EEG Signal Analysis for Multimodal Simple Concepts Decoding

Sergio Guillén, Lorenzo J. Tardón, Ana M. Barbancho, and **Isabel Barbancho**
ATIC Research Group, ITIS Software
Universidad de Málaga, Málaga, Spain
{sergiogj,lorenzo,abp,ibp}@ic.uma.es

Abstract

In this paper, we explore the use of a feature extraction model for the detection of basic decision-making concepts such as "yes" and "no" in several communication channels using electroencephalography (EEG) signals. Power topographic distribution of both concepts are explored, showing similar pattern activation in all communication channels chosen. Bi-LSTM model was used for the classification of the feature matrices extracted from EEG trials when transformed using real cepstrum, achieving, on average, 81% of accuracy across all subjects. This could help people with disabilities improve their quality life by enabling communication even when vocal communication is not possible.

1 Introduction

Electroencephalography (EEG) has prompted researcher's attention for years since allowing to record brain activity non-invasively with high temporal resolution (Luck, 2014). Language, as we commonly understand it, is the sole province of humans (Patel, 2007). Phonemes and syllables of human language are acoustically complex entities to produce (Patel, 2007), which historically led researchers to focus on the processing of individual words (lexical items) (Petersen et al., 1988). By locating and understanding the role that each brain area plays in language processing could help people with severe neurological impairments, including communication, such as pure dyslexia or aphasia (Petersen et al., 1988). Historically, their location was obtained with functional neuroimaging techniques such as functional Magnetic Resonance Imaging (fMRI), while their timing was captured by using electromagnetic techniques such as electroencephalography (EEG) (Price, 2012). The use of this non-invasive technique, could provide a means of communication with impaired people when using monosyllable words or simple concepts (Lazarou et al., 2018). By identifying brain activation patterns associated with different communication channels and classifying opposing simple concepts, such as "yes" and "no", when presented in all of them, it becomes possible to establish a direct communication pathway allowing their decoding into words by utilizing electrodes placed in specific brain regions, offering individuals the ability to express themselves without the need for physical speech.

Language processing has been widely explored by researchers, and it has shown to involve several brain regions, with the most popular neural model of language being based on the writings of Broca, Wernicke and Lichtheim at the end of the 19th Century and Geschwind in the mid 20th Century (Price, 2012). The Broadman areas named after them are classically related to language production and processing (Hall and Hall, 2021). However, in recent years, it has been proven the role of more neural structures in language processing (Nizara, 2018). In both (Rezazadeh Sereshkeh et al., 2017; Choi and Kim, 2019), the decoding of "yes" and "no" concepts from EEG signals was achieved through a feature extraction stage, followed by classification. In the former, a multilayer perceptron (MLP) was employed, achieving 63.17% accuracy scores on average, while the latter used a support vector machine (SVM) with a 86.03% attained when combining multiple time-frequency subwindows.

Historically, cepstrum has its roots in the general problem of signals deconvolution (Childers et al., 1977), but it has proved it usefulness, not only in speech signal processing, but also in EEG signal processing (Sen et al., 2023; Han et al., 2024). Inspired from the success of cepstral features we propose a feature extraction model using EEG signal analysis to discern between two monosyllable words with opposite meanings, "yes" and "no", that could enhance communication possibilities for persons with some kind of motor/neurological disability supported by the use of a Bidirectional Long-Short Term Memory (Bi-LSTM) Neural Network model.

This manuscript is organized as follows: In Sec. 2, the methodology followed to record the EEG signal analyzed is presented. Sec. 3 describes the main pre-processing stages considered and Sec. 4 shows the features extracted that build the features matrix used in the classification stage. Results and conclusions are presented in Secs. 5 and 6, respectively.

2 Methods

In this section, the methodology followed to record EEG signals is described, including the participant's details, experiment description, and the specific equipment used.

2.1 Participants

We recorded data from a heterogeneous group of 15 healthy participants (3 male participants, mean age = 24.00; 12 female participants, mean age = 28.16), most of them members of Universidad de Málaga (UMA), whose participation was entirely voluntary with no monetary or any other kind of compensation. All of them received detailed oral information about the experiment before providing their written informed consent to take part in the experiment.

This study was conducted in accordance with the Declaration of Helsinki and was approved by Comité Ético de Experimentación de la Universidad de Málaga, Reg. CEUMA: 61-2021-H. The privacy and confidentiality of the participants were strictly protected throughout the study.

