# Visual Brain Decoding for Short Duration EEG Signals

Rahul Mishra
*Indian Institute of Technology Mandi*
Mandi, INDIA
d16043@students.iitmandi.ac.in

Krishan Sharma
*Vehant Technologies Pvt. Ltd.*
Noida, INDIA
krishans@vehant.com

Arnav Bhavsar
*Indian Institute of Technology Mandi*
Mandi, INDIA
arnav@iitmandi.ac.in

***Abstract*—In this work, we propose a CNN-based approach for classification of short duration EEG signals for visual brain decoding. These signals are captured for a visual perception task by showing digit images on a computer screen, and the task involves classification of the EEG signals into 10 classes, corresponding to the digits shown. The captured EEG signals are of very short duration (approx. 2sec), which are typically very noisy. We use a correlation based technique for the removal of highly noisy samples. Further, a sample refinement approach for the selection of relevant channels is also proposed. Both these steps constitute the data refinement process, which we demonstrate has a significant effect on the CNN classification performance. We validate the proposed approach on a publicly available MindBigData (The "MNIST" of Brain Digits) dataset.**

***Index Terms*—CNN, Correlation coefficients, EEG Classification, visual brain decoding**

## I. Introduction

Brain decoding is based on the premise that the brain activity as measured by devices such as Functional Magnetic Resonance Imaging (fMRI), Electroencephalography (EEG), Magnetoencephalography (MEG) etc. can contain signatures of underlying neural processes or mental states corresponding to different emotions, intentions to perform a task (e.g. in case of motor imagery), attention, eye movements, decision making in certain task, mental load, stress etc. With progress in brain computer interfaces and machine learning or deep learning techniques, encouraging attempts are being reported to draw inferences about such mental states based on external measurements of electrical activity (via EEG) or magnetic fields and their effects (via fMRI, MEG).

Among the above methods for measuring neuronal signals, EEG which is the most common one due to some advantages such as high temporal resolution and relatively low cost. While, traditionally an important application of EEG was in the clinical setting for seizure detection and analysis [1], its uses have flourished in various cognitive science applications involving brain decoding and brain computer interfaces.

A relatively new sub-domain, which can be termed as perceptual brain decoding (PBD), involves identifying an external perceptual (e.g. visual, audio) stimulus, using the responses from brain (evoked by such stimuli). Perceptual brain decoding has benefits from both the cognitive and clinical perspectives.

Recent research in the PBD have shown some motivating outcomes [2], [3] that indicate the presence of the discriminative information in the brain recordings corresponding to the visual perception task. Despite the above motivating results, PBD is still considered as a challenging task. There are few factors which may play a crucial role in a brain decoding task, e.g., the recording time of EEG signals and the quality of captured EEG samples. Considering that the task of PBD is relatively new, there is little clarity about aspects such as sufficient time to observe and imagine an image, and the corresponding time to capture EEG signals. While there have been some encouraging attempts for classification of long duration (about 10s per instance) EEG signals evoked by visual inputs [4], a similar classification for short duration EEG signals needs more exploration. In order to explore this direction we are focusing on short duration (2s per instance) EEG signals, which we believe can be noisier, and may need refinement. The author in [5] focus on the requirement of good quality EEG samples for any human cognition task. The authors also compare the signal quality of different devices used for recording brain activities.

Thus, in this work, we consider a classification task for EEG signals, where the classes correspond to stimuli evoked by images of ten different digits (from 0-9)[1]. Our contributions are as follows: 1) Considering the requirement of good quality EEG samples for short duration EEG signals, in this work we first propose an approach to select, arguably, more discriminative EEG channels, followed by the selection of good quality EEG signals from these channels. 2) The above process amounts to an automatic refining of the EEG data, and we demonstrate that the refined dataset yields significantly better classification results than using the unrefined data. To support the classification performance, we provide a t-SNE visualization (a method of visualizing a high-dimensional data in a low-dimensional space [6]) of the discrimination across the classes before and after the refinement. 3) In addition to the data refinement approach, we also propose a novel 1D CNN based classification method, which involves convolutions across time and channel samples.

The rest of the paper is organized as follows. In section 2, we discuss recent works related EEG classification (for both medical as well as perceptual decoding). In section 3, we provide the details of the dataset which is publicly available.

