ORIGINAL ARTICLE

Improving the Classification of Real-World SSVEP Data in Brain-Computer Interface Speller Systems Using Deep Convolutional Neural Networks

Elham Rostami, Farnaz Ghassemi *, Zahra Tabanfar

Department of Biomedical Engineering, Amirkabir University of Technology, Tehran, Iran

*Corresponding Author: Farnaz Ghassemi
Email: ghassemi@aut.ac.ir

Received: 30 August 2021 / Accepted: 03 November 2021

Abstract

Purpose: Brain-Computer Interface (BCI) Speller systems help people with mobility impairments improve their cognitive and physical abilities. Steady-State Visual Evoked Potential (SSVEP) signals have been used to build high-speed BCI speller systems. SSVEP signals are a subtype of Visual Evoked Potential (VEP), a form of co-frequency, and the harmonics response elicited by a specific frequency stimulus. Noise and artifacts are critical issues for target detection in SSVEP-based BCI systems.

Materials and Methods: Thus, it is essential to provide target detection techniques that operate well in the presence of noises. One solution for overcoming the noise impact is to employ approaches that automatically extract the appropriate features for target detection from the training data. Deep Convolutional Neural Network (DCNN) was utilized in this study to automatically extract features from SSVEP data in noisy conditions. Moreover, the BETA database, which contains SSVEP data from 70 individuals collected outside of the electromagnetic shielding room, was used. In this regard, a suitable DCNN structure for target stimulus frequency identification was first designed. The network was pre-trained with part of the data from the BETA database. Finally, at the single-subject level, this pre-trained network was retrained and evaluated.

Results: The results showed that after retraining, the accuracy and Information Transfer Rate (ITR) increased (p-value < 0.01) for all participants.

Conclusion:.The enhancement in accuracy and ITR are 25.72% and 43.10 bpm, respectively.

Keywords: Brain Computer Interface Speller; Steady State Visual Evoked Potentials; Deep Convolutional Neural Network; Electroencephalogram.

DOI: https://doi.org/10.18502/fbt.v9i4.10382

1. Introduction

Brain-Computer Interface (BCI) speller based on Steady-State Visual Evoked Potentials (SSVEP) is a subtype of BCI spellers used for rehabilitation and assisting people with mobility impairments [1, 2]. In SSVEP-based BCI speller systems, a series of flickering visual stimuli at specific frequencies are shown to users, and the resultant evoked SSVEP is detected to determine the user's command [3-6].

One of the challenges in SSVEP research is to reduce the destructive effect of noise and artifacts. There are several reasons for the appearance of noise and artifacts in SSVEP signals (e.g., participants' movement, electrode displacement, poor electrode connection, blinking, eye movement, and Electrocardiography (ECG) and Electromyography (EMG) effect) [7]. In addition, signal recording environments are often contaminated by the effects of high-current cables, Wi-Fi, wireless signals, and other electrical equipment [8]. Therefore, there is a need for methods that provide good results in the mentioned environments.

There are several methods for identifying the target frequency in SSVEP signals. For example, the Canonical Correlation Analysis (CCA) [9], the Filter Bank Canonical Correlation Analysis (FBCCA) [10], the Task-Related Component Analysis (TRCA) [11] and the Canonical Correlation Analysis of Task Related Components (CCAoTRC) [12] can be mentioned. These methods are traditional hand-crafted feature extraction methods [13]. Also, Deep Convolutional Neural Networks (DCNNs) are used for target detection in BCI-Speller systems [13, 14]. It should be noted that using the DCNNs has surpassed traditional hand-crafted feature extraction methods because the feature extraction process is done automatically in deep layers of the network [15]. Automatic feature extraction is an important ability in networks. Podmore *et al.*, for example, introduced a DCNN called PodNet to decode high-class SSVEP-based BCI systems (40 targets). This network consists of 5 subunits called Pods which were designed in the same way. Details of PodNet implementation are provided in the Materials and Methods section. Podmore *et al.* used the benchmark database in their research. The results showed that the accuracy and Information Transfer Rate (ITR) in this method are 77% and 101 bpm for a 2-second window length [16]. Guney *et al.* also introduced a Deep Neural Network (DNN) in their research to improve ITR in multi-class classifications [17]. The proposed network structure contained convolutional, fully connected, Rectified Linear Unit (RELU), and drop-out layers. They used the benchmark and BETA databases to evaluate their method. The results showed that considering 0.4 seconds of stimulation, the ITR for the benchmark and BETA databases was 265.23 bpm and 196.59 bpm, respectively. Safari *et al.* developed a DCNN with single-channel Electroencephalogram (EEG) as input [18]. There are Convolution, Batch Normalization, Average Pooling, Drop Out, fully connected and softmax layers in the network structure. This study [18] showed that the average accuracy and ITR were 74.30% and 57.51 bpm, respectively. Also, channel O1 had the best performance among other channels.

