Enhancing the EEG signals visualization

André Furlan
UNESP - Universidade Estadual Paulista Júlio de Mesquita Filho
São José do Rio Preto, André Furlan, Brazil
ensismoebius@gmail.com

ABSTRACT

Electroencephalogram (EEG) data has inherent complexity, including a large volume, diverse frequencies (e.g., alpha, beta, theta, sigma), and numerous emitting sources (neurons). Additionally, potential electromagnetic interferences from external sources pose challenges during EEG analysis. This work aims to propose effective visualization techniques for EEG data, drawing insights from recent research on EEG visualization. The goal is to create informative visualizations that aid in EEG research by showing detailed representations of the data without the use of misleading gradients. The visualization will include interactive features for subject and stimulus selection, providing a comprehensive understanding of EEG signals with high-contrast color schemes to enhance perceptual clarity.

ACM Reference Format:
André Furlan. 2023. Enhancing the EEG signals visualization. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Electroencephalogram (EEG) data is widely used in neuroscience research to study brain activity. EEG records electrical signals generated by brain neurons during a task making it a valuable tool in various applications, including Brain-Computer Interfaces (BCI) and clinical diagnosis of neurological states [5]. Analyzing EEG data, however, poses challenges due to its complexity, such as a vast number of emitting sources (neurons) and potential electromagnetic interferences from external sources like cell phones, electric motors, and magnetic fields from wires [5]. To facilitate EEG data interpretation and visualization, it is crucial to develop effective and efficient techniques that allow researchers to gain valuable insights from the data. This paper proposes a visualization method for EEG data, combining line graphs and energy visualizations with a display of the electrode positions in the 10-20 system [12]. The proposed visualization will be interactive, allowing users to select subjects and stimuli and explore signal details. The goal is to provide researchers with an informative and comprehensive view of EEG data to support their investigations.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2 DATA ACQUISITION

To demonstrate the proposed visualization technique, EEG data was acquired from an existing dataset. The chosen dataset is the "Open access database of EEG signals recorded during imagined speech" [2]. This dataset contains EEG recordings during the imagined pronunciation of vowels and commands, as well as EEG and audio recordings during the actual pronunciation. The dataset includes recordings from 15 Argentinian volunteers between the ages of 24 and 28, with an equal representation of males and females [2].

As shown in Figure 1, the data was acquired using the 10-20 system [12] which consist of electrodes spaced by 20% of the head size and a margin of 10% in relation to *nasion*, *inion* and to the preauricular point. In this specific database only the sensors *F*3, *F*4, *C*3, *C*4, *P*3 and *P*4 was used.

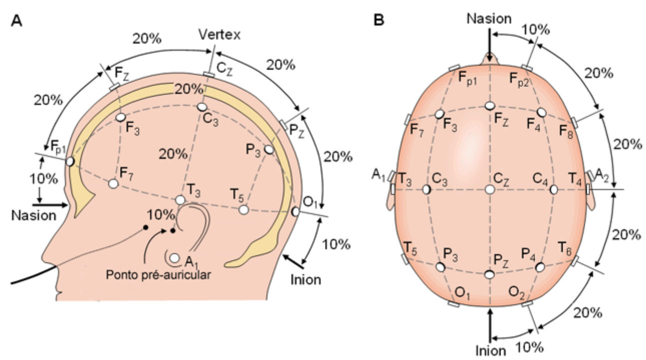

Figure 1: system 10-20: The electrodes are evenly separated across the scalp of the subject. The A1 sensor is the preauricular point. Source: [12]

Although this dataset serves as the initial data source for the proposed visualization, our future research intends to develop a custom EEG database in collaboration with local health institutions.

3 DATABASE STRUCTURE

Each row in the dataset corresponds to a subject-stimuli-modality set, and the data is arranged as follows: EEG signals from different electrodes, modalities (imagined or pronounced), and stimulus codes representing different categories of stimuli (e.g., letters and commands). The modalities distinguish between imagined and pronounced speech, while the stimulus codes identify specific categories of stimuli. Figure 2 shows a sample of the initial rows in the database, and Figure 3 displays the last rows.

