24th ABCM International Congress of Mechanical Engineering
December 3-8, 2017, Curitiba, PR, Brazil

COBEM-2017-1839
CLASSIFICATION OF EEG SIGNALS USING GENETIC PROGRAMMING

Felipe Rebelo Lopes
Marco Antonio Meggiolaro
Pontifícal Catholic University of Rio de Janeiro, Marquês de São Vicente, 225, Gávea
felipelopes_9@hotmail.com

Abstract. *Brain Computer Interface systems are based on an analysis of electroencephalogram signals, associated with an intention of a human being only based on thought. Feature extraction and classification are still a challenge. This paper presents the genetic programming technique as an alternative to pattern classifiers. Publicly available BCI competition IV dataset I, a multichannel 2-class motor-imagery dataset, is used for this purpose. The Wavelets Transform method is applied to decompose the signal in frequency sub-bands. In addition, different features are calculated. Measurement of time, statistics, and information theory are the most important features of this type of signal. Moreover, to classify features into two classes (imaginary movement of the right hand or feet), a multi-gene genetic programming is used to remove noise. It is shown that the training performance reaches a minimum error in 20 generations. The model is validated from its high rate of successfully classified commands.*

Keywords: *Genetic Programming, BCI systems, Wavelet transform, ECG Signals.*

1. INTRODUCTION

Brain Computer Interface (BCI) systems are a type of pattern recognition system because they receive the stimulus of the brain activity, extract the characteristics required for an output, and classify them to send signals to a controller. This system is capable of translating human intentions using a helmet that extracts EEG signals and then controls external devices. These signals are recordings of the brain electrical activity, capturing the voltage variation of the neurons' internal movement. Often the EEG is used to detect epilepsy in which the signal has some abnormalities.

The system performance depends on the feature extraction and classification algorithms employed. Hence, several works have been proposed to perform feature extraction, such as Fast Fourier Transform (FFT), Short-time Fourier Transform (SFT), Discrete Wavelets Transform (DWT), Genetic Algorithms (GA), as well as Adaptive Autoregressive (AAR) methods (Güler and Übeyli, 2005; Subasi, 2007; Xu, et al., 2009; Corralejo, et al., 2011; Pattnaik, et al., 2016; Werner and Hanka, 2016).

To classify the patterns, several studies have been realized, and many algorithms evaluated to obtain the best results, such as the Linear Discriminant Analysis (LDA) (Llera et al., 2014; Scherer et al., 2004), Support Vector Machine (SVM) (Qiu et al, 2017; Xu, et al., 2009; Lotte et al., 2007; Adjed et al., 2016), Bayesians classifier (Miao et al., 2017), K-nearest Neighbor (k-NN) (Kayikcioglu and Aydemir, 2010) and Artificial Neural Network (ANN) (Coyle et al., 2010; Hazrati and Erfanian, 2010; Lekshmi et al., 2014; Huan and Ramaswamy, 2004; Barbosa et al., 2010; Barbosa et al., 2009; Barbosa et al., 2013).

Basically the LDA uses hyperplanes to separate the data into classes. It seeks one projection that maximizes the distances between means of the classes and another that minimizes the variance between them. Its disadvantage is the linearity hypothesis in signals of more complex non-linear EEGs or with noise. Like the LDA, the SVM separates data into hyperplanes but it maximizes the margins, which is the distance between the nearest training point and the hyperplane. Even so, it is also very sensitive to dynamic noise.

Artificial neural networks are widely used, inspired by the structure of a human brain. In Barbosa et al. (2013), a BCI system was developed using a Probabilistic Neural Network (PNN) and a Multilayer Perceptron Neural Network (MLP), resulting in a high rate of successfully classified commands, about 90%. However, this training process usually consumes a lot time, and needs enough training samples to obtain a good performance. These techniques can be slower than required for practical real-time applications, as well as costly in terms of hardware resources, which establish a clear bottleneck (Tan, et al., 2016).

Recently, Tan, et al. (2016) discuss a new method called Extreme Learning Machine (ELM). This method looks like a feed-forward neural network, but with a higher speed and better generalization compared to the standard neural network and the SVM classifier.

Concerning Genetic Programming, works that classify EEG signals into subjects with epilepsy have been performed in Guo et al. (2011) and in Bhardwaj et al. (2016). Other papers show the power of the technique to perform classification (La Cava et al, 2017; Nag and Pal, 2016; Tran et al, 2016).

