Features Selection for EEG During Mental Arithmetic Task – Brain-Computer Interface

A. Nirmala Devi
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, India,
na7293@srmist.edu.in
av.nirmaladevi@gmail.com

R. Rathna
Department of Information Technology, SRM Institute of Science and Technology, Ramapuram, Chennai, India
rathnar@srmist.edu.in

ABSTRACT

Our objective is to find an efficient technique for obtaining features that can describe the continuous and underlying temporal dynamics of electroencephalography (EEG) data during the performance of mental tasks. The statistical technique known as the paired t-test is proposed to evaluate each feature. From the alpha, beta, and gamma bands of EEG, the features like mean, root mean square, skewness, mode, data range, interquartile range (IQR), and three Hjorth parameters are extracted to differentiate a signal before-during mental arithmetic task. The results suggest how essential the feature selection procedure is for the recognition task. Hjorth parameters seen between datasets during and before the arithmetic exercise have shown a significant statistical difference (p0.05). Studying an efficient way to distinguish mental task performance based on the extracted characteristics is of tremendous relevance and practical value because they accurately capture the continuity and internal dynamic changes of the EEG signals.

CCS CONCEPTS

• **Human-centered computing** → Human computer interaction (HCI).

KEYWORDS

Brain-Computer Interface, EEG (Electroencephalography), Feature Extraction, Feature selection, Attention

ACM Reference Format:
A. Nirmala Devi and R. Rathna. 2022. Features Selection for EEG During Mental Arithmetic Task – Brain-Computer Interface. In *International Conference on Information Management & Machine Intelligence (ICIMMI 2022), December 23, 24, 2022, Jaipur, India.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3590837.3590885

1 INTRODUCTION

A brain-computer interface (BCI) is created to identify and assess certain parts of brain signals. It helps users to convert these values into the necessary device commands in real-time and with feedback all at once [1]. Signal features, or simply features are the brain-signal qualities that are utilized for this purpose.The process of separating the relevant signal characteristics from unimportant stuff and displaying them in a condensed and/or understandable form that can be understood by a human or machine is known as feature extraction [2]. The EEG waves can be divided into beta, alpha, theta, delta, and gamma waves depending on the type of activity and, consequently, the frequency range it falls within [3]. Their following frequency ranges as given below

Table 1: The frequency ranges of EEG Waves

Waveform	Frequency Range	Activity
Beta	13 - 30Hz	extremely active interactions and brain activity
Alpha	8 – 13Hz	extremely calm widening the meditation
Theta	4 – 8Hz	drowsy, falling asleep, and dreaming
Delta	0.1 – 4Hz	A sound slumber without dreams
Gamma	30 – 100Hz	excessive brain activity

The goal of the current study is to analyze various features extracted from the EEG's alpha, beta, and gamma frequency bands and choose the one that can distinguish between two patterns of brain activity, one occurring before and one occurring during calculations.

2 RELATED WORK

The literature review relates various approaches for EEG signal analysis.

A mental arithmetic task detection technique requiring solving math exercises (serialized subtraction of two integers), was presented by Binish Fatimah et al., [6]. It made use of the Fourier transform to comprehend the brain reaction from a single lead EEG signal. Women between the age of 16 and 21 and men between the age of 17 and 26 participated. The SVM was used to classify the decomposed signals after entropy, and variance features were filtered from them.

Multifractal analysis was used by Qiang Wang et al., [7] to categorize real-time EEG signals for the recognition of mathematical tasks. Features including power spectrum density and an autoregressive model were used in this study. Support Vector Machine

ICIMMI 2022, December 23, 24, 2022, Jaipur, India

ACM ISBN 978-1-4503-9993-7/22/12…$15.00
https://doi.org/10.1145/3590837.3590885

(SVM) was used to classify the characteristics and define the fractal dimension.

To research the brain-computer interface, Biswarup Ganguly et al., [8] established a classification of mental arithmetic tasks based on EEG (BCI). 36 participants' EEG signals were captured, and eight features were taken out of each electrode. The stacked long-short-term memory (LSTM) architecture was fed to these features to improve and construct the brain-computer interface model.

