in 3.093% and 3.119% relative error for the LR and ε-SVR methods, respectively; however, the error of the IQ estimation of 11 subjects was increased to 3.749% and 3.690% in the LR- and ε-SVR-based regression models, when we used the fNIRS features for the IQ prediction. Therefore, the use of a combination of selected EEG and fNIRS features yields a better performance for the IQ estimation compared to only fNIRS features.

4. Discussion

The current study indicated that the information extracted from the fNIRS and EEG data was valuable for predicting the IQ of individuals. Due to the small-sample-size and high-dimension features, we applied two dimensionality reduction methods including LDA and PCA to the original feature space. Based on the results in Fig. 7, PCA algorithm had a better performance compared to LDA for feature selection of this task. According to Fig. 7, the performances of LR and ε-SVR methods were similar. Here, the best models based on EEG-fNIRS and only fNIRS attributes were obtained for ε-SVR and LR, respectively.

The simultaneous implementation of EEG and fNIRS imaging modalities provided a 3.093% relative error in the IQ estimation of 11 subjects based on the leave-one-out approach. On the other hand, utilization of the selected fNIRS features for the IQ estimation resulted in 3.690% relative error. Therefore, the EEG based features proved their benefit as complementary attributes for the fNIRS data in the IQ estimation task. Note that the noise and artifact reduction of EEG signal and its processing are more complex and time-consuming than those of fNIRS signal. Also, the simultaneous recording of two modalities (like EEG and fNIRS) is more challenging and time-consuming, leading to a limited dataset. Therefore, considering the small difference of errors (Tables 2 and 3), the use of the fNIRS imaging modality alone in IQ estimation facilitates the process of recording and reducing computational costs.

Previous works [2,13,61] have used 14-channel EEG and abbreviated 15-item version of Raven's test to study the brain dynamics and identify the inter- and intra-individual variability in the way the brain worked [13]. However, they did not use EEG information to estimate the IQ value of the subjects. To the best of our knowledge, our study is the first step in utilizing the EEG and fNIRS information to predict the IQ value of individuals participating in the experiment.

5. Conclusion

To the best of our knowledge, prior studies have simply focused only on clustering subjects into low- and high-intelligence groups with tests lasting more than 10 min. They also estimated the intelligence level of subjects based on their answers to the intelligence test while recording signals. However, their algorithms were unable to determine the IQ value of subjects. In the present study, we filled these gaps through simultaneous recording and processing of two modalities, i.e., EEG and fNIRS, and revising the experimental paradigm. As we do not consider any time limitation for answering each question, the stress of subjects decreased significantly compared to that reported in the previous works. Further, the real IQ of subjects was computed before the

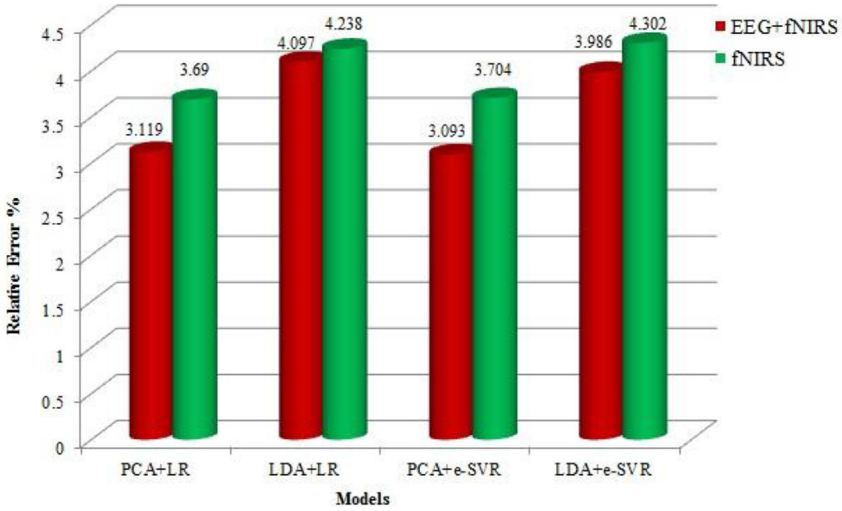


Fig. 7. The percentage of the relative error of different IQ estimation models based on the EEG-fNIRS (red) and fNIRS features (green).

Table 2
The estimated IQ of 11 subjects using the selected EEG-fNIRS and fNIRS features based on the principal component analysis and linear regression method.

