Acquisition, Pre-Processing, and Feature Extraction of EEG Signals to Convert it into an Image Classification Problem

Ishita Goyal[1], Anmol Mehta[2]

[1]*Maharaja Agrasen Institute of Technology, Rohini sector -22*
[2]*Maharaja Agrasen Institute of Technology, Rohini sector -22*

Abstract – *This paper essentially centers around the acquisition, pre-processing, feature extraction and conversion of Electroencephalogram (EEG) signals into spectrograms so that they can be used in any image classification/deep learning model easily. In our study, we worked on detecting speech but this methodology can be extended for different types of human conduct such as eye movement, lip movement, remembrance, attention, hand clenching, etc.*

Key Words: EEG, acquisition, pre-processing, signal, feature extraction, classification, speech, brain wave.

1. INTRODUCTION

The human brain has millions of neurons that play a significant role in controlling the behavior of the human body. These neurons will go about as data transporters between the body and cerebrum. Understanding the psychological conduct of the cerebrum should be possible by examining either signals or pictures from the mind. Human conduct can be visualized from states such as eye movement, lip movement, remembrance, attention, hand clenching, etc. These states are related to particular signal frequencies which assist with understanding the functional behavior of the brain.

Electroencephalography (EEG) is a productive methodology that assists with getting brain signals corresponding to various states from the scalp surface area. These signals are classified as delta, theta, alpha, beta, and gamma depending on signal frequencies ranging from 0.1 Hz to more than 100 Hz. An EEG can determine changes in a brain's activity that may help diagnose brain disorders, particularly epilepsy or a seizure. An EEG may likewise be useful for diagnosing and treating these problems: A brain tumor, brain damage from a head injury, brain dysfunction that can have a variety of causes (encephalopathy), inflammation of the brain (encephalitis), stroke, sleep issues. An EEG may likewise be utilized to affirm brain death in somebody in a persistent coma. A continuous EEG is used to help find the correct level of anesthesia for somebody in a medically induced coma.

This paper focuses on EEG signals and the complete end-to-end process of collecting them and predicting features from them. In this study, speech is used as a reference which can be substituted by any activity one wants to extract from it.

2. METHODOLOGY

To address the research objectives, it is important to establish an appropriate and effective methodology to ensure that an accurate and reliable outcome is obtained through this work. The EEG signals are first collected and then cleaned to remove noise present in the data using different pre-processing techniques. Data is extracted in form of images from these pre-processed signals by making use of Feature Extraction.

2.1 Signal Acquisition

The measurement of the neurophysiologic state of the brain, is known as signal acquisition. In an EEG test, electrodes (at metal discs) are placed onto the scalp using a sticky substance. These electrodes pick up the electrical signals from the brain and send them to an EEG machine, which will record the signals as wavy lines onto paper or on a computer. The EEG machine records your brain's electrical activity as a series of traces; each trace corresponds to a different region of the brain. The acquisition of the signal is done by a non-invasive method of EEG. In the noninvasive technique, medical scanning devices or sensors are mounted on caps or headbands and they help to read the brain signals.

A total of four subjects/volunteers participated in the EEG experiment. Each subject was first asked to listen to 5 different natural utterances and then speak out loud the utterances that they listened to. Every sentence was pre-recorded for 1 second each and repeated 40 times in a single audio recording with a gap of 1 second between them. For the speaking data, a sound of 0.5 seconds was played for the subject to distinguish the gap time and the speaking time of 1.5 seconds. The EEG was recorded in parallel while they were listening to the utterances as well the EEG was recorded in parallel while 10 they were speaking out the utterances that they listened to.

The natural utterances that the subjects listened to were predefined. Annotations (labeling of events in EEG) were added at the exact instance when the audio was played while recording the EEG signal. The five natural utterances that the subjects listened to were; "NEED HELP", "IT'S TOO HOT", "I AM HUNGRY", "GIVE ME WATER", "FEELING LONELY". As shown in fig 1 this is the EEG recording observed while they were listening to the utterances. We collected 80 Speech -EEG recordings per subject for each sentence and 40 Listening-EEG recordings per subject for each sentence.