Despite the benefits that come along with DCNNs, the network's performance for detecting stimulus frequency decreases as noise and artifacts increase [16]. Therefore, DCNNs need to be improved for use in noisy data. This study aims to develop a deep learning-based methodology for identifying target frequency in SSVEP-based BCI spellers when the data is acquired in the real world (outside the laboratory condition). To accomplish this goal, DCNN's capability, i.e., automatic feature extraction has been used. The structure of PodNet [16] has been taken as the basic structure of this research and the BETA database [19] is used as the data. This research provides a strategy to make the PodNet network more resistant when data is recorded outside the laboratory, where it is typically contaminated by ambient noise.

2. Materials and Methods

2.1. Database

The BETA database containing EEG signals from 70 participants (42 males, age: 9 to 64 years old) was used [19]. This database is collected outside the laboratory and without any electromagnetic shield. The BETA database contains real-world data properties due to its out-of-laboratory recording. There are 40 stimulus frequencies in this database, set from 8 Hz to 15.8 Hz with a 0.2 Hz interval. The task is designed in 4 blocks with 40 trials corresponding to 40 targets in each block. The stimulation period is 2 seconds for the first to the fifteenth participant and 3 seconds for the other participants. The 3-second epochs were shortened to 2 seconds to equalize trial length.

Also, like most SSVEP studies, only ten channels of occipital and parietal lobes (O_1, O_z, O_2, PO_z, PO_3, PO_4, PO_5, PO_6, PO_7, PO_8) were selected for processing.

At this stage, 9 participants with low accuracy percentages in SSVEP target frequency detection using the CCA approach were found. Because the CCA approach has a low noise resistance [9], the proposed method was tested on these 9 participants (4, 11, 17, 26, 31, 41, 55, 64, 69).

2.2. PodNet Structure

The PodNet is a DCNN that extracts stimulus-related features embedded in EEG signals for SSVEP target frequency detection. This network was proposed for the first time by Podmore *et al.* [16]. Each PodNet is made up of some structural units termed Pods. Convolutional, drop-out (50%), batch normalization, Rectified Linear Unit, and max-pooling layers are included in these Pods. In this research, due to the length of the trials (2 seconds), this network was created in 4 separate Pods (with a few modifications) using the Keras library (Tensorflow Backend). In addition, following the last Pod, there is a fully connected layer with softmax operation. Figure 1 depicts the network's detailed structure.

2.3 . Processing Methods for Evaluation

In the present study, the EEGNet [14], the DeepConvNet [20], and some traditional methods such as CCA, FBCCA, and TRCA were utilized to compare the results of the PodNet network. More details about the structure of the two networks and traditional methods can be found in the mentioned references.

3. Results

The PodNet was trained on 61 participants (except the before mentioned nine participants). 43 out of 61 participants were randomly selected for training, 9 participants for