[1] http://www.mindbigdata.com/opendb

ISBN: 978-9-0827-9706-0

Section 4 deals with our approach for filtering, removal of noisy samples and channel selection followed by classification on refined EEG data.

## II. Related Work

A large fraction of literature on EEG classification is focused on clinical applications such as seizure detection [1], [7]. However, apart from these, there are various other application domains involving EEG based analysis such as event-related potential detection for an EEG based image annotation system [8], mental workload of a person [9], emotion classification [10], [11], sleep cycle information extraction [12], and motor imagery task classification [13] etc. The authors in [14] review the significant current approaches for EEG classification using deep learning approaches.

Schirrmeister et al. [15] study distinct CNNs architectures specifically designed for decoding of imagined stimulus from EEG signals. The work also highlights the potential of CNNs for the brain decoding task. Bashivan et al. [16] present a novel approach to learn effective representations of EEG signals from raw EEG data with the help of topology-preserving multi-spectral images and LSTM based features. Some of the recent research also includes analyzing brain activity of a person performing a visual task [3], [17]. However, a very limited number of methods have been developed [18], [19] to address the problem of decoding the EEG signals associated with the task of visual perception.

Tirupattur et al. [4] proposed a deep learning network for the classification of long duration EEG signals while performing a visual perception task on ThoughtViz dataset [4]. In a recent work the authors [20] proposed an LSTM based deep learning network for the task of EEG classification based on digits based visual stimuli on a short duration dataset from the same source (MindBigData) that we have used in this work. Another method on the same dataset involving a GRU based deep network has been proposed by the authors of [21]. As such the MindBigData is recorded using 4 devices, from which the authors of both of the above works have used EEG signals from a 4 channel MUSE device (details is in the subsequent section). However, the classification accuracy for these works is in the range of 11% to 30%. On the other hand, in this work, while we also work with the MindBigData, as indicated in the next section, we have chosen to use the EEG signal captured with the Emotiv Epoc device which has 14 channels, considering that there is a large scope of improvement of the classification performance on this dataset.

## III. *MindBigData* - MNIST of Brain Digits

MindBigdata[2] is a publicly available dataset for perceptual brain decoding. This dataset is a collection of EEG signals which are obtained by exposing a subject to a visual stimuli of numerical digits (as shown in Fig. 2) multiple times. The image appears on the screen and the subject is asked to imagine the appeared digit for 2 seconds. The EEG signals is captured for this imagined time. This dataset is prepared by capturing EEG signals using four commercial EEG devices. These, along with their channel information are as follows:

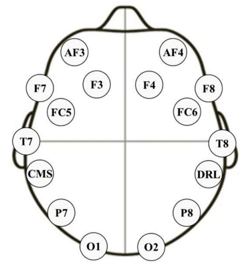


Fig. 1. Electrode locations in Emotiv Epoc (reproduced from [22])

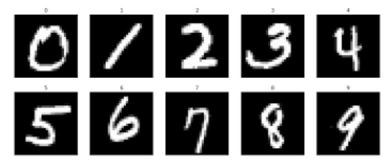

Fig. 2. Samples of MNIST based visual stimuli [23]

- NeuroSky MindWave ( “FP1”)
- Emotiv EPOC (“AF3”, “F7”, “F3”, “FC5”, “T7”, “P7”, “O1”, “O2”, “P8”, “T8”, “FC6”, “F4”, “F8”, “AF4”)
- Interaxon Muse (“TP9”, “FP1”, “FP2”, “TP10”)
- Emotiv Insight (“AF3”, “AF4”, “T7”, “T8”, “PZ”)

(All locations are w.r.t. the standard (10/20) locations.)

It may be noted that the parietal lobe of human brain is mainly responsible for the functions like perception, object classification, knowledge of numbers etc [24]. Therefore we are using Emotive Epoc device data for digit classification task as electrodes locations (“P7”, “P8”) of parietal lobe is available only in it. This device uses 14 electrode as shown in Fig. 1 with a sampling rate of 128 Hz. To our knowledge, this is the only publicly available dataset for digit classification using short duration EEG signals, captured with this device.

## IV. The proposed approach

Broadly, there are two primary components in our work.

- Data refinement
- EEG classification

### A. Data Refinement

The data refinement is an important aspect in this work, which we demonstrate, is responsible for a significant improvement in the classification performance. Below we discuss the steps in this process.