2.2 Experiment Description

At present, it is not clear that comprehension of a word necessarily entails activation of a detailed perceptual representation of the object to which it refers, at least not to the same degree as that evoked by the object itself (Binder et al., 2009). Inspired by that, in order to detect each word (lexical item/concept) processing, subjects were presented with blocks of "yes" and "no" words in the same language (Spanish), presented in different ways, we will refer to them from now on as communication channels. These are aimed to cover all possible ways of communication, including these scenarios:

- **Read:** Words were displayed as text on the screen, so participants were instructed to read them.
- **See:** Words were displayed with a representative symbol on the screen, so participants were instructed to look at them.
- **Listen:** Words were played through the speakers, so participants were instructed to listen to them.
- **Say:** Words were displayed as text on the screen, so participants were now instructed to read them out loud.
- **Think:** Words were shown as text on the screen, so participants were now instructed to think of the word displayed.

In the two last scenarios, each word was first presented and then, the action was performed, expecting to avoid the mixture of different cognitive stimuli. Note that although the appearance of these blocks was not randomized, the presentation of the words within them, lasting 7 seconds on average, was.

Recall that all participants recruited were healthy subjects. Because of this, the experiment is designed so that no specific feedback is necessary from subjects during the experiment's recording. Also, some resting time is allocated between sub-experiments.

Experiments were conducted in a separate room with soundproof windows, and curtains to avoid external noise. Non-essential electronic devices were turned off to reduce electromagnetic interference. Participants were instructed to avoid unnecessary movements.

2.3 Equipment

BrainVision's actiChamp-Plus and acti-CAP were used (Brain Products, 2016) in this work. 64 active electrodes were arranged according to the $10-20$ system (American, 1994; Klem et al., 1999), which provide high-quality recordings with low background noise. Among them, FCz and FPz are used as reference and ground channels, respectively, while FT9 and FT10 electrodes are displaced to record vertical (VEOG), and horizontal (HEOG) ocular activity. Iz electrode is used for low-quality audio capture ($f_s = 2500Hz$), leading to a final count of 61 electrodes used to measure EEG signals. The maximum impedance measured across all participants was kept under $10k\Omega$, and balanced during every recording session (Sanei and Chambers, 2013). Figure 1 shows the electrode positioning configuration employed.

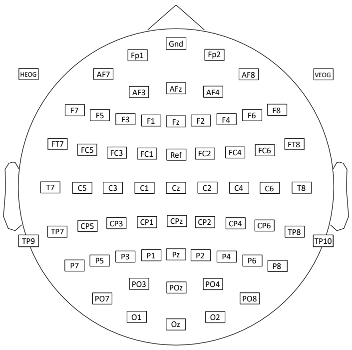

Figure 1: Electrode locations used.

BrainVision Recorder (Brain Products, 2021) was used to capture EEG data supported by E-Prime (W. Schneider and Zuccolotto, 2016) running on an external computer for stimuli presentation, and sending timestamps to the EEG recorder. Speakers connected to the E-Prime computer allowed for audio stimuli presentation. Their volume was kept constant to avoid differences between participants.

3 EEG data Pre-processing

This section details the pre-processing steps taken using MATLAB R2022a (MathWorks, 2022), along with Fieldtrip Toolbox (Oostenveld et al., 2010) for EEG data handling.

EEG signal's small amplitude requires them to be amplified in order to be properly analyzed (Luck, 2014). This causes several noise sources, both correlated and uncorrelated, to be also amplified, potentially masking the neural activity of the brain (Cohen, 2014). To address this issue, both a high-pass filter and low-pass filter, with cut-off frequencies of $0.1Hz$ and $45Hz$, respectively, were applied to eliminate slow drifts and non-cognitive signals, also proving a baseline correction for each subject. Also, considering muscle artifacts falling within the $30 - 100Hz$ frequency range (Luck, 2014), the chosen cut-off frequencies should effectively mitigate them (Hassan and Hussain, 2023).

After filtering, EEG signals are down-sampled to $f_s' = 100Hz$ to reduce the computational cost without compromising the results.