Each line consist in a series of numbers organized as follows:

- F3 - 0:4095

- F4 - 4096:8191
- C3 - 8192:12287
- C4 - 12288:16383
- P3 - 16384:20479
- P4 - 20480:24575
- modality - 24576:24576
- estimuli - 24577:24577
- artifact - 24578:24578

Modality can be 1 for imagined speech or 2 for pronounced speech. Stimuli assumes 1: A, 2: E, 3: I, 4:O and 5:U and 6:*Arriba*, 7: *Abajo*, 8: *Adelante*, 9: *Atrás*, 10: *Derecha*, 11: *Izquierda*.

	0	1	2	3	4	5	6	7	8	9	...
0	-45.027295	-44.902287	-44.495685	-43.828843	-42.930203	-41.833993	-40.578754	-39.205796	-37.757699	-36.276644	...
1	-10.482894	-9.311390	-8.069559	-6.784121	-5.482265	-4.190737	-2.934883	-1.737725	-0.619207	0.404488	...
2	29.254169	28.537506	27.732951	26.864038	25.959221	25.050858	24.173944	23.364693	22.658914	22.090382	...
3	20.516124	19.929198	19.218417	18.392212	17.465399	16.458635	15.397435	14.310845	13.229902	12.185891	...
4	-5.097755	-5.770929	-6.527508	-7.369527	-8.296717	-9.306164	-10.392103	-11.545732	-12.755099	-14.005145	...
...	...	...	...	...	...	...	...	...	...	...	...
683	20.776462	21.249038	21.844552	22.558554	23.379841	24.290637	25.267153	26.280516	27.298048	28.284797	...
684	7.054769	7.039320	7.049230	7.079761	7.126160	7.184204	7.250641	7.323584	7.402738	7.489561	...
685	2.071435	2.343421	2.726060	3.209852	3.780173	4.418085	5.101340	5.805475	6.505064	7.174842	...
686	-11.204902	-11.467989	-11.691680	-11.880146	-12.036192	-12.161035	-12.254149	-12.313226	-12.334364	-12.312443	...
687	-1.571957	-1.814884	-2.240675	-2.819675	-3.515513	-4.287240	-5.091722	-5.886054	-6.629877	-7.287520	...

688 rows × 24579 columns

Figure 2: Sample of initial rows in the EEG database

...	24569	24570	24571	24572	24573	24574	24575	24576	24577	24578
...	26.359719	25.260140	24.151324	23.092276	22.140159	21.347616	20.760296	1.0	9.0	2.0
...	-6.613661	-6.842419	-7.089494	-7.353440	-7.631651	-7.920954	-8.218081	1.0	11.0	2.0
...	-7.599078	-8.424428	-9.183207	-9.845966	-10.388099	-10.790734	-11.041316	1.0	9.0	2.0
...	14.119956	13.507964	12.736278	11.814361	10.760473	9.600974	8.369264	1.0	8.0	2.0
...	-12.909528	-12.038187	-11.221590	-10.498695	-9.905959	-9.475264	-9.231932	1.0	11.0	2.0
...	...	...	...	...	...	...	...	...	...	...
...	-0.111416	0.975065	1.833694	2.449819	2.817991	2.941838	2.833562	1.0	4.0	1.0
...	40.836727	41.055821	41.298617	41.573870	41.890841	42.258662	42.685654	1.0	2.0	1.0
...	-6.096871	-8.136745	-10.007141	-11.671469	-13.100744	-14.274493	-15.181258	1.0	3.0	2.0
...	12.364592	11.908984	11.384433	10.802266	10.172038	9.501039	8.794078	1.0	1.0	2.0
...	0.951228	0.148025	-0.692121	-1.564288	-2.462504	-3.379946	-4.309173	1.0	5.0	2.0

Figure 3: Sample of last rows in the EEG database

4 PREPROCESSING

To facilitate data manipulation and visualization, pre-processing techniques are applied to the EEG data. The data is grouped and organized to ensure better interactions and efficient representation. Figure 4 demonstrates the pre-processed and grouped data, ready for visualization.

5 CASE STUDIES

The proposed visualization method drew inspiration from case studies that utilized real EEG data. These case studies focused on various subjects and stimuli, showcasing their effectiveness and utility. Accordingly, certain features will be incorporated, while others will be discarded.