Subjects	Real IQ	IQ: EEG and fNIRS	IQ: fNIRS	Relative error (EEG + fNIRS)	Relative error (fNIRS)
S1	104	106.805	108.813	3.030	4.628
S2	109	110.637	106.405	0.567	2.381
S3	140	141.158	142.774	1.418	1.981
S4	109	116.380	114.363	5.374	4.920
S5	134	137.942	131.239	1.962	2.060
S6	104	107.438	109.782	5.146	5.560
S7	119	121.853	115.084	2.851	3.291
S8	128	120.104	118.464	5.276	7.450
S9	100	94.065	95.163	4.567	4.837
S10	138	137.751	139.852	1.986	1.342
S11	124	122.395	121.336	2.124	2.148

Table 3
The estimated IQ of 11 subjects using the selected EEG-fNIRS and fNIRS features based on the principal component analysis and epsilon-support vector regression method.

Subjects	Real IQ	IQ: EEG and fNIRS	IQ: fNIRS	Relative error (EEG + fNIRS)	Relative error (fNIRS)
S1	104	106.805	108.496	2.697	4.323
S2	109	110.637	110.247	1.501	1.144
S3	140	141.158	136.459	0.827	2.530
S4	109	116.380	113.173	6.770	3.829
S5	134	137.942	132.511	2.941	1.111
S6	104	107.438	110.018	3.306	5.787
S7	119	121.853	124.415	2.398	4.550
S8	128	120.104	119.256	6.169	6.831
S9	100	94.065	93.628	5.935	6.372
S10	138	137.751	141.923	0.180	2.842
S11	124	122.395	126.381	1.294	1.920

experiment; so, these values for IQ were more reliable than the IQ calculation procedure during the experiment since the neural behavior of the subjects was not affected by the signal recording and stressful conditions.

Several models were extracted according to two subsets of features including EEG-fNIRS and fNIRS information using various dimensionality reduction methods (i.e., PCA and LDA) and regression analysis (i.e., LR and ε-SVR).

The best IQ determination models were obtained using the PCA method. The performances of the best LR and ε-SVR-based models were slightly different. The minimum relative error of EEG-fNIRS- and fNIRS-based systems was achieved by 3.093% and 3.690% for ε-SVR and LR regression models, respectively. Therefore, we preferred to use the fNIRS-based model because of the lower cost and complexity of its processing. Furthermore, 17 significant fNIRS channels were identified using *t*-test. These channels were placed on the prefrontal, frontal, and parietal lobes. Neuropsychological research also suggests that the prefrontal, frontal, and parietal lobes are very important in the cognitive problem-solving process [1] and do not need conscious learning. The findings of the present study also verified this theory.

In the future, we can improve this IQ estimation algorithm based on fNIRS via increasing the number and variety of subjects, because recording fNIRS signal is much simpler in the areas limited to the significant channels compared to the EEG-fNIRS or EEG recording and processing across the entire head.

Conflicts of interest

None declared.

Acknowledgements

Sh. Firooz would like to appreciate the Iranian National Brain Mapping Lab (NBML), Tehran, Iran for providing data acquisition service for this research work.

This study was carried out in accordance with the recommendations NBML and Medical University of Iran, Tehran, Iran with written informed consent from all subjects. The Medical University of Iran and NBML approved the protocol. Moreover, this research did not receive specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] P. Carter, IQ and Psychometric Tests, second ed., Replica Press, 2007.
[2] F.M. Cordova, M.H. Diaz, F. Cifuentes, L. Canete, F. Palominos, Identifying problem solving strategies for learning styles in engineering students subjected to intelligence test and EEG monitoring, Procedia Comput. Sci. 55 (2015) 18–27 https://doi.org/10.1016/j.procs.2015.07.003.
[3] J.-S. Kang, A. Ojha, G. Lee, M. Lee, Differences in brain activation patterns of individuals with high and low intelligence in linguistic and visuo-spatial tasks: an EEG study, Intelligence 61 (2017) 41–55 https://doi.org/10.1016/j.intell.2017.01.002.
[4] C.A. Paynter, K. Kotovsky, L.M. Reder, Problem-solving without awareness, Neuropsychologia 48 (10) (2010) 3137–3144 https://doi.org/10.1016/j.neuropsychologia.2010.06.029.
[5] T. Hinault, P. Lemaire, N. Phillips, Aging and sequential modulations of poorer