Independent Component Analysis (ICA) is then applied for visual artifact rejection (Sanei and Chambers, 2013). This technique decomposes electrode signals, $\boldsymbol{x}(k) = [x_1(k), x_2(k), \ldots, x_e(k)]^T$, into statistically independent components (IC), through and unmixing matrix, $\boldsymbol{W}$, according to:

$$\boldsymbol{i}(k) = \boldsymbol{W} \cdot \boldsymbol{x}(k) \tag{1}$$

where $\boldsymbol{i}(k)$ are statistically independent signal components. Electrooculogram (EOG) and audio channels were removed for IC extraction. Note that k refers to the samples of the signal within the excerpt considered.

Based on the approach presented in (Villena et al., 2019), a threshold process based on the correlation coefficients was applied to detect potential artifactual components and discard them before mixing the remaining ICs back. Both the ICs obtained and the EEG signal are pre-epoched in the segments of interest (trials), and the Pearson Correlation Coefficients (ρ) were obtained by following:

$$\rho(e,n) = \frac{\sum_{k=1}^{K}(x_e(k)-\overline{x_e})\left(i_n(k)-\overline{i_n}\right)}{\sqrt{\sum_{k=1}^{K}(x_e(k)-\overline{x_e})^2 \sum_{k=1}^{K}\left(i_n(k)-\overline{i_n}\right)^2}} \tag{2}$$

where e refers to each EOG channel considered, and n to the n-th component extracted.

Note that this step is performed for each selected trial, and components are considered as artifactual when they surpass the defined threshold, in its absolute value, in at least the 80% of the trials considered. To avoid erasing cognitive information, this threshold process was supported by visual supervision of the detected components, ensuring that only artifactual information was removed from the original signal (non epoched) $x_e(k)$, to obtain the reconstructed signal, $x_e'(k)$.

Since in this experiment we are interested in assessing neural processing of different words or concepts, EEG signals are now epoched to contain the two words considered, i.e., "yes" and "no" in all the communication channels assessed. With this in mind, the power topographic distribution of all subjects is compared in all the scenarios assessed. Figure 2 shows the scalp power distribution averaged for all participants available when presented with "yes" and "no" words in each communication channel. Note that for each channel of communication, the topographic distribution of both words are displayed using the same normalized color axis.

In this figure, it can be observed how some activation areas are present in all communication channels chosen when processing these two concepts. Primary, the main activity is focused on the frontal

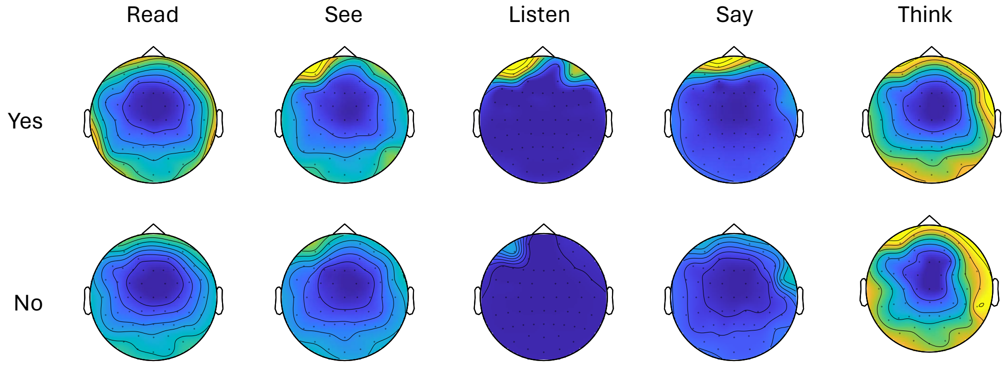

Figure 2: Average scalp power distribution when processing "yes" and "no" words/concepts when different communication channels are used.

region rather than parietal-occipital and temporal, depending on the task performed.

As for the former, mainly a left-laterallyzed activity over the frontal cortex is consistent with what is stated in (Petersen et al., 1988), where task calling for semantic processing of individuals words activation was observed over the frontal region. Later neuro-psychological experiments showed that damages in frontal or inferior parietal areas in this hemisphere caused deficits in tasks that required speech syllables identification, which highlighted the possible role of a fronto-parietal circuit in the perception of speech (Hickok and Poeppel, 2007). However, note that is not always restricted to the left hemisphere. This is not odd since previous researches indicates that there is probably, at least, one pathway in each hemisphere related with speech processing (Hickok and Poeppel, 2007). This behavior is also consistent with the spatial distribution observed in (Choi and Kim, 2019), where right frontal region exhibited the most useful features for the discrimination task.

