Brain Signal Analysis for Mind Controlled Type-Writer Using a Deep Neural Network

Rohini Das, Sayan Goswami, Sayantani Ghosh, Mousumi Laha, Chandrima Debnath and Amit Konar
Artificial Intelligence Lab, Electronics and Telecommunication Engineering Department, Jadavpur University, Kolkata, India
rohinidas.907@gmail.com, sayan.goswami.106@gmail.com, sayantani.sonrisa25@gmail.com,
lahamou@gmail.com, chandrima5debnath@gmail.com, konaramit@yahoo.co.in

***Abstract*—The prime objective of this work is to develop a notable methodology for modeling a mind controlled type writing system to cater the needs of individuals suffering from various communication related disorders using EEG signal analysis technique. First, the EEG signals are procured from twelve subjects who were involved in mental utterance of seven vowel sounds. The eLORETA analysis of the acquired signals confirms the involvement of occipital, parietal and pre-frontal lobes for this cognitive activity. The procured signals are then filtered to circumvent the effects of various artifacts and are transferred to a novel deep multilayer perceptron classifier for categorization of seven class labels. Performance analysis undertaken confirms that the proposed classifier is able to distinctly categorize the seven class levels with a very high precision level. A coding mechanism has also been proposed to represent the consonants by the concoction of two vowel sounds segregated by a space. Thus, the proposed scheme can be effectively utilized as a mind controlled type writing system to serve the basic requirements of disabled individuals. Additionally, this technique can also be used in certain military scenarios that demand non-verbal communication as a secure option and in various gaming applications that will aid in exhilarating the player's experience.**

***Index Terms*—EEG, eLORETA, vowel sound imagery, activation function, deep multilayer perceptron**

I. Introduction

The daily lives of individuals suffering from various communication oriented ailments, especially the ones with advanced stages of Amyotropic Lateral sclerosis (ALP) [1], are crippled due to their inaccessibility to common conveyance modalities (writing , speech and/or gestures). Thus, the development of some form of technology is quite necessary to cater the basic needs of such diseased individuals. The present research attempts to achieve this goal by modelling a Mind Driven Type Writer (MDTW) whose prime function involves the abstraction of imagined vowel sounds from the brain response of subjects and its conversion to written/printed form of words or sentences.

Current research under this domain [2]–[5] deals only with the classification of vowel sound imageries using EEG induced analysis technique with little/no emphasis regarding the modelling of any neuro-prosthesis system. Moreover, the classifiers used here are quite primitive and provide fair results. The present work seeks to negate this void.

Our research begins with the exploration of active brain regions using eLORETA analysis method from the procured EEG signals of volunteers involved in mental enunciation of seven vowel sounds (A, AA, EE, AeA, O, Ae, UU) related to Bengali language semantics. In the next step, the signals are filtered to eliminate the effects of various artifacts and noise. The filtered signals are then transferred to an Independent Component Analysis (ICA) [6] module to select the most important channels involved in this cognitive activity. Finally, the selected signals are classified using our proposed Deep Multi-Layer perceptron based classifier to accurately categorize seven distinct class labels. This classifier has been chosen for its automatic feature extraction capability and precision in classifying multiple class labels [7]. In this model, we investigate the use of a new activation function, Swish [8] as a strategy to enhance the accuracy of the classifier module. The prime motivation behind the choice of this function is that it is quite efficient in tackling the vanishing gradient issue [9] and also allows better information propagation during training phase. Additionally, a code book has been developed to encode the consonants by coalescence of two vowel sounds separated by a space. For example, the letter B is represented as A_A, the letter C is denoted as A_EE and suchlike. So, the combination of two vowel sounds will be able to represent $\binom{7}{2} = \frac{7!}{2!(7-2)!} = 21$ consonants. Thus, this encoding scheme will be of immense help to the diseased subjects for communicating words or sentences constituting of both vowels and consonants.

Experimental results obtained infer that the proposed classifier module is able to classify seven vowel sound imageries with an enhanced accuracy level. Moreover, the proposed methodology has also shown supercilious performance when tested as a MDTW for both vowels and consonants by utilizing the classification of mental vowel sound imageries. Thus, this methodology can act as an efficacious tool to convey the messages of disabled individuals. Besides, the proposed technique can be utilized in military applications where non-verbal communication can provide a better and secure possibility under grave scenarios like war. Additionally, this system can play a vital role in numerous BCI based gaming applications.

