

RESEARCH

Open Access



A deep learning-based brain-computer interaction system for speech and motor impairment

Nader A. Rahman Mohamed^{1*}

*Correspondence:
nader_mohamed@hotmail.com;
nader.shaaban@must.edu.eg

¹ Biomedical Engineering
Department, Faculty
of Engineering, Misr University
for Science and Technology
(MUST), Giza, Egypt

Abstract

Some people may experience accidents, strokes, or diseases that lead to both motor and speech disabilities, making it difficult to communicate with others. Those with paralysis face daily challenges in meeting their basic needs, particularly if they have difficulty speaking. Individuals with dysarthria, amyotrophic lateral sclerosis, and similar conditions may find it challenging to understand speech.

The proposed system for automatic recognition of daily basic needs aims to improve the quality of life for individuals suffering from dysarthria and quadriplegic paralysis. The system achieves this by recognizing and analyzing brain signals and converting them to either audible voice commands or texts that can be sent to a healthcare provider's mobile phone based on the system settings.

The proposed system uses a convolutional neural network (CNN) model to detect event-related potentials (ERPs) within the EEG signal to select one of six basic daily needs while displaying their images randomly. Ten volunteers participated in this study, contributing to the creation of the dataset used for training, testing, and validation. The proposed approach achieved an accuracy of 78.41%.

Keywords: Electroencephalogram (EEG), Brain-Computer Interface (BCI), Event Related Potential (ERP), Convolutional Neural Network (CNN)

Introduction

A Brain-Computer Interface (BCI) system utilizes EEG signals to make decisions and take actions [1]. EEG signals typically have sinusoidal patterns with varying frequency ranges. Their amplitudes range from 0.001 to 0.01 mV and their bandwidth from 0.5 to 40 Hz. EEG activity can be recorded within seconds after a stimulus is presented, making it a highly time-sensitive measure.

Event-Related Potentials (ERPs) are brief changes in EEG signals that are time-locked to specific sensory, cognitive, or motor events, such as the presentation of a visual stimulus or the onset of a movement. ERPs are characterized by positive and negative peaks, each with a distinct latency and scalp distribution, reflecting different stages of information processing in the brain. The most commonly studied ERP

components include P1, N1, P2, N2, and P3, among others [2]. These components are classified and named based on their latencies and amplitudes [3, 4].

1. P50 Wave “P1 Wave”:

It is a major positive peak between 40 and 75 ms after presenting a stimulus.

2. N100 “N1 Wave”:

It is a major negative valley showing between 90 and 200 ms after presenting a stimulus.

3. P200 “P2 Wave”:

It is the major positive peak around 100-250 ms after presenting a stimulus.

4. N200 “N2 Wave”:

It is a major negative valley showing up around 200 ms after presenting a stimulus.

5. P300 “P3 Wave”:

It is the major positive peak around 250-400 ms after presenting a stimulus.

Figure 1 shows an enhanced and filtered EEG signal showing different Event Related Potentials (ERPs) waveforms, after presenting a stimulus.

The P300 component is believed to reflect the allocation of cognitive resources and the updating of attentional resources to the task-relevant stimulus. It has been extensively studied in cognitive psychology and neuroscience, and has been found to be sensitive to various factors, including stimulus probability, task difficulty, and

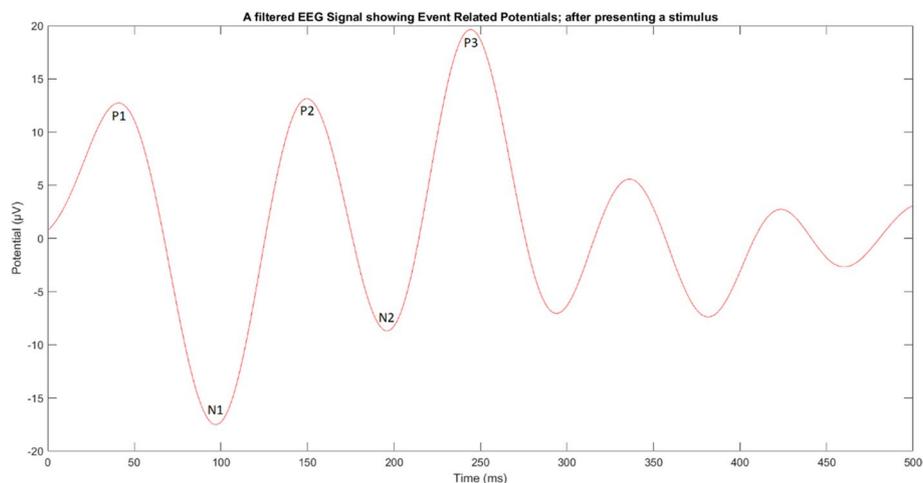


Fig. 1 A filtered EEG signal showing different Event Related Potentials (ERPs) waveforms after presenting a stimulus

stimulus salience. In addition, research has explored the clinical applications of P300 detection, particularly in the diagnosis and monitoring of neurological disorders [4].

For further information on electrode placement and positions on the scalp required for recording brain electrical activity (EEG), please refer to [5, 6].

Related work

Several research studies have focused on detecting the P300 signal for real-time applications. Researchers have used either existing datasets or created their own through experimental setups. The first proposed use of P300 was by Farwell and Donchin, who used a 6×6 matrix with 26 alphabet letters and 10 control commands. Their study involved four healthy volunteers, and the stepwise linear discriminant analysis algorithm was used to detect the P300 component [7].

Guan et al. found that the lower the probability of highlighting a character within the system matrix, the easier it is to detect the P300 component. They proposed a single-character random flashing spelling system based on the Farwell and Donchin system [8].

A. Rokotomamonj and V. Guigue won the BCI Competition III for the P300 speller paradigm using Dataset II, achieving an accuracy of 73.5% using a linear support vector machine [9]. Krusienski et al. compared the performance of various P300 detection algorithms and agreed that SWLDA is suitable for P300 detection [10].

Gerardo et al. proposed a method that uses empirical mode decomposition (EMD) and support vector machine (SVM) for detecting the P300 signal and achieved an accuracy of 53.4%, while Salvaris and Sepulveda examined the ease of usability of the spelling system based on a row-column random flashing configuration [11, 12]. Townsend et al. proposed an alphabet spelling system based on the checkerboard paradigm, achieving an accuracy of 77%, and Wang et al. developed a method to detect the P300 signal using Fisher distance [13, 14].

Cecotti et al. proposed a method based on convolutional neural network (CNN) to detect the P300 signal, and Haghghatpanah et al. proposed a detection method based on independent component analysis (ICA) and linear discriminant analysis (LDA) for P300 detection, achieving an accuracy of 65%. They also proposed a two-stage P300 detection algorithm that involved deionizing raw EEG data using wavelet transform, extracting features using independent component analysis, and using a neural network classifier with an accuracy of 71.5% [15–17].

Two methods for detecting the P300 signal were proposed by Wang et al. and Ravi et al. Wang's method used ICA, Fisher distance, and Wavelet Transform, while Ravi's method utilized LSSVM and five channels for EEG signal acquisition, achieving an accuracy of 71.5% [18, 19].