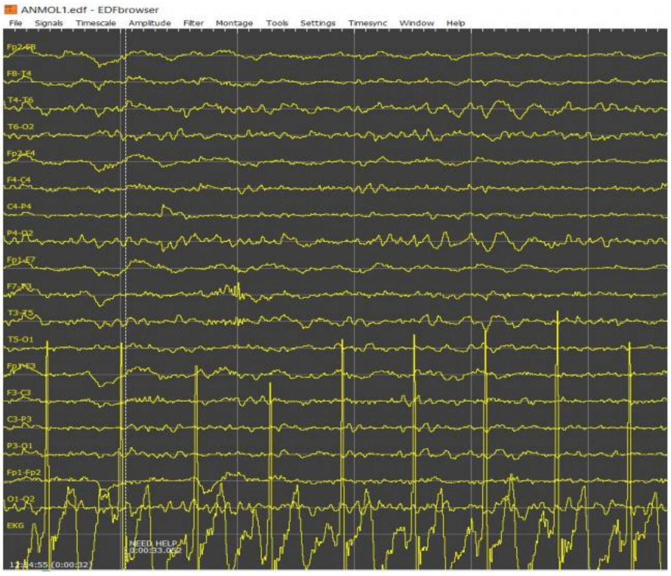

Fig -1: EDF dataset

2.2 Signal pre-processing

Non-invasive BCI has less signal clarity but is used as it is considered to be the safest. Less signal clarity in non-invasive BCIs is because the electrodes cannot be placed directly on the desired part of the brain. The removal of noise is done in this process. This stage includes decomposing or de-noising of the captured signal to remove noise and to enhance the EEG signal. A common technique used for the removal of noise from many different types of signals is the use of filters. The recorded signals as described in the above section need to be further processed to get rid of the noise embedded in them. The signals thus obtained are again used to extract only signals that were generated on each subject when a sentence was spoken/when the sentence was played from the total length of the signal. Finally, we extract several features from those signals. So, all these processes are described under a different heading in the following steps.

2.2.1 EEG bandpass filtering:

As far as the noise embedded in the recordings is concerned, such as superimposed artifacts from various sources, they can be effectively reduced by appropriate bandpass filtering. More particularly, the influence of eye blinking is most dominant below 4 Hz; heart-functioning causes artifacts around 1.2 Hz, whereas muscle artifacts affect the EEG spectrum above 30 Hz. Non-physiological artifacts caused by power lines are around 50-60 Hz.

Another reason to choose the bandpass filtering is due to a particular interest in the area of EEG frequency range. EEG signals can be isolated in 5 different frequency bands where each specific frequency band is more prominent in certain states of mind. Based on this fact, the two frequencies we choose are Alpha (8-12 Hz) and Beta (12-30 Hz). So to cater to the need of both removing the artifacts while retaining the signals within the particular band of interest, i.e. frequencies within the Alpha (8-13 Hz) and Beta (13-30 Hz) bands, we apply the 10th order "Butterworth bandpass filter".

Consequently, by extracting the Alpha and Beta frequency bands only from the acquired EEG recordings, we make sure to remove most of the physiological and non-physiological artifacts. We chose the 10th order because high order filters provide greater roll-off rates between pass band and stop band, and can be necessary to achieve the required levels of stopband attenuation or sharpness of cutoff. As shown in fig 2 we have applied the "Butterworth bandpass filter".

The lists below are some advantages of the Butterworth filter which leads us to choose it:

- Maximally flat magnitude response in the pass-band.
- Good all-around performance.
- Pulse response is better than Chebyshev.
- Rate of attenuation better than Bessel.

In EDFbrowser by choosing the filter from the top menu we were able to apply the filter we wanted to have on our signal. The result of this filtering will be new signals with frequencies between 8 Hz and 30 Hz.

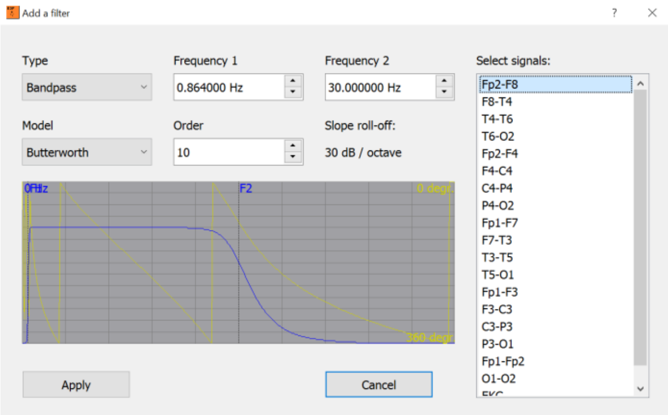

Fig -2: Signal pre-processing

2.2.2 EEG data segmentation

The filtered signals we get from the above step is further needed to be processed to get the signals that are generated when a sentence was spoken/when the sentence was played from the total length of the signal. That is, we are here motivated in retaining 40 segments of signals each of length 1 seconds for each subject for which one sentence was played and 80 segments of signals each of length 1.5 seconds for each subject for which one sentence was spoken.

