IOP Conference Series: Materials Science and Engineering

PAPER • OPEN ACCESS

Solving classification problems of visual evoked potentials for the brain-computer interfaces

IOP Conf. Series: Materials Science and Engineering **862** (2020) 052051 doi:10.1088/1757-899X/862/5/052051

Solving classification problems of visual evoked potentials for the brain-computer interfaces

V A Bulanov[1], A V Zakharov[2] and S S Chaplygin[2]

[1] IT Universe LLC, 3, Eroshevskogo str., Samara, 443086, Russia
[2] Samara State Medical University, 89, Chapaevskaya str., Samara, 443099, Russia

E-mail: zakharov1977@mail.ru

Abstract. Development of the electroencephalogram-based neurocomputer interfaces requires application of the efficient algorithms for signal analysis. One of the methods of neurocomputer interface development is based on using single visual evoked potentials for characteristics control. However, it is a difficult task, requiring a combination of various methods of signal processing such as Blind Source Separation method, machine learning method and other modern mathematical and computational tools. In this paper, we drew a comparison between various classifiers for the visual evoked potentials recognition problem. The electroencephalogram records analyzed in this paper were published in the public domain.

1. Introduction

Neurotechnologies, such as neurocomputer interfaces (NCI), are an important branch of modern science. One of the main concepts for the development of neurocomputer interfaces (NCI) using electroencephalogram (EEG) is a visual evoked potentials (VEP) recognition that occurs as a human brain response to the presence of external visual stimuli [1]. A number of ongoing studies are dedicated to the VEP classification algorithms improvement [2-6].

For researchers, it is becoming a common practice to share in the public domain the data received by using expensive lab equipment or in complex experiment conditions. This might be the next step in science development. In this paper we used the data published in the public domain by a group of scientists L. Vareka, P. Bruha, R. Moucek [7].

2. Objective of research

The objective of the research is to compare various classifiers for the target and non-target single visual evoked cortical potential recognition.

3. Materials and methods

In this research, the data published in the public domain by a group of scientists was used [8].

The database contains electroencephalograms of 19 test subjects that were presented with visual stimuli and tested under the three-stimulus paradigm – a type of the odd-ball paradigm [9]. EEG were recorded by electroencephalograph (the device make is not specified) with 19 channels and presented in *BrainVision* format. Flashes of three (red, green and yellow) LED indicators were used as visual stimuli. Stimuli were presented in a random order with a predefined frequency distribution: 83% of non-target stimulus presentations (red LED flash), 13.5% of target stimulus presentations (green LED), 3.5%

of distracting stimulus presentations (yellow LED). The stimulus duration was 500 ms, the inter-stimulus interval (*ISI*) was 1000 ms. The experiment was conducted in three phases with a short break, each test subject was presented with 30 target, 184 non-target and 8 distracting stimuli.

The detailed description of the experiment is provided in work of researchers [7].

In our study datasets with numbers 76, 85, 86, 87, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 104, 105, 106 were used. The classification *features* were selected according to the method suggested in the paper [2]. The EEG-signal from seven leads (*P3, P4, O1, O2, T5, T6, Pz*) was used. The features vector was formed as follows: three time-windows were defined in the EEG-signal of each lead from each epoch with a length of 1500 ms from the stimulus presentation. Time-windows interval from the beginning to the end equaled to 80-250, 250-600, 600-800 ms from the beginning of the epoch. This selection method is supported by general characteristics of VEPs and averaged epoch analysis for all test subjects on the lead *Pz*. The amplitude values for each time-window were averaged by each lead. The averaged values were classification features vector components. Consequently, the vector consisted of 21 values (3 windows * 7 channels). The features vector and further classification were created in the statistical data processing *R* software version 3.3.1 (*https://www.r-project.org/*).

For comparison, the most efficient for the binary classification problem classifiers were selected based on the research data provided in the work [10]. The classifiers belonged to different groups – neural networks, decision trees, Bayes method, nearest neighbors, support vectors method. All the presented algorithms are described in detail and have been used in various applications [11-13]. The modifications of algorithms selected for our research purposes were the following:

averaged NNet – Neural Networks Using Model Averaging, an ensemble method for averaging the operation of the same neural network.

ELM – Extreme Learning Machines, a simple neural network with one hidden level. Weights connecting inputs to the hidden level are assigned randomly and are not updated.

GBM – Gradient Boosting Models, an ensemble machine learning method. The method represents one of the decision trees implementations.

k-NN – k-Nearest Neighbors algorithm, classification based on the distance for *k* nearest neighbors. Can be applied with different metrics.

LDA – Fisher's*Linear Discriminant Analysis*. This well-known method of classification is based on the application of Bayes theorem.

Naive Bayes – implementation of the classification method based on the application of Bayes theorem. Implies the independence of features in the analyzed dataset.