*1) Filtering:* We follow the standard filtering paradigm by removing the DC (0 Hz frequency) component from the data. Further, an additional band-pass filter (3 Hz to 30 Hz) is applied to extract the relevant EEG bands (Theta band, alpha band and beta band).

[2]http://www.mindbigdata.com/opendb

*2) Removal of Noisy Samples:* Since the EEG signals are highly prone to noise, the first step after filtering is to remove the noisy channels. Let $X = \{X^0, ..., X^i, ..., X^9\}$ denote the total training EEG data where $X^i = \{X^i_1, .., X^i_j, ..., X^i_{14}\}$ denotes the EEG data corresponding to $i^{th}$ digit ($i \in \{0 \text{ to } 9\}$). $X^i_j \in \mathbf{R}^{d \times n_j}$ is the $i^{th}$ digit data corresponding to $j^{th}$ channel, where $d$ and $n_j$ denote the signal length and number of signals, respectively.

We make use of event-related potential (ERP) signal to remove the very noisy samples, which may not contribute for classification. An ERP signal is the measured brain response that is the direct result of a specific sensory, cognitive, or motor event [25]. ERPs are measured by taking the means of EEG signals as

$$\boldsymbol{\mu^i_j} = \frac{1}{n_j}[X^i_j \mathbf{1}], \tag{1}$$

where $\boldsymbol{\mu}^i_j \in \mathbb{R}^d$ denotes the mean of $X^i_j$, and $\mathbf{1}$ denotes a column vector of all-ones of length $n_j$.

We then measure the correlation coefficient between each signal $\boldsymbol{x}$, i.e., columns of $X^i_j$, and ERP signal $\boldsymbol{\mu}^i_j$ as

$$z = \frac{d(\sum x\mu^i_j) - (\sum x)(\sum \mu^i_j)}{\sqrt{[d(\sum x^2) - (\sum x)^2][d(\sum {\mu^i_j}^2) - (\sum \mu^i_j)^2]}} \tag{2}$$

The value of correlation coefficient is in between -1 to +1. If the value of the correlation coefficient is high, that means the signal $\boldsymbol{x}$ is closer to the ERP signal and can be considered as less noisy. We select only those signals that are having a correlation coefficient greater than a certain threshold.

*3) Selection of relevant EEG channels:* Since all the electrode channels are not contributing significantly to the PBD task, selection of discriminative channels is necessary. In order to find out the discriminative channels, we follow an approximate greedy strategy. More specifically, we perform a binary classification task for each channel data independently. The data obtained after filtering using correlation coefficient is used for one-vs-one digit classification for each of the channels. In total, there are 45 such pairs for binary digit classification. For each channel, the average EEG classification accuracy for all the possible pairs of digits for all the channels are listed in the given Table I.

It is clear from this Table I that channel no. 5 (T7), 6 (P7), 9 (P8) and 10 (T7) carries more discriminative information than the other remaining channels. It is in accordance with our expectations as channel no. 6 and 9 represent the electrode locations of parietal lobes and results in meaningful EEG representation for this particular task. So, for further analysis, we consider only these 4 channels. Thus, after this step, the actual data has only 4 channels instead of 14. Interestingly, the earlier works which use the data from the MUSE device also work with 4 channels. However, the important difference is that in our case, the 4 channels are selected automatically from the original 14 channels, based on their potential importance to the classification task, and significantly contribute towards the same.

TABLE I
AVERAGE CLASSIFICATION ACCURACY OF BINARY DIGIT CLASSIFICATION USING K-NN FOR ALL ELECTRODE CHANNELS

| Channel No | Average Classification Acc. |
|---|---|
| *CH 1* | 0.5212 |
| *CH 2* | 0.545 |
| *CH 3* | 0.521 |
| *CH 4* | 0.528 |
| *CH 5* | 0.641 |
| *CH 6* | 0.634 |
| *CH 7* | 0.525 |
| *CH 8* | 0.549 |
| *CH 9* | 0.64 |
| *CH 10* | 0.566 |
| *CH 11* | 0.536 |
| *CH 12* | 0.527 |
| *CH 13* | 0.537 |
| *CH 14* | 0.5152 |

*4) Channel Interpolation:* Due to the removal of noisy samples, it is quite possible that some signals from some samples corresponding to these 4 channels may be removed. Therefore we introduce the notion of channel interpolation by interpolating the missing sample data of a particular channel. Considering the presence of the noisy samples in the dataset, we have chosen only those samples which have at least 3 signals with a correlation coefficient greater than a certain threshold. The channel interpolation can be described as below:

- Find out the channel with signal having low correlation (less than a certain threshold) from the training sample of a class.
- Interchange this less correlated channel signal with a high correlated signal from the same channel of other sample of the same class.