strategy effects: an EEG study in arithmetic problem solving, Brain Res. 1630 (2016) 144–158 https://doi.org/10.1016/j.brainres.2015.10.057.
[6] R. Stanescu-Cosson, P. Pinel, P.F. van De Moortele, D. Le Bihan, L. Cohen, S. Dehaene, Understanding dissociations in dyscalculia: in a brain imaging study of the impact of number size on cerebral networks for exact and approximate calculation, Brain: J. Neurol. 123 (11) (2000) 2240–2255 https://doi.org/10.1093/brain/123.11.2240.
[7] S. Micheloyannis, V. Sakkalis, M. Vourka, C.J. Stam, P.G. Simos, Neural networks involved in mathematical thinking: evidence from linear and non-linear analysis of electroencephalographic activity, Neurosci. Lett. 373 (3) (2005) 212–217 https://doi.org/10.1016/j.neulet.2004.10.005.
[8] B. De Smedt, R.H. Grabner, B. Studer, Oscillatory EEG correlates of arithmetic strategy use in addition and subtraction, Exp. Brain Res. 195 (4) (2009) 635–642 https://doi.org/10.3389/fpsyg.2012.00428.
[9] Y. Ku, B. Hong, X. Gao, S. Gao, Spectra-temporal patterns underlying mental addition: an ERP and ERD/ERS study, Neurosci. Lett. 472 (1) (2010) 5–10 https://doi.org/10.1016/j.neulet.2010.01.040.
[10] P. Zarjam, J. Epps, F. Chen, N.H. Lovell, Estimating cognitive workload using wavelet entropy-based features during an arithmetic task, Comput. Biol. Med. 43 (12) (2013) 2186–2195 https://doi.org/10.1016/j.compbiomed.2013.08.021.
[11] M.I. Nunez-Pena, Effects of training on the arithmetic problem-size effect: an event-related potential study, Exp. Brain Res. 190 (1) (2008) 105–110 https://doi.org/10.1007/s00221-008-1501-y.
[12] W.B. Bilker, J.A. Hansen, C.M. Brensinger, J. Richard, R.E. Gur, R.C. Gur, Development of abbreviated nine-item forms of the Raven's standard progressive matrix test, Assessment 19 (3) (2012) 354–369 https://doi.org/10.1177/1073191112446655.
[13] M.H. Diaz, F.M. Cordova, L. Canete, F. Palominos, F. Cifuentes, C. Sanchez, M. Herrera, Order and chaos in the brain: fractal time series analysis of the EEG activity during a cognitive problem solving task, Procedia Comput. Sci. 55 (2015) 1410–1419 https://doi.org/10.1016/j.procs.2015.07.135.
[14] N. Jasovec, Differences in cognitive processes between gifted, intelligent, creative, and average individuals while solving complex problems: an EEG study, Intelligence 28 (3) (2000) 213–237 https://doi.org/10.1016/S0160-2896(00)00037-4.
[15] R.J. Haier, B.V. Siegel, C. Tang, L. Abel, M.L. Buchsbaum, Intelligence and changes in regional cerebral glucose metabolic rate following learning, Intelligence 16 (3–4) (1992) 415–426 https://doi.org/10.1016/0160-2896(92)90018-M.
[16] U. Basten, K. Hilger, C.J. Fiebach, Where smart brains are different: a quantitative meta-analysis of functional and structural brain imaging studies on intelligence, Intelligence 51 (2015) 10–27 https://doi.org/10.1016/j.intell.2015.04.009.
[17] U. Basten, C. Stelzel, C.J. Fiebach, Intelligence is differentially related to neural effort in the task-positive and the task-negative brain network, Intelligence 41 (5) (2013) 517–528 https://doi.org/10.1016/j.intell.2013.07.006.
[18] M. Dadgostar, S.K. Setarehdan, S. Shahzadi, A. Akin, Functional connectivity of the PFC via partial correlation, Optik 127 (11) (2016) 4748–4754 https://doi.org/10.1016/j.ijleo.2016.01.139.
[19] F.M. Noori, N. Naseer, N.K. Qurreshi, H. Nazeer, R.A. Khan, Optimal feature selection from fNIRS signals using genetic algorithms for BCI, Neurosci. Lett. 647 (2017) 61–66 https://doi.org/10.1016/j.neulet.2017.03.013.
[20] F. Wallois, M. Mahmoodzadeh, A. Patil, R. Grebe, Usefulness of simultaneous EEG–NIRS recording in language studies, Brain Lang. 121 (2) (2012) 110–123 https://doi.org/10.1016/j.bandl.2011.03.010.
[21] C.-M. Lu, Y.-J. Zhang, B.B. Biswal, Y.-F. Zang, D.-L. Peng, C.-Z. Zhu, Use of fNIRS to assess resting state functional connectivity, J. Neurosci. Methods 186 (2) (2010) 242–249 https://doi.org/10.1016/j.jneumeth.2009.11.010.
