# A Novel Algorithm for Detection of Social Joint Attention from single-trial EEG signals of Autistic Spectrum Disorder (ASD)

Mahdie Ghane Ezabadi
*Department of biomedical Engineering*
*Amirkabir University of Technology*
*Tehran, Iran*
Mahdie_ghane@aut.ac.ir

Mohammad Hassan Moradi
*Department of biomedical Engineering*
*Amirkabir University of Technology*
*Tehran, Iran*
mhmoradi@aut.ac.ir

2021 28th National and 6th International Iranian Conference on Biomedical Engineering (ICBME) | 978-1-6654-9984-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICBME54433.2021.9750300

*Abstract* – **Autism Spectrum Disorder (ASD) is a developmental disability that shows its symptoms during second year of life or later. People with ASD have poor performance in different abilities such as social joint attention. In recent years, scientists has developed different methods and technologies to help ASD people to have a better life. Brain-Computer Interface (BCI) is one of these technologies that plays an important role in the rehabilitation of patients with neurological disorders. BCI is a technology that makes a bridge between brain and computer which let brain signals like EEG be recorded for processing. On the other hand, it has been shown that social joint attention can be detected in EEG signals using P300 which is one of the most popular components of Event-Related Potential (ERP). Therefore, recording EEG signals with the use of BCI for processing them with a reliable algorithm can be a great step toward understanding the helping people with ASD to improve their social joint attention. In this study, a novel algorithm based on a Convolutional Neural network will be introduced for detection attention in single-trial EEG signals more precisely. As a dataset, IFMBE MEDICON 2019 challenge dataset will be used in which autistic adults were learning social joint-attention with the help of the BCI system. Results show that it can improve the performance of detecting P300 from single-trial EEG signals effectively in comparison to other algorithms. This method increased final target detection accuracy from 92.37% to 94.85%.**

***Index Terms - Convolutional Neural Network (CNN), EEG, Attention, Single-trial P-300 detection, Classification.***

## I. Introduction

Autism Spectrum Disorder (ASD) is a neurological and developmental disorder that can be diagnosed at any age. However, its symptoms appear in childhood, generally in the first two years of life [1]. This disorder affects how a person acts and interacts with others, communicates, and learns. One of the common problems in people with ASD is joint attention which means they cannot pay attention to an object that someone else is pointing [1], [2]. Therefore, it hurts some of their ability like social performance since they cannot share their experience and learning process as they cannot get other's points.

Recently, it has been proved that Virtual Reality (VR) P300-based Brain-Computer Interference (BCI) systems is a suitable method for improving neurologically-based disorders [3]. Therefore, Amaral et al developed a promising BCI protocol using VR for improving the social joint attention of the youngest with ASD. On the hypothesis, individuals with ASD can learn joint attention skills by internalizing the response to a given social cue. Therefore, they created a task where attentional mental state is measured through the detection of oddballs, which in turn induces a P300 signal [5], [6]. This training tool has been demonstrated to be beneficial and very promising in the field of social cognition training in ASD individuals [7]. This study records Electroencephalography (EEG) signals during their learning process and tries to decode their attention and its progress using it.

Electroencephalography (EEG) is a well-known technique to capture brain signals which are produced by neurons, using non-invasive electrodes placed on different locations over the scalp. When someone pays attention, his EEG signal pattern changes and it peaks 300ms after stimulus onset. This component is named P300. It can be detected while EEG signals are averaged because averaging intensify P-300 peaks and reduce EEG noise [4]. However, to measure attention during learning, it is needed to find a way for detecting P-300 in single-trial EEG. To develop this tool, it is necessary to design and implement a fully functional BCI that can identify the P300 component in the EEG signal during online phase. Different studies have tackled this specific problem for BCI applications which improved the accuracy and speed of P-300-based BCIs [8], [9]. One effective way to increase accuracy for P-300 detection is replacing traditional methods like finding features of signals by defined function and trying to classify with classification methods such as LDA and SVM with more state-of-art methods like deep learning. Considering this idea in mind, deep learning and neural networks could be very useful in both extracting features and classification attitude.