### B. EEG Classification

For classification of EEG signals we have used convolution neural networks. The motivation for using CNN comes from the ability of CNN to learn contextual information of the data. In this work it is important to capture the context information in two direction (across time axis as well as across channel axis). Thus, we have used 1-D CNN across time followed by 1-D CNN across channels. This configuration of CNN architecture enables us to capture the neighbouring information in the two required directions.

*1) CNN Network:* The details of base deep learning model is given below:

The input data is of the dimension (4 x 249) (i.e. 4 channels and 249 samples)

- Application of 1D CNN on each channel axis to capture neighbouring information across the time axis
- Application of 1D CNN on channel axis to consider neighbourhood information across channel axis.
- Application of maxpool layer to provide robustness against intra-class variation.
- Application of 1D CNN on time axis

- Fully connected dense layers followed by an output layer with softmax activation.

The numbers of convolution filters for each block are 32, 25 and 32 respectively. The fully connected layers connected layers have 128, 64 and 32 neurons. The final softmax layer is of the size equal to the number of classes. ReLU activation has been used after each of the internal layers. We train the classifiers with Adam optimizer with a batch size of 100 and learning rate of 1e-4. The network was trained from scratch.

## V. Experiment & Results

Here, we provide the results of our experiments with the MindBigData dataset using the proposed 1-D CNN based deep network. The ratio of training and test data is roughly 90:10. More specifically, for each class the training partition consists of EEG data corresponding to trials involving 90% of the images displayed, while in test data we have EEG data corresponding to trials involving the remaining (10%) set of images. In the subsections below, we discuss the results considering different components of our approach.

### *A. Results showing the effect of the removal of noisy samples, and discriminative channel selection*

The comparison of classification accuracy for cases without removal of noisy samples, and without channel selection is shown in table II. For the experiments without channel selection, we have used a similar network for all the classification tasks except that the size of 1-D CNN on channel axis is 14x1. For the case of 14 channel classification with data refinement, we have chosen all those samples in which atleast 10 signals have a correlation coefficient value greater than a certain threshold. Column three shows the results for the case with selection of 4 channels but choosing random samples equal to the amount of data that is chosen based on our threshold strategy. This comparison shows that refinement strategy (selection of good EEG samples) helps to significantly improve the performance even on 14-channels (Col 1 vs Col 2). While only a reduction of channel without the correlation based selection does not yield much gains (Col 1 vs Col 3), reducing the channels further on this cleaner, as a part of our overall process, again improves the results considerably (Col 2 vs Col 4).

TABLE II
COMPARISON OF CLASSIFICATION ACCURACY (%) WITHOUT REMOVAL OF NOISY SAMPLES & WITHOUT CHANNEL SELECTION

| EEG data for 14 channels without removing noisy samples | Data refinement with 14 channels (z>0.1) | Random selection of samples for 4 channels (Without ) (noise removal) | Refined with 4 channels and interpolation (z>0.1) |
|---|---|---|---|
| 11.5 | 32.4 | 13.9 | 57.3 |

### *B. Results for different values of correlation coefficient, and the effect of interpolation*

The comparison of classification accuracy with different values of correlation constant is given in table III. We provide the results with the interpolation of channels and without interpolation of channels in training data. The results are showing that as we increase the value of correlation coefficient for sample refinement, the classification accuracy is increasing. These results validate our approach that selection of good quality samples (removal of noisy samples), using even conservative correlation values, along with the channel interpolation in training data, helps in providing a good classification performance..