	F3	F4	C3	C4	P3	P4	Modalidade	Estímulo	Artefatos
0	[-45.02729511371338, -44.902286619666974, -44....	[-42.88673441931802, -43.131941456247525, -43....	[-49.081541886780045, -48.70818433782669, -47....	[-25.96819622975463, -26.706540702987603, -27....	[-25.551885656945796, -26.19597534183959, -26....	[-16.606480259592328, -17.409928194632304, -18...	1.0	9.0	2.0
1	[-10.482893531611804, -9.311389950651217, -8.0...	[-8.607144876748036, -7.515336983751907, -6.33...	[-10.615747285195177, -9.012588534088582, -7.3...	[-11.151471661119817, -10.131651675975537, -8....	[-16.492283434988245, -16.493624120757733, -16...	[-9.593184540158994, -9.443670181177227, -9.12...	1.0	11.0	2.0
2	[29.254169285191605, 28.53750638291766, 27.732...	[38.6208778575885, 38.61044263420951, 38.57428...	[13.206548177609838, 12.081677067132, 10.90945...	[18.697176197728652, 19.274349000422525, 19.93...	[7.153843993446792, 6.20305936017795, 5.236629...	[10.661000328893696, 10.315621167511502, 9.970...	1.0	9.0	2.0
3	[20.51612432799457, 19.929198299796106, 19.218...	[17.1059887215891, 16.720299849930846, 16.0859...	[5.406977135437583, 4.9887770666504485, 4.4574...	[11.562747054462589, 11.706945284789308, 11.62...	[5.739807221949882, 5.624339796526538, 5.38861...	[9.708160257114608, 9.587486324335973, 9.25310...	1.0	8.0	2.0
4	[-5.097754669221089, -5.770928813671626, -6.52...	[-6.131912881669465, -7.002213340085349, -7.92...	[-3.0303351241006156, -3.1053122748990125, -3....	[-0.53054417784955, -0.7941560501285959, -1.12...	[3.219428974753273, 2.96058002456934, 2.697398...	[4.421831032863709, 4.51171592537299, 4.555835...	1.0	11.0	2.0
...	...	...	...	...	...	...	...	...	...

Figure 4: Pre-processed and Grouped EEG Data: Note the columns named after its respective source (i.e., the sensors) and the last three ones "modalidade," "Estímulo," and "Artefatos," which are, respectively, the modality, stimuli and artifacts.

5.1 Visualization Inspired by [3]

EPviz uses line graphs to visualize EEG data [3] and provide a general overview of the data. The visualization by lines is efficient because it shows the high and low frequency components. In the example shown in Figure 5, the EPViz program visualizes the signals captured by all the electrodes, providing a good overview. However, it lacks the detailed view that would allow a more accurate analysis of the displayed curves.

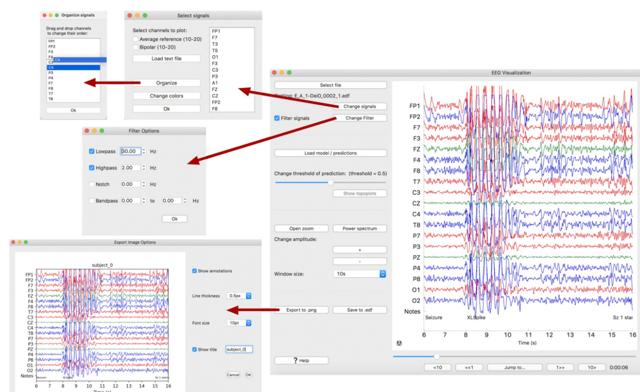

Figure 5: EPviz - Line graph visualization

5.2 Visualization Inspired by [4]

Another relevant methodology uses topological color maps to indicate signal energy at each sensor. This visualization approach shown in Figure 6 provides an informative representation of regions with intense brain activity during the corresponding time period. However, the view makes a interpolation of measured values leading to potential misinterpretations of signal intensities according to gradients.

5.3 Visualization Inspired by [5]

The methodology shown in Figure 7 simplifies the visualization to display brain activity at four different moments. However, similar to the previous visualization, the gradient issue persists, potentially leading to incorrect interpretations.