There is strong evidence that posterior middle temporal regions are involved in accessing lexical and semantic information. Although it's not present with the same magnitude in all cases assessed, activation in left temporal region directly relates with Broca's and Wernicke's areas, both intimately related with language processing (Petersen et al., 1988; Hickok and Poeppel, 2007). Semantic processing has been particular observed over inferotemporal and posterior inferior parietal regions, as stated in (Binder et al., 2009). Activation in these regions are somehow expected since posterior areas are involved in visual feature extraction and more anterior areas are involved in lexico-semantic processing of the whole word (Price, 2012).

Language processing is widespread and occupy a large proportion of the cortex in the human brain (Binder et al., 2009), with its neural organization, being task dependent, as stated in (Hickok and Poeppel, 2007).

At the view of these power topographic distributions, we hypothesized that the processing of both concepts is equivalent regardless of the type of medium chosen for their presentation. Based on this, they will be treated equally without discerning between them in further stages.

For concept processing detection, the real cepstrum is applied to the original EEG signal. Its used is based on previous results as presented in sec. 1. The real cepstrum of a signal is obtained as the Inverse Fourier Transform of the logarithm of the magnitude of the spectrum (Shourie, 2016). In our research, the real cepstrum analysis process has been applied to the EEG signals to use their coefficients as a parameter vector to characterize signals and analyze them improving the results obtained when compared with the use of the EEG signal directly. The real cepstrum can be computed as:

$$c_{e,t}(k) = \frac{1}{2\pi} \int log|X_{e,t}(w)|e^{jwk}dw \quad (3)$$

where $c_{e,t}(k)$ are the real-valued coefficients of the cepstrum for the e-th electrode, and the t-trial, and $X_{e,t}(w)$ the Fourier transform of the input signal, $x'_{e,t}(k)$.

4 Feature Extraction

This step allows for reducing the complexity of the classification step (Alghamdi et al., 2023; Danyal et al., 2023). To this end, each feature was extracted following a single-trial method, i.e., within each subject, trial, and electrode (Grierson and Kiefer, 2014).

Based on previous EEG research (Hassan and Hussain, 2023; Al-Qazzaz et al., 2023), once individually tested, both statistical and Power Spectral Density (PSD) features-based are chosen to maximize the results obtained when combined:

- **Standard Deviation:** This feature measures the dispersion of the signal around its mean value:

$$\sigma_{e,t} = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K} |c_{e,t}(k) - \bar{c}_{e,t}(k)|^2}, \tag{4}$$

  with $\bar{c}_{e,t}(k)$ the sample mean:

$$\bar{c}_{e,t}(k) = \frac{1}{K}\sum_{k=1}^{K} c_{e,t}(k) \tag{5}$$

  and K the sample length of the excerpt considered.

- **Root Mean Square:** Square root of the averaged squared values of the signal excerpt:

$$RMS_{e,t} = \sqrt{\frac{1}{K}\sum_{k=1}^{K} |c_{e,t}(k)|^2} \tag{6}$$

- **Absolute Power Value:** Overall power was extracted from the PSD obtained, following previous research (Sammler et al., 2007)(Stancin et al., 2021), through Welch method, first proposed by (Welch, 1967).

  A Hamming window ($W(l), l = 0, ..., L-1$), with a 50% overlap was used for segmenting input signal into $S = 8$ segments of length L. The absolute power spectral value of the e-th electrode at the t-th trial is obtained by summing up the spectral estimation over all the frequency bins, $p(f_n)$, of the PSD obtained for the EEG excerpt considered, as follows:

$$P_{e,t} = \sum_{n=0}^{L/2} p(f_n) \tag{7}$$

- **Averaged Spectral Flux:** Rate of change of the PSD of the input signal averaged over time. It is calculated by using:

$$F_{e,t} = \frac{1}{S}\sum_{s=1}^{S} \sqrt{\sum_{n=0}^{L/2} |p(f_{n+1}) - p(f_n)|^2} \tag{8}$$

Following this approach, each trial outputs a feature matrix of dimensions (ExF), where $E = 61$ are the active electrodes used, and $F = 4$ are the features considered.

The aforementioned features have already shown compelling results for EEG signal characterization in musical mode detection (Guillén et al.). Note that all of them are energy-based, thus the process outlined can be viewed as a form of data augmentation. In contrast to typical EEG approaches where augmentation involves altering the dataset through noise addition or geometric transformations (Lashgari et al., 2020; George et al., 2022), the approach presented aims to optimize the information encompassed in the original dataset without altering it.