Several authors have proposed a hybrid BCI spelling system that combines P300 with least square support vector machine (SSVEP) [20–22]. Fernández-Rodríguez et al. focused on stimulus types that could enhance the performance of the BCI speller [23].

In addition, Ron-Angevin et al. demonstrated that a medium font size in the row-column random flashing configuration spelling system had the highest usability, and Li et al. introduced a classification algorithm based on an improved CNN to detect P300 signals [24, 25].

Overall, detecting the P300 component is a valuable application of ERPs that has been utilized in the development of BCIs for spelling and communication in individuals with severe motor impairments. However, to achieve accurate and efficient spelling with a P300-based system, several design factors need optimization, such as the stimulus presentation rate, size and layout of the matrix, and classification algorithm used to detect the P300. These methods vary in their accuracy, speed, and complexity [26]. Despite the challenges, P300-based BCIs are a promising area of research and development that can enhance communication and quality of life for with severe motor impairments.

Motivation and aim of the study

Although Event Related Potentials (ERPs) have been widely used in several applications addressing the development of P300-based spelling BCIs, it is evident that there is potential for continued improvement and innovation in this field. While there are still challenges to be addressed, such as the need for robust and reliable signal acquisition, processing, and classification, the potential benefits for individuals with severe motor impairments are significant, and the field is likely to continue to grow and advance in the future.

In addition, the challenge of separating functionally meaningful event-related potentials from EEG signals that occur simultaneously and may also be entirely or partially overlapped still remains a challenge for most researchers in this field.

There are numerous methods for classifying EEG signals, most of which involve two main steps. The first step entails using an algorithm to measure the dynamic temporal distortion as a means of differentiating between the EEG signals that need to be classified. Alternatively, one can employ mathematical tools like simple statistics or advanced mathematical techniques to represent the EEG signals as feature vectors. In the second step, an algorithm is used to classify this data utilizing techniques such as k-nearest neighbors, neural networks, support vector machines, etc. However, all these methods require some form of feature development as a separate step before they can classify the EEG signal. Convolutional Neural Networks (CNNs) have several significant advantages over other classification methods due to their ability to extract time-independent, information-rich deep features that are highly noise-tolerant [27, 28].

This paper presents a novel approach for extracting highly informative features from EEG signals using CNNs to detect ERPs within an EEG signal, providing a potential communication tool for paralyzed people with motor and speech impairments who struggle with their inability to communicate with those around them. In the study, participants were asked to announce a certain need out of six basic daily needs while their pictures flashed in random order.

Methods

Proposed system design

The proposed system comprises three phases:

The first phase is the data acquisition phase, in which the user focuses on one of the six basic needs while images are randomly flashed. The wireless EEG headset

acquires the EEG signal, which is then amplified, sampled at 2048 Hz, and passed on digitally to the second phase.

The second phase is the signal processing phase, which consists of two sub-phases. The first is "Signal Preprocessing and Digital Filtering," where the signal undergoes filtration. The second sub-phase is "Feature Extraction," where Convolutional Neural Networks (CNNs) process the signal to extract highly informative deep features that are time-independent concerning the presence of ERP patterns. A decision is made during the third phase based on these features.

The third phase is the classification phase, where the desired need is selected out of the six basic daily needs. This selection is announced based on the system settings, whether by audible voice command or by a text message sent to the healthcare provider's mobile phone.

Figure 2 shows the system block diagram, and Fig. 3 shows the images for the six basic daily needs.

Data acquisition and pre-processing

Brain electrical activity was captured using a 14-channel wireless EEG headset with electrodes positioned based on the 10–20 International Electrode Positioning System. Figure 4 displays the 14-channel EMOTIV EPOC + Wireless EEG Headset, which was used to acquire the EEG signal. It is worth noting that it is a dry EEG headset with self-placement function for electrodes.

The acquired EEG signal was amplified since the amplitudes of the signals range from 0.5 to 100 μ V. The instrumentation amplifier, which is a part of the acquisition system, must be suitable for the nature of the EEG bio-signal. Therefore, its common-mode-rejection ratio should be greater than 100 dB, the input impedance should be greater than 50 G Ω , and the frequency response should be from 0.3 to 35 Hz. The signal was then digitized using A/D with a sampling frequency of 2048 Hz.

As raw EEG data contains high-frequency perturbations and random noise, it is often challenging to extract useful information from this raw data with the presence of these noises. Therefore, it was crucial to digitally filter the upcoming EEG signal using a 6th

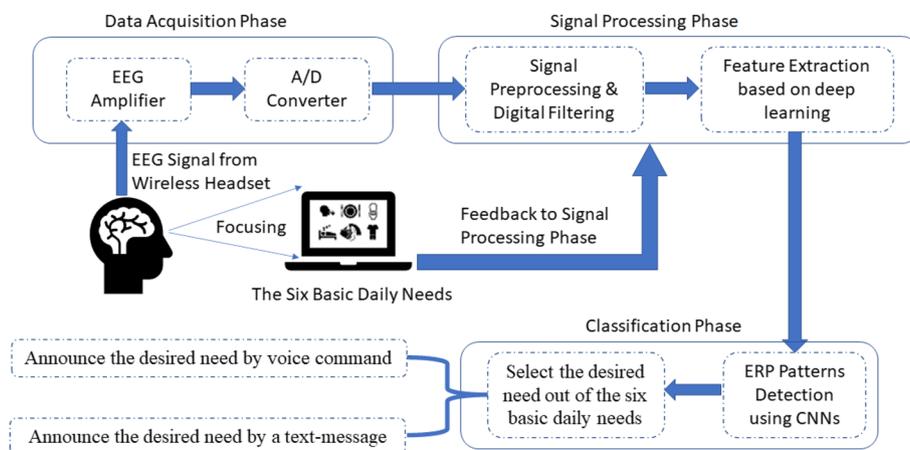


Fig. 2 A system block diagram

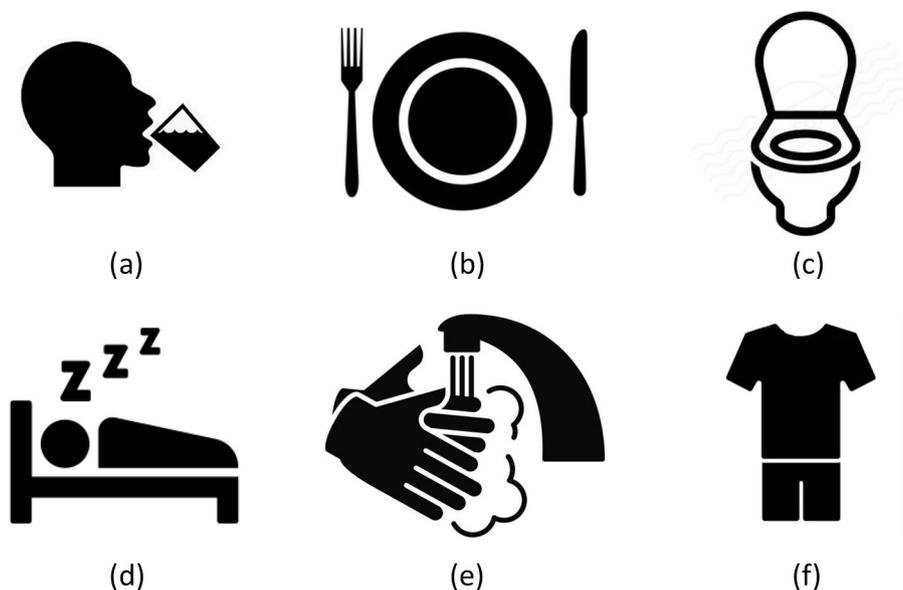


Fig. 3 The images for the six basic daily needs; **a** the need to drink, **b** the need to eat, **c** the need to go to the toilet, **d** the need to sleep, **e** the need to wash, and **f** The need to change clothes

