Linear vs Nonlinear Classification of Social Joint Attention in Autism Using VR P300-Based Brain Computer Interfaces

Lucia de Arancibia¹✉, Patricia Sánchez-González^{1,2} ,
Enrique J. Gómez^{1,2} , M. Elena Hernando^{1,2} ,
and Ignacio Oropesa¹

¹ Biomedical Engineering and Telemedicine Centre, ETSI Telecomunicación,
Center for Biomedical Technology, Universidad Politécnica de Madrid,
Madrid, Spain

lucia.dearancibia.casillas@alumnos.upm.es, {psanchez,
egomez, mariaelena.hernando, ioropesa}@gbt.tfo.upm.es

² Centro de Investigación Biomédica en Red, Biomateriales y Nanomedicina,
Madrid, Spain

Abstract. VR P300-based BCI has proven to be a suitable method for training social attention skills in youngsters with autism spectrum disorder (ASD). In this study, we present a method that could be used in such an application to identify which object the user is paying attention to in a virtual environment by means of EEG recordings only. Temporal and time-frequency features were explored. Furthermore, the prediction accuracy of linear and nonlinear classification methods was assessed and compared, along with their computational times and complexity, and linear discriminant analysis (LDA) yielded the best overall performance (82%). The successful predictions and low computational times demonstrate the feasibility of the proposed solution for a VR-BCI neurorehabilitation tool.

Keywords: BCI · P300 · Autism Spectrum Disorder ·
Linear discriminant analysis · Support vector machines · Virtual reality ·
Neurorehabilitation

1 Introduction

Autism Spectrum Disorder (ASD) is currently one of the most prevalent serious developmental conditions in developed countries. In Europe, average prevalence is of 1 in 89 children aged 7–9 years, and estimated direct annual costs range from 1,534€ to 22,378€ per individual [1]. Moreover, predictions suggest that by 2025 the total costs derived from ASD in the USA will rise to over \$450,000 million [2].

ASD symptoms are present from early childhood and limit everyday activities throughout patients' life. A distinctive symptom is impaired social joint attention: failure to share the focus of attention on an object with other individuals. This poses greater difficulties in understanding social cues and thus interacting with others [3].

Amaral et al. [4] introduced a promising, novel VR P300-based BCI paradigm for the neurorehabilitation of youngsters with ASD. On the hypothesis that subjects with ASD can learn joint attention skills by internalizing the response to a given social cue, they created a task where attentional mental state is measured through the detection of oddballs, which in turn induces a P300 signal. This training tool has demonstrated to be beneficial and very promising in the field of social cognition training in ASD [5]. In order to develop it, it is necessary to design and implement a fully functioning BCI that can correctly identify the P300 component of the EEG. Different studies have tackled this specific problem for BCI applications.

The first P300-based BCI made use of temporal information of filtered EEG data such as peak values, area under the curve and covariance evaluation with a P300 template to discriminate between target and non-target signals [6]. More recent studies include more complex features based on temporal, time-frequency or spatial domains as input to machine learning classifiers [7–11]. Additionally, in Krusienski et al. [12] one can find a comparison of several classification algorithms for P300 detection. Their results suggest that linear classifiers may be sufficient to correctly identify P300 signals. However, classification accuracies vary widely among subjects and thus it is not clear whether simple classifiers are able to provide good generalization.

In this study, we propose a combination of temporal and frequency features to compare performance of linear and nonlinear classification approaches to the problem and describe our best solutions.

2 Methods

In order to carry out this study, data from clinical trial NCT02445625 was used, as described in [4]. The dataset includes the EEG recordings from 15 participants that underwent 7 sessions each throughout 4 months.

The paradigm proposed by Amaral et al. reduces an 8-group classification problem to a binary discrimination problem, i.e., the discrimination between target events (1 in 8) and non-target events (remaining 7 events in a run). For our predictions, we averaged all EEG events corresponding to the same object in the same block and obtained 8 averaged signals per block (number of signals averaged is equal to the number of runs per block). We then extract features from the averaged signals and calculated the likelihood of them comprising a P300 component. The object whose corresponding signal yields maximum probability is chosen as the predicted target object of the block.