The remaining segments of the paper have been compiled as follows. Section II describes the architecture of the proposed classifier. Section III provides the details of the experiments undertaken. Section IV discusses the performance analysis of the proposed model. The inference has been portrayed in Section V.

978-1-7281-5284-4/20/$31.00 ©2020 IEEE

II. SYSTEM OVERVIEW

This section introduces a summarized description of all principles and methodologies adopted to classify vowel sound imageries of subjects from their brain response. Fig.1 provides a detailed description of the overall system. The EEG signals are captured from the scalp of the subjects using 19 electrodes when they are involved in imagination of vowel sounds. The procured EEG signals are then fed to the exact Low Resolution Topographic Analysis (eLORETA) software [10], to identify the active brain regions responsible for this work. Next, to eradicate the effects of numerous artifacts from the raw EEG data, the signals are filtered using an Elliptical Band pass filter of order 10. The filtered signals are then transferred to an ICA (Independent Component Analysis) module to eradicate the artifacts concomitant in the pass band of the filter. The artifact free signals are transmitted to a novel deep based Mulilayer perceprton (MLP) network to classify seven vowel sounds imageries.

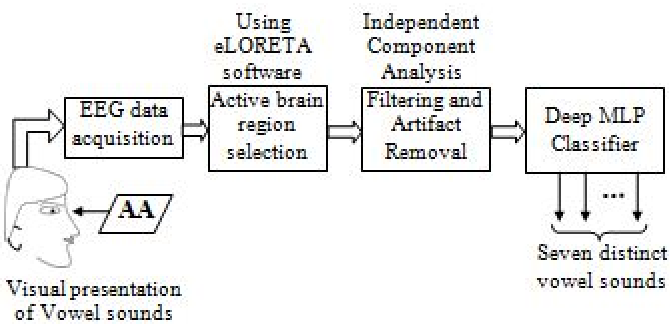

Fig. 1. Block diagram of the entire experimental framework

III. PROPOSED ARCHITECTURE OF CLASSIFIER

This section illustrates the architecture the proposed deep based classifier module. The diagrammatic representation of this classifier has been depicted in Fig.2.

The first layer represents the input layer consisting of $X = (x_1, ...x_i, ...x_d)^T$ flattened EEG feature vectors of vowel sound imageries where d denotes the dimension of the vector. The next four layers comprises of four hidden layers consisting of p, q, r, t number of neurons in the first, second, third and fourth layers respectively. The weights assigned for the first hidden layer are depicted as $W^{H_1} = (w_{11}^{H_1}, ..., w_{kl}^{H_1}, ...w_{pd}^{H_1})^T$, $1 \leqslant l \leqslant d$, $1 \leqslant k \leqslant p$. Similarly the weights assigned to the rest of the hidden layers are denoted as $W^{H_2} = (w_{11}^{H_2}, ..., w_{mk}^{H_2}, ...w_{qp}^{H_2})^T$, $W^{H_3} = (w_{11}^{H_3}, ..., w_{nm}^{H_3}, ...w_{rq}^{H_3})^T$, $W^{H_4} = (w_{11}^{H_4}, ..., w_{sn}^{H_4}, ...w_{tr}^{H_4})^T$ respectively. The output of the four hidden layers are represented as $H^1 = (h_1, ..., h_k, ...h_p)^T$, $H^2 = (h_1, ..., h_m, ...h_q)^T$, $H^3 = (h_1, ..., h_n, ...h_r)^T$, $H^4 = (h_1, ..., h_s, ...h_t)^T$ respectively. $W^Y = (w_{11}^Y, ..., w_{os}^Y, ...w_{zt}^Y)^T$,$1 \leqslant o \leqslant z$, are the weights assigned for the output layer, where z represents the class labels. The output vector is represented as $Y = (y_1, ...y_o, ...y_z)^T$. The weights in this network are trained using an adaptive learning rate optimization methodology called Adam [11] optimizer. The technique is easy to implement, utilizes less memory capacity and is computationally efficient in handling data comprising of noisy and sparse gradients.