Random Forest – an ensemble machine learning method. Well-known machine learning algorithm.

SVM – Support Vector Machines. In this paper, the classifier was applied with two *kernels*: linear and Gaussian. They are denoted by *SVM linear kernel* and *SVM rbf kernel* respectively.

In the purpose of classifiers training the package *caret* version 6.0 (*https://cran.r-project.org/web/packages/caret/index.html*) was used, which can set parameters automatically for a wide range of algorithms. During the training all classifiers were tested by *5-fold cross validation* method established in the work [11].

Each classifier had been trained 30 times (cycles). In each training cycle, 380 target and non-target epochs were randomly selected from the entire array of epochs (~ 12,680 epochs). After the training was over, the accuracy of classifier recognition was tested on each of the 19 datasets. In order to assess the performance of the classifiers, the accuracy indicators (for reliability of attributing an epoch to a target or non-target category) and the *Cohen's kappa* (for consistency between true-positive, true-negative, false-positive, and false-negative recognition cases) were selected.

4. Results and discussion

The research results are presented in figures 1, 2 and tables 1, 2.

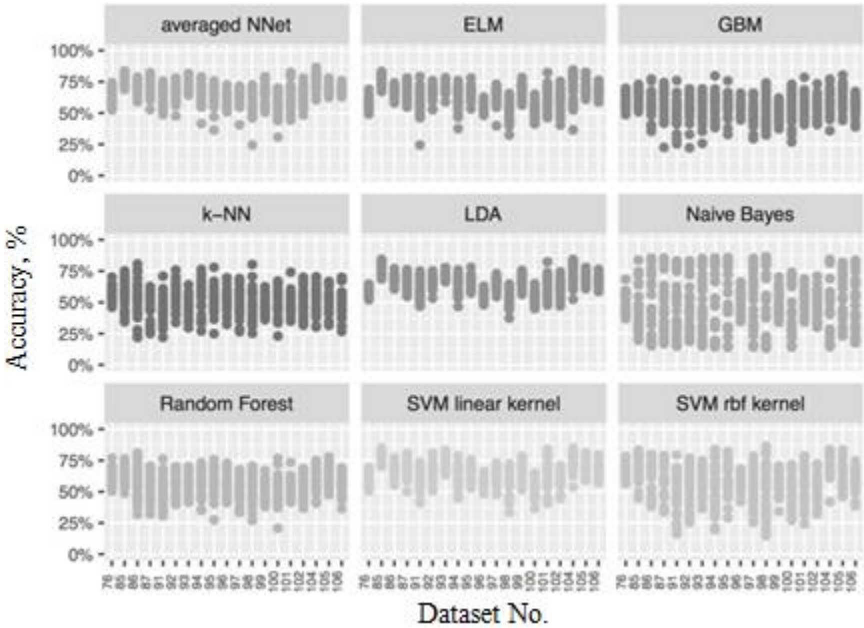

Figure 1. Classification accuracy of recognition target and non-target single VEPs by nine different classifiers. For each dataset, the accuracy is presented for each of the 30 cycles of classification.

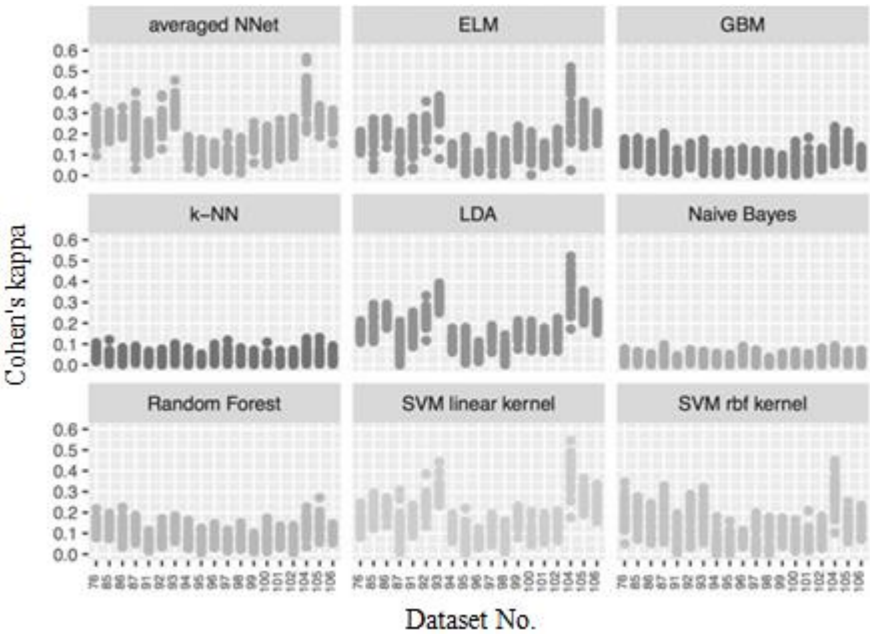

Figure 2. Cohen’s kappa for classification with nine different classifiers. For each dataset, the kappa is presented for each of the 30 cycles of classification.