In recent years, the emergence of neural networks caused conspicuous performance in areas such as machine learning, which is due to learning high-level features automatically from raw data [9]. Convolutional Neural Network is one of the deep neural network architectures that its usage in extracting features from raw signals, especially bio-signals has been raised in recent years. In this way, this structure attracts much attention as a feature extractor of EEG signals due to the complex behavior of EEG signals and its low signal-to-noise ratio (SNR). Thus, researchers tried to benefit from this architecture for having a better result. Cecotti et al was the first person who developed a simple structure of CNN consisting of five layers for P-300 detection [10]. After that, Lawhern et al proposed a more sophisticated lightweight network with more layers which was named EEGNet for detecting P-300 from single-trial EEG signals [11]. Although CNN's are proper for extracting high-level features, their performance fails when approaches training long-term high-level distributed features [11], [12]. As a result, in this research, CNN's are only used for feature extraction. After that, a combination of three classification algorithms with the highest accuracy folding with weighted voting is used [13], [14]. The rest of the paper is organized as follows: In the following sections, first, the dataset is described and after that, the proposed structure will elucidate for learning. In the end, obtained results are discussed followed by a comparison with previous learning algorithms.

## II. Methodology

The summary of data and methods is depicted in Fig. 1.

### *A. Data*

The dataset has used in this work is an open-access dataset described in [7]. Methods for data acquisition and the description of the experiment and its goal are precisely described in [7], [8]. However, these methods would be briefly mentioned here for more investigation.

In this study, MEDICON 2019[1] for the IFMBE challenge dataset is used to evaluate the proposed method's performance. This dataset consists of 7 sessions of EEG signal during 4 months from 15 participants with Autism Spectrum Disorder (ASD). Each session is divided to train (offline phase) and test (online phase) sections. The train set includes 20 blocks with a specific target and each block contains 80 trials, so training data consists of 1600*7=11200 trials for each participant. At the beginning of each block, the participants were asked to look at a specific object in the scene that has been shown, while in each trial, one of the eight objects in the scene was randomly highlighted. So each object would be highlighted ten times in each block. The final challenge's purpose is to determine the target object for each block. Because the test data label wasn't accessible, train dataset split to test, validation and train. To be more accurate, at first all of the data's blocks shuffled, then seven blocks were picked up as a test dataset and consider the rest for the train and validation phase. To make the result reliable to make comparison with other methods, shuffling process was repeated four times, and every time, test data validated with trained model using train data. Then, obtained accuracy for all of test set averaged and averaged accuracy reported as a final test accuracy.

### *B. Signal Pre-processing:*

According to the challenge description file, the EEG signals were recorded using eight electrodes (C3, Cz, C4, CPZ, P3, P4, POZ). Also, signals have been filtered between 2 and 30 HZ using high-pass and low-pass filters. The sampling rate was set to 250HZ. In addition, since there were no missed data, no data deleted and all of the data has been used. Therefore, as mentioned in [7], data is completely clean and ready to use and no noise rejection or filtering is needed. To be more sure for having a clear data, for every subject noise rejection classifier has been used. In this way, first, all of the trials placed after the other and then, EEGLAB algorithm for noise classification used. Finally, it turned out there is no noticeable noise to be considered. Also, for removing artifacts, Fast ICA has been applied. Since independent component analysis is one of the techniques for blindly separating task-dependent irrelevant sources, it can be used for artifacts source removing. Therefore, the Fast ICA algorithm was applied to EEG signals separately for each subject using an open-source EEGLAB toolbox. By doing so, signals transformed from 8 electrode space to 8 Independent Component (IC) space. It should be mentioned that as ICA is a method for artifact removing, this process is a part of feature extraction, and output is not interpreted as any meaningful EEG source.

For the next step, since the every trial has been recorded 200ms before stimulus and has been continued 1200ms after stimulus, the data from 0 to 1000ms selected [12]. Then, each signal normalized between 0 and 1. After these steps, each trial transformed to the two dimensions matrix (8,250). To make every signals as an inputs of the proposed Neural Network, its shapes changed to be (250, 1, 8). In the next step, train data needed to become balanced since it is highly unbalanced dataset. On the other words, it's every run contains seven zero labels (non-target) and only one target label. Therefore, it's clear that the model fails to train properly while there isn't enough target data to be trained through the network. To mention this problem, different methods of balancing algorithm including over-sampling, under-sampling and resampling such as SMOTE, has been tried and finally, it turns out that changing cost function makes the best result and highest averaged accuracy and recall score on the validation data. The new cost function is defined in Eq.1:

$$\textbf{\textit{Defined CF}} = \text{MSE (FP)} + 7 * \text{MSE (FN)} \qquad (1)$$

This cost function can be modeled as a repetition over-sampling method that means repeating signals of minority class many times to make its size equal to the majority class's size. Therefore, here minority class needs to be repeated seven times. By doing so, the neural network tries to learn the features of target data to make its loss function less.