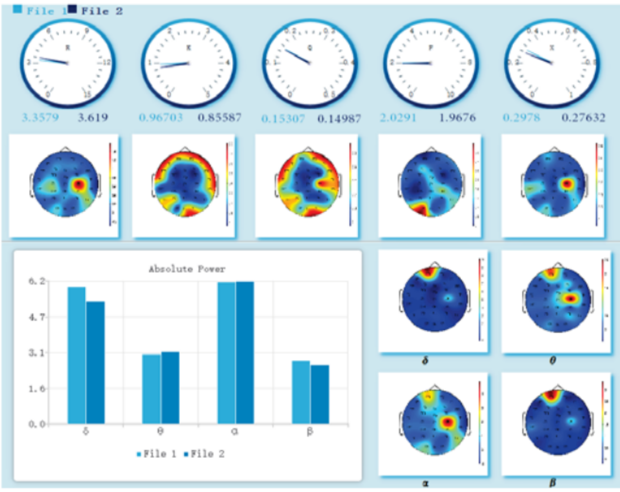

Figure 6: Topological visualization

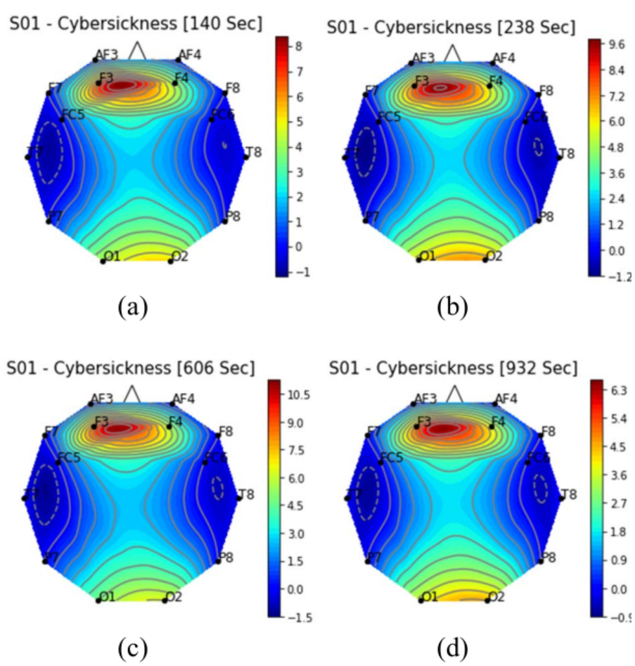

Figure 7: Topological visualization over time

6 PROPOSAL

The proposed visualization method includes line graphs and energy visualizations combined with a non-gradient representation of electrode positions in the 10-20 system. The interactive features allow users to select subjects and stimuli, explore specific signal details, view the signal's origin, and scale up or down the displayed curves. Additionally, the visualization will include bars showing the total energy of the displayed signal interval. The use of high-contrast colors will aid in perceptual differentiation. Figures 8 and 9 illustrate the general and specific views proposed.