Using single-channel EEG data, Fatema Nasrin et al., [10] suggest a way for determining the functional connectivity between the frontal lobe and pre-frontal lobe regions when young adults (aged 16 to 20) engage in mental arithmetic (subtraction). He concluded from the findings of his investigation that the precision of the Bidirectional Long Short-Term Memory (BLSTM) design was significant. This was able to identify the correct state of mental arithmetic in a 5-second time window with a mean accuracy of 75.88% across 23 channels.

Hoda Edris Abadi et al., [11] novel approach involved the extraction of many geometric features from Poincare design analysis, which employed the crucial comparison t-test to identify variations in brain activity, with a significance threshold of less than 0.05 in the states of mental calculation and face repose. In the two approaches, autonomous learning and diagnosis have also been accomplished using an artificial neural network (ANN).

Lakhan Dev Sharma et al., [12] identified the brain's reaction to stress stimulus had been proposed utilizing an effective method for the characterization of mental load using an electroencephalogram (EEG) signal and Bayesian optimized K-Nearest Neighbor (BO-KNN). Following entropy-based feature extraction, F-score-based feature selection was carried out to improve the classification accuracy

The contents of the paper are as follows: Section III: Methodology, which covers details on data collection, features that were extracted, and the techniques that were used to evaluate hypotheses. In Section IV, the results and discussion are covered, and the work is wrapped up in section V.

3 METHODOLOGY

3.1 Data Collection

The required dataset was retrieved from the EEGMAT database. It had EEG signals captured from 36 subjects [4] with a duration of 180 seconds before the mental arithmetic task and 60 seconds during the mental arithmetic task. There were 21 channels in each recording as shown in Figures 1 and 2.

In the chosen data set the EEG procedure had been done with Neurocom EEG 23-channel system (Ukraine, XAI-MEDICA), the electrodes positioned to gather the most signals possible as shown in Figure.3.

In the pre-processing stage, a 50 Hz notch filter has been utilized to remove the reference DC component. The extraction of characteristics from the alpha, beta, and gamma frequency bands is a step further in the study. Chebyshev bandpass filters with predetermined frequency ranges had been used for this. From these filtered data, we have retrieved the features listed below.

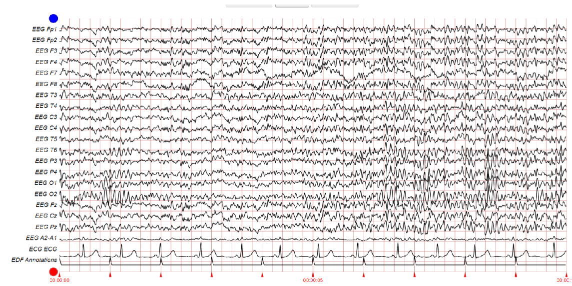

Figure 1: EEG signals for a 10-second frame of before mental arithmetic calculation [4]

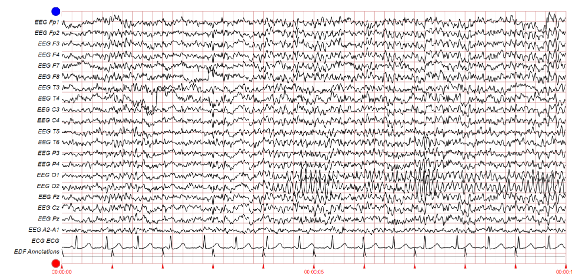

Figure 2: EEG signals for a 10-second frame during mental arithmetic calculation [4]

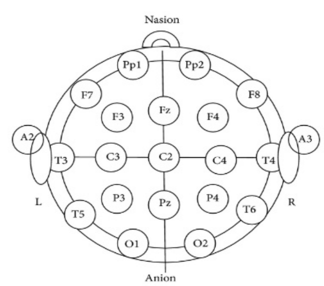

Figure 3: 10/20 International System of electrode placement. [13]

- Mean: The average of all the samples is provided by the signal's mean values.

$$x_{mean} = \frac{1}{N} \sum_{i=1}^{N} x(i) \tag{1}$$

where N is the signal's overall sample count.