In this paper, the wavelet transform method is applied to the extraction of characteristics and genetic programming for an EEG classification to obtain the BCI output.

2. DATA PROCESSING

2.1 Dataset

The dataset IV of the BCI Competition (dataset 1), available in http://www.bbci.de/competition/iv/#dataset1, is provided by the Berlin BCI group from TU Berlin and Fraunhofer FIRST. It was used to evaluate the classification performance of classifiers in Blankertz et al. (2007). All EEG data were collected during a motor imagery task without feedback. For each subject, two classes of motor imagery were selected from the three classes: imaginary movement of the left hand, right hand, or feet (both feet).

According to Blankertz et al. (2007), the EEG was sampled at a 100 Hz sampling rate and was filtered between 0.05 and 200 Hz. The motor imagery tasks were cued by soft acoustic stimuli (words left, right, and foot) for periods of varying time lengths between 1.5 and 8 seconds. The end of the motor imagery period was indicated by the word stop. Intermitting periods had also a varying duration of 1.5 to 8s.

The data shows 59 channels of each measured point. In each signal there is a marker showing which class is the cue (right / left or feet) and the position which that measurement is. In this case, there are 200 markers, 100 of the right hand and 100 of the feet.

2.2 Discrete Wavelets Transform

In this work, a time-frequency scheme based on wavelet transforms is used to extract features, chosen because the EEG is a non-stationary signal.

The discrete wavelet transform (DWT) analyzes the signal in different frequency bands at different resolutions, helping to create unique features. Then, it is possible to decompose this signal into several levels of signals with more information. An efficient way to implement the DWT was developed by Mallat (1989). Decomposing a signal, each stage consists of two high-pass and low-pass digital filters. The output of the high pass filter is information D1 and the output of the low pass filter is A1. So, A1 is decomposed by two more digital filters and so on until the required decomposition level is reached. The choice of the best level decomposition to be used is important as well because each one corresponds to a different frequency.

According to Xu et al. (2009), the rhythm μ and the rhythm β with origin in the sensorimotor cortex were considered to be good signal features for EEG-based BCI, so the sub-bands D2 and D3 were used for their classification.

2.3 Feature Extraction

From the sub-bands of DWT, only D2 and D3 sub-bands are used for feature extraction. From each band, five measures are commonly used in the EEG signal analysis, namely: Mean, Standard deviation, Energy, Curve length and Skewness, as shown respectively in Eqs. 1-5. These measures are selected from time, statistics and information theory, and reveal the most important characteristics of an EEG signal (Guo et al., 2011).

With these five values of each sub-band, it is possible to create a vector with ten characteristics.

$$\frac{1}{N}\sum_{i=1}^{N} S_i \tag{1}$$

$$\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(S_i-\mu)^2} \tag{2}$$

$$\frac{1}{N}\sum_{i=1}^{N} S_i^2 \tag{3}$$

$$\sum_{i=1}^{N} |S_{i+1}-S_i| \tag{4}$$

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{S_i-\mu}{\sigma}\right)^4-3 \tag{5}$$

2.4 Genetic Programming

Genetic programming (GP) is a technique based on the principle of Darwin's natural selection, where each gene has a computer program matching with its peers. It was presented in the 1980s in the work of Koza (1992), in symbolic regression. This technique is actually an extension of genetic algorithms. The difference between them is the representation given in the solution. In the genetic algorithm the solution is a number, whereas in genetic programming the solutions are computer programs represented in tree form.

In GP, a random initial population formed by computer programs is generated. The functions are chosen randomly from the established set of functions (these functions can be basic arithmetic operations, Boolean logic or any other mathematical function). Terminals are chosen randomly in the set of terminals and the maximum depth tree is an initially controlled parameter. For fitness, each program is evaluated in relation to their tasks, i.e. if the error produced is the smallest, resulting in the (locally) best program. Then the genetic operators are selected and applied to create a new program population. The best program found in any generation defines the output of genetic programming. There is a variation of GP called multi-gene genetic programming (MGGP). The MGGP is a weighted linear combination of the outputs from a number of GP trees. Each of the trees may be considered to be a gene.

Several architectures of traditional GP and multi-gene GP (MGGP) are used here, and different parameters of population size, number of generations, maximum number of genes allowed in an individual, maximum tree depth, tournament size, and mutation operators are also used. The proposed model of the system is shown in Fig. 1.