[1]https://www.medicon2019.org/scientific-challenge/

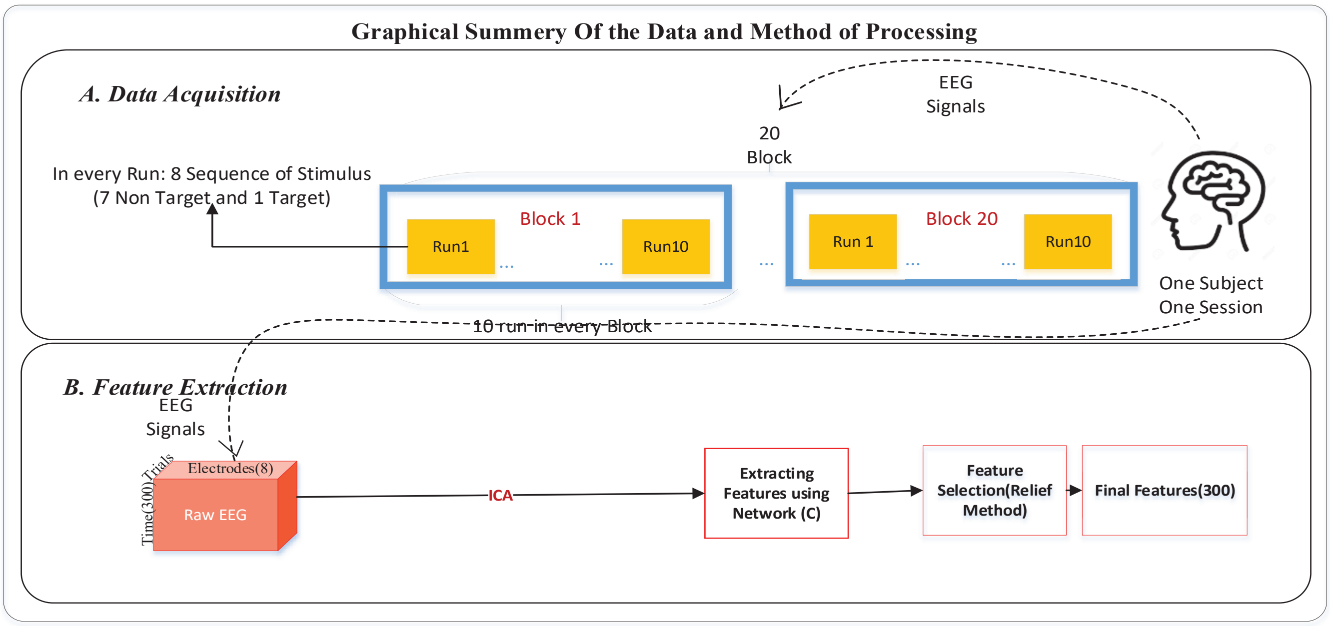


**Figure 1: Summery of Used Data and Proposed Method for detecting P-300 from single-trial EEG signals.**

### *C. Feature Extraction*

Convolutional Neural Network (CNN) is one kind of neural network that uses convolution filter for extracting features from input which are mostly images. In recent years, using a convolutional filter for extracting features from signals, especially bio-signals has been increased. After the convolutional filter, other layers such as batch-normalization layer, pooling layer, and a non-linear function layer (Relu or soft-max) connect to convolutional layer. Batch-normalization layer normalize mini-batch of each layer which helps network to be trained in less epoch as long as it helps network to be more stabilized [16]. One of the CNN's behavioral that makes it so strong for analyzing signals is its adaptability, which means that its structure can be designed due to signals properties. In this study, Two-branch CNN has been applied to extract features to detect P-300 from single-trial EEG. Two branches of the proposed structure consisting small and large filters that are responsible for extracting high frequency and low-frequency patterns. The right-side branch works well for extracting shallow features and the left-side branch works well for extracting deep features. Suggested architecture is shown in Fig.2. Firs, right-sided branch will be interpreted. This branch consists of two convolutional layers and two pooling layers as in has been shown in Fig.2. The first convolutional layer is purely spatial, in other words, it extends through all channels at each time point separately and combines electrodes with each other. After that, the pooling layer reduces extracted features' dimension, and next, the second convolutional layer arrives which acts as an averaging filter through time. The second branch of the structure can be used for extracting temporal-spatial features of EEG signals. This branch has consist of filters with different size for extracting features in different frequencies. In this way, the first layer's filter is 65, meaning that extracting features theta band, considering sampling rate 250 HZ, which is one of the important EEG sub-bands when investigating about attention concept. After the first layer, a pooling layer acts like a down-sample operation and reduce the size of outputs from convolutional layer by 8. Similar to the previous convolution layer, the next convolutional filter extracts features from the alpha band, another important band for investigating attention in EEG signals as discussed. Also, the RelU function has been used for activation function after convolutional layers. At the end of the structure, dropout (0.5) is used to accelerate training and regularization of the network. Also, a fully connected layer as the last layer, not only reduces features but also combines them and makes better features.