TABLE III
CLASSIFICATION FOR DIFFERENT VALUES OF CORRELATION COEFFICIENT THRESHOLD, AND EFFECT OF INTERPOLATION

| Correlation Threshold(z) | Classification Accuracy(%) without Interpolation | Classification Accuracy(%) with Interpolation |
|---|---|---|
| z>0.1 | 48.2 | 57.3 |
| z>0.15 | 51.5 | 60.5 |
| z>0.2 | 55 | 70.1 |

The detailed confusion matrix for z>0.2 is given below in Fig. 3. From the figure it is clear that the individual category classification accuracy for all digits are consistent except for some drop in a couple of classes. The reason for this may be due to the similarity in the digits with other category (like the similarity between 1 and 7 and in between 4 and 9).

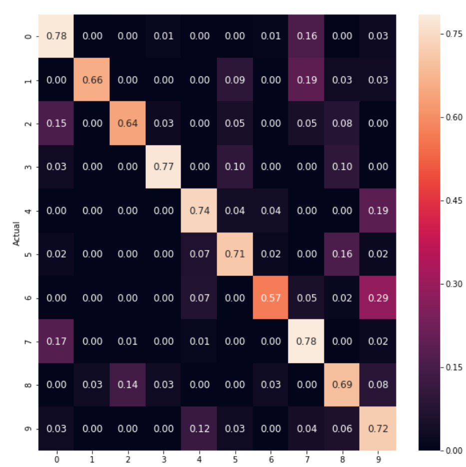


Fig. 3. Confusion matrix for all classes (z >0.2)

### *C. Visualization of seperability using t-SNE representation*

In figure 4 we show the t-SNE plots of the train and test data. t-Distributed Stochastic Neighbor (t-SNE) Embedding is a non-linear dimensionality reduction technique used for visualizing high-dimensional data in a low dimensional space [26]. It is clear that the raw data is highly overlapped (both in case of training data (a) and testing data (c)). We then show the t-SNE of the samples after training on the refined data (with z>0.2

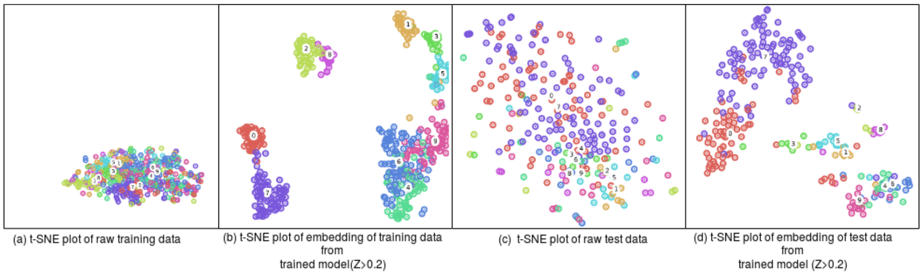


Fig. 4. Visualization of training and test data using t-Stochastic Neighbor Embedding

and with interpolation), for which we obtained the embeddings of the train and test data from the last layer of our network (before softmax output layer). The plot of t-SNE of these data embeddings is given in Fig. 4 (c) and (d) which clearly shows the discrimination in the train & test data embedding.

## VI. Conclusion

In this work we have proposed a correlation based approach for the EEG data refinement, in a visual brain decoding task, involving classification of visually evoked EEG signals. We believe that the extremely noisy samples can deteriorate the performance of any EEG based classification and hence removal of such noisy samples should be the first step for any EEG based PBD task. After data refinement we have proposed a 1-D CNN based deep learning network for classification. The proposed approach is showing promising classification performance of around 70% on MindBig data, and the effect of data refinement on the separability of the classes is also convincingly visualized via t-SNE representation.

## References


[1] G. Chen, "Automatic EEG seizure detection using dual-tree complex wavelet-fourier features," *Expert Systems with Applications*, vol. 41, no. 5, pp. 2391–2394, 2014.

[2] A. G. Huth, T. Lee, S. Nishimoto, N. Y. Bilenko, A. T. Vu, and J. L. Gallant, "Decoding the semantic content of natural movies from human brain activity," *Frontiers in Systems Neuroscience*, vol. 10, p. 81, 2016.

[3] A. G. Huth, W. A. De Heer, T. L. Griffiths, F. E. Theunissen, and J. L. Gallant, "Natural speech reveals the semantic maps that tile human cerebral cortex," *Nature*, vol. 532, no. 7600, pp. 453–458, 2016.