- Root Mean Square (RMS): The RMS value of XRMS is calculated for a signal x(t) as follows

$$x_{RMS} = \sqrt{\frac{1}{T} \int_{t-T}^{T} x(t)^2} \tag{2}$$

- Skewness: The asymmetry of the data distribution relative to the central mean is measured. More data is to the left of the mean if the skewness number is negative, and to the right, if it is positive. It is said that the skewness parameter is

$$s = \frac{E(x - \mu)^3}{\sigma^3} \tag{3}$$

Where the sample's mean and standard deviations are μ and σ, respectively (t). 'E' stands for 'expectation'.

- Mode: The value that appears the most frequently in a dataset is called its mode, or x(t).
- Data Range: It gives back the difference between the sample dataset's maximum and minimum values x(t).

$$\text{Data range} = x_{max} - x_{min} \tag{4}$$

- Interquartile Range: The interquartile range of the data samples in a time series object is returned by IQR. IQR refers to the difference between an array of random values in upper and lower quartiles (Q3 and Q1, respectively). The quartile measures the median by considering an even dataset of 2n values or an odd dataset of 2n+1 values. The median of the first quartile's values is in Q1, while the median of the third quartile's values is in Q3.

$$IQR = Q_3 - Q_1 \tag{5}$$

- Kurtosis: Kurtosis is a term used to describe a measure of a random variable's probability distribution.

$$k = \frac{E(x - \mu)^4}{\sigma^4} \tag{6}$$

- Hjorth Parameters: A time-domain signal's statistical characteristics are shown by Hjorth parameters. Activity, Mobility, and Complexity are the three different parameters [9]. These variables are quite useful when analyzing EEG signals.
- Hjorth Activity: It shows the variance of any signal x(t)

$$\text{Activity} = \text{var}\,(x\,(t)) \tag{7}$$

- Hjorth Mobility: Mobility is equal to the square root of the ratio of the variance of the signals and the first derivatives x(t) signals. This variable is proportional to the spectrum's standard deviation.

$$\text{Mobitlity} = \sqrt{\frac{\text{var}\left(\frac{dx(t)}{dt}\right)}{\text{var}\,(x\,(t))}} \tag{8}$$

- Hjroth Complexity: The degree to which the signal shape resembles a pure sinusoid wave is determined by this parameter. If the signal is more like the sine signal, the complexity converges to unity.

$$\text{complexity} = \frac{\text{mobility}\left(\frac{d(x(t))}{dt}\right)}{\text{mobility}\,(x\,(t))} \tag{9}$$

3.2 Hypothesis Testing:

Statistical inference is used in statistical hypothesis testing to determine whether an assumption or hypothesis has been confirmed or rejected. These tests establish whether there is a difference between two sets of data. The t-test is one of many tests used in statistics for assessing hypotheses. To establish the statistical significance for paired observations, the two-sample t-test looks at the difference between the population means.