We validate the impact of different models components by omitting one component or replacing it with other components, retraining the model, and comparing results. There, four more additional models have been trained:

- **NOT_Branched:** we tried to keep just a branch with a deep structure and small filters.
- **NO_BN:** batch normalization omitted.
- **NO_FullyConnectedLayer:** We omitted the fully connected layer. Also, we changed the number of neurons in the fully connected layer.
- **NO_Dropout: dropout:** was omitted.

When features are extracted, a feature selection algorithm is applied to reduce their dimension. We used the Relief-f algorithm to rank features, then features with positive weight were chose.

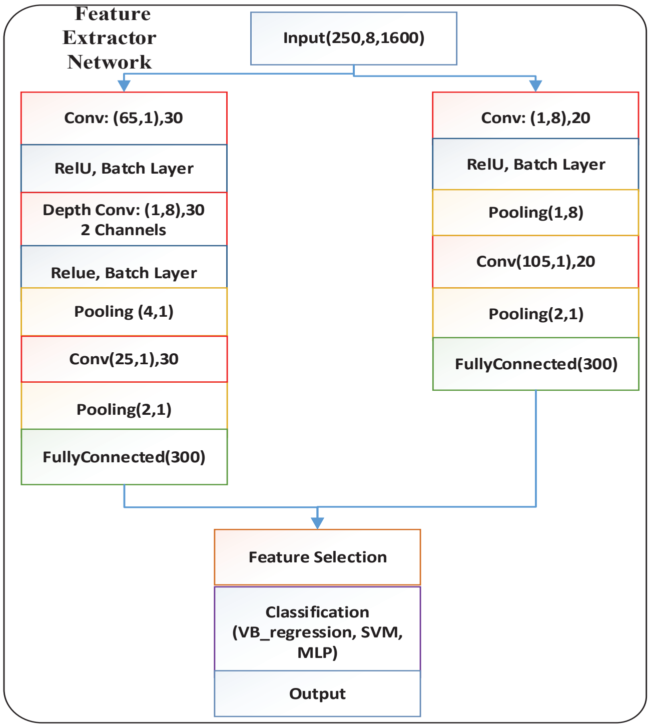


**Figure 2: Feature Extraction Network.**

*D. Classification*

CNN's do not work well in the training and classification of long-term high-level distributed features. As a result, we used three classifiers with the best results including VB-regression, Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP) for classification. Variational Beyesian regression with Automatic Relevance Determination (VB-ARD) is a new logistic regression algorithm that benefits the Variational Bayesian method for finding hyper-parameters. This method works well for sparse features [16]. It's clear that method is logistic regression and for having a classifier from this method, specific threshold is needed. Like most of the time, it has been considered as 0.5. Given result from training data has shown that the (VB-ARD) regression induces high accuracy with low recall. SVM classifier results in medium accuracy and recall, and MLP results in high recall and low accuracy. To dig out in the introduce method and find out each classification method's features and why their combination makes best result, at first, each of classification methods would be elucidate with more detail.

*Support Vector Machine:* Support Vector Machine is one of the most robust machine learning algorithm which is supervised. In this algorithm, instead of all data, most effectiveness data which has been called Support Vector would be trained to set a model. More detail about the method and how support vector data are determined could be find in [17].

*Multi-Layer Perceptron:* Multi-Layer perceptron is another supervised learning algorithm which is a class of feedforward artificial neural network. MLP consists of input, hidden and output layers which each layer except of input layer is a node that uses non-linear activation function. MLP also utilize a backpropagation for training its neurons. More detail about this methods could be find in [18].