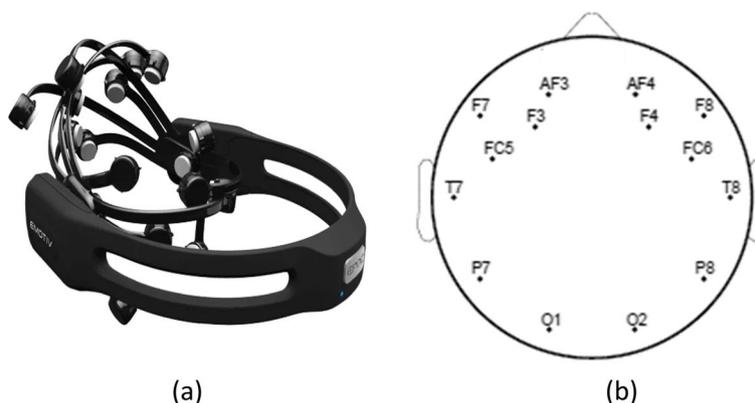


Fig. 4 a 14-Channels EMOTIV EPOC + Wireless EEG Headset. **b** Electrodes' locations map on the scalp

order Butterworth band-pass digital filter with a low cutoff frequency and high cutoff frequency of 1 Hz and 12 Hz, respectively.

The importance of this band-pass digital filter (1–12 Hz) is to remove noise and preserve the ERP pattern information, which will increase the Signal-to-Noise Ratio (SNR) by removing noise caused by non-brain activities, such as motion, and removing artifacts induced by the monitor refresh rate.

Images grid layout design and the communication system

The communication layout is a 2-by-3 grid matrix, with each cell containing an image representing one of the six basic daily needs, as shown in Fig. 3. During the flashing process, the user is asked to focus on a specific image representing their desired need, while the images are flashed in a random order. The user silently counts how many times the target image appears, as a way to keep their attention focused on the screen. The system then checks for

the presence of ERP patterns in the acquired EEG signal within 500 ms latency time. Only ERP patterns corresponding to the desired need image being flashed will show up.

Flashing timeline analysis

Variations in the electrical brain activity signal (EEG) occur in response to random and unpredictable stimuli, and these waveforms appear in a time latency window ranging from 40 to 400 ms after the stimulus is presented. Therefore, the flashing timeline was designed such that the images were flashed in a random order, one image at a time, with each flash of an image lasting for 100 ms. The Inter Stimulus Interval was 100 ms, and during the following 400 ms, none of the images were flashed, resulting in a Stimulus-to-Stimulus interval (SSI) of 500 ms. Figure 5 illustrates the flashing timeline analysis of the proposed system.

EEG dataset collection

The dataset was collected with the aid of ten volunteers-eight males and two females-who were involved in the data collection process.

The user focuses on a specific image that represents their desired need while the images are flashed in a random order. Simultaneously, the software displays the image sequence randomly on the screen while acquiring the brain electrical activity from the 14-channel EMOTIV EPOC + Wireless EEG Headset.

The flashing timeline is designed so that the images are displayed in a random order, one at a time, with an Inter-Stimulus Interval (ISI) of 100 ms and a Stimulus-to-Stimulus Interval (SSI) of 500 ms.

The distance of the subject from the screen does not affect the performance as long as the images representing the daily needs are clearly visible. However, during the study, the users performed the experiment in front of the computer while wearing the headset, and the distance between the user and the screen was kept between 80-100 cm.

Feature extraction

A novel approach has been introduced to extract highly informative features from EEG signals using Convolutional Neural Networks (CNNs), which have several advantages over other classification methods. They are highly noise-tolerant and can extract time-independent, informative deep features.

The proposed approach works with the original EEG data sequence and its down-sampled versions at different time scales. It does this by applying a set of 1-D convolution kernels with variable length l , which corresponds to the number of time steps or window size, and fixed width $k = 14$, which corresponds to the number of EEG channels of the wireless headset. Each kernel performs convolution on the EEG signal by moving in one direction from the beginning of the EEG signal towards its end.

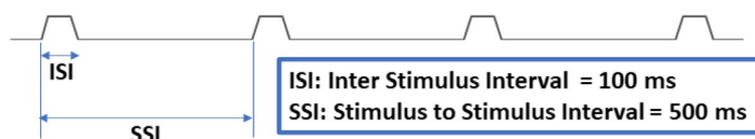


Fig. 5 Flashing timeline analysis

Suppose that the original EEG signal captured from the 14-channel wireless EEG headset is of length $n = 1024$, which corresponds to the duration of 500 ms, with a sampling frequency of 2048 Hz.

$$\begin{aligned}
 EEG_1 &= \{D_{11}, D_{12}, D_{13}, D_{14}, \dots, D_{1n},\} \\
 EEG_2 &= \{D_{21}, D_{22}, D_{23}, D_{24}, \dots, D_{2n},\} \\
 EEG_3 &= \{D_{31}, D_{32}, D_{33}, D_{34}, \dots, D_{3n},\} \\
 &\vdots \\
 EEG_k &= \{D_{k1}, D_{k2}, D_{k3}, D_{k4}, \dots, D_{kn},\}
 \end{aligned}
 \tag{1}$$

Where $n = 1024, k = 14$.

The moving average works by converting these EEG data sequences into new data sequences of different frequencies.

$$EEG^l = \frac{1}{l} \sum_{i=1}^l D_i \text{ where } l \text{ is the moving average window size}
 \tag{2}$$

With varying values of l , multiple EEG data sequences of different frequencies will be generated for each channel EEG data sequence with a length of n . Each of these new EEG data sequences will have a length of $(n - l + 1)$, as shown in Fig. 6.

The next step is max pooling, which involves aggregating the largest values from the convolution layer vectors. The largest value from each vector is selected, and a new vector (the maxima vector) is formed from these values.

Figure 7 illustrates how convolution kernels with variable length l and fixed width k work by moving in one direction from the beginning of the EEG signal towards its end.

The same procedure will be applied to the down-sampled versions of the original EEG data sequence. The down-sampled versions can be generated by choosing different down-sampling rates k and retaining every $k - th$ data point in the new down-sampled EEG data sequence.