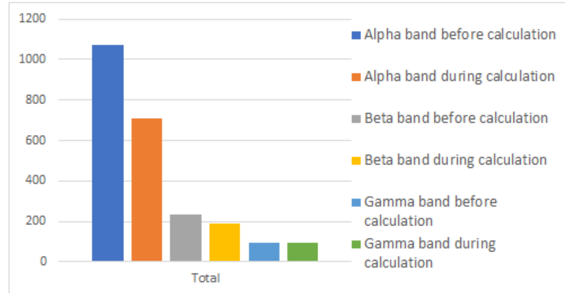

Figure 4: Comparison of Hjorth activity feature in alpha, beta, and gamma bands

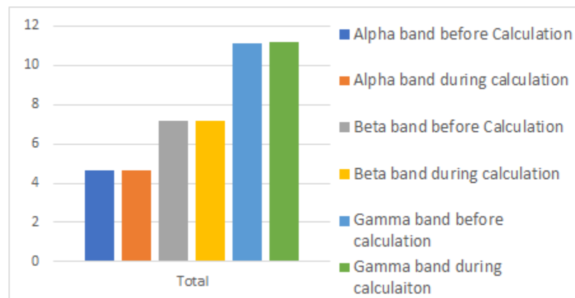

Figure 5: Comparison of Hjorth mobility feature in alpha, beta, and gamma bands

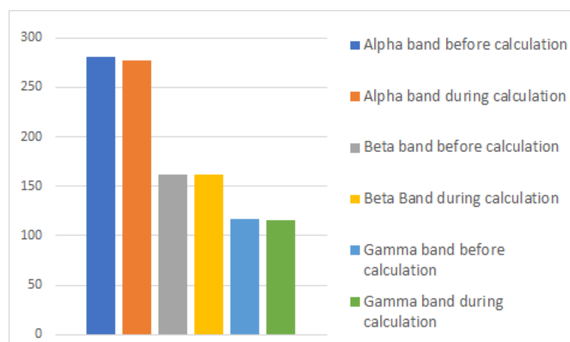

Figure 6: Comparison of Hjorth complexity feature in alpha, beta, and gamma bands

4 RESULT AND DISCUSSION

These features are computed from the EEG's alpha, beta, and gamma frequency bands and averaged across all 36 participants' brain regions. Participants' brain activity during the Mental Arithmetic calculation task is compared to their before and during brain activity in the alpha, beta, and gamma frequency bands, as illustrated below figures.

The averaging of Hjorth characteristics across all participants in three different frequency bands is depicted in the above figures. These results show that the gamma frequency band dominates the

Table 2: Results of the two-sample t-test

Feature	Alpha	Beta	Gamma
Mean	0.3131	6.91E-04	0.4471
Root Mean Square	0.0098	0.1495	0.7294
Skewness	0.3849	3.99E-04	0.4436
Mode	1.76E-05	7.91E-04	0.031
IQR	0.0256	2.21E-01	0.673
Kurtosis	0.0229	1.10E-02	0.3863
Hjorth Parameters			
Activity	0.0173	1.62E-01	0.1785
Mobility	0.0311	0.0148	0.0227
Complexity	0.0247	0.0231	0.0265

functions of the brain. To determine the extent of a significant difference between the feature values of mental arithmetic task-induced brain activity, a two-sample t-test is conducted for each extracted feature. Table 1 displays the results. The p-value obtained from a two-sample t-test is used to determine whether there is evidence of a difference between the two population means. Stronger evidence supporting different population means is defined by a lower p-value. A p-value of less than 0.05 confirms the existence of a significant difference between the two data sets when the significance level is set at 5%.

Based on Table 1's findings, it can be concluded that RMS and mode features distinguish themselves in the alpha band, but they do not do so in the beta or gamma bands. On the other end, the skewness feature is dominant exclusively in the beta band (p0.05). Although the IQR feature accepts the null hypothesis and exhibits distinct differences in the alpha band, it fails in the beta and gamma bands since its scores are more than 0.05 in both bands. Other characteristics such as mean, utterly fail to distinguish the task-induced EEG signals. In kurtosis, it shows the distinction in both alpha and beta bands but fails in the gamma band. The mobility and complexity features are consistent in separating the data sets in all the frequency bands taken into consideration when Hjorth parameters are analyzed. These two Hjorth parameters are therefore proven to be the most useful among all other attributes computed and displayed above. Additionally, when performance within frequency bands is examined, most gamma bands show a significant difference (p 0.05) from the other two frequency bands. Consequently, the t-test results confirm this theory.

5 CONCLUSION

The EEG signal is extremely complicated, non-stationary, and nonlinear. Choosing the parameters that should be used to represent such a complicated signal is challenging. Relevant features can be extracted from an EEG signal using feature extraction techniques. The characteristics of the high-frequency gamma-band have demonstrated a significant difference (p 0.05) when compared to other aspects in the most promising EEG frequency bands. This study investigated a straightforward and reliable method of feature selection. With less computational time, the procedure can help to increase computational efficiency. Future research can use effective machine learning techniques to achieve its goals.