*Variational Beyesian Regression with Automatic Relevance Determination (VB-ARD):* Variational Bayesian Regression is also another supervised learning method that has been introduced by *Drugowitsch et al* recently. This method tries to combine traditional Bayesian Decision algorithm and consider a variational bond for it as an optimization method to find the best parameter. And as it has been shown in [19] this method is good for sparse data. To summarize this method, ARD tries to eliminate irrelevant features by optimizing the feature-related weights. Therefore, this algorithm is able to assign an individual hyper-prior to each regression coefficient separately and find the most relevant features.

It showed in Fig.4 that each algorithms works well in keeping one score high while it fails to keep all of the validating score like recall and accuracy high enough. Therefore, the voting algorithm was applied to predict the final label. In the voting algorithm, the predicted label of every classifier intensified with a weight which has obtained with the try and error approach. Then, weighted predicted label sum to each other, and in every run, epoch with highest summed label assumed as attended epoch and all other epochs assumed as a not-attended one.

*E. Target Detection:*

To compare our results with the previous approach, the target of each block was calculated. Target detection was the main goal of the challenge. To detect the target of each block, first for each classifier, probability of being a target for each $i^{th}$ object in $j^{th}$ run of $k^{th}$ block calculated using Eq.2:

$$P_{(i)1}^{(k)} = \sum_{j=0}^{N-1} P_{(i,j)1}^{(k)} \qquad (2)$$

So, the object with the most probability can be assumed as a block's label. If for all of the classifiers, the weighted approach like what is used in finding binary labels will be investigated. It means the formula for calculation probability change as Eq.3:

$$P_{(i)1}^{(k)(Final)} = 1.1P_{(i)1}^{(k)(SVM)} + 0.9P_{(i)1}^{(k)(VB-Regression)} + 0.8P_{(i)1}^{(k)(MLP)} \qquad (3)$$

Then, object with maximum probability will be chose as a final label of one block.

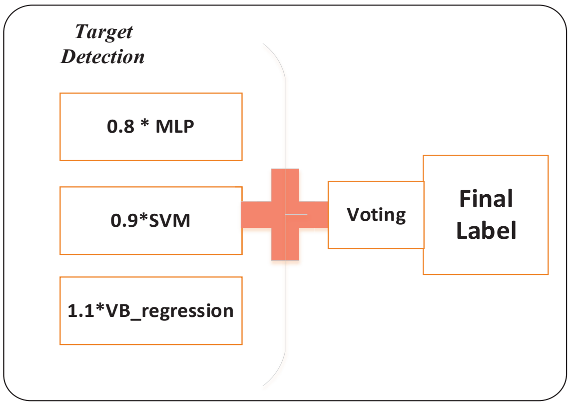


**Figure 3: Target detection method.**

## III. RESULTS AND DISCUSSION

Detailed accuracy is shown in Fig.2. As can be observed in Fig.4, obtained accuracy from VB-regression is higher than MLP, and MLP's accuracy is higher than SVM. The averaged accuracy through all subjects obtain from SVM, VB-regression, and MLP is 72.4%, 83.7%, 76.8%, respectively. Also, the average recall through all participants is 65.9%, 63.3%, and 71.7%, respectively.

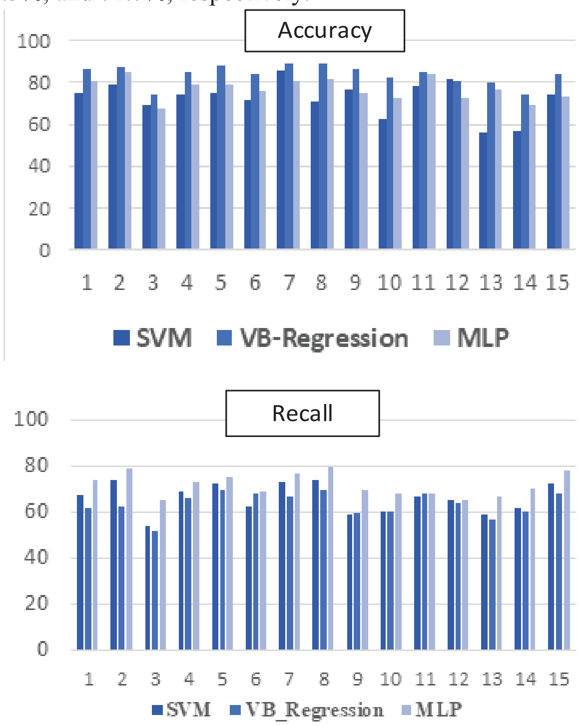


**Figure 4: Comparison of accuracy and recall score between three methods**

The final accuracy which has calculated using the voting approach has been shown in Fig.5. It can be observed that with this approach, the average of all subject's accuracy reaches 91.13% while its recall reaches 64.51%.