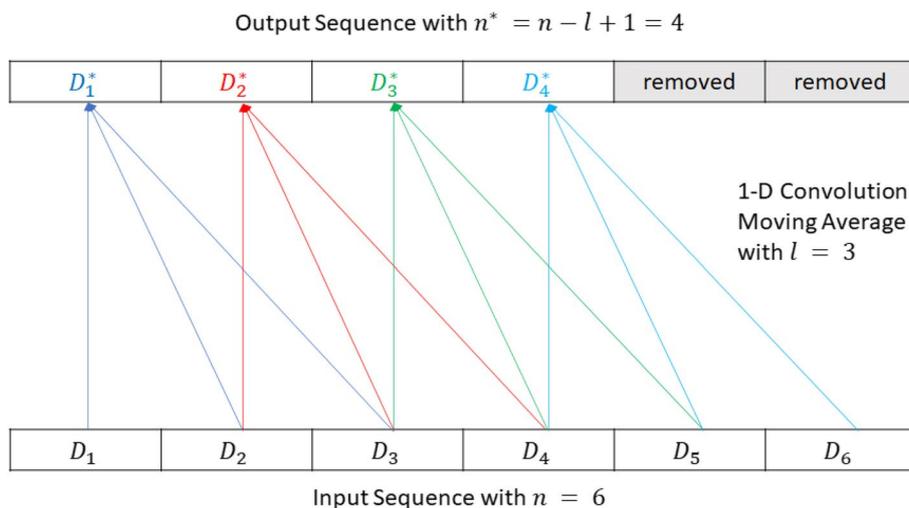


Fig. 6 Illustration of 1-D Convolution moving average kernel

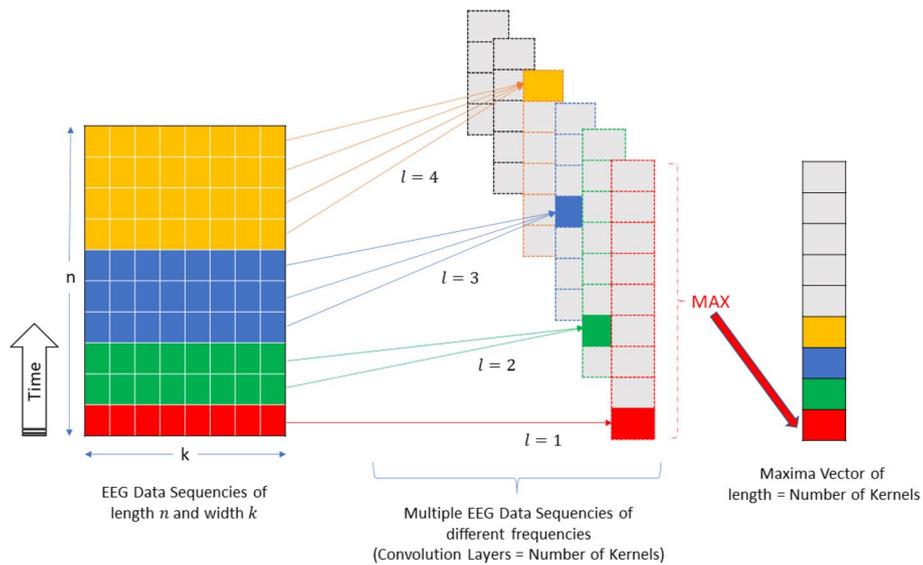


Fig. 7 How convolution kernels with variable length l and fixed width k works

$$EEG^k = \bigcup_{i=0}^{\frac{n-1}{k}} D_{k*i+1} \tag{3}$$

With different k , multiple EEG data sequences of different time scales will be generated for each channel's EEG data sequence with a length of n . Each generated sequence will have a length of $Round(n/k)$ rounded to the nearest integer, as illustrated in Fig. 8.

All Maxima Vectors from all stages are then concatenated into one (Concatenated Maxima Vector), which will then undergo processing by another layer of a set of 1-D convolution kernels with variable length l and fixed width $k = 1$. The output of this layer is a (Final Maxima Vector), which is the final feature vector holding highly

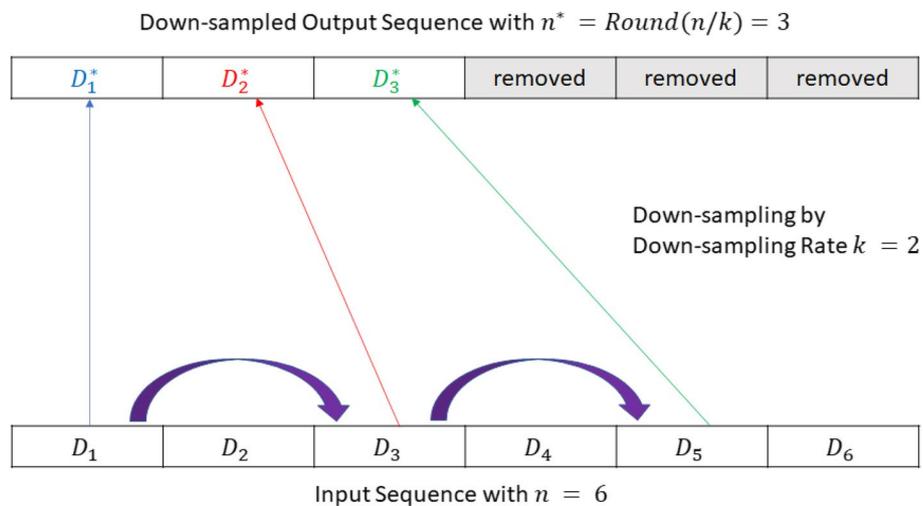


Fig. 8 Illustration of down-sampling

informative deep complex features that can be used as input to the Artificial Neural Network, as illustrated in Fig. 9.

Results

The collected dataset contained 4440 EEG signals labeled as P300, which were showing ERP patterns, especially the P300 pattern, and 22200 normal EEG signals labeled as NoneP300, which were not showing any ERP patterns, as shown in Fig. 10.

Since 5/6 of the EEG signals in the dataset were NoneP300, the classifier might learn that it can achieve high accuracy by simply classifying all signals as NoneP300. Consequently, the reported accuracy of 83.34% is misleading due to the imbalance between P300 and NoneP300 signals in the dataset.