We converted these filtered signals into ASCII format using the EDF browser. As the EEG signals were recorded at the frequency of 125 Hz. We had 125 rows for a single second of data. Using the EDFbrowser, we could read the annotations and be able to identify the exact time at which the recording started. Time and its corresponding labels were added in the ASCII file which was then converted to the csv file. For each listening EEG recording, we thus had 5000 resting data and 5000 listening data per person and for speaking EEG recording we had 4650 resting data and 14100 speaking data per person. Fig 3 depicts EEG data segmentation where the EDF file is converted to the csv file.

Time	Label	Fp2-F8	F8-T4	T4-T6	T6-O2	Fp2-F4	F4-C4	C4-P4	P4-O2	Fp1-F7
41	Resting	10.22	-9.92	14.8	0.31	10.68	5.04	-1.53	1.53	4.27
41.008	Resting	8.39	1.98	5.95	-1.68	8.85	4.27	0.92	0.76	3.51
41.016	Resting	12.21	16.94	-11.75	-4.73	8.39	0.92	3.05	0.46	6.41
41.024	Resting	9.46	5.65	-2.29	-2.59	8.09	0	1.53	0.46	9.92
41.032	Resting	7.02	-6.87	12.36	0.76	6.87	3.51	0.92	1.98	9.31
41.04	Resting	18.01	1.53	3.97	-1.22	8.09	6.26	2.14	5.95	11.6
41.048	Resting	24.72	12.97	-5.49	-3.2	13.73	6.26	3.2	5.95	19.07
41.056	Resting	19.53	10.68	3.2	-1.98	17.24	7.63	3.81	2.75	20.75
41.064	Resting	20.9	9.31	5.8	-2.59	15.72	7.63	5.04	5.34	17.4
41.072	Resting	27.16	11.6	-1.68	-3.81	15.11	4.73	4.88	8.54	17.7
41.08	Resting	22.74	-2.14	10.38	-1.37	19.07	3.05	2.29	5.19	17.7
41.088	Resting	14.5	-12.21	22.74	0.46	19.68	4.88	-1.07	1.68	16.02
41.096	Resting	15.11	1.83	7.02	-2.75	10.83	8.85	-2.44	3.97	16.63
41.104	Resting	20.75	14.19	-9.46	-5.8	4.73	8.7	0.31	5.95	16.94
41.112	Resting	22.58	8.54	-5.8	-4.73	11.29	3.05	2.44	3.66	15.41
41.12	Resting	19.07	0.31	1.37	-3.51	16.63	-1.22	0	1.83	16.63
41.128	Resting	16.02	-3.2	3.66	-3.81	10.99	1.37	-1.53	1.98	17.4
41.136	Resting	15.41	-1.53	2.9	-1.68	7.02	4.88	1.83	1.68	15.11
41.144	Resting	14.34	1.98	0.61	2.44	11.9	3.36	3.66	0.76	12.66
41.152	Resting	11.44	2.9	-2.29	2.75	14.04	-3.51	2.14	2.14	12.21
41.16	Resting	7.78	1.98	-0.76	-0.31	9.16	-3.05	-0.61	2.9	14.04
41.168	Resting	5.49	-2.29	6.56	-0.31	5.65	7.32	-3.97	0.31	15.41
41.176	Resting	7.48	-4.43	6.71	2.29	5.49	6.87	-1.07	0.46	11.9
41.184	Resting	8.54	1.22	-1.37	4.43	6.26	-0.76	3.2	3.97	7.63
41.192	Resting	5.19	9.16	-3.51	4.58	6.41	3.81	1.22	3.51	6.71
41.2	Resting	3.36	12.97	-2.14	1.53	5.19	8.54	-0.46	2.29	3.97
41.208	Resting	5.34	6.56	-2.59	-0.31	5.95	0.15	-0.15	2.9	-0.76
41.216	Resting	4.43	-4.12	0.76	-0.15	5.8	-4.58	-2.29	1.98	-1.98
41.224	Resting	1.37	-2.9	0	-1.68	1.07	-1.98	-2.59	0.31	-2.59
41.232	Resting	-0.46	5.34	-6.87	-3.97	-1.98	0.31	-3.05	-1.22	-4.88
41.24	Resting	-3.05	3.81	-6.87	-2.75	-0.76	1.98	-5.95	-4.12	-5.34
41.248	Resting	-7.48	0.76	-5.19	1.53	-0.92	-1.22	-2.75	-5.49	-5.8
41.256	Resting	-11.6	4.43	-7.78	2.75	-2.75	-6.71	1.68	-4.73	-9.31
41.264	Resting	-10.38	0.76	-2.75	-2.44	-3.2	-4.73	-1.98	-5.04	-12.97
41.272	Resting	-6.26	-2.75	1.22	-6.87	-3.36	-1.68	-3.97	-5.65	-12.21
41.28	Resting	-7.93	8.39	-7.32	-6.41	-4.88	-2.44	-0.76	-5.34	-8.24