[4] P. Tirupattur, Y. S. Rawat, C. Spampinato, and M. Shah, "Thoughtviz: Visualizing human thoughts using generative adversarial network." New York, NY, USA: Association for Computing Machinery, 2018. [Online]. Available: https://doi.org/10.1145/3240508.3240641

[5] T. Radüntz, "Signal quality evaluation of emerging EEG devices," *Frontiers in Physiology*, vol. 9, p. 98, 2018. [Online]. Available: https://www.frontiersin.org/article/10.3389/fphys.2018.00098

[6] G. Hinton and S. T. Roweis, "Stochastic neighbor embedding," in *NIPS*, vol. 15. Citeseer, 2002, pp. 833–840.

[7] R. J. Oweis and E. W. Abdulhay, "Seizure classification in EEG signals utilizing hilbert-huang transform," *Biomedical engineering online*, vol. 10, no. 1, p. 38, 2011.

[8] V. Parekh, R. Subramanian, D. Roy, and C. Jawahar, "An EEG-based image annotation system," in *National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics*. Springer, 2017, pp. 303–313.

[9] G. Di Flumeri, G. Borghini, P. Aricò, N. Sciaraffa, P. Lanzi, S. Pozzi, V. Vignali, C. Lantieri, A. Bichicchi, A. Simone *et al.*, "EEG-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings," *Frontiers in human neuroscience*, vol. 12, p. 509, 2018.

[10] J. Chen, P. Zhang, Z. Mao, Y. Huang, D. Jiang, and Y. Zhang, "Accurate EEG-based emotion recognition on combined features using deep convolutional neural networks," *IEEE Access*, vol. 7, pp. 44 317–44 328, 2019.

[11] Y. Gao, H. J. Lee, and R. M. Mehmood, "Deep learninig of EEG signals for emotion recognition," in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE, 2015, pp. 1–5.

[12] H. Ghimatgar, K. Kazemi, M. S. Helfroush, and A. Aarabi, "An automatic single-channel EEG-based sleep stage scoring method based on hidden markov model," *Journal of neuroscience methods*, p. 108320, 2019.

[13] Y. He, D. Eguren, J. M. Azorín, R. G. Grossman, T. P. Luu, and J. L. Contreras-Vidal, "Brain–machine interfaces for controlling lower-limb powered robotic systems," *Journal of neural engineering*, vol. 15, no. 2, p. 021004, 2018.

[14] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: a review," *Journal of neural engineering*, vol. 16, no. 3, p. 031001, 2019.

[15] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[16] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent-convolutional neural networks," *arXiv preprint arXiv:1511.06448*, 2015.

[17] S. Nishimoto, A. T. Vu, T. Naselaris, Y. Benjamini, B. Yu, and J. L. Gallant, "Reconstructing visual experiences from brain activity evoked by natural movies," *Current Biology*, vol. 21, no. 19, pp. 1641–1646, 2011.

[18] A. Kapoor, P. Shenoy, and D. Tan, "Combining brain computer interfaces with vision for object categorization," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–8.

[19] N. Bigdely-Shamlo, A. Vankov, R. R. Ramirez, and S. Makeig, "Brain activity-based image classification from rapid serial visual presentation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, no. 5, pp. 432–441, 2008.

[20] J. J. Bird, D. R. Faria, L. J. Manso, A. Ekárt, and C. D. Buckingham, "A deep evolutionary approach to bioinspired classifier optimisation for brain-machine interaction," *Complexity*, vol. 2019, 2019.

[21] B. L. K. Jolly, P. Aggrawal, S. S. Nath, V. Gupta, M. S. Grover, and R. R. Shah, "Universal EEG encoder for learning diverse intelligent tasks," in *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)*. IEEE, 2019, pp. 213–218.

[22] K. Stytsenko, E. Jablonskis, and C. Prahm, "Evaluation of consumer EEG device emotiv epoc," in *MEi: CogSci Conference 2011, Ljubljana*, 2011.

[23] L. Deng, "The mnist database of handwritten digit images for machine learning research [best of the web]," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.

[24] A. D. Wagner, B. J. Shannon, I. Kahn, and R. L. Buckner, "Parietal lobe contributions to episodic memory retrieval," *Trends in Cognitive Sciences*, vol. 9, no. 9, pp. 445–453, 2005.

[25] S. J. Luck, *An introduction to the event-related potential technique*. MIT press, 2014.

[26] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.