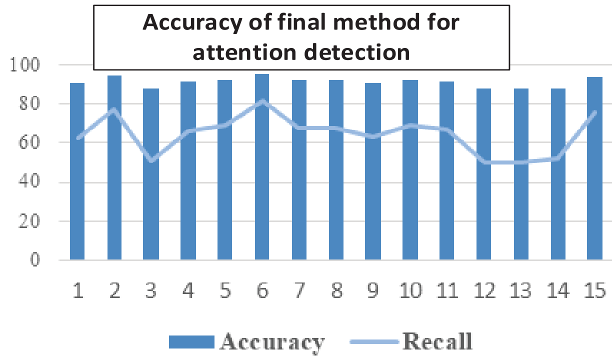


**Figure 5: With using voting method, accuracy of detecting attention has increased.**

At the other sight, we can investigate the accuracy of each session through all participants. Fig.6 shows the plot describing accuracy from all sessions. From this plot, it can be proposed that the accuracy of every session is higher than the previous one, except for sessions five and seven. Decreased accuracy for session five can be interpreted as a reduction in the subject's performance since there is a long interval between sessions five and four (four first sessions were held weekly and the next sessions held monthly).

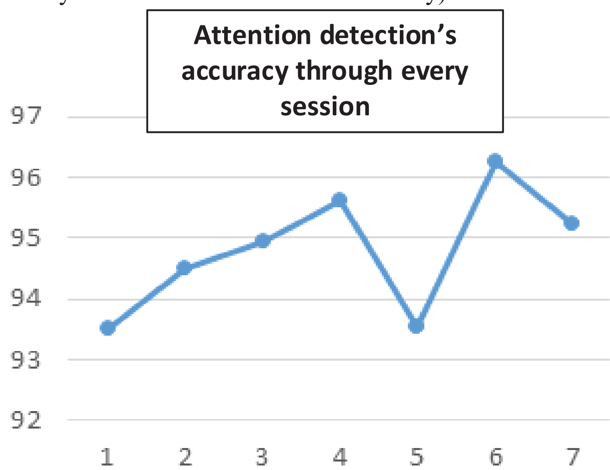


**Figure 6: Attention's detection accuracy in every session.**

To obtain the goal of the competition, a target of blocks is calculated. The obtained accuracy of target detection for each subject per session is shown in Fig.7. This figure shows that the introduced approach could reach an accuracy of 94.85%. To compare our method with previous method, we consider target detection's accuracy. Considering previous work, it can be assumed that our approach could be reach higher accuracy since the best result without our method is 92.38% which comes from combination of EEGNet network, which is a famous network for interpreting EEG signals, and GRU network [17]. In addition, results from [18], [19] shows that the accuracy of other CNN and network architecture that had a good performance in other research is so less than our developed approach. Therefore, it can be concluded that our method is able to detect attention in EEG signals with higher accuracy than other previous approaches.

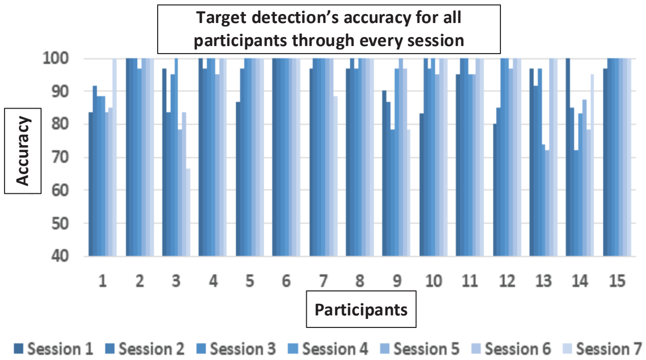


**Figure 7: Accuracy of detecting final target object using final voting method.**

## IV DISCUSSION AND CONCLUSION

In this paper, we proposed a novel algorithm based on extracting features from Convolutional Neural Network and voting method for classifiers. By using our method, we could reach an accuracy of 94.85% for detecting the target's label of every block. It is noticeable that we didn't access test records, so we split train data to train, validation and test, and validate our method with 4-fold cross-validation. Further works are needed for understanding the performance of our method while the number of repetitions in every run decrease.