[22] K.-S. Hong, H.-D. Nguyen, State-space models of impulse hemodynamic responses over motor, somatosensory, and visual cortices, Biomed. Opt. Express 5 (6) (2014) 1778–1798 https://doi.org/10.1364/BOE.5.001778.
[23] A.P. Buccino, H.O. Keles, A. Omurtag, Hybrid EEG-fNIRS asynchronous brain-computer interface for multiple motor tasks, PLoS One 11 (1) (2016), https://doi.org/10.1371/journal.pone.0146610.
[24] K.-S. Hong, M. Jawad Khan, M.J. Hong, Feature extraction and classification methods for hybrid EEG-fnirs brain-computer interfaces, Front. Hum. Neurosci. 12 (246) (2018), https://doi.org/10.3389/fnhum.2018.00246.
[25] M. Balconi, L. Cortesi, D. Crevelli, Motor planning and performance in transitive and intransitive gesture and imagination. Does EEG (RP) activity predict hemodynamic (fNIRS) response? Neurosci. Lett. 648 (2017) 59–65 https://doi.org/10.1016/j.neulet.2017.03.049.
[26] M. Balconi, E. Grippa, M.E. Vanutelli, What hemodynamic (fNIRS), electrophysiological (EEG) and autonomic integrated measures can tell us about emotional processing, Brain Cogn. 95 (2015) 67–76 https://doi.org/10.1016/j.bandc.2015.02.001.
[27] F.M. Al-Shargie, T.B. Tang, N. Badruddin, S.C. Dass, M. Kiguchi, Mental stress assessment based on feature level fusion of fNIRS and EEG signals, https://doi.org/10.1109/ICIAS.2016.7824060, (2016).
[28] R.J. Sternberg, Encyclopedia of Human Intelligence, first ed., Macmillan, 1994.
[29] D.J. Barthalomew, Measuring Intelligence: Facts and Fallacies, Cambridge University Press, 2004, https://doi.org/10.1017/CBO9780511490019.
[30] A. Comelli, A. Stefano, S. Bignardi, G. Russo, M.G. Sabini, M. Ippolito, S. Barone, A. Yezzi, Active contour algorithm with discriminant analysis for delineating tumors in positron emission tomography, Artif. Intell. Med. 94 (2019) 67–78 https://doi.org/10.1016/j.artmed.2019.01.002.
[31] R.B. Cattell, Culture-Free Intelligence Test, Scale 1, Handbook, Institute of Personality and Ability, Champaign, Illinois, 1949.
[32] R.B. Cattell, S.E. Krug, K. Barton, Technical Supplement for the Culture Fair Intelligence Tests, Scales 2 and 3, Institute for Personality and Ability Testing, Champaign, 1973.
[33] G. Domino, M.L. Domino, Psychological Testing: an Introduction, Cambridge University Press, 9781139455145, 2006.
[34] M. Kunda, K. McGreggor, A.K. Goel, Taking a look (Literally!) at the raven's intelligence test: two visual solution strategies, Proceedings of the 32nd Annual Conference of the Cognitive, Science Society, 2010, pp. 1961–1969.
[35] J. Raven, The Raven's progressive matrices: change and stability over culture and time, Cogn. Psychol. 41 (1) (2000) 1–48 https://doi.org/10.1006/cogp.1999.0735.
[36] H.H. Jasper, The ten twenty system of the international federation, Electroencephalogr. Clin. Neurophysiol. 10 (1958) 371–375.
[37] M. Rudner, C. Signoret, The role of working memory and executive function in communication under adverse conditions, Front. Psychol. 7 (2016), https://doi.org/10.3389/fpsyg.2016.00148.
[38] R. Xu, Ch. Zhang, F. He, X. Zhao, H. Qi, P. Zhou, L. Zhang, D. Ming, How physical activities affect mental fatigue based on EEG energy, connectivity, and complexity, Front. Neurol. 9 (915) (2018), https://doi.org/10.3389/fneur.2018.00915.
[39] A. Mognon, J. Jovicich, L. Bruzzone, M. Buiatti, ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features, Psychophysiology 48 (2) (2011) 229–240 https://doi.org/10.1111/j.1469-8986.2010.01061.x.
[40] F. Tian, V. Sharma, F.A. Kozel, H. Liu, Functional near-infrared spectroscopy to investigate hemodynamic responses to deception in the prefrontal cortex, Brain Res. 1303 (2009) 120–130 https://doi.org/10.1016/j.brainres.2009.09.085.
[41] D.A. Freedman, Statistical Models: Theory and Practice, Cambridge University Press, 2009.