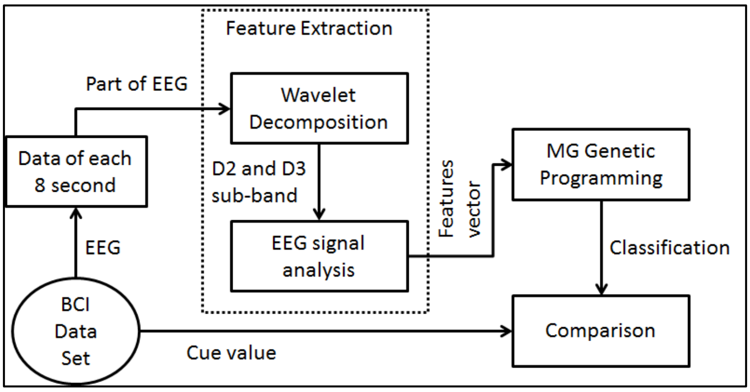

Figure 1. Proposed structure of the classification system.

3. RESULTS AND DISCUSSION

The proposed GP and MGGP as classifiers are implemented in MATLAB (R2016a), with the package GPTIPS (Searson, et al., 2010). The algorithms are applied to the EEG database using an Intel Core i3-4005U 1.70GHz with 4GB of RAM. The results are presented in Table 1, as well as the respective GP parameters.

Table 1. GP Parameters of the Best Result.

Parameters	Values
Population size	500
Number of generations	40
Tournament size	8
Elite fraction	15%
Max tree depth	5
Max nodes per tree	inf
Using function set	(x,-,+,/,sin,cos)
Number of inputs	10
Max genes	5
Range	[-10 10]
Fitness function	Regression_fitfun.m

The decomposition of EEG signals using the wavelet transform extracted five characteristics, resulting in a 10 by 200 matrix. This signal characteristics matrix is the input of the genetic programming. For the target vector, the same dimension is obtained from the BCI data set, which selects a class with values 1 (right hand) or 0 (feet). This way, a line vector 1 x 200 is created, with the first hundred values equal to 1, and the last hundred equal to 0. This form of input to genetic programming already helps since the classes are separated. For the training set, a 10 by 180 matrix is separated, and for the test a 10 x 20 matrix.

The results of genetic programming can be seen in Fig. 2, which shows the performance evolution of the MGGP, presenting a decrease in the value of RMSE as the number of generations increases. Notice that in 17 generations the stopping criterion (0.01) is reached. The upper graph represents the fitness of the best individual and the lower graph shows the mean performance of the population.

The computational processing takes 1.3 minutes. Figure 3 shows that the proposed model prediction on the training data is accurate with an R^2 of 0.99994, and for the test data 0.99958. These values show a good generalization of the proposed model. The prediction model created by the MGGP results in a good separation of the classes, thus the performance of the model is satisfactory, see Fig. 3. Finally, a part of the constructed model (a gene) is shown in Fig. 4.