## ACKNOWLEDGMENT

We thank Carlos P.Amaral and his team for providing us with their data. Also, the authors would like to thank the organizers of the MEDICON IFBME challenge for providing us with other method's result.

## REFERENCES

[1] Sutton, S., Tueting, P., Zubin, J., John, E.R.: Information delivery and the sensory evoked potential. Science 155(3768), 1436–1439 (1967)
[2] Wolpaw, J., Wolpaw, E.: Brain-Computer Interfaces: Principles and Practice. Oxford University Press, USA (2012)
[3] ASDEU. Autism Spectrum Disorders in the European Union. Executive summary (2018)
[4] Leigh, J.P., Du, J.: Brief report: forecasting the economic burden of Autism in 2015 and 2025 in the United States. J. Autism Dev. Disord. 45, 4135–4139 (2015)
[5] World Health Organization. Autism spectrum disorders. https://www.who.int/news-room/ fact-sheets/detail/autism-spectrum-disorders. Accessed 20 June 2019
[6] Amaral, C.P., Simões, M.A., Mouga, S., Andrade, J., Castelo-Branco, M.: A novel brain computer interface for classification of social joint attention in autism and comparison of 3 experimental setups: a feasibility study. J. Neurosci. Methods 290, 105–115 (2017)
[7] Amaral, C., Mouga, S., Simões, M., Pereira, H.C., Bernardino, I., Quental, H., Playle, R., McNamara, R., Oliveira, G., Castelo-Branco, M.: A feasibility clinical trial to improve social attention in autistic spectrum disorder (ASD) using a brain computer interface. Front. Neurosci. 12, 477 (2018)
[8] Kothe, C.A., Makeig, S.: BCILAB: a platform for brain–computer interface development. J. Neural Eng. 10, 056014 (2013). https://doi.org/10.1088/1741-2560/10/5/056014
[9] Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K.R.: Single-trial analysis and classification of ERP components-a tutorial. Neuroimage (2011). https://doi.org/10.1016/j. neuroimage.2010.06.048
[10] Zander, T.O., Kothe, C., Welke, S., Roetting, M.: Utilizing secondary input from passive brain-computer interfaces for enhancing human-machine interaction. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2009). https://doi.org/10.1007/978-3-642-02812-0_86
[11] Ferrez, P.W., Del, R., Millán, J.: Error-related EEG potentials generated during simulated brain-computer interaction. IEEE Trans. Biomed. Eng. (2008). https://doi.org/10.1109/TBME.2007.908083
[12] Haibo He H and Garcia E 2009 Learning from imbalanced data IEEE Trans. Knowl. Data Eng. 21 1263–84
[13] Hoffmann U, Vesin J M, Ebrahimi T and Diserens K 2008 An efficient P300-based brain–computer interface for disabled subjects J. Neurosci. Methods 167 115–25
[14] Wainer, A.L., Ingersoll, B.R.: The use of innovative computer technology for teaching social communication to individuals with autism spectrum disorders (2011). https://doi.org/10.1016/j.rasd.2010.08.002
[15] Treder, M.S., Blankertz, B.: (C)overt attention and visual speller design in an ERP-based braincomputer interface. Behav. Brain Funct. (2010). https://doi.org/10.1186/1744-9081-6-28
[16] Drugowitsch, J.: Variational Bayesian inference for linear and logistic regression. ArXive-prints (2013)
[17] Zhao, H., Yang, Y., Karlsson, P., & McEwan, A. (2020). Can recurrent neural network enhanced EEGNet improve the accuracy of ERP classification task? An exploration and a discussion. Health and Technology, 10(4), 979–995. https://doi.org/10.1007/s12553-020-00458-x
[18] David Borra, Silvia Fantozzi, Elisa Mogosso. "Convolutional neural network for a P300 brain-computer interface to improve social attention in autistic spectrum disorder" Mediterranean Conference on Medical and Biological Engineering and computing
[19] Eduarro Santamirra, Victor Martinez, Javior Gomez, Roberto Honero. "Deep Learning Based on the Combination of Convolutional and Recurrent Layers for ERP_Based Brain-Computer Inference " Mediterranean Conference on Medical and Biological Engineering and computing