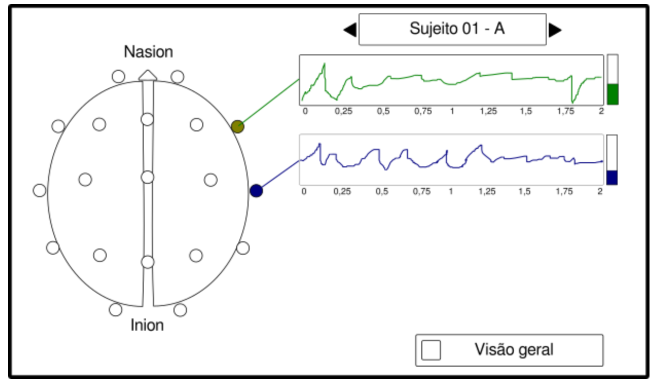

Figure 8: Specific view of the proposed visualization

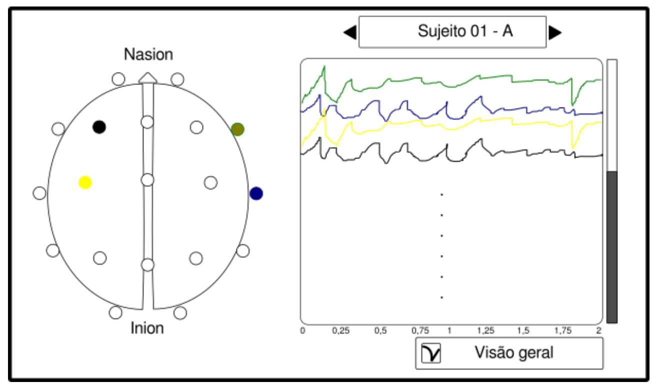

Figure 9: General view of the proposed visualization

Both the **subject and stimulus selection** and the checkbox are integral components of the **visualization cycle** [7]. When modifying the type of vision, stimuli and/or the analyzed subject, the view must promptly respond by updating the displayed data.

The use of **high contrast** colors aids in easily locating the origin of the data [9], preventing the common problem of differentiating data when very similar colors are employed.

The **magnification** feature enables a detailed view of each signal [10], facilitating the visualization of information linked to low and high-frequency signals. This provides the identification of **objects according to their "silhouette"** [11].

It's worth noting that the attributes utilized in the proposal have a **sequential** [6] order for the amplitude and frequency of the signals over time (line graphs), while simultaneously being **categorical** [6] because the frequency analysis can reveal valuable insights about the subject being analyzed.

The energy attribute (vertical bar on the side of the line graph) is **quantitative**.

Regarding **categorical** attributes, the region to the left of the graphs also deserves attention. This section indicates which brain regions are emitting specific signals, allowing, for example, the inference of the most active brain regions during thought execution and/or action.

In terms of **semiotic categories** [8] it is possible to highlight:

- firstness: Colors that stand out and immediately call the user's attention.
- secondness: The action/reaction applied to controls and images.
- thirdness: The consequent interpretation of the image presented as well as the reasoning necessary for understanding what is shown.

7 IMPLEMENTATION

The visualization was implemented using the Python3 [1] programming language due to its extensive set of data visualization and manipulation libraries. It is important to note that the final result differs slightly from the proposal, however, this does not compromise the quality of the final product, as can be seen in Figure 10 and 11.

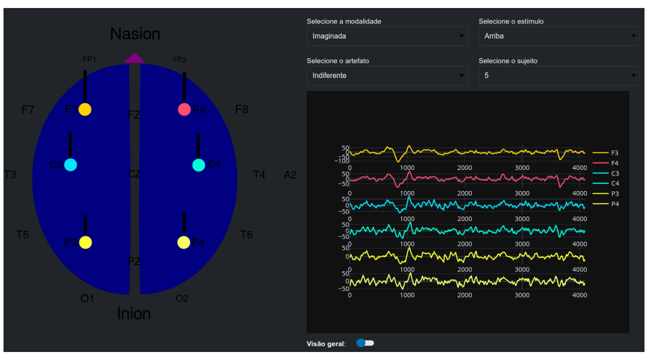

Figure 10: The application screenshot in specific mode: All signals are represented separately.

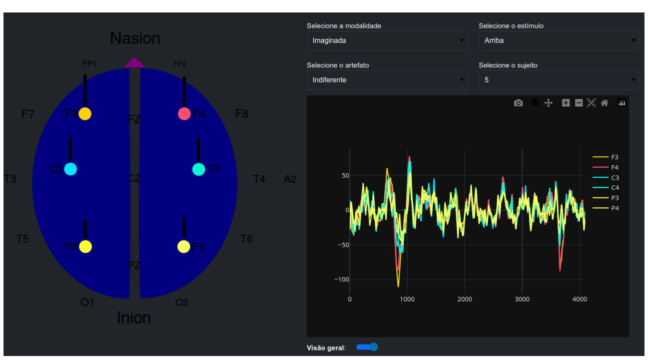

Figure 11: The application screenshot in general mode: All signals are represented in a single plot.

7.1 User interface discussion

Focusing on the left side of the visualization: Like is shown in Figure 12, the sensors that are being used are represented by circles with vivid and contrasting colors which matches with its respectively plot on the right side, while the unused sensors are represented only by their names. At the top of each circle, there is a bar representing the corresponding signal's energy captured by that sensor. It is noteworthy that each signal may have different energy levels.