5 Results

This work proposes a feature-based model for the characterization and classification of EEG trials when processing "yes" and "no" words when presented in different communication channels such as text, symbol, sound, speech, or thought, with the final aim of helping the communication possibilities of people with some type of motor/neural disability that may diverge into communication difficulties. For the classification, a type of Long Short-Term Memory (LSTM) is used, reviewing both intra-, for each participant, and inter-subject, for all participants combined, scenarios. In (Rezazadeh Sereshkeh et al., 2017), a LSTM model was used for the decoding of "yes" and "no" as stated in section 1.

LSTM algorithms have shown their effectiveness in automatically predicting timeline properties (Algarni et al., 2022). Bi-LSTM classifier consists of an input sequence layer of $E = 61$ inputs, each input a vector built upon the features previously stated. Then, a bidirectional LSTM layer, built up of a forward layer and a backward layer, with 20 hidden units is used to learn the bidirectional long-term dependencies between sequence data flows. A fully-connected layer with two possible states ("yes" and "no" classes) is placed prior to outputting the label chosen for the data classified using a non-linear softmax layer supported by the cross entropy loss. A count of 1500 epochs is chosen to reach model convergence. Figure 3 shows the architecture of the model chosen.

This model is similar to the one used in (Ariza et al., 2022), though specific changes were done to adopt its structure to the task at hand, as described.

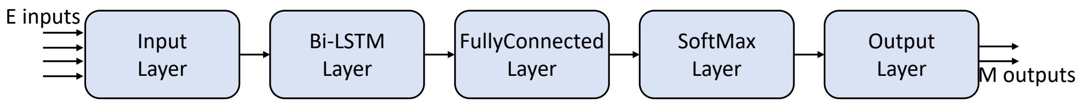

Figure 3: Diagram of the configuration of the Bi-LSTM network model used.

To avoid biases, and over-fitting in the classification step, and considering the dataset sample size, a 3-fold cross-validation process is carried out, so the results presented in further sections are the average of the 3 folds considered.

Using the data obtained after the experiment, matrices are labeled after the subject's response during each trial, depending on the concept processed: "yes" if the ongoing concept was "yes", or "no", otherwise. With this, binary classification was carried out. Table 1 shows Accuracy, Recall, Precision, and F1-score metrics obtained for each subject confusion matrix, along with their average value.

Table 1: Results (%) of binary classifications: "yes" or "no" word/concept processed.

Subject ID	Accuracy	Precision	Recall	F1-score
S_1	73.13	66.25	76.81	71.14
S_2	71.88	56.25	81.82	66.67
S_3	85.00	70.00	100.00	82.35
S_4	80.63	100.00	72.07	83.77
S_5	75.00	65.00	81.25	72.22
S_6	77.50	91.25	71.57	80.22
S_7	83.13	85.00	81.93	83.44
S_8	85.63	86.25	85.19	85.72
S_9	75.63	96.25	68.14	79.79
S_{10}	85.63	95.00	80.00	86.86
S_{11}	88.75	91.25	86.90	89.02
S_{12}	74.38	82.50	70.97	76.30
S_{13}	87.50	90.00	85.71	87.80
S_{14}	93.75	93.75	93.75	93.75
S_{15}	81.25	90.00	76.60	82.76
Averaged	81.25	83.92	80.85	81.45

In this table it can be observed that, although each subject outputs different results, the model's performance is consistent between them, and manages to surpass 81% of accuracy on average, proving this method to successfully discern when participants are processing one word or another regardless of the communication channel employed according to (Perelmouter and Birbaumer, 2000; Müller-Putz et al., 2008). These averaged values can be also observed in Figure 4, where the confusion matrix obtained by summing up all confusion matrices of subjects is presented. Note that the results attained are in line with (Choi and Kim, 2019) where a SVM model was used obtaining 86.03% of accuracy score, and surpass results from (Rezazadeh Sereshkeh et al., 2017) where a LSTM model was used, reaching up to 63.17%

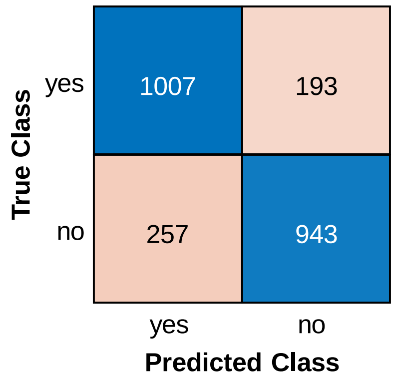

Figure 4: Average confusion matrix obtained by summing up all subject's confusion matrices.