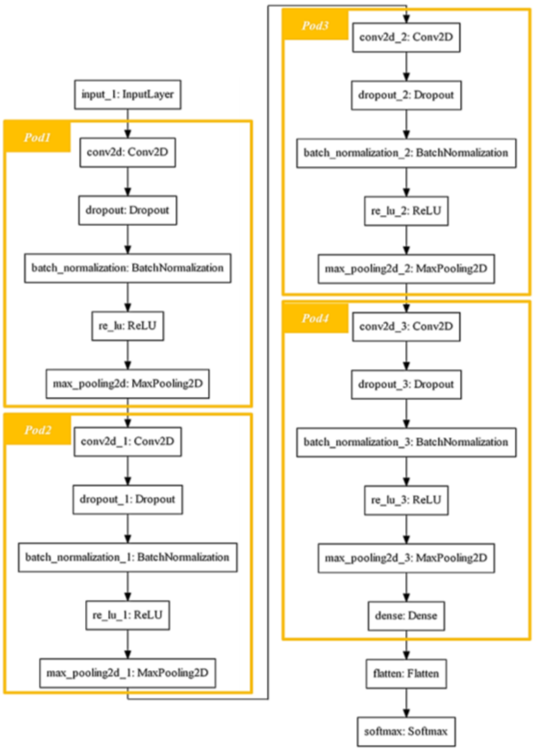

Figure 1. The details of the PodNet structure

validation, and 9 for testing. The accuracy and ITR acquired from this training are 72.29% and 90.17 bpm for validation data, respectively, and 73.19% and 91.89 bpm for testing data. This model is then used separately in two different conditions for each of the 9 selected participants. The first condition is without model retraining (only in the test block) and the second condition is with model retraining. Two blocks were selected for training in each participant, one block for validation and one block for testing. The testing block is the same between the two conditions. The summary statistics of these approaches for the testing block are shown in Table 1. The results demonstrate that after retraining the PodNet, the accuracy and ITR on participants with low accuracy improved (Wilcoxon signed-rank statistical test, P-value = 0.0039). In Figure 2, the classification accuracy (%) in the first and second conditions are presented for 9

Table 1. Summary Statistics of accuracy (%) and ITR (bpm) (in test block) for two methods: 1- First condition (without model retraining), 2- Second condition (with model retraining)

Methods	Evaluation Criteria	Mean and Standard Deviation	Median and Interquartile Range	Maximum	Minimum
First condition (without model retraining)	Accuracy ITR	48.44 ± 25.93 53.56 ± 41.06	47.50 ± 43.75 46.66 ± 69.61	85 117.58	12.5 4.61
Second condition (with model retraining)	Accuracy ITR	74.16 ± 16.00 96.67 ± 32.52	72.50 ± 23.75 90.60 ± 50.08	95 143.14	45 42.66

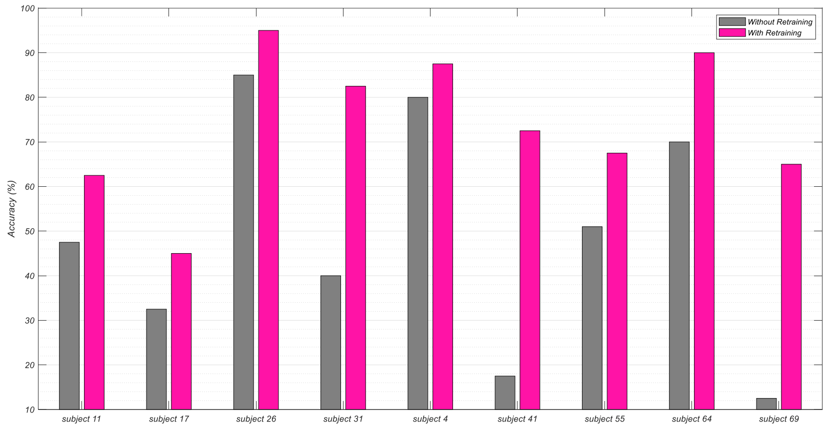

Figure 2. Accuracy (%) in participants with low accuracy for both conditions of with and without PodNet retraining

participants with low accuracy. According to the findings, the network appears to perform efficiently in real-world data by learning the specific features of the data in its deep layers.

Podmore in [16] used common inter-subject information for single-subjects with low accuracy. In other words, the approach of Podmore *et al.* was Single-Subject Optimization, while in the present study, this approach has been used to reduce the noise effect.

Figure 3 depicts the output signal's Power Spectral Density for all 100 filters in the convolution layer (Participant No. 4, Class 1 and Pod 1). After inspecting the diagrams, it was discovered that 26 of the 100 output signals contain high-frequency components in their power spectrum (noisy components). The finding implies that the network can learn to decompose noisy components in its convolution filter of the first layer and potentially reduce the impact of noise on subsequent phases. As an example, in Figure 4, two signals were selected from 100 output signals, one of which contains high-frequency components in its power spectrum.