2.1 Feature Extraction

- **Features based on time domain.** The P300 peak typically appears around 200 ms to 300 ms after the stimulus. For each channel in the subset, the 250 samples corresponding to the time slot from stimulus onset (0 ms) until 1000 ms later were extracted and downsampled by average (decimation factor = 10), and the resulting samples from each channel were concatenated obtaining a single vector of 200

samples from each observation. Downsampling allowed us to reduce the number of features and decrease classification time without missing important information.

- **Features based on continuous wavelet transform (CWT).** The CWT was computed with a Mexican Hat wavelet on scales corresponding to the delta (0.5 to 4 Hz) and theta (4 to 8 Hz) bands, following [7]. To further reduce computation time as well as extract the most differential features and eliminate redundancies, the t-CWT technique [8] was applied. This technique applies Student's t-statistics to find points of maximum statistical difference between average P300 and non-P300 in the time–frequency domain. In our case, 128 points were selected for each of the 8 channels and concatenated, adding a total of $128 * 8 = 1024$ time-scale features.
- **Feature reduction techniques.** The temporal and frequency features make a total of 1224 features per signal to be used as input to a classifier. Such a large number can slow down the training process and cause overfitting. Therefore, principal component analysis (PCA) reduction was introduced. The first 120 principal components were used, since they are enough to cover 99% of the variance in the data.

2.2 Classification Algorithms and Training Considerations

To solve the EEG classification problem, we explored linear machine learning approaches such as linear discriminant analysis (LDA) and support vector machines with linear (LSVM) kernel, and a more complex support vector machine with radial (RSVM) kernel. For the latter, hyperparameters gamma and C were selected for each patient during training through bayesian optimization with 5-fold cross validation. The search was evaluated for values ranging from 10^{-3} to 10^3 for both hyperparameters.

The main problem found for classification of the EEG signals was that the data is highly imbalanced: only one in every 8 EEG segments contains a P300. Thus, by classifying all inputs as non-target, an accuracy of 87.5% would be achieved. To avoid this problem, oversampling and boosting methods were explored.

In order to train the classifiers, two approaches were attempted and evaluated. At first, inter-subject classification (train a single classifier using all the training data from all subjects' calibration phases) was approached. However, this alternative ignored subject variability, and the obtained models yielded lower overall accuracy results. To test whether this was due to excessive generalization, we approached the problem as a subject-specific classification problem (train an individual classifier for every subject using only his/her individual calibration data). This approach accounts only for intra-subject variability and fits to the patient's particularities. It demonstrated better overall accuracies and was used for the final predictions.

In order to validate the implemented algorithms and choose the best method for phase II predictions, they were validated on phase I online (test) data. The features and classifiers that yielded best performances (measured in terms of low training times and accurate target object detection) were then used to make predictions for phase II.

3 Results and Discussion

3.1 Phase I Predictions and Validation

Table 1 shows the training, P300 classification and target object prediction performances of the different methods used and tested with phase I online data.

Table 1. Prediction of phase I online data. F₁ score refers to average value for P300 detection (binary classification problem), while online accuracy refers to average target object detection accuracy. In parenthesis maximum and minimum values per session are given. Training and prediction times are an average value per session. In bold: best result (LDA)

Features	Classification	Other methods	Training time	F ₁ score	Online accuracy (%)
Temporal, t-CWT features and PCA	LDA	–	0.037 s	0.76 (0.38–0.94)	80 (38–100)
	LSVM	–	2.82 s	0.70 (0.41–0.91)	75 (40–94)
	RSVM	–	23.6 s	0.76 (0.5–0.96)	79 (50–94)
	LDA	Oversampling	0.1 s	0.72 (0.28–0.91)	78 (32–94)
	LDA	Boosting	5.2 s	0.73 (0.34–0.94)	79 (34–100)

For subject-specific training, LDA was the most accurate as well as the fastest classification method. Non-linear RSVM did not yield better performance than simpler linear classifiers. However, it did improve the accuracy for subject 13, something that was not achieved through other proposed methods. This suggests that linear models might be too generalist and not capable of particularizing in the most complex cases.