AUC-ROC curves are drawn to support these metrics, which have shown to be directly correlated with the accuracy, but also considering the miss-classification cost and giving an indication of the amount of "work done" by the classification scheme evaluated (Bradley, 1997). Figure 5 shows the AUC-ROC curve obtained per user (colored and dotted), and the average curve obtained for all of them (black and continuous) in this scenario.

In this figure it can be observed how the model manages to convergence for all subjects, with an average AUC-ROC value of 92.29%, proving the model chosen to successfully discern between the words studied.

Inter-subject scenario was also assessed, but due to the experiment configuration, the model chosen did not manage to attain compelling results under the same training options chosen. This was somehow expected since in (Price, 2012) was stated that intra-operative stimulation showed diversity in location of language functions and morpho-metrical imaging studies based on diversity of brain shape and gyral patterns.

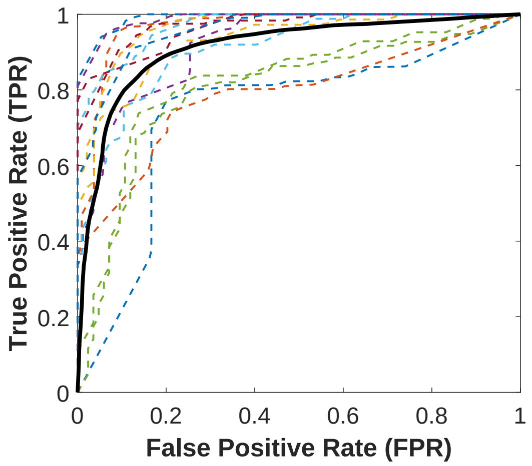

Figure 5: AUC-ROC curves comparison between all subjects curves (dotted-colored curves) and average curve (continuous-black curve) for binary classification using EEG data of all subjects when processing “yes” and “no” words/concepts.

6 Conclusions

In this paper, an analysis of brain responses to words/concepts with opposite meanings, i.e., “yes” and “no”, has been carried out using EEG signals. The final purpose of this work was to discern between them aiming to help the communication of people with some disability. For this purpose, the subjects participating in the experiment were presented the stimuli through several communication channels, i.e., text, symbol, sound, speech, and thought, trying to cover all the possible processing methods of the concepts. Note that all of them were considered since the control group was composed of healthy subjects. Nevertheless, it should be expected that when some kind of disability is present, at least one of them is still possible.

The scalp power distribution of all scenarios mentioned was reviewed, showing in all cases similar activation patterns, consistent with previous studies in language processing, which may indicate that both words/concepts are processed similarly regardless of the presentation medium. Based on this finding, all trials from different communication channels are used jointly in the processing and classification of signals.

Real cepstrum is used to characterize EEG signals once pre-processed. Then, energy-based features are extracted on a single-trial basis for each electrode individually.

Intra- and inter-subject scenarios are explored. The Bi-LSTM Neural Network model chosen successfully discerns between “yes” and “no” words/concepts regardless of the communication channel chosen in the former, attaining an average 81.25% accuracy value in the intra-subject binary classification scheme supported by an average AUC-ROC value of 92.29%, showing an improvement in the discrimination task when compared with previous researches. Inter-subject scenario was also assessed, but no compelling results were obtained maybe due to the variability of the multi-modal communication scheme considered and physiological differences across subjects.

Based on the results obtained, the processing scheme described in this work stands as a valuable tool to explore the possibility of enhancing the communication capabilities of people with some motor/neural disability by detecting simple words/concepts of opposite meaning. These findings encourage enlarging the dataset and continue the research.

Limitations

The main limitation of our study lies in the sample size at our disposal. This is a common limitation in EEG experiments where participants recruitment, specially without monetary compensation, is limited. However, although working with a reduced dataset might led to possible misinterpretations, sample size does not necessarily affect the validity of the research outcome, allowing the results obtained to be considered valid (Vozzi et al., 2021). Nevertheless, a larger sample could provide a more comprehensive and representative perspective. To overcome this drawback, it is expected to expand our dataset in further research stages.