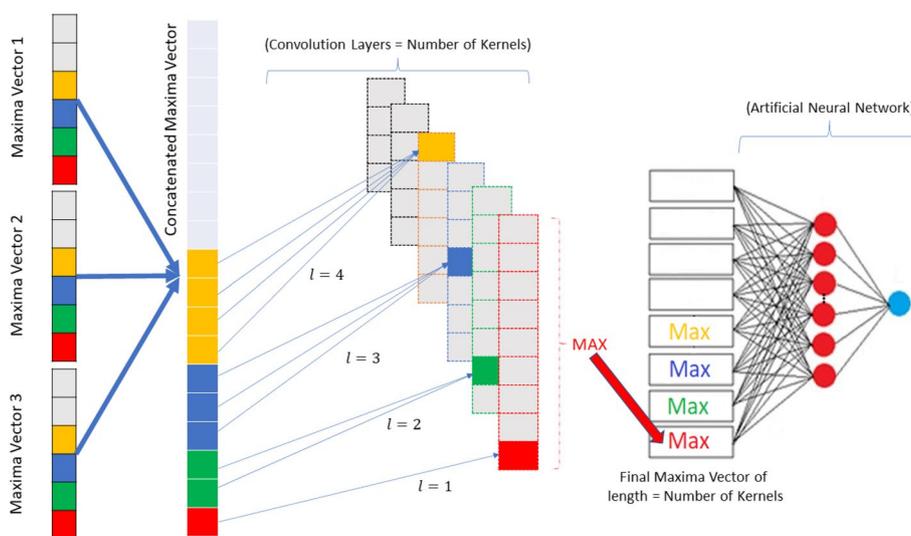


Fig. 9 Maxima vectors concatenation, and generation of the final maxima vector as an input layer to the artificial neural network

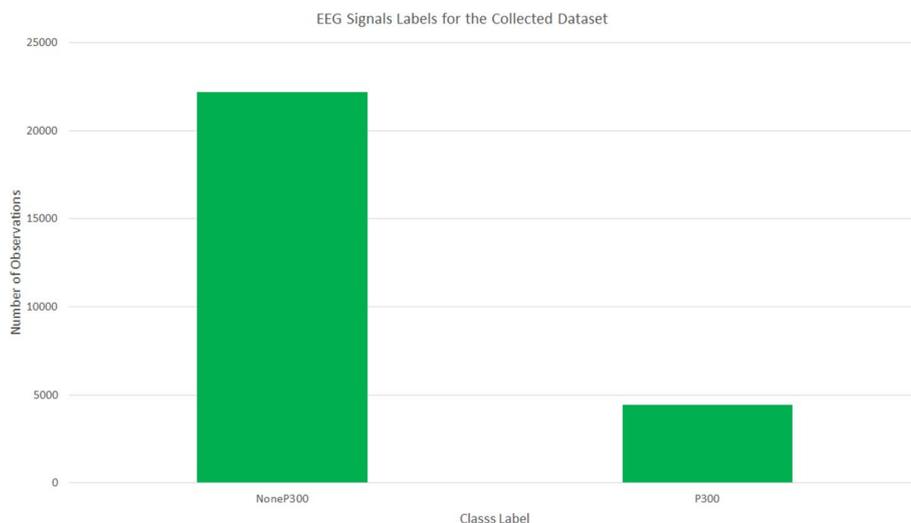


Fig. 10 Summary of EEG signals labels, for the collected dataset

To avoid this bias and solve the class imbalance problem, it was crucial to either over-sample the P300 EEG signals to be equal to the NoneP300 EEG signals or down-sample the NoneP300 EEG signals to be equal to the P300 EEG signals. In this paper, the second choice was chosen to down-sample the NoneP300 EEG signals to avoid redundancy of data that could affect accuracy. Additionally, the dataset was randomly divided into two groups for training and validation purposes with a ratio of 70% to 30%, respectively. Figure 11 shows the summary of training and validation classes used for the proposed deep learning convolutional neural network.

The first deep learning architecture proposed convolutional neural network (CNN) with three sets of 1-D convolutional layers and ReLU layer function that perform a threshold operation such that any input value less than zero is set to zero, as shown in Eq (4). In each set of layers, there is a normalization layer between the convolutional



(a)



(b)

Fig. 11 Summary of training and validation classes for the proposed deep learning convolutional neural network. **a** Summary of training classes. **b** Summary of validation classes

layer and threshold layer to speed up training of the convolutional neural network and reduce sensitivity to network initialization. After the completion of the three sets, there is a 1-D global averaging layer that returns the average values of rectangular regions of its input for down-sampling and reducing the number of connections to the fully connected neural network. This is followed by a SoftMax activation layer, which is a normalized exponential as shown in Eq (5), and acts as the output unit activation function. Figure 12 illustrates the architecture of the proposed deep learning convolutional neural network layers.

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{4}$$

$$y_r(x) = \frac{\exp(a_r(x))}{\sum_{j=1}^k \exp(a_j(x))}, \text{ where } 0 \leq y_r \leq 1 \text{ and } \sum_{j=1}^k y_j = 1 \tag{5}$$

Figure 13 shows the training accuracy process for all volunteers, with a maximum accuracy of 96.30% and a final trend accuracy of 78.32%.

Table 1 is summarizing the training accuracy for the training process of each volunteer from the ten volunteers.

Figure 14 is showing both the training process and the validation process with a validation accuracy of 78.41%, and the confusion matrix for its validation process.

The confusion matrix shown provides a clearer picture of the classification accuracy. The vertical axis represents the true class labels, while the horizontal axis represents the predicted class labels.

The number of signals that actually contained ERP patterns (P300 label) and were correctly predicted as such by the CNN classifier is known as True Positive (TP), with a value of 770. The number of signals that did not contain ERP patterns (NoneP300 label) and were correctly predicted as such by the CNN classifier is known as True Negative (TN), with a value of 1046.

The number of signals that actually contained ERP patterns (P300 label) but were mistakenly predicted as Non-ERP Signal (NoneP300 label) by the CNN classifier is called False Negative (FN), and it equaled 112. The number of signals that actually did not

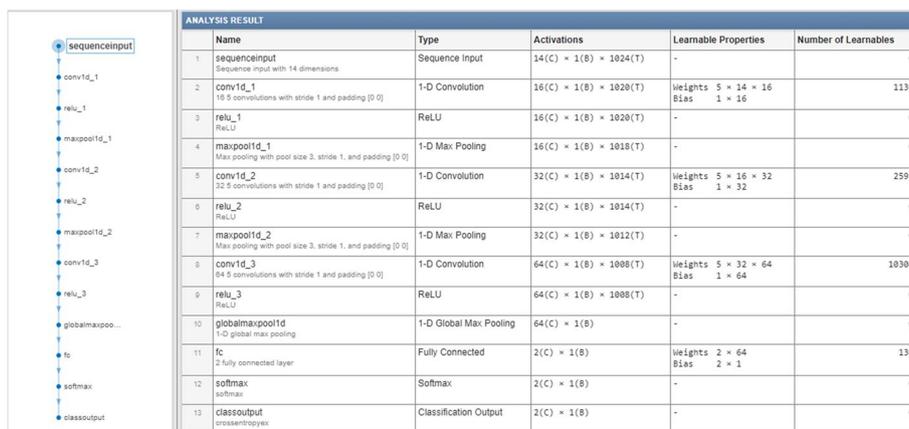


Fig. 12 Convolutional neural network layers analysis

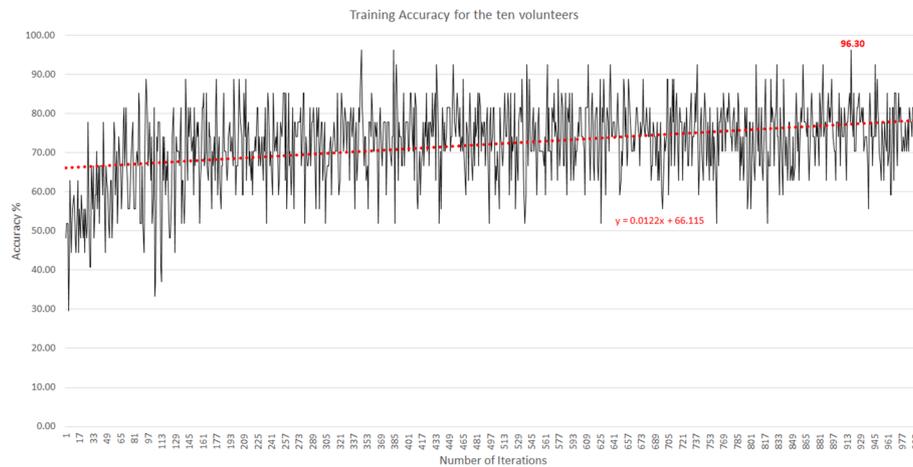


Fig. 13 Training accuracy for all volunteers showing a maximum accuracy of 96.30%, and a final trend accuracy of 78.32%