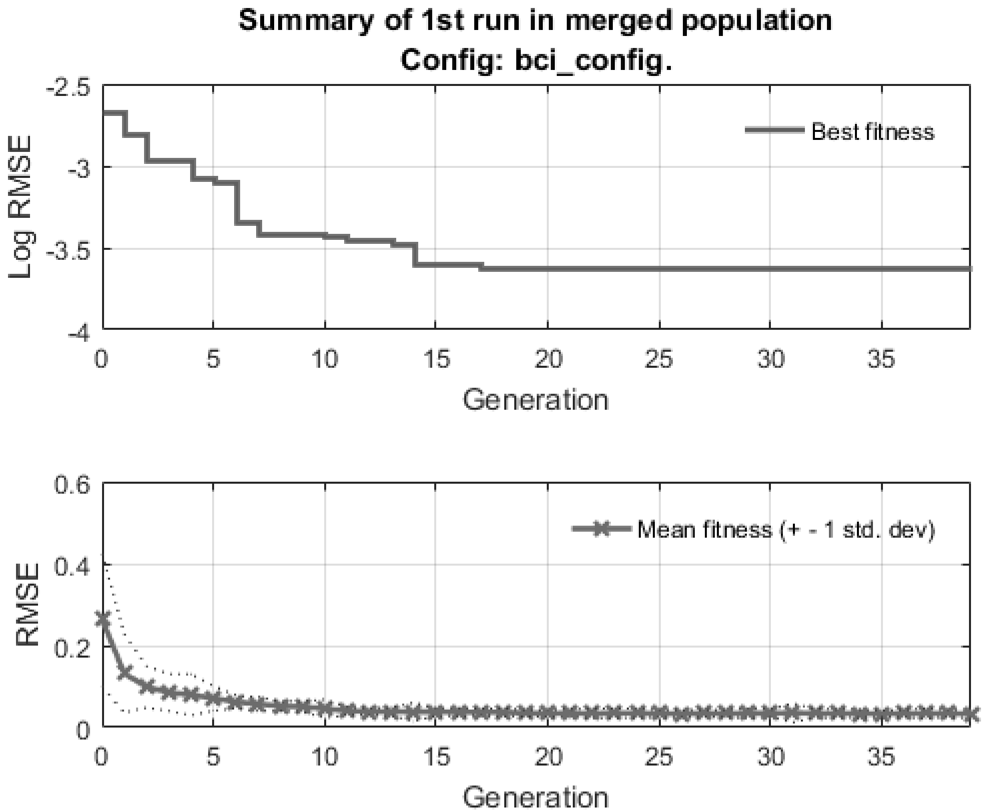

Figure 2. MGGP summary.

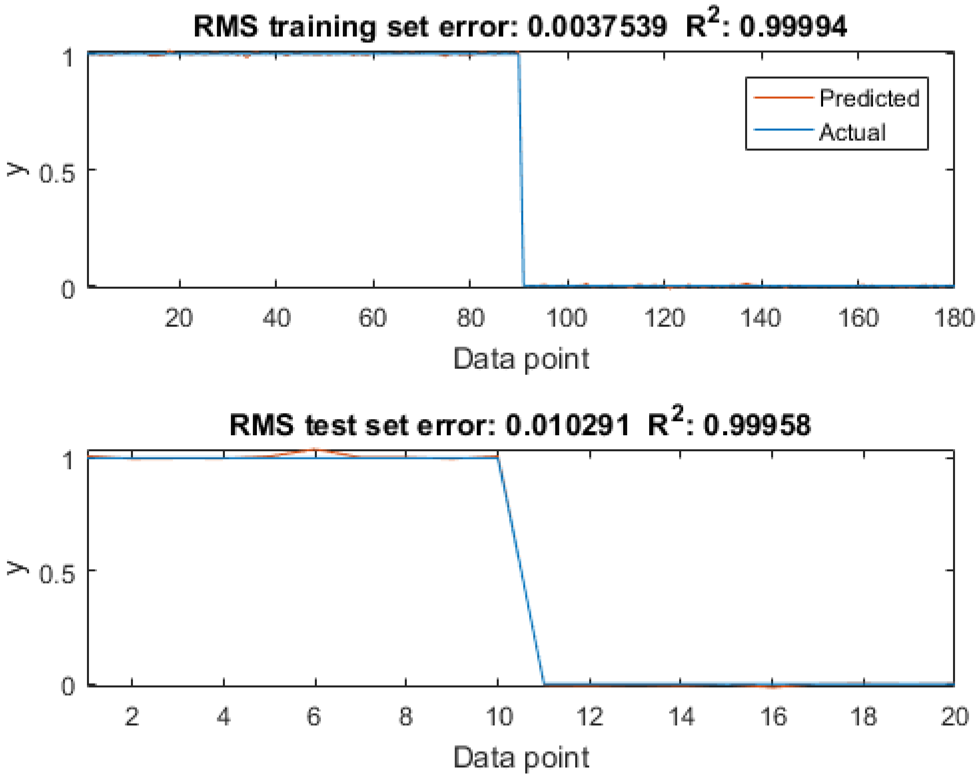

Figure 3. Prediction model.

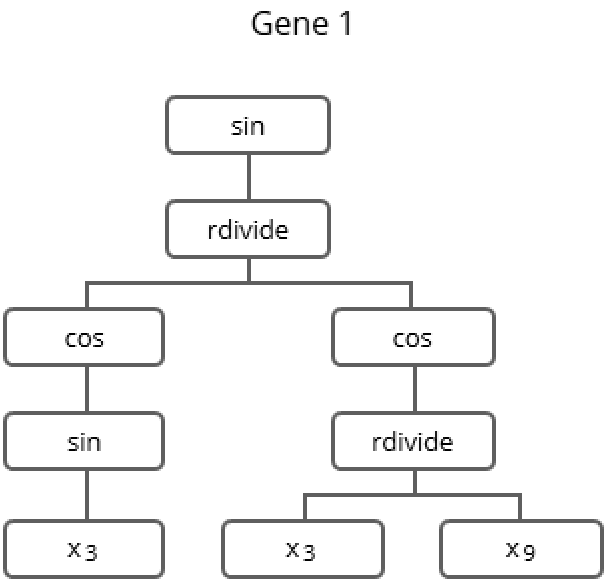

Figure 4. Gene 1 of the MGGP model, where X_3 is the energy of sub-band D2 and X_9 is the curve length of the sub-band D3.