Ethics Statement

The authors of this article declare that they have no known conflicts of interest and received no funding or financial support from any organization that could potentially bias the work reported in this paper.

Acknowledgements

This publication is part of project PID2021-123207NB-I00, funded by MCIN/AEI/10.13039/501100011033/FEDER, UE. This work was done at Universidad de Málaga, Campus de Excelencia Internacional Andalucía Tech.

References

Noor K. Al-Qazzaz, Reda J. Lafta, and Maimonah A. Khudhair. 2023. Estimations of emotional synchronization indices for brain regions using electroencephalogram signal analysis. *Advances in Non-Invasive Biomedical Signal Sensing and Processing with Machine Learning*, pages 315–344.

Mona Algarni, Faisal Saeed, Tawfik Al-Hadhrami, Fahad Ghabban, and Mohammed Al-Sarem. 2022. Deep learning-based approach for emotion recognition using electroencephalography (EEG) signals using bi-directional long short-term memory (bi-LSTM). *Sensors*, 22(8).

Mawadda Alghamdi, Saeed M. Qaisar, Shahad Bawazeer, Faya Saifuddin, and Majed Saeed. 2023. Brain-computer interface (BCI) based on the EEG signal decomposition butterfly optimization and machine learning. *Journal of medical signals and sensors*.

Electroencephalographic Society American. 1994. Guideline thirteen: Guidelines for standard electrode position nomenclature. *Journal of Clinic Neurophysiology*, 11(1):111–113.

Isaac Ariza, Lorenzo J. Tardón, Ana M. Barbancho, Irene De-Torres, and Isabel Barbancho. 2022. Bi-lstm neural network for EEG-based error detection in musicians' performance. *Biomedical Signal Processing and Control*, 78.

Jeffrey R. Binder, Rutvik H. Desai, William W. Graves, and Lisa L. Conant. 2009. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*, 19(12):2767–2796.

Andrew P. Bradley. 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145–1159.

GmbH Brain Products. 2016. ActiCAP (64 channels) [apparatus]. Gilching, Germany. https://www.brainproducts.com/solutions/acticap/.

GmbH Brain Products. 2021. Brainvision recorder (version 2.2.2) [software]. Gilching, Germany. https://www.brainproducts.com/solutions/recorder/.

D.G. Childers, D.P. Skinner, and R.C. Kemerait. 1977. The cepstrum: A guide to processing. *Proceedings of the IEEE*, 65(10):1428–1443.

Jeong Woo Choi and Kyung Hwan Kim. 2019. Covert intention to answer "yes" or "no" can be decoded from single-trial electroencephalograms (EEGs). *Computational Intelligence and Neuroscience*, (1):4259369.

Mike X Cohen. 2014. *Analyzing Neural Time Series Data*, 1st edition. The MIT Press.

Mahmood Danyal, Riaz H. Naseem, and Nisar Humaira. 2023. Introduction to non-invasive biomedical signals for healthcare. *Advances in Non-Invasive Biomedical Signal Sensing and Processing with Machine Learning*, pages 1–24.

Olawunmi George, Roger Smith, Praveen Madiraju, Nasim Yahyasoltani, and Sheikh I. Ahamed. 2022. Data augmentation strategies for EEG-based motor imagery decoding. *Heliyon*, 8(8).

Mick Grierson and Chris Kiefer. 2014. Contemporary approaches to music BCI using P300 Event Related Potentials. *Guide to Brain-Computer Music Interfacing*, pages 43–59.

Sergio Guillén, Lorenzo J. Tardón, Ana M. Barbancho, and Isabel Barbancho. Neural processing of musical mode through EEG signals. *Engineering Applications of Artificial Intelligence*. (under review).

John E. Hall and Michael E. Hall. 2021. *Textbook of Medical Physiology*, 14th edition. Elsevier.

Siqi Han, Chao Zhang, Jiaxin Lei, Qingquan Han, Yuhui Du, Anhe Wang, Shuo Bai, and Milin Zhang. 2024. Cepstral analysis-based artifact detection, recognition, and removal for prefrontal EEG. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 71(2):942–946.

Fatima Hassan and Syed F. Hussain. 2023. Review of eeg signals classification using machine learning and deep-learning techniques. *Advances in Non-Invasive Biomedical Signal Sensing and Processing with Machine Learning*, pages 159–183.