Fig -3: Dataset in CSV format

2.3 Feature Extraction

Now, we have the signals in which the noise and artifacts have been removed from it and we should make a decision on which features to be extracted from this signal to make some data sets that will be used as input to a neural network. We know that EEG signals are classified based on their frequencies into the delta, theta, alpha, beta, and gamma. As shown in fig 4 describes how EEG signals are classified based on their frequencies.

Now to extract these features from the time signal we plotted the FFT graph using the Numpy library. The signal was recorded at 125 Hz which allowed us to extract the frequency in a usable range up to 60 Hz. Now to further analyze our data in both the time and frequency domain we use the matplotlib function specgram. Using this function we generate a plot of 224*224 with an NFFT value of 16 with an overlap of 8, which meant that for every 16 value points we are calculating an FFT with an overlap of 8 which helps to reduce the variance of the estimate of the power density spectrum. We save the plots in a grayscale format which will be used as an input in the neural network.

Spectrogram is a feature that helps analyze the EEG signal in both frequency and power domain, which is essential since the EEG have different frequencies at different points of time. In Fig 5 spectrogram helps visualize the EEG signal in frequency, power, and time domain.

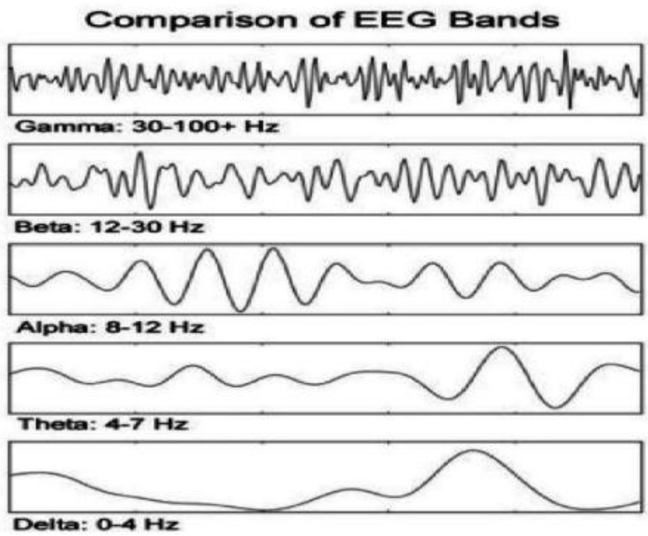

Fig -4: Comparison of EEG bands

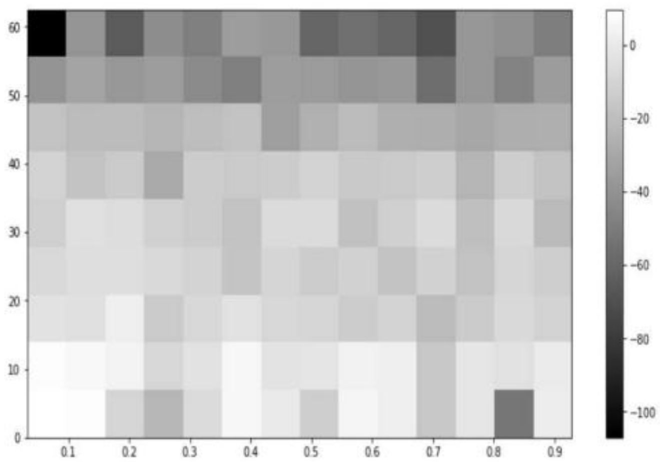

Fig -5: Spectrogram

2. RESULTS

The models can be implemented on 2 different forms of the spectrogram dataset. In the first, the input can be a channel-wise (channels of EEG) distribution, and in the second, it can be distributed sentence-wise, with spectrograms of different channels merged into a single image (using Image module of Pillow/PIL).