The utilization of ReLU (Rectified Linear Unit) [12] as an activation function for introducing non-linearity in the hidden layer of neurons is a general convention. This function has the capability to boost up the model convergence [13] and helps to introduce sparsity amongst the hidden neurons. This function is represented by equation (1)

$$\sigma_{ReLU}(u) = max(0, u) = \begin{cases} u_i, & \text{if } u_i \geqslant 0 \\ 0, & \text{if } u_i < 0 \end{cases} \tag{1}$$

However, we utilize a new activation function Swish in our proposed architecture [8] instead of the classical ReLU since it suffers from certain drawbacks leading to over fitting issues. This is a hybrid function formed by the coalescence of Sigmoid [14] and input components. The major characteristics of this function involves better information propagation during training and effective management of vanishing gradient issue. Moreover, this function provides low computational complexity and have shown to outperform all the primitive techniques by a large margin [8]. This function is depicted by equation (2)

$$\sigma_{swish}(u) = u \times sigmoid(u) = \frac{u}{1 + e^{-u}} \tag{2}$$

In the final layer, the probabilities are computed after passing through the Softmax [15] activation function which provides the probability of each class label. This function maps the input within the range of $[0, 1]$ and thus the combination of all the Softmax values add up-to unity. This function is depicted by equation (3)

$$\sigma_{softmax}(u_i) = \frac{e^{u_i}}{\sum_k e^{u_k}} \tag{3}$$

IV. EXPERIMENT AND RESULTS

A. Experimental setup

A 21-channel Nihon-Kohden EEG equipment has been utilized to perform this experiment as illustrated in Fig.3. All the available channels have been used to acquire data to explore the consequences and results of this experiment. The FP_z channel is regarded as the ground and the earlobe channels A_1 and A_2 are used as reference electrodes. The EEG signals are procured from frontal $(F_3, F_4, F_7, F_8, F_z)$, prefrontal (FP_1, FP_2), parietal (P_3, P_4, P_z), occipital (O_1, O_2), motor cortex (C_3, C_4, C_z) and temporal $(T_1, T_2, T_3, T_4, T_5, T_6)$ electrodes.

The structure of stimuli is depicted in Fig. 4. 12 healthy/normal subjects volunteered for this experiment. They have been instructed to sit in a comfortable position with their arms lying on the armrest to alleviate the effects of various