Gregory Hickok and David Poeppel. 2007. The cortical organization of speech processing. *Nature reviews. Neuroscience*, 8:393–402.

George H. Klem, Hans Lüders, Herbert H. Jasper, and Christian Erich Elger. 1999. The ten-twenty electrode system of the international federation. the international federation of clinical neurophysiology. *Electroencephalography and clinical neurophysiology. Supplement*, 52:3–6.

Elnaz Lashgari, Dehua Liang, and Uri Maozv. 2020. Data augmentation for deep-learning-based electroencephalography. *Journal of Neuroscience Methods*, 346.

Iolietta Lazarou, Spiros Nikolopoulos, Panagiotis C. Petrantonakis, Ioannis Kompatsiaris, and Magda Tsolaki. 2018. EEG-based brain–computer interfaces for communication and rehabilitation of people with motor impairment: A novel approach of the 21st century. *Frontiers in Human Neuroscience*, 12(14).

Steven J. Luck. 2014. *An Introduction to the Event-Related Potential Technique*, 2nd edition. The MIT Press. ISBN: 978-0-262-52585-5.

Inc. MathWorks. 2022. MATLAB version: 9.12.0.2039608 (r2022a update 5) [software]. Natick, Massachusetts, United States. https://www.mathworks.com/.

Gernot Müller-Putz, Reinhold Scherer, Clemens Brunner, Robert Leeb, and Gert Pfurtscheller. 2008. Better than random? a closer look on bci results. *International Journal of Bioelektromagnetism*, 10:52–55.

El Imrani Nizara. 2018. Una revisión de la neuroanatomía y neurofisiología del lenguaje. *Revista Neuro-Psiquiatría*, 8(3):196–202.

Robert Oostenveld, Pascal Fries, Eric Maris, and Jan Mathijs Schoffelen. 2010. Fieldtrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011:9.

Aniruddh D. Patel. 2007. *Music, Language and the Brain*, 1st edition. Oxford University Press USA.

J Perelmouter and N Birbaumer. 2000. A binary spelling interface with random errors. *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, 8(2):227—232.

Steve Petersen, P Fox, Michael Posner, Mintun MA, and Marus Raichle. 1988. Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, 331:585–9.

Cathy J. Price. 2012. A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, 62(2):816–847.

Alborz Rezazadeh Sereshkeh, Robert Trott, Aurélien Bricout, and Tom Chau. 2017. EEG classification of covert speech using regularized neural networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(12):2292–2300.

Daniela Sammler, Maren Grigutsch, Thomas Fritz, and Stefan Koelsch. 2007. Music and emotion: Electrophysiological correlatees of the processing of pleasant and unpleasant music. *Psychophysiology*, pages 293–304.

Saeid Sanei and Jonathon A Chambers. 2013. *EEG signal processing*. John Wiley & Sons.

Ovishake Sen, Anna M. , Pranay R. Raman, Kabir S. Khara, Adam Khalifa, and Baibhab Chatterjee. 2023. Machine-learning methods for speech and handwriting detection using neural signals: A review. *Sensors*, 23(12):5575.

Nasrin Shourie. 2016. Cepstral analysis of EEG during visual perception and mental imagery reveals the influence of artistic expertise. *Journal of medical signals and sensors*, 6(4).

I. Stancin, M. Cifrek, and A. Jovic. 2021. A review of EEG signal features and their application in driver drowsiness detection systems. *Sensors*, 21.

Alejandro Villena, Lorenzo J. Tardón, Isabel Barbancho, Ana M. Barbancho, Elvira Brattico, and Niels T. Haumann. 2019. Preprocessing for lessening the influence of eye artifacts in EEG analysis. *Applied Sciences*, 9(9).

Alessia Vozzi, Vincenzo Ronca, Pietro Aricò, Gianluca Borghini, Nicolina Sciaraffa, Patrizia Cherubino, Arianna Trettel, Fabio Babiloni, and Gianluca Di Flumeri. 2021. The sample size matters: To what extent the participant reduction affects the outcomes of a neuroscientific research. a case-study in neuromarketing field. *Sensors*, 21(18).

A. Eschman W. Schneider and A. Zuccolotto. 2016. E-prime 3.0 user's guide [software]. Pittsburgh: Psychology Software Tools, Inc.. https://pstnet.com/products/e-prime/.

P. Welch. 1967. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73.