3. CONCLUSION

This study can be useful for many purposes. Firstly, researchers who are working on EEG signals to extract features can use this approach to help achieve their objectives. Secondly, it can be used to capture signals and try alternative ways to help many individuals in day-to-day life who cannot express themselves by allowing them to communicate with others. Third, it has converted a complex signal classification problem into an easy to work on an image classification problem.

REFERENCES

[1] Gautam Krishna, Yan Han, Co Tran, Mason Carnahan, and Ahmed H Tewfik, "State-of-the-art speech recognition using EEG and towards decoding of speech spectrum from eeg," arXiv preprint arXiv:1908.05743, 2019

[2] Gopala K Anumanchipalli, Josh Chartier, and Edward F Chang, "Speech synthesis from neural decoding of spoken sentences," Nature, vol. 568, no. 7753, pp.493, 2019.

[3] Gautam Krishna, Co Tran, Jianguo Yu, and Ahmed Tewfik, "Speech recognition with no speech or with noisy speech," in Acoustics, Speech, and Signal Processing (ICASSP), 2019 IEEE International Conference on. IEEE, 2019

[4] Gautam Krishna, Co Tran, Mason Carnahan, and Ahmed Tewfik, "Advancing the speech recognition with no speech or with noisy speech," in 2019 27th European Signal Processing Conference (EUSIPCO). IEEE, 2019.

[5] Katrin Kirchhoff, Gernot A Fink, and Gerhard Sagerer, "Combining acoustic and articulatory feature information for robust speech recognition," Speech Communication, vol. 37, no. 3-4, pp. 303–319, 2002.

[6] Yichuan Liu and Hasan Ayaz, "Speech recognition via fnir sbased brain signals," Frontiersin Neuroscience, vol. 12, pp. 695, 2018.

[7] Gautam Krishna, Co Tran, Mason Carnahan, and Ahmed Tewfik, "Advancing speech recognition with no speech or with noisy speech," in 2019 27th European Signal Processing Conference (EUSIPCO). IEEE, 2019 27

[8] NF Ramsey, E Salari, EJ Aarnoutse, MJ Vansteensel, MB Bleichner, and ZV ZV Freudenburg, "Decoding spoken phonemes from sensorimotot cortex with high-density ecog grids," Neuroimage, 2017.

[9] Stephanie Martin, Peter Brunner, In˜aki Iturrate, Jose´ del R Milla´n, Gerwin Schalk, Robert T Knight, and Brian N Pasley, "Word pair classification during imagined speech using direct brain recordings," Scientific reports, vol. 6, pp. 25803, 2016.

[10] Natsue Yoshimura, Atsushi Nishimoto, Abdelka- der Nasreddine Belkacem, Duk Shin, Hiroyuki Kam- bara, Takashi Hanakawa, and Yasuharu Koike, "Decoding of covert vowel articulation using electroencephalography cortical currents," Frontiers in neuroscience, vol. 10, pp. 175, 2016.

[11] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, SherjilOzair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets," in Advances in neural information processing systems, 2014, pp. 2672–2680.

[12] John Kominek, Tanja Schultz, and Alan W Black, "Synthesizer voice quality of new languages calibrated with mean mel cepstral distortion," in Spoken Languages Technologies for Under-Resourced Languages, 2008.

[13] Tian Tan, Yanmin Qian, Hu Hu, Ying Zhou, Wen Ding, and Kai Yu, "Adaptive very deep convolutional residual network for noise-robust speech recognition," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol26, no8, pp.1393–1405, 2018.

[14] Jianguo Yu, Konstantin Markov, and Tomoko Matsui, "Articulatory and spectrum features integration using generalized distillation framework" in Machine Learning for Signal Processing (MLSP), 2016 IEEE 26th International Workshop on. IEEE, 2016, pp. 1–6.

[15] Tanja Schultz, Michael Wand, Thomas Hueber, Dean J Krusienski, Christian Herff, and Jonathan S Brumberg, "Bio signal-based spoken communication: A survey," IEEE/ACM Transactions on Audio, Speech and Language Processing, vol25, no.12, pp 2257–2271,2017.

[16] Mashael M AlSaleh, Mahnaz Arvaneh, Heidi Chris- tensen, and Roger K Moore, "Brain-computer interface technology for speech recognition: A review," in Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2016 Asia-Pacific. IEEE, 2016, pp. 1–5