In the following, the EEGNet and The DeepConvNet networks were used for comparison. Following primary training on 61 individuals, the two networks were used in the first and second conditions for 9 participants with low accuracy. In both networks, the accuracy and ITR results in the second condition were significantly better than the first condition (Wilcoxon signed-rank statistical test, p-value < 0.05 for both accuracy and ITR). For comparison, traditional approaches (CCA, FBCCA, and TRCA) were also applied. Table 2 compares the performance of CCA, FBCCA, TRCA, EEGNet, and DeepConvNet techniques to that of PodNet.

The Friedman statistical test was used to compare the outcomes of the PodNet (second condition) in two categories. The first category was a comparison with traditional methods (the CCA, the FBCCA, and the TRCA) and the second category was a comparison with DCNNs (the EEGNet and the DeepConvNet (in the second condition)). The test findings demonstrate that in both categories, at least two of the approaches have a significant difference (first category: p-value = 3.92×10^{-4}, degree

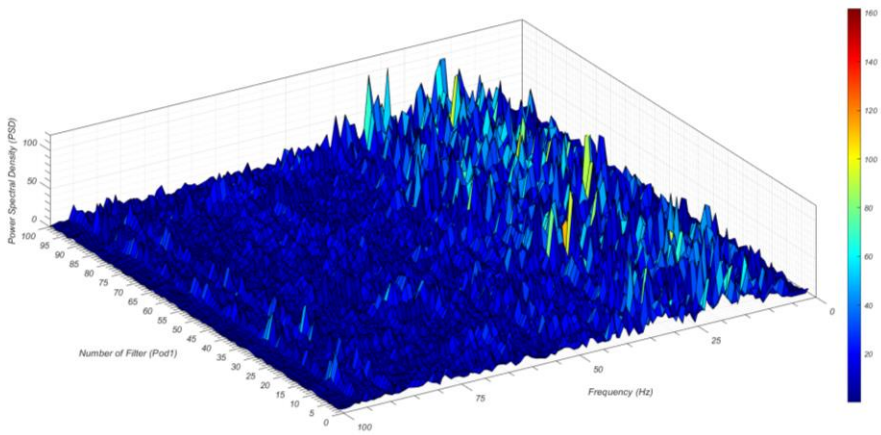

Figure 3. the Power Spectral Density of the output signal for all 100 filters in the convolution layer (Participant No. 4, Class 1 and Pod 1)

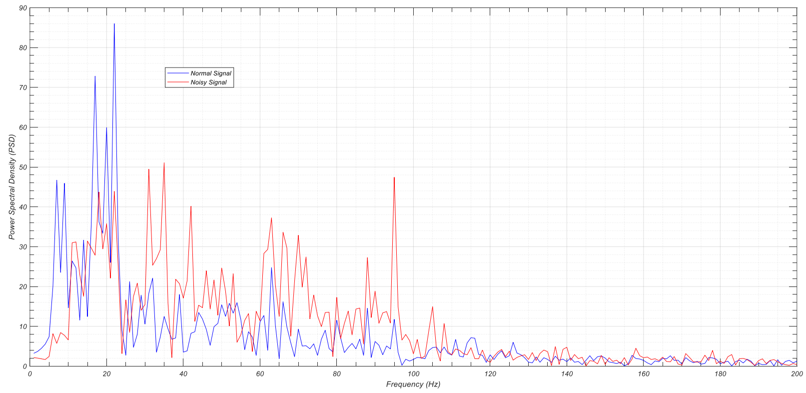

Figure 4. Power Spectral Density of the two output signals of the convolution filter in Pod 1

of freedom (df) = 3 and Chi-sq = 18.24, second category: p-value = 0.0015, degree of freedom (df) = 2 and Chi-sq = 13.00). The Tukey-Kramer post-hoc test was also employed to identify significant differences across groups.

The post-hoc test findings (p-values) for the two categories are shown in Table 3 and Table 4, respectively. According to Table 3 and Table 4, PodNet in the second condition significantly outperforms the CCA, the FBCCA, the TRCA, the EEGNet (in the second condition), and the DeepConvNet (in the second condition).

4. Discussion

Overcoming the noise effect is one of the challenges of BCI systems. In this research, the PodNet structure was utilized to detect the target stimulus frequency of SSVEP signals in the BETA database. The fundamental goal of this research is to increase the PodNet performance by reducing the effect of noise. To this end, pre-trained PodNet (on a substantial portion of the BETA database) was retrained on the part of single-subject data. The number of the final single-subjects is nine, and they were chosen based on the CCA method's accuracy. According to the results, PodNet performance was increased by retraining the network on the part of a single-subject from the BETA database. Retraining the pre-trained EEGNet and the pre-trained DeepConvNet on single-subject data enhanced accuracy as well.