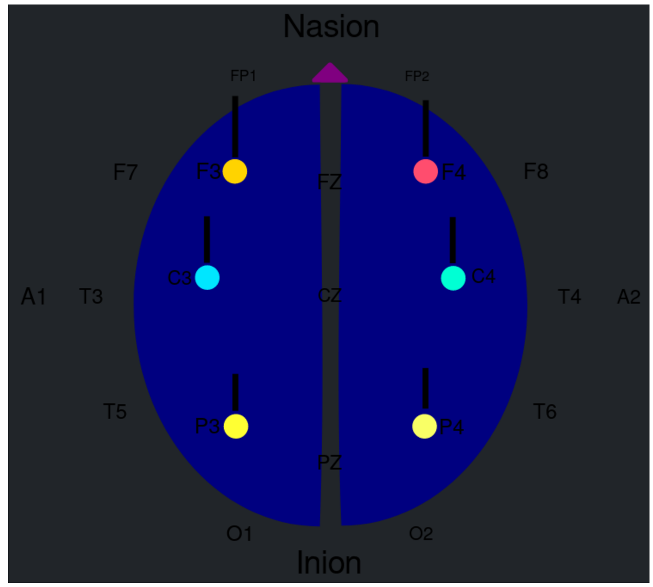

Figure 12: The sensors and its captured energy are represented by the colored circles and their top bars respectively.

The top section shown in Figure 13 is where the data filters reside: The first one "Selecione a modalidade" is a selector of the modality of the speech: Imagined or pronounced. Just setting this value alone does not affect the visualization because it needs the type of stimuli too. At the "Selecione o estímulo" selector is possible to set the type of stimuli which the visualization are needed. Considering that the previous selector is already set, this will trigger the change on the left and right parts of the screen showing the data visualization for the first subject. The subject can be changed using the "Selecione o sujeito" selector. Finally, the "Selecione o artefato" selector filters if data with artifacts are shown or not shown.

Figure 13: The four selectors of the visualization: "Selecione a modalidade", "Selecione o estímulo", "Selecione o artefato" and "Selecione o sujeito"

The right section (Figure 14) displays the plots of the signals captured by the sensors. Each plot has a color matching to its respective sensor, and to avoid misinterpretations, there are labels on the right. This section changes when new filters are set and when the switch "Visão geral" depicted in Figure 15 is changed. Here, it is possible to zoom in on data, hide plots by clicking on labels, and even save the results to a file.

Figure 14: The right section of the visualization: The signals may be shown separately or in a single plot.

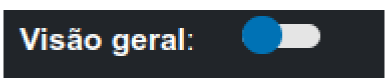

Figure 15: This button switches between "General" and "Specific" modes.

7.2 Source code

The code is designed to load EEG data from a MATLAB file, preprocess it. This visualization uses Panel, Plotly, SciPy, Pandas, BeautifulSoup and NumPy pyhton libraries. The code consists of two main classes: `Dashdata` and `Dashboard`.

The `Dashdata` class is responsible for loading and processing EEG data. The main tasks of this class are as follows:

```
def hello_world():
        print("Hello, World!")

hello_world()
```

- **loading EEG Data**: The `Dashdata` class loads EEG data from a MATLAB file using the `scipy.io.loadmat` function. The EEG data is stored in a MATLAB struct format, and the class uses `squeeze_me` and `mat_dtype` options to convert the data into a more usable format.
- **data preprocessing**: After loading the data, the class creates a Pandas DataFrame called `dataframe` to store the EEG data. The EEG signals for specific sensors are combined into arrays, and the DataFrame is updated accordingly. This step simplifies data handling and allows easier access to EEG signal data for visualization.
- **setting up Properties**: The class sets up properties to store essential information, such as the relationship between EEG sensors and their corresponding colors. This mapping is useful for visualizing EEG signals with different colors representing different sensors.
- **mapping Modalities, Stimuli, and Artifacts**: The `Dashdata` class provides dictionaries to map descriptive names of modalities, stimuli, and artifacts to their respective numeric codes. This mapping facilitates user interaction, as select widgets can be populated with descriptive options.
- **loading brain Visualization**: The class loads an SVG image of the brain, which can be used for visualizing the locations of EEG sensors on the brain. This SVG image is used in the dashboard to enhance the user experience and provide context to the EEG data.