The complete model created by the MGGP is:

$$y = 25.6 * \sin(\cos(x_3)) - 21.7 * \cos(\sin(x_3)) - 0.481 * x_3 - 0.112 * \sin\left(\frac{\cos(\sin(x_3))}{\cos\left(\frac{x_3}{x_9}\right)}\right) + 0.481 * \sin(x_1) - \frac{1.37 * \sin(x_3)\sin(x_5)(x_4 + 2.88)}{\cos\left(\frac{x_3}{x_9}\right)} + 0.255 \quad (6)$$

where x_1 is the mean of sub-band D2, x_3 is the energy of sub-band D2, x_4 is the curve length of sub-band D2, x_5 is the skewness of sub-band D2, and x_9 is the curve length of sub-band D3. Using Eq. 6, the results are shown in Fig. 5, with an accurate classification of the cue (right hand or feet).

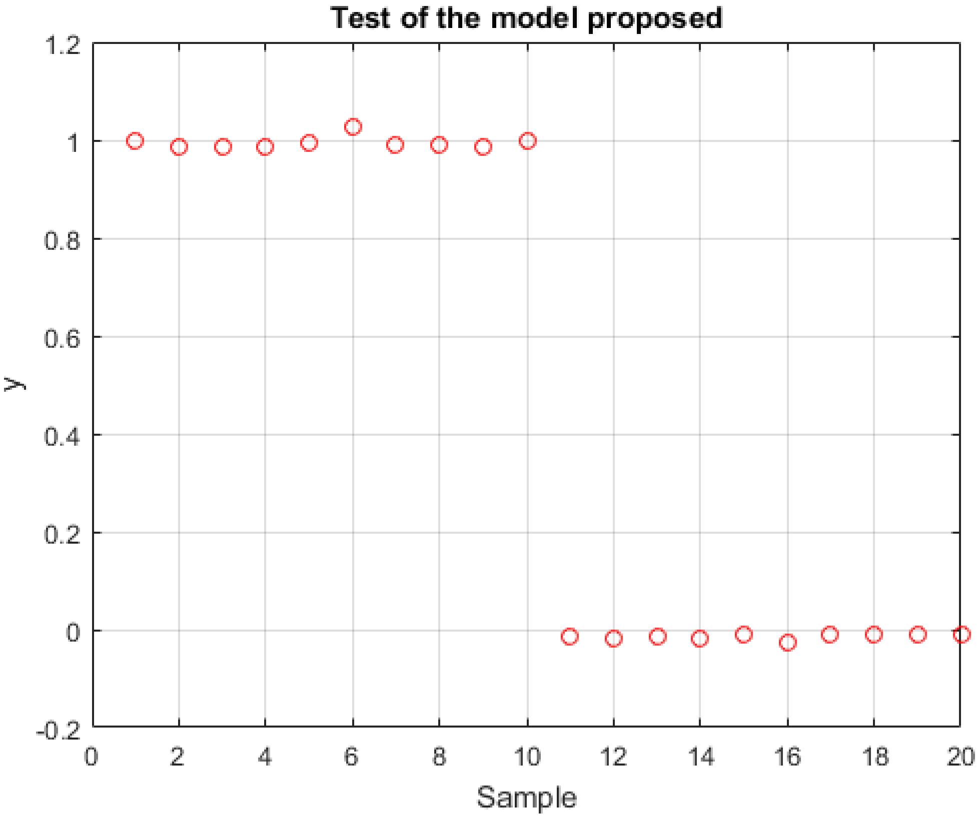

Figure 5. Test of completed model.