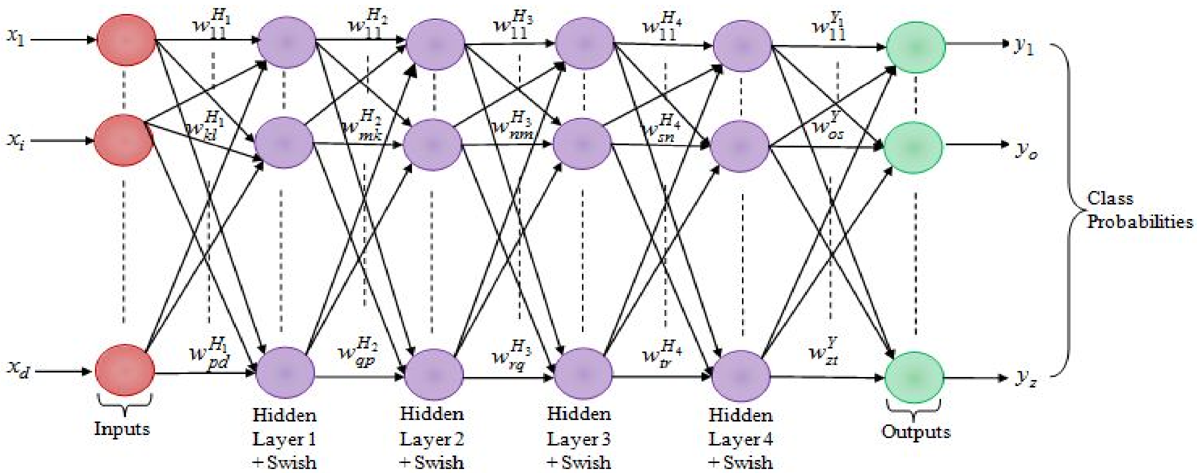

Fig. 2. Architecture of Proposed Classifier

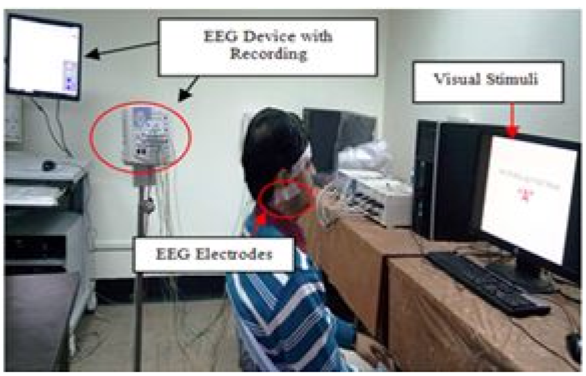

Fig. 3. Experimental Setup

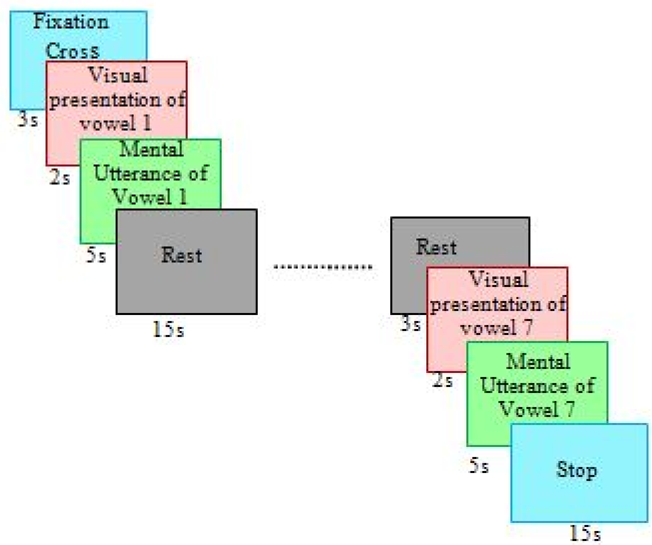

Fig. 4. Structure of stimuli used for the experiment

artifacts. The experiment comprises 5 sessions, where each session includes 5 trials. In each trial, the subject receives visual stimulus of 5 seconds duration containing instruction to mentally utter the vowel sound presented on the computer screen. A sufficient time-gap of 15 seconds is maintained between 2 consecutive presentation to avoid residual effect of the previous stimulus. Consequently, for twelve subjects 12 x 7 vowel sounds x 5 session/stimulus x 5 trials/sessions = 2100 training instances are generated for the said purpose. Fig.4 illustrates one sequential instance of the stimulus presented to each subject.

B. Experiment 1: Source localization by utilization of eLORETA

In this experiment, the highly active brain regions are selected using exact low resolution brain electromagnetic topographic analysis (eLORETA) technique [16], [17]. EEG data has been collected from the scalp of the subject for 5 seconds (= 5000 milliseconds) duration when the he/she is involved in mental utterance of vowel sounds. eLORETA software segments this duration into a fixed 640 time frames. Thus, topographic solutions are provided during a period of 7.81 milliseconds for each time frame. Fig.5 illustrates the eLORETA solution of vowel sound imagery. It is evident from the figure that the occipital region is highly active for the first four time-frames, representing the visual perception of vowel sound stimuli for approximately 32 milliseconds duration. The parietal and pre-frontal regions are highly active bilaterally, for the remaining time-frames, which signify the involvement of vowel sound imagination for the mind-driven type-writer experiment.

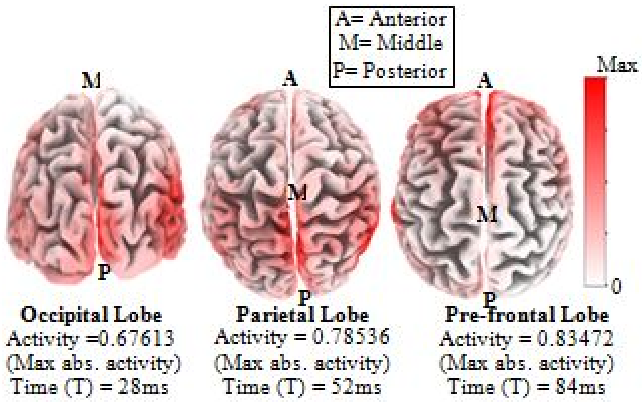

Fig. 5. 3D surface plot for vowel sound imagery

C. Experiment 2: Pre-processing and Elimination of Artifact

This experiment aims at removing physiological artifacts [18] such as eye blinking, respiration, muscles artifacts etc. In this experiment, an Infinite Impulse Response (IIR) filter has been selected for the present study instead of a Finite Impulse Response (FIR) design as it utilizes a meagre number of filter coefficients for a particular filter order. This methodology is implemented by utilizing an Elliptical Band pass filter of order 10 with pass band frequency in the range of 3 to 13 Hz. The Elliptical Band pass filter has been chosen for its better roll-off and stop band attenuation as compared to other standard filters [19].