The `Dashboard` class is designed to create an interactive EEG Data Visualization Dashboard using Panel and Plotly libraries. It utilizes the `Dashdata` class to access preprocessed EEG and visual elements data.

- **widgets and UI Layout** The dashboard creates various widgets for user interaction, such as selects for modalities, stimuli, artifacts, and subjects. Additionally, a switch widget is used to toggle between general and detailed views of the EEG data. The dashboard's layout is arranged using rows and columns to organize the visual elements effectively.
- **event handling** The class defines callback methods to handle user interactions with the widgets. These methods update the visualization based on the user's selections for modalities, stimuli, artifacts, subjects, and the general view toggle. The EEG data is filtered and plotted accordingly to provide an interactive and dynamic visualization experience.
- **data visualization** The dashboard uses Plotly to visualize the EEG data. It creates line plots representing EEG signals for different sensors. Depending on the user's selection, the dashboard can display single-panel or multiple-panel plots. The latter provides a more detailed view of the EEG signals.

8 FUTURE WORK

As the research progresses, the proposed visualization will be further developed to display the hierarchical decomposition of captured waves into their respective components (α, β, θ, σ). Additionally, the visualization will show the categories and classes to which each signal belongs.

9 CODE REPOSITORY

To access the source code, please visit:
https://github.com/ensismoebius/VisualizacaoDeInformacao

10 CONCLUSION

This work proposes a novel visualization method for EEG data, incorporating line graphs and energy visualizations with a non-gradient representation of electrode positions in the 10-20 system. The proposed visualization aims to facilitate EEG data analysis and interpretation by offering interactive features for subject and stimulus selection, as well as specific signal exploration. High-contrast colors will be employed to enhance perceptual clarity. Future research will extend the proposed visualization to include hierarchical representations of signal components and their associated categories and classes.

ACKNOWLEDGMENTS

The author would like to thank the Universidade Estadual Paulista Júlio de Mesquita Filho for its support and resources during this research.

REFERENCES

[1] [n. d.]. Python3 Programming Language. https://www.python.org/. Accessed: July 16, 2023.
[2] Germán A. Pressel Coretto, Iván E. Gareis, and H. Leonardo Rufiner. 2017. Open access database of EEG signals recorded during imagined speech.
[3] Danielle Currey, Jeff Craley, David Hsu, Raheel Ahmed, and Archana Venkataraman. 2023. EPViz: A flexible and lightweight visualizer to facilitate predictive modeling for multi-channel EEG. *Plos one* 18, 2 (2023), e0282268.
[4] Zhou Fang, Liuye Yao, and Zhiyu Qian. 2019. Research on Multi-parameter Visualization Technology of Brain Function Based on EEG. In *2019 International Conference on Medical Imaging Physics and Engineering (ICMIPE)*. 1–4. https://doi.org/10.1109/ICMIPE47306.2019.9098189
[5] Khaitami, Adhi Dharma Wibawa, Supeno Mardi, Susiki Nugroho, and Alfi Zuhriya Khoirunnisaa. 2019. EEG Visualization for Cybersickness Detection During Playing 3D Video Games. In *2019 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. 325–330. https://doi.org/10.1109/ISITIA.2019.8937083
[6] Andy Kirk. 2012. *Data Visualization: A Successful Design Process.* SAGE Publications Ltd.
[7] Tamara Munzner. 2014. *Visualization Analysis & Design.* CRC Press.
[8] Lúcia Santaella. 2017. *O que é semiótica.* Brasiliense.
[9] Robert Spence. 2014. *Information Visualization: An Introduction.* Springer.
[10] Edward Tufte. 1983. *The Visual Display of Quantitative Information.* Graphics Press.
[11] Edward Tufte. 2006. *Beautiful Evidence.* Graphics Press.
[12] Edrin Vicente. 2023. Sistema 10-20 para localizar alvos terapêuticos em EMT. https://www.kandel.com.br/post/como-localizar-os-pontos-de-estimulacao-para-estimulacao-magnetica-transcraniana