Table 1 Summary of the training process accuracy

Training accuracy for:	Maximum training accuracy	Final training accuracy
Volunteer 1	92.59%	88.89%
Volunteer 2	85.19%	66.67%
Volunteer 3	85.19%	74.07%
Volunteer 4	85.19%	69.95%
Volunteer 5	88.89%	74.37%
Volunteer 6	92.59%	82.63%
Volunteer 7	81.48%	65.34%
Volunteer 8	92.59%	74.18%
Volunteer 9	88.89%	80.64%
Volunteer 10	92.59%	74.34%
All the ten Volunteers	96.30%	78.32%

contain ERP patterns (NoneP300 label) but were mistakenly predicted as ERP Signal (P300 label) by the CNN classifier is called False Positive (*FP*), and it equaled 388.

The efficiency (accuracy) performance of the CNN classifier is the number of correctly classified predictions (*TP + TN*) out of the total number of datasets, which can be expressed by Eq (6).

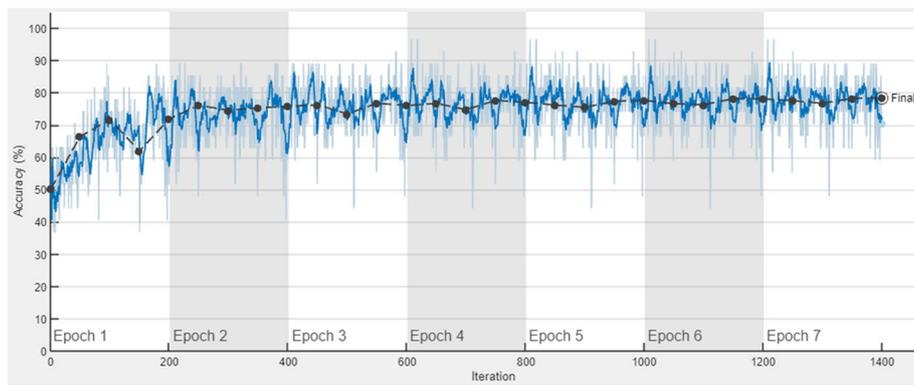
$$Efficiency (Accuracy) = \frac{TP + TN}{TP + FN + TN + FP} \tag{6}$$

and it equaled 78.41%.

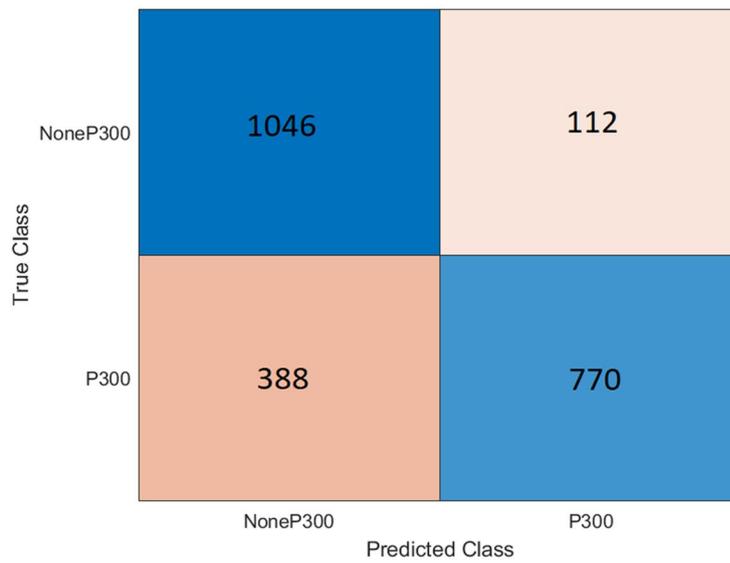
A second design based on a long short-term memory (LSTM) network was tested to compare results, and it showed an accuracy of 69.69%, which is lower than the proposed deep learning convolutional neural network (CNN).

Figure 15 shows the architecture of the long short-term memory neural network layers.

Figure 16 displays the training and validation processes of the LSTM, which achieved a lower validation accuracy of 69.69% compared to the proposed deep learning convolutional neural network (CNN).



(a)



(b)

Fig. 14 a Convolutional neural network training process and validation process showing a validation accuracy of 78.41%. b Confusion matrix for validation process

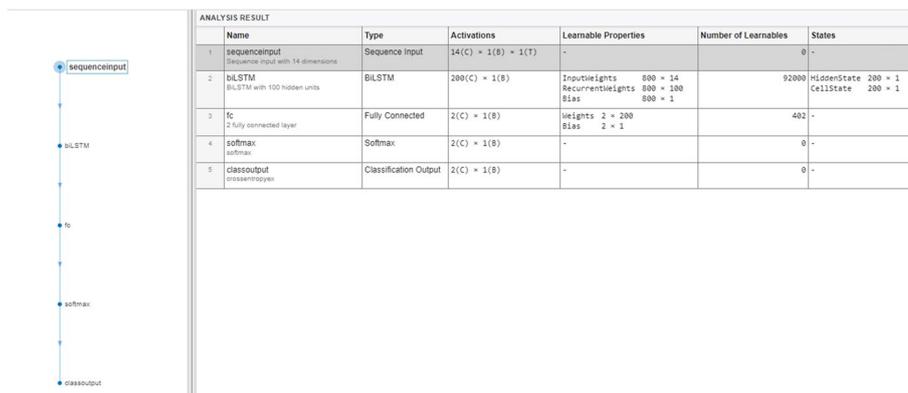


Fig. 15 Long short-term memory neural network layers analysis

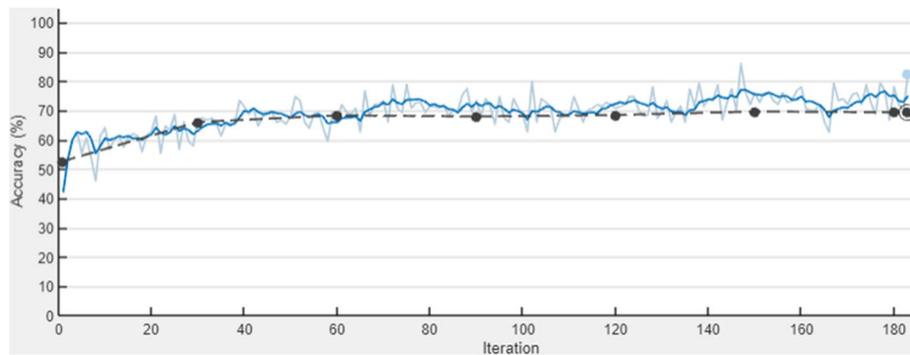


Fig. 16 Long short-term memory neural network training process showing a lower accuracy of 69.69%

Discussion

The results show that the collected dataset contained 4440 EEG signals labeled as P300 and 22200 normal EEG signals labeled as NoneP300. The reported accuracy of 83.34% is misleading due to the class imbalance problem, where 5/6 of the EEG signals in the dataset were NoneP300. To solve this problem, the dataset was down-sampled to avoid redundancy of data that could affect accuracy. Additionally, the dataset was randomly divided into two groups for training and validation purposes.