According to the findings, when PodNet learns proprietary features, the network's ability to identify stimulus frequencies in participants with low CCA accuracy improves. In other words, by retraining the network on a subset of the

Table 3. The results of the post-hoc test (Comparison with traditional methods, Accuracy)

Group 1	Group 2	p-value
The PodNet (Second condition)	The CCA method	0.0003***
	The FBCCA method	0.0064**
	The TRCA method	0.0421*

* Indicates p-value < 0.05, ** Indicates p-value < 0.01 and *** Indicates p-value < 0.001.

Table 2. Summary Statistics of the accuracy (%) and ITR (bpm) (in test block) for six different methods

Evaluation Criteria	Methods	Mean and Standard Deviation	Median and Interquartile Range	Maximum
Accuracy	The CCA method	9.72 ± 5.61	8.12 ± 7.03	20.62
	The FBCCA method	10.69 ± 6.18	10.62 ± 6.57	23.12
	The TRCA method	25.21 ± 24.24	13.75 ± 34.53	68.75
	The EEGNet (Second condition)	56.11 ± 20.00	57.50 ± 30.62	90
	The DeepConvNet (Second condition)	64.72 ± 20.02	60.00 ± 28.75	90
	The PodNet (Second condition)	74.16 ± 16.00	72.50 ± 23.75	95
ITR	The CCA method	2.71 ± 0.64	1.78 ± 1.23	11.76
	The FBCCA method	3.33± 0.86	3.28 ± 1.02	14.35
	The TRCA method	16.63 ± 15.56	5.57 ± 27.95	83.23
	The EEGNet (Second condition)	60.39 ± 11.15	62.76 ± 22.98	129.73
	The DeepConvNet (Second condition)	75.62± 11.17	67.10 ± 20.72	129.73
	The PodNet (Second condition)	93.95 ± 7.43	90.60 ± 15.03	143.14

Table 4. The results of the post-hoc test (Comparison with DCNNs, Accuracy)

Group 1	Group 2	p-value
The PodNet (Second condition)	The EEGNet (Second condition)	0.0332*
	The DeepConvNet (Second condition)	0.0014**

* Indicates p-value < 0.05 and ** Indicates p-value < 0.01.

data (training blocks), the noise effect for other subsets of the same data can be reduced (validation and test blocks). Deep Neural Networks automate the feature extraction process. When the network is retrained on the portion of data containing the noise, it can recognize noise as a common component in all cases (train, validation, and test data). As a result, it is possible to exclude that component in algorithms aiming at detecting stimulus frequency. The effect of noise on the system is thereby decreased, and this new network will be more resistant to noise and artifacts.

Furthermore, the results reveal that the accuracy of the PodNet (second condition) is significantly better than the CCA, FBCCA, and TRCA as a traditional hand-crafted feature extraction method (p-value < 0.05). The results show that PodNet, as a DCNN, performs better than traditional approaches in SSVEP signal processing, and this result is consistent with [15, 18].

References

1- Zahra Tabanfar, Farnaz Ghassemi, and Mohammad Hassan Moradi, "Estimating brain periodic sources activities in steady-state visual evoked potential using local fourier independent component analysis." *Biomedical Signal Processing and Control*, Vol. 71p. 103162, 2022/01/01/ (2022).

2- Shakib Sara Sepehri and Ghassemi Farnaz, "Investigating the Effect of Stimulus Type on Electroencephalogram Signal in a Brain-Computer Interface System with Interaction Error." *Frontiers in Biomedical Technologies*, Vol. 9 (No. 2), 03/04 (2022).

3- Kang-min Choi, Seonghun Park, and Chang-Hwan Im, "Comparison of Visual Stimuli for Steady-State Visual Evoked Potential-Based Brain-Computer Interfaces in Virtual Reality Environment in terms of Classification Accuracy and Visual Comfort." *Computational Intelligence and Neuroscience*, Vol. 2019p. 9680697, 2019/07/01 (2019).