Table 1. Classification accuracy of recognition target and non-target single VEPs by nine different classifiers. Averaging of 19 datasets.

Classifier	Average accuracy and standard deviation, %	Cohen's kappa

averaged NNet	66.42±8.65	0.19
ELM	64.15±9.05	0.16
GBM	55.45±9.44	0.08
k-NN	50.71±10.48	0.03
LDA	64.91±8.29	0.17
Naive Bayes	46.1±21.01	0.01
Random Forest	55.62±9.67	0.09
SVM linear kernel	64.86±9.41	0.17
SVM rbf kernel	58.44±13.69	0.13

The analysis of the obtained results shows that all indicators are very close to a normal distribution. Therefore, the sample can be described by both average value and standard deviation. Table 1 presents averaged accuracy and Cohen's kappa classifications. Table 2 presents detailed data about classification accuracy for each dataset.

LDA and its regularized modifications *swLDA* and *shrinkage-LDA* are used in the work [2]. For further development of the presented method, a new *averaged NNet* classifier is offered, having higher recognition accuracy than *LDA* and potential for further accuracy increase.

Averaged NNet has higher accuracy and Cohen's kappa than *LDA* classifier. Whereas, *averaged NNet* shows a slight accuracy increase in comparison to *LDA* (on average +1,51%), it shall be noted that *averaged NNet* has more training options. Consequently, it provides an opportunity to increase the classification accuracy through the effective parameters selection. Due to some of the specifics of its implementation, *LDA* has almost zero variable parameters, so the adjustability is not possible.

SVM linear kernel classifier also performed well during the tests. Its accuracy is slightly lower than in *averaged NNet* (on average –1,56%) and *LDA* (on average –0,05%). The application of *SVM* group classifiers for the purpose of VEP classification is very common and fully studied, the comparison between *LDA* and *SVM* classifiers is presented in the work [14].

Our classifiers accuracy study contributes to and extends the well-known studies on classifiers comparison. For example, some authors [15] compare the accuracy of *Fisher LDA, swLDA, PCM (Pearson's Correlation Method), SVM linear, and SVM Gaussian* classifiers. The comparison between *Bayesian LDA, Fisher LDA, swLDA, SVM linear, SVM Gaussian, multilayer feed-forward neural network, feature extraction method* classifier is represented in the study [16].

5. Conclusion

Obtaining high classification accuracy was not one of the objectives considered in this paper. The study just compared classification accuracy between several classifiers on the same input data. For further development of the EEG pattern recognition method based on the linear *LDA* classifiers, presented in the works [2, 3, 17-19], we suggested using *averaged NNet* classifier based on the simple neural network, because it performed well on solving binary classification problems for target and non-target VEP. The accuracy of *averaged NNet* classifier can be increased, if the EEG signal pre-processing, spatial filtration and various methods of reducing the dimension of classifying features, as well as setup of the classifier's parameters, are applied.

IOP Conf. Series: Materials Science and Engineering **862** (2020) 052051 doi:10.1088/1757-899X/862/5/052051

Table 2. Classification accuracy of recognition target and non-target single VEPs by nine different classifiers. The accuracy (in percent) is presented for each of the 19 datasets. The best accuracy rate in a row (i.e. for each dataset) is marked with bold font. The best accuracy rate in a column (i.e. for each classifier) is marked with underlined italic font. Note: result 77,64% is the best for the dataset 85 and for the classifier SVM linear kernel.

Dataset No.	aver-aged NNet	ELM	GBM	k-NN	LDA	Naive Bayes	Random Forest	SVM linear kernel	SVM rbf kernel
76	63.54	58.62	61.14	<u>*56.55*</u>	58.80	50.31	63.55	58.40	**66.97**
85	<u>*77.38*</u>	<u>*77.08*</u>	<u>*63.70*</u>	56.29	<u>*77.45*</u>	45.96	<u>*64.43*</u>	<u>**77.64**</u>	69.89
86	**71.57**	68.35	56.23	54.57	68.97	42.42	57.90	68.67	62.70
87	**70.74**	67.05	55.49	48.67	67.31	41.02	54.86	68.64	60.21
91	**62.76**	57.81	49.72	46.74	59.38	41.42	50.50	59.39	51.05
92	**67.99**	67.04	54.37	52.99	67.76	40.98	54.92	66.79	59.92
93	**72.38**	70.99	54.00	49.23	71.97	43.47	55.46	71.82	59.85
94	**66.04**	62.21	56.11	51.05	63.88	50.41	55.13	63.84	55.53
95	64.83	64.66	55.96	51.47	65.19	48.18	57.85	**65.21**	59.03
96	**59.75**	56.34	55.50	52.77	56.54	50.74	54.61	57.21	56.66
97	63.36	63.16	50.35	50.81	**64.26**	45.22	51.31	63.76	49.96
98	**58.78**	54.17	51.81	48.32	54.86	43.51	53.26	55.20	52.71
99	65.80	66.98	50.97	47.54	**67.26**	49.67	51.21	66.31	55.87
100	**55.91**	52.43	53.47	47.28	53.33	43.92	52.05	51.57	49.35
101	**63.00**	62.13	54.56	49.21	62.82	51.39	53.44	62.44	53.10
102	**63.46**	61.10	54.11	48.93	61.95	47.83	52.92	61.48	54.63
104	**75.53**	71.66	59.32	51.01	73.99	42.04	58.24	74.82	<u>*72.45*</u>
105	**70.17**	69.24	61.17	50.76	69.41	<u>*52.37*</u>	60.02	69.90	60.56
106	68.98	67.75	55.57	49.34	68.20	45.14	55.06	**69.30**	59.98