REFERENCES

[1] J. N. Mak and J. R. Wolpaw, "Clinical applications of brain-computer interfaces: current state and prospects," Biomedical Engineering, IEEE Reviews in, vol. 2, pp. 187–199, 2009.
[2] L. Shaw and A. Routray, "Statistical features extraction for multivariate pattern analysis in meditation EEG using PCA," 2016 IEEE EMBS International Student Conference (ISC), 2016, pp. 1-4, DOI: 10.1109/EMBSISC.2016.7508624.
[3] Sanei, Saeid and Jonathon A. Chambers. "EEG Signal Processing." (2007).
[4] Zyma, I., Tukaev, S., Seleznov, I., Kiyono, K., Popov, A., Chernykh, M., & Shpenkov, O. (2019). Electroencephalograms during Mental Arithmetic Task Performance. Data, 4(1), 14. MDPI AG. Retrievedfrom http://dx.doi.org/10.3390/data4010014
[5] Fatimah B., Pramanick D., Shivashankaran P., "Automatic detection of mental arithmetic task and its difficulty level using EEG signals," in Proc. 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 1–3 July 2020, pp. 1–6, DOI: 10.1109/ICCCNT49239.2020.9225647.
[6] Fatimah B., Javali A., Ansar H., Harshitha B., Kumar H., "Mental Arithmetic Task Classification using Fourier Decomposition Method," in Proc. 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 28–30 July 2020, pp. 46–50, DOI: 10.1109/ICCSP48568.2020.9182149.
[7] Zyma, Qiang Wang, Olga Sourina, "Real-Time Mental Arithmetic Task Recognition from EEG Signals," in Proc. 2013 IEEE Transactions on Neural Systems and Rehabilitation Engineering, issue 2, vol. 21,pp. 225-232, DOI: 10.1109/TNSRE.2012.2236576
[8] Ganguly B., Chatterjee A., Mehdi W., Sharma S., Garai S, "EEG Based Mental Arithmetic Task Classification Using a Stacked Long Short-Term Memory Network for Brain-Computer Interfacing," in Proc. 2020 IEEE Vlsi Device Circuit and System (VLSI DCS), Kolkata, India, 18–19 July 2020, pp. 89–94, doi: 10.1109/VLSIDCS47293.2020.9179949.
[9] Oh, Seung-Hyeon & Lee, Yu-Ri & Kim, Hyoung-Nam. (2014). A Novel EEG Feature Extraction Method Using Hjorth Parameter. International Journal of Electronics and Electrical Engineering. 2. 106-110. 10.12720/ijeee.2.2.106-110.
[10] F. Nasrin and N. I. Ahmed, "Predicting the Correctness of Mental Arithmetic Task From EEG Using Deep Learning," 2021 International Conference on Science & Contemporary Technologies (ICSCT), 2021,pp.1-5,
[11] DOI: 10.1109/ICSCT53883.2021.9642567.
[12] Abadi HE, Moridani MK, Mirzakhani M. Mental arithmetic task detection using geometric features extraction of EEG signal based on machine learning. Bratisl Lek Listy. 2022;123(6):408-421. DOI: 10.4149/BLL_2022_064. PMID: 35576542.
[13] Sharma, L.D., Chhabra, H., Chauhan, U. *et al.* Mental arithmetic task load recognition using EEG signal and Bayesian optimized K-nearest neighbor. Int. j. inf. tecnol. 13, 2363–2369 (2021). https://doi.org/10.1007/s41870-021-00807-7
[14] Rojas, Gonzalo & Alvarez, Carolina & Montoya Moya, Carlos & de la Iglesia Vaya, Maria & Cisternas, Jaime & Gálvez, Marcelo. (2018). Study of Resting-State Functional Connectivity Networks Using EEG Electrodes Position As Seed. Frontiers in Neuroscience. 12. 10.3389/fnins.2018.00235.