4- Aya Rezeika, Mihaly Benda, Piotr Stawicki, Felix Gembler, Abdul Saboor, and Ivan Volosyak, "Brain–Computer Interface Spellers: A Review." *Brain Sciences*, Vol. 8 (No. 4), (2018).

5- Aung Aung Phyo Wai *et al.*, "Towards a Fast Steady-State Visual Evoked Potentials (SSVEP) Brain-Computer Interface (BCI)." *ArXiv*, Vol. abs/2002.01171(2020).

6- Elham Rostami and Farnaz Ghassemi, "Investigation of the Effect of Attention in BCI systems based on Steady_State Visually Evoked Potential." Presented at the *Conference: 8th Basic and Clinical Neuroscience Congress*, Iran University of Medical., (1398).

7- No-Sang Kwak, Klaus-Robert Müller, and Seong-Whan Lee, "A convolutional neural network for steady state visual evoked potential classification under ambulatory environment." *PLOS ONE*, Vol. 12 (No. 2), p. e0172578, (2017).

8- Eisuke Hanada and Takato Kudou, "Electromagnetic noise in the clinical environment." *Peoceedings of ISMICT2009*, 02/25 (2009).

9- Zhonglin Lin, Changshui Zhang, Wei Wu, and Xiaorong Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs." *IEEE transactions on biomedical engineering*, Vol. 53 (No. 12), pp. 2610-14, (2006).

10- Xiaogang Chen, Yijun Wang, Shangkai Gao, Tzyy-Ping Jung, and Xiaorong Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface." *Journal of Neural Engineering*, Vol. 12 (No. 4), p. 046008, 2015/06/02 (2015).

11- M. Nakanishi, Y. Wang, X. Chen, Y. T. Wang, X. Gao, and T. P. Jung, "Enhancing Detection of SSVEPs for a High-Speed Brain Speller Using Task-Related Component Analysis." (in eng), *IEEE transactions on bio-medical engineering*, Vol. 65 (No. 1), pp. 104-12, Jan (2018).

12- Elham Rostami, Farnaz Ghassemi, and Zahra Tabanfar, "Canonical Correlation Analysis of Task Related Components as a noise-resistant method in Brain-Computer Interface Speller Systems based on Steady-State Visual Evoked Potential." *Biomedical Signal Processing and Control*, Vol. 73p. 103449, 2022/03/01/ (2022).

13- Nik Aznan, Stephen Bonner, Jason Connolly, Noura Al Moubayed, and Toby Breckon, On the Classification of SSVEP-Based Dry-EEG Signals via Convolutional Neural Networks. (2018).

14- Nicholas Waytowich *et al.*, "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials." *Journal of Neural Engineering*, Vol. 15 (No. 6), p. 066031, 2018/10/31 (2018).

15- Nouri Ali and Azizi Kyanoosh, "Introducing a Convolutional Neural Network and Visualization of its Filters for Classification of EEG Signal for SSVEP Task." *Frontiers in Biomedical Technologies*, Vol. 7 (No. 3), 09/30 (2020).

16- Joshua Podmore, Toby Breckon, Nik Aznan, and Jason Connolly, "On the Relative Contribution of Deep Convolutional Neural Networks for SSVEP-based Bio-Signal Decoding in BCI Speller Applications." *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. PPpp. 1-1, 03/13 (2019).

17- Osman Berke Guney, Muhtasham Oblokulov, and Huseyin Ozkan, "A Deep Neural Network for SSVEP-based Brain-Computer Interfaces." (in eng), *IEEE transactions on bio-medical engineering*, Vol. PP2021/09// (2021).

18- Saffari Farzad and Khadem Ali, "Single-Channel Selection for Detecting Steady-State Visual Evoked Potentials in a Brain-Computer Interface Speller." *Frontiers in Biomedical Technologies,* Vol. 8 (No. 3), 09/01 (2021).

19- Bingchuan Liu, Xiaoshan Huang, Yijun Wang, Xiaogang Chen, and Xiaorong Gao, BETA: A Large Benchmark Database Toward SSVEP-BCI Application. (2019).

20- Robin Tibor Schirrmeister *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization." *Human Brain Mapping,* https://doi.org/10.1002/hbm.23730 Vol. 38 (No. 11), pp. 5391-420, 2017/11/01 (2017).