The first deep learning architecture proposed, a convolutional neural network (CNN), achieved a training accuracy ranging from 65.34% to 88.89% for individual volunteers, and an overall training accuracy of 78.32% for all volunteers.

The CNN classifier achieved an efficiency (accuracy) of 78.41% using the validation group, with a confusion matrix showing that the number of signals that actually contained ERP patterns (P300 label) and were correctly predicted as such by the CNN classifier was 770 (True Positive), while the number of signals that did not contain ERP patterns (NoneP300 label) and were correctly predicted as such by the CNN classifier was 1046 (True Negative).

However, the study also tested a second deep learning architecture, a long short-term memory (LSTM) network, which showed lower accuracy of 69.69% compared to the proposed CNN architecture. Overall, the study demonstrates the effectiveness of using deep learning architectures, particularly CNN, for classifying EEG signals labeled as P300 and NoneP300.

Conclusions

The study proposes a deep learning CNN-based approach for the classification of P300 signals from EEG signals. The study finds that the class imbalance between P300 and NoneP300 signals in the dataset could lead to a biased classifier, and therefore, it is crucial to balance the dataset. The proposed CNN architecture achieves an accuracy of 83.34%, with the training process of each volunteer yielding an accuracy ranging from 65.34% to 88.89%. The study also compares the performance of the proposed CNN with that of an LSTM network and finds that the CNN outperforms the LSTM.

In fact, the classification accuracy strongly depends on participants' correct performance. Therefore, pre-orientation and adaptation for using the system are crucial for those who will use the system.

In conclusion, developing a communication channel between the human brain and the computer by using electrical changes caused by mental activity and turning it into a control signal can be very important, especially for paralyzed patients who are unable to speak and communicate with their environment. The proposed system analyzes electrical EEG brain signals generated by mental activity and detects Event-Related Potential (ERP) patterns to facilitate easy communication between patients and those around them. As the aim was to provide acceptable performance and higher accuracy, I think that the proposed system has successfully achieved that, as it shows an average accuracy of 78.41% (with a maximum accuracy of 96.30%).

Therefore, the approach of using Convolutional Neural Networks (CNNs) results in extracting highly informative deep features from EEG signals that are capable of detecting the presence of Event-Related Potential (ERP) patterns, providing acceptable performance and higher accuracy for the proposed system. This suggests that the proposed CNN-based approach is a promising technique for the classification of P300 signals from EEG signals, and it could potentially be used in applications such as brain-computer interfaces and cognitive neuroscience research.

Abbreviations

BCI	Brain Computer Interface
SVM	Support Vector Machine
EEG	Electroencephalogram
ERP	Event Related Potential
SWLD	Stepwise Linear Discriminant Analysis
SWLDA	Stepwise Linear Discriminant Analysis Algorithm
EMD	Empirical Mode Decomposition
ICA	Independent Component Analysis
LDA	Linear Discriminant Analysis
VEP	Visual Evoked Potential
SSVEP	Steady State Visual Evoked Potential
SNR	Signal-to-Noise Ratio
A/D	Analogue-to-Digital Converter
ISI	Inter Stimulus Interval
SSI	Stimulus-to-Stimulus Interval
FFT	Fast Fourier transform
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
TP	True Positive
TN	True Negative
FP	False Positive (Type I Error)
FN	False Negative (Type II Error)

Acknowledgements

Author expresses his sincere gratitude to all volunteers and participants who shared in collecting data during the Data Collection Phase for their helpful support. Also, author would like to thank (Biomedical Engineering Department, Faculty of Engineering, Misr University for Science and Technology, Egypt) for any advice or discussion that improved the study.

Author's contributions

The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation. The author(s) read and approved the final manuscript.

Funding

The author declares that this study was not funded.

Availability of data and materials

The dataset used and analyzed during this study are available from the author on reasonable request.

Declarations

Ethics approval and consent to participate

All procedures performed in studies involving human participants were approved by Institutional Review Board at Misr University for Science and Technology (MUST-IRB) (Approval Number: 2022/0032 in December 28, 2022)—MUST-IRB is registered at the office for Human Research Protections, US Department of Health and Human Services and Operates under Federal Wide Assurance No. FW A00025577—and were in accordance with the ethical standards and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards:

1. The study was neither therapeutic nor clinical research.
2. The study was risk free and not harmful for life and health of all participants.
3. Placement of electrodes and EEG signal acquire was obtained by a qualified technician, and under the supervision of a qualified doctor.
4. Both nature and purpose of this study were explained to all participants in this study.
5. All participants in this study were mature and adult, and were in such a mental, physical and legal state as to be able to exercise fully their power of choice.
6. An informed consent was obtained from all participants in this study.
7. It was free to all participants in this study to withdraw permission for research to be continued, at any time during this study.

This paper does not contain any studies with animals.

Consent for publication

Author declares that all images are entirely unidentifiable and there are no details on individuals reported within the paper.

Competing interests

The author declares that he has no competing interests.