[42] N. Naseer, K-Sh. Hong, fNIRS-based brain-computer interfaces: a review, Front. Hum. Neurosci. 9 (3) (2015), https://doi.org/10.3389/fnhum.2015.00003.
[43] P. Pinti, F. Scholkmann, A. Hamilton, P. Burgess, I. Tachtsidis, Current status and issues regarding pre-processing of fnirs neuroimaging data: an investigation of diverse signal filtering methods within a general linear model framework, Front. Hum. Neurosci. 12 (505) (2019), https://doi.org/10.3389/fnhum.2018.00505.
[44] R. Mankiewicz, The Story of Mathematics, vol. 158, Princeton University Press, 9780691120461, 2004.
[45] V.K. Adikarapatti, Optimal EEG Channels and Rhythm Selection for Task Classification, Master Thesis Write State University, 2007.
[46] A. Nakate, P.D. Bahirgonde, Feature extraction of EEG signal using wavelet transform, Int. J. Comput. Appl. 124 (2) (2015).
[47] S.G. Firooz, F. Almasganj, Y. Shekofteh, Improvement of automatic speech recognition systems via nonlinear dynamical features evaluated from the recurrence plot of speech signals, Comput. Electr. Eng. 58 (2017) 215–226 https://doi.org/10.1016/j.compeleceng.2016.07.006.
[48] N. Erfanian-Saeedi, F. Almasganj, F. Torabinejad, Support vector wavelet adaptation for pathological voice assessment, IEEE. Trans. Computers in Biology and Medicine, 41 2011, pp. 822–828 (9), https://doi.org/10.1016/j.compbiomed.2011.06.019.
[49] P.A. Bromiley, N.A. Thacker, E. Bouhova-Thacker, Shannon Entropy, Renyi Entropy, and Information, The University of Manchester, 2010.
[50] S.S. Kesić, S.Z. Spasić, Application of Higuchi's fractal dimension from basic to clinical neurophysiology: a review, Comput. Methods Progr. Biomed. 133 (2016) 55–70 https://doi.org/10.1016/j.cmpb.2016.05.014.
[51] F. Cervantes-De la Torre, J. González-Gutiérrez, J.L. Carrillo-Estrada, J.C. Ruiz-Suárez, Fractal dimension algorithms and their application to time series associated with natural phenomena, J. Phys. Conf. Ser. 475 (1) (2003), https://doi.org/10.1088/1742-6596/475/1/012002.
[52] N. Naseer, F.M. Noori, N.K. Qureshi, K-Sh. Hong, Determining optimal feature-combination for LDA classification of functional near-infrared spectroscopy signals in brain-computer interface application, Front. Hum. Neurosci. 10 (237) (2016), https://doi.org/10.3389/fnhum.2016.00237.
[53] G. Roffo, Feature Selection Library (Matlab Toolbox), Technical Report, 2018.
[54] L. Ladha, T. Deepa, Feature selection methods and algorithms, Int. J. Adv. Trends Comput. Sci. Eng. 3 (5) (2011) 1787–1797.
[55] L.J.P. van der Maaten, E.O. Postma, H.J. van den Herik, Dimensionality Reduction: A Comparative Review, Tilburg University, 2009 Technical Report, TiCC-TR 2009-005.
[56] L. Agnello, A. Comelli, S. Vitabile, Feature Dimensionality Reduction for Mammographic Report Classification, Springer, Computer Communications and Networks, 2016 Ch. 15 https://doi.org/10.1007/978-3-319-44881-7_15.
[57] R. Lyman Ott, M. Longnecker, An Introduction to Statistical Methods and Data Analysis, sixth ed., Texas A&M University, 2003 Ch. 11 & 12.
[58] M. Minghui, Z. Chuanfeng, Application of support vector machines to a small-sample prediction, Adv. Petrol. Explor. Dev. 10 (2) (2015) 72–75 https://doi.org/10.3968/7830.
[59] Ch-Ch. Chang, Ch-J. Lin, LibSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (3) (2011) 6263–6282 https://doi.org/10.1145/1961189.1961199.
[60] V. Cherkassky, Y. Ma, selection of meta-parameters for support vector regression, Lecture Notes in Computer Science vol. 2415, Springer, 2002, pp. 687–693 https://doi.org/10.1007/3-540-46084-5_112.
[61] H. Diaz, F.M. Cordova, L. Canete, F. Plaminos, F. Cifuentes, G. Rivas, Inter-channel correlation in the EEG activity during a cognitive problem solving task with an increasing difficulty questions progression, Procedia Comput. Sci. 55 (2015) 1420–1425 https://doi.org/10.1016/j.procs.2015.07.136.