Oversampling and boosting were applied to see whether they improved results of the best classifier (LDA). Both methods slightly improved the accuracy for some sessions but did not improve the overall performance for machine learning classifiers.

3.2 Prediction of Phase II Online Data

The above-mentioned features with PCA and an LDA classifier trained with subject-specific data yielded a final overall phase II online data prediction accuracy of 82%. Although for some sessions accuracy was as low as 50%, we accomplished an accuracy above 80% for 41 sessions (out of 60).

The results for P300 signal classification are above average for some comparable state-of-the-art studies such as [9], which achieves a mean classification accuracy of 48.15% with 5 EEG epochs; [10], that reached accuracies up to 79%; or [11], which reports a P300 detection F1 score of 0.52.

Moreover, we achieve an improvement in target object detection with respect to the study by Amaral et al. [4]. Their highest accuracy was 72% for subjects with ASD compared with ours of 100%, accomplished by averaging only 5 EEG events. These results are also higher than their accuracies for healthy subjects (mean accuracy of 80%).

Our study suggests that subject-specific classification outperforms inter-subject models. This implies that the BCI should be trained for every new patient. However, our methods accomplish classification and object detection in a considerably short time. Around 30 s are needed to train a classifier with enough calibration data. Less than 0.1 s are needed to classify an EEG event. A total of around 0.1 s are needed to detect the target object of a block. This means the user needs to wait less than 1 s since the block is over until he/she receives feedback. Such a result is very advantageous for the pursued neurorehabilitation application for patients with ASD.

In conclusion, even though the proposed methods can be improved, the best overall results demonstrate their feasibility. The desirable next step would be to test it live in a real situation.

References

1. ASDEU. Autism Spectrum Disorders in the European Union. Executive summary (2018)
2. 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)
3. World Health Organization. Autism spectrum disorders. <https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders>. Accessed 20 June 2019
4. Amaral, C.P., Simões, M.A., Mougá, 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)
5. Amaral, C., Mougá, 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)
6. Farwell, L.A., Donchin, E.: Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr. Clin. Neurophysiol.* **70**, 510–523 (1988)
7. Demiralp, T., Ademoglu, A., Schürmann, M., Basar-Eroglu, C., Basar, E.: Detection of P300 waves in single trials by the wavelet transform (WT). *Brain Lang* (1999)
8. Bostanov, V., Kotchoubey, B.: The t-CWT: A new ERP detection and quantification method based on the continuous wavelet transform and Student's t-statistics. *Clin. Neurophysiol.* **117**(12), 2627–2644 (2006)
9. Agapov, S.N., Bulanov, V.A., Zakharov, A. V., Sergeeva, M.S.: Wavelet algorithm for the identification of P300 ERP component. [arXiv:1611.00033](https://arxiv.org/abs/1611.00033) [q-bio.NC] (2016)

10. Manor, R., Geva, A.B.: Convolutional neural network for multi-category rapid serial visual presentation BCI. *Front. Comput. Neurosci.* **9**, 146 (2015)
11. Liu, M., Wu, W., Gu, Z., Yu, Z., Qi, F., Li, Y.: Deep learning based on batch normalization for P300 signal detection. *Neurocomputing* **275**, 288–297 (2018)
12. Krusienski, D.J., Sellers, E.W., Cabestaing, F., Bayoudh, S., McFarland, D.J., Vaughan, T. M., Wolpaw, J.R.: A comparison of classification techniques for the P300 speller. *J. Neural Eng.* **3**(4), 299 (2006)