Acknowledgement
This paper is completed with financial support provided by the Ministry of Science and Higher Education of the Russian Federation, unique project ID RFMEFI60418X0208.

References
[1] Guger C, Krausz G, Allison B Z and Edlinger G 2012 Comparison of dry and gel based electrodes for P300 brain-computer interfaces *Front. Neurosci* **6** 60
[2] Blankertz B, Lemm S, Treder M, Haufe S and Müller K R 2011 Single-trial analysis and classification of ERP components - a tutorial *Neuroimage* **56(2)** 814–25
[3] Cecotti H and Ries A J 2017 Best practice for single-trial detection of event-related potentials: Application to brain-computer interfaces *International Journal of Psychophysiology* **111** 156-69
[4] Gao J, Tian H, Yang Y, Yu X, Li C and Rao N 2014 A novel algorithm to enhance P300 in

single trials: Application to lie detection using F-score and SVM *PLoS One* **9(11)** e109700
[5] Stankevich V F, Pantiukhin Y I and Spitsyn V G 2016 Towards P300 based brain-computer interface: different approaches of classification *Fundamental Research* **5(3)** 507-14
[6] Stokes M and Spaak E 2016 The Importance of Single-Trial Analyses in Cognitive Neuroscience *Trends in Cognitive Sciences* **20(7)** 483–6
[7] Vareka L, Bruha P and Moucek R 2014 Event-related potential datasets based on a three-stimulus paradigm *Gigascience* **3(1)** 35
[8] Agapov S N, Bulanov V A, Zakharov A V and Sergeev M S 2017 Comparison of classifiers in the problem of classification of unique visual called potentials of the human brain *Pavlov Magazine of higher nervous activity* **67(4)** 521-6
[9] Dudáček K, Mautner P, Mouček R and Novotný J 2011 Odd-ball protocol stimulator for neuroinformatics research *Proc. Int. Conf. on Applied Electronics, AE 2011* (Pilsen, Czech Republic: University of West Bohemi) pp 107–10
[10] Fernández-Delgado M, Cernadas E, Barro S and Amorim D 2014 Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? *J. Mach. Learn. Res.* **15** 3133–81
[11] Lantz B 2013 *Machine Learning with R* (Birmingham, UK: Packt Publishing Ltd.)
[12] James G, Witten D, Hastie T and Tibshirani R 2013 *An Introduction to Statistical Learning*. (Berlin, Germany: Springer)
[13] Ripley B D 1995 *Pattern Recognition and Neural Networks* (Cambridge, UK: Cambridge University Press) p 403
[14] Kuhn M and Johnson K 2013 *Applied Predictive Modeling* (Berlin, Germany: Springer)
[15] Krusienski D J, Sellers E W, Cabestaing F, Bayoudh S, McFarland D J, Vaughan T M and Wolpaw J R 2006 A comparison of classification techniques for the P300 speller *J. Neural Eng.* **3(4)** 299–305
[16] Manyakov N V, Chumerin N, Combaz A and Van Hulle M M 2011 Comparison of classification methods for P300 brain computer interface on disabled subjects *Comput. Intell. Neurosci.* **2011(2)** 1–12
[17] Bobrov P D, Korshakov A V, Roshchin V Yu and Frolov A A 2012 Bayes method for implementation of the brain-computer interface based on the movement presentation *Journ. Vyssh. Nerv. Deiat. Im.* **62(1)** 1-11
[18] Ganin I P and Kaplan A Y 2014 The P300-based brain-computer interface: presentation of the complex "flash + movement" stimuli *Journ. Vyssh. Nerv. Deiat. Im.* **64(1)** 32-40
[19] Frolov A, Húsek D, Bobrov P 2011 Comparison of four classification methods for brain-computer interface *Neural Network World* **21(2)** 101