Next, in this experiment the filtered EEG signals are further evaluated using Independent Component Analysis to restore 19 independent components of the 19 EEG signals. Out of the 19 independent channels, 7 are chosen for further analysis as depicted in Fig.6

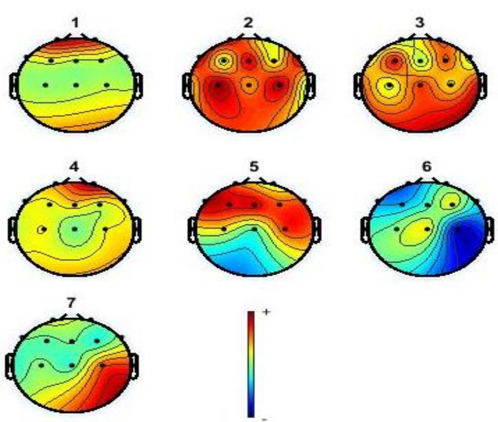

Fig. 6. Scalp maps of selected channels

V. Classifier Performance Evaluation

The proposed classifier used for the current application comprises of 8192, 1024, 256, 64 neurons in the fist, second, third and fourth hidden layers respectively. The classifier model has been trained using the Adam optimizer with a learning rate of 0.0001 for 10 epochs.

A. Relative Performance Analysis

The performance of the proposed classifier has been evaluated on the basis of three metrics- Classification Accuracy (CA), Sensitivity and Specificity which are defined as by equations (4),(5),(6).

$$CA = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{5}$$

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

where, TP,TN,FP,FN indicates the number of true positives, true negatives, false positives and false negatives respectively.

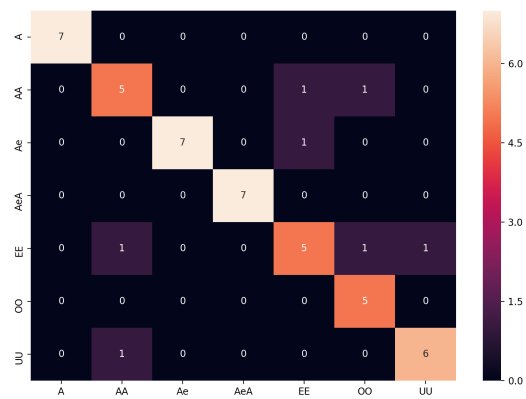

Fig. 7. Confusion Matrix for a single trial

TABLE I
AVERAGE (STANDARD DEVIATION) OF PERFORMANCE METRICS FOR THE PROPOSED CLASSIFIER MODEL

Vowel Sounds	CA (%)	Sensitivity	Specificity
A	93.15	0.95	0.92
AA	92.38	0.88	0.97
AeA	94.60	0.90	0.98
EE	93.89	0.97	0.90
O	92.97	0.89	0.94
Ae	96.15	0.94	0.97
UU	93.79	0.95	0.92

Figure 7 depicts confusion matrix for the classifier for a single trial. Table I illustrates average classification metrics for the seven vowel imagery classes. It is apparent from these two tables, that the proposed classifier is successful in classifying the seven class labels with high precision level.

TABLE II
COMPARATIVE STUDY OF CLASSIFICATION ACCURACY WITH RESPECT TO VARYING ACTIVATION FUNCTIONS FOR VOWEL SOUND AE

Activation Function	CA (%)
Sigmoid [14]	85.96
Hyperbolic Tangent [20]	88.63
ReLU [12]	90.12
Leaky ReLU [21]	91.27
ELU [22]	91.98
SELU [23]	92.56
Swish [8]	96.15

B. Comparative Analysis on the basis of Activation Functions

To further analyse the performance of the proposed classifier, the classification accuracy of the model is evaluated by varying the activation function in the hidden layers. This comparison is depicted in Table II. It can be clearly inferred from the table that the classification accuracy of the proposed Swish induced network outperforms the same model comprising of classical activation functions by a significant margin.

VI. CONCLUSION

This work provides an innovative approach to model and design a mind driven type writing system by utilizing the EEG signals acquired from subjects during the mental enunciation of seven vowel sounds. Experimental analysis performed by eLOERTA software infers the participation of occipital, prefrontal and parietal lobes for this cognitive activity. The acquired signals are pre-processed then transmitted to the proposed classifier model which is able to successfully classify the seven class labels but also yields a high classification accuracy. Moreover, a codebook has been developed to encode the consonants by utilizing two vowel sounds separated by a space. Thus, the proposed methodology involving the classification and decoding scheme can jointly contribute towards the complete design of mind controlled type writer to aid patients suffering from various communication related impairments. Additionally, this technique can be utilized in military applications under precarious circumstances where non-verbal communication is most desirable. Besides, the proposed approach can also be employed in various gaming applications to elate the experience of the users.