Received: 1 March 2023 Accepted: 2 May 2023

Published online: 09 May 2023

References

1. Brain vision UK blog, The brief history of brain computer interface, notes <http://www.brainvision.co.uk/blog/2014/04/the-brief-history-of-brain-computer-interfaces>.
2. A. Amcalar and M. Cetin, Design, Implementation and Evaluation of a Real-Time P300-based Brain-Computer Interface System. 20th International Conference on Pattern Recognition, 23–26 Aug. 2010. pp. 117–120 <https://doi.org/10.1109/ICPR.2010.37>.
3. Fuerst DR, Gallinat J, Boutros NN (2007) Range of sensory gating values and test–retest reliability in normal subjects. *Psychophysiology* 44(4):620–626 <https://doi.org/10.1111/j.1469-8986.2007.00524.x>.
4. Blackwood D, Muir W (1990) Cognitive brain potentials and their application. *Br J Psychiatry* 157(S9):96–101 <https://doi.org/10.1192/S0007125000291897>.
5. Niedermeyer E, Lopes da Silva FH (1993) *Electroencephalography: Basic principles, clinical applications and related fields*, 3rd edition, Lippincott. Williams & Wilkins, Philadelphia
6. Trans Cranial Technologies, 10/20 System Positioning Manual, vol. 1, 2012, https://trans-cranial.com/docs/10_20_pos_man_v1_0_pdf.pdf
7. Farwell LA, Donchin E (1988) Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr Clin Neurophysiol* 70(6):510–523 [https://doi.org/10.1016/0013-4694\(88\)90149-6](https://doi.org/10.1016/0013-4694(88)90149-6).
8. Cuntai Guan, M. Thulasidas and Jiankang Wu. High performance P300 speller for brain-computer interface. IEEE International Workshop on Biomedical Circuits and Systems: Singapore; 2004. pp. 13–16 <https://doi.org/10.1109/BIOCAS.2004.1454155>.
9. "BCI Competition III". 2005, <http://www.bbc.de/competition/iii/results/>.
10. Krusienski DJ, Sellers EW, Cabestain F, Bayouh S, McFarland DJ, Vaughan TM, Wolpaw JR (2006) A comparison of classification techniques for the P300 speller. *J Neural Eng* 3(4):299–305 <https://doi.org/10.1088/1741-2560/3/4/007>. IOP Publishing
11. T. Solis-Escalante, G. G. Gentiletti and O. Yanez-Suarez. Single Trial P300 detection based on the Empirical Mode Decomposition. International Conference of the IEEE Engineering in Medicine and Biology Society; 2006. pp. 1157–1160 <https://doi.org/10.1109/IEMBS.2006.260589>.
12. Salvaris M, Sepulveda F (2009) Visual modifications on the P300 speller BCI paradigm. *J Neural Eng* 6(4):1–8 <https://doi.org/10.1088/1741-2560/6/4/046011>.
13. Townsend G, LaPallo BK, Boulay CB, Krusienski DJ, Frye GE, Hauser C, Schwartz NE, Vaughan TM, Wolpaw JR, Sellers EW (2010) A novel P300-based brain–computer interface stimulus presentation paradigm: Moving beyond rows and columns. *Clin Neurophysiol* 121(7):1109–1120 <https://doi.org/10.1016/j.clinph.2010.01.030>.
14. Wang P, Ji-Zhong S, Jin-He S (2010) P300 detection algorithm based on fisher distance. *Int J Modern Educ Comp Sci* 2:9–17 <https://doi.org/10.5815/ijmecs.2010.02.02.002>.
15. Cecotti H, Graser A (2011) Convolutional Neural Networks for P300 Detection with Application to Brain-Computer Interfaces. *IEEE Trans Pattern Anal Mach Intell* 33(3):433–445 <https://doi.org/10.1109/TPAMI.2010.125>.
16. N. Haghighatpanah, R. Amirfattahi, V. Abootalebi and B. Nazari. A single channel-single trial P300 detection algorithm. 21st Iranian Conference on Electrical Engineering (ICEE); 2013. pp. 1–5 <https://doi.org/10.1109/IranianCEE.2013.6599576>.

17. N. Haghghatpanah, R. Amirfattahi, V. Abootalebi and B. Nazari. A two stage single trial P300 detection algorithm based on independent component analysis and wavelet transforms. 19th Iranian Conference of Biomedical Engineering (ICBME). 2012. pp. 324–329 <https://doi.org/10.1109/ICBME.2012.6519702>.
18. Y. Wang, J. Shen, J. Liang and Y. Ji. Research of P300 Feature Extraction Algorithm Based on ICA and Wavelet Transform. Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics. 2014. pp. 41–45 <https://doi.org/10.1109/IHMSC.2014.18>.
19. Raju VN, Ra I, Sankar R (2015) A P300-Based BCI Classification Algorithm Using Least Square Support Vector Machine. *Int J Softw Eng Appl* 9:247–254
20. Wang M, Daly I, Allison B, Jing Jin Yu, Zhang LC, Wang X (2015) A new hybrid BCI paradigm based on P300 and SSVEP. *J Neurosci Methods* 244:16–25 <https://doi.org/10.1016/j.jneumeth.2014.06.003>.
21. Yin E, Zhou Z, Jiang J, Chen F, Liu Y, Hu D (2014) A Speedy Hybrid BCI Spelling Approach Combining P300 and SSVEP. *IEEE Trans Biomed Eng* 61(2):473–483 <https://doi.org/10.1109/TBME.2013.2281976>.
22. Yin E, Zeyl TJ, Saab R, Chau T, Dewen Hu, Zhou Z (2015) A Hybrid Brain-Computer Interface Based on the Fusion of P300 and SSVEP Scores. *IEEE Trans Neural Syst Rehabil Eng* 23(4):693–701 <https://doi.org/10.1109/TNSRE.2015.2403270>.
23. Fernandez-Rodriguez A, Velasco-Alvarez F, Medina-Julia MT, Ron-Angevin R (2019) Evaluation of emotional and neutral pictures as flashing stimuli using a P300 Brain- Computer Interface speller. *Neural Eng* 16(5):1–12 <https://doi.org/10.1088/1741-2552/ab386d>.
24. R. Ron-Angevin, L. Garcia, A. Fernandez-Rodriguez, J. Saracco, J. M. Andre and V. Lespinet-Najib. (2019) Impact of speller size on a visual P300 Brain-Computer Interface (BCI) system under two conditions of constraint for eye movement. *Comp Intelligence Neuroscience* 1–16 <https://doi.org/10.1155/2019/7876248>.
25. Li F, Li X, Wang F, Zhang D, Xia Yi, He F (2020) A novel P300 classification algorithm based on a principal component analysis-convolutional neural network. *Appl Sci* 10(4):1546 <https://doi.org/10.3390/app10041546>.
26. Li Q, Lu Z, Gao N, Yang J (2019) Optimizing the performance of the visual P300-Speller through active mental tasks based on color distinction and modulation of task difficulty. *Front Hum Neurosci* 13(130):1–14 <https://doi.org/10.3389/fnhum.2019.00130>.
27. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, Santamaría J, Fadhel MA, Al-Amidie M, Farhan L. (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data* 8(53) <https://doi.org/10.1186/s40537-021-00444-8>.
28. Taye MM (2023) Theoretical understanding of convolutional neural network: concepts, architectures, applications, future directions. *Computation* 11(3):52 <https://doi.org/10.3390/computation11030052>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Nader A. Rahman Mohamed Nader A. Rahman Mohamed (PhD. in System & Biomedical Engineering) was born in Egypt in 1973. He received the B.S. and M.S. degrees in Systems and Biomedical Engineering from Faculty of Engineering – Cairo University in June 1996 and June 2002, respectively. He received the Ph.D. degree in Systems and Biomedical Engineering from Faculty of Engineering – Cairo University in June 2009. He worked as a Biomedical Engineering Consultant in the field of Ophthalmology, Medical Lasers, and Endoscopy in Center for Advanced Software & Biomedical Engineering Consultation (CAS-BEC), Faculty of Engineering, Cairo University. He joined the department of Biomedical Engineering – Faculty of Engineering – Misr University for Science and Technology, in September 2009. His current research interests include image

and signal processing, artificial intelligence, neural network, biomedical equipment design and control, biomechanics and biomaterials.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