ACKNOWLEDGMENT

The authors are immensely thankful for the financial support provided Ministry of Human Resource Development (MHRD) sponsored JU-RUSA 2.0 project bestowed upon Jadavpur University, Kolkata - 700032, India.

REFERENCES

[1] L. C. Wijesekera and P. N. Leigh, "Amyotrophic lateral sclerosis," *Orphanet journal of rare diseases*, vol. 4, no. 1, p. 3, 2009.

[2] C. S. DaSalla, H. Kambara, M. Sato, and Y. Koike, "Single-trial classification of vowel speech imagery using common spatial patterns," *Neural networks*, vol. 22, no. 9, pp. 1334–1339, 2009.

[3] S. Iqbal, Y. U. Khan, and O. Farooq, "Eeg based classification of imagined vowel sounds," in *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*. IEEE, 2015, pp. 1591–1594.

[4] B. M. Idrees and O. Farooq, "Vowel classification using wavelet decomposition during speech imagery," in *2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN)*. IEEE, 2016, pp. 636–640.

[5] N. Hashim, A. Ali, and W.-N. Mohd-Isa, "Word-based classification of imagined speech using eeg," in *International Conference on Computational Science and Technology*. Springer, 2017, pp. 195–204.

[6] A. Kachenoura, L. Albera, L. Senhadji, and P. Comon, "Ica: a potential tool for bci systems," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 57–68, 2007.

[7] S. Wan, Y. Liang, Y. Zhang, and M. Guizani, "Deep multi-layer perceptron classifier for behavior analysis to estimate parkinson's disease severity using smartphones," *IEEE Access*, vol. 6, pp. 36 825–36 833, 2018.

[8] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," 2017.

[9] R. Pascanu, T. Mikolov, and Y. Bengio, "Understanding the exploding gradient problem," *CoRR, abs/1211.5063*, vol. 2, p. 417, 2012.

[10] R. D. Pascual-Marqui, D. Lehmann, T. Koenig, K. Kochi, M. C. Merlo, D. Hell, and M. Koukkou, "Low resolution brain electromagnetic tomography (loreta) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia," *Psychiatry Research: Neuroimaging*, vol. 90, no. 3, pp. 169–179, 1999.

[11] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014.

[12] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010, pp. 807–814.

[13] M. D. Zeiler, M. Ranzato, R. Monga, M. Mao, K. Yang, Q. V. Le, P. Nguyen, A. Senior, V. Vanhoucke, J. Dean *et al.*, "On rectified linear units for speech processing," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 3517–3521.

[14] J. Turian, J. Bergstra, and Y. Bengio, "Quadratic features and deep architectures for chunking," in *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers*, 2009, pp. 245–248.

[15] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[16] M. Hata, H. Kazui, T. Tanaka, R. Ishii, L. Canuet, R. D. Pascual-Marqui, Y. Aoki, S. Ikeda, H. Kanemoto, K. Yoshiyama *et al.*, "Functional connectivity assessed by resting state eeg correlates with cognitive decline of alzheimer's disease–an eloreta study," *Clinical Neurophysiology*, vol. 127, no. 2, pp. 1269–1278, 2016.

[17] L. Canuet, R. Ishii, R. D. Pascual-Marqui, M. Iwase, R. Kurimoto, Y. Aoki, S. Ikeda, H. Takahashi, T. Nakahachi, and M. Takeda, "Resting-state eeg source localization and functional connectivity in schizophrenia-like psychosis of epilepsy," *PloS one*, vol. 6, no. 11, 2011.

[18] P. Manoilov, "Eeg eye-blinking artefacts power spectrum analysis," in *Proceedings of the International Conference on Computer Systems and Technology, Veliko Tarnovo, Bulgaria*, 2006, pp. 15–16.

[19] A. Widmann, E. Schröger, and B. Maess, "Digital filter design for electrophysiological data–a practical approach," *Journal of neuroscience methods*, vol. 250, pp. 34–46, 2015.

[20] B. Karlik and A. V. Olgac, "Performance analysis of various activation functions in generalized mlp architectures of neural networks," *International Journal of Artificial Intelligence and Expert Systems*, vol. 1, no. 4, pp. 111–122, 2011.

[21] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. icml*, vol. 30, no. 1, 2013, p. 3.

[22] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," *arXiv preprint arXiv:1511.07289*, 2015.

[23] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-normalizing neural networks," in *Advances in neural information processing systems*, 2017, pp. 971–980.