Concerning the complexity of the equation generated by the MGGP, other simple models were founded like Eq. 7 but with lower accuracy (R^2 of 0.941).

$$2.66 * \sin(\sin(x_3)) - 0.232 \tag{7}$$

4. CONCLUSIONS

This article proposed a classification of BCI signals using genetic programming. The feature extraction was performed with a wavelet transform, which proved to be efficient in obtaining information from the EEG signals. The characteristics obtained from measures of time, statistics, and information theory, were able to separate the two classes in question. The adopted genetic programming quickly obtained a model and, with respect to the training data, presented a good selection of classes. The test with actual user data validated the MGGP model, with a high rate of successful classifications reached.

5. ACKNOWLEDGEMENTS

The authors thank the Brazilian Government Agency CAPES for the support of this research.

6. REFERENCES

Adjed, F., Faye, I., Ababsa, F., Gardezi, S. J., and Dass, S. C., 2016. “Classification of skin cancer images using local binary pattern and SVM classifier”. In *AIP Conference Proceedings*, Vol. 1787, n. 1, p. 080006).

Barbosa, A. O. G., Diaz, D. R. A., Vellasco, M. M. B., Meggiolaro, M. A., and Tanscheit, R., 2009. “Mental tasks classification for a noninvasive bci application”. In *International Conference on Artificial Neural Networks*, p. 495-504.

Barbosa, A. O. G., Freitas, D. Z., de Moura Guedes, J. Q., and Meggiolaro, M. A., 2013. "Implementation of a Wheelchair Control Using a Four-Command Brain Computer Interface". In *Proceedings of the 22st International Congress of Mechanical Engineering -COBEM2013,* p. 5046-5057.

Barbosa, A. O., Achanccaray, D. R., and Meggiolaro, M. A., 2010. "Activation of a mobile robot through a brain computer interface". In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, p. 4815-4821.

Bhardwaj, A., Tiwari, A., Krishna, R., and Varma, V., 2016. "A novel genetic programming approach for epileptic seizure detection". *Computer methods and programs in biomedicine*, Vol. 124, p.2-18.

Blankertz, B., Dornhege, G., Krauledat, M., Müller, K. R., and Curio, G., 2007. "The non-invasive Berlin brain–computer interface: fast acquisition of effective performance in untrained subjects". *NeuroImage*, Vol. 37, n.2, p.539-550.

Corralejo, R., Hornero, R., and Alvarez, D., 2011. "Feature selection using a genetic algorithm in a motor imagery-based Brain Computer Interface". In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, p. 7703-7706.

Coyle, D., McGinnity, T. M., and Prasad, G., 2010. "Improving the separability of multiple EEG features for a BCI by neural-time-series-prediction-preprocessing". *Biomedical Signal Processing and Control*, Vol. 5, n.3, p.196-204.

Güler, I., and Übeyli, E. D., 2005. "Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients". *Journal of neuroscience methods*, Vol. 148, n. 2, p.113-121.

Guo, L., Rivero, D., Dorado, J., Munteanu, C. R., and Pazos, A., 2011. "Automatic feature extraction using genetic programming: an application to epileptic EEG classification". *Expert Systems with Applications*, Vol. 38, n.8, p.10425-10436.

Hazrati, M. K., and Erfanian, A., 2010. "An online EEG-based brain–computer interface for controlling hand grasp using an adaptive probabilistic neural network". *Medical engineering & physics*, Vol. 32, n.7, p.730-739.

Huan, Nai-Jen, and Ramaswamy Palaniappan., 2004. "Neural network classification of autoregressive features from electroencephalogram signals for brain? computer interface design." *Journal of neural engineering*, Vol. 1, n.3, p.142.

Kayikcioglu, T., and Aydemir, O., 2010. "A polynomial fitting and k-NN based approach for improving classification of motor imagery BCI data". *Pattern Recognition Letters*, Vol. 31, n.11, p.1207-1215.

Koza, J. R., 1992. *Genetic programming: on the programming of computers by means of natural selection (Vol. 1)*. MIT press.

La Cava, W., Silva, S., Vanneschi, L., Spector, L., and Moore, J., 2017. "Genetic programming representations for multi-dimensional feature learning in biomedical classification". In *European Conference on the Applications of Evolutionary Computation*, p. 158-173.

Lekshmi, S. S., Selvam, V., and Rajasekaran, M. P., 2014. "EEG signal classification using Principal Component Analysis and Wavelet Transform with Neural Network". In *Communications and Signal Processing (ICCSP) International Conference on*, p. 687-690.

Llera, A., Gómez, V., and Kappen, H. J., 2014. "Adaptive multiclass classification for brain computer interfaces". *Neural computation*, Vol. 26, n. 6, p.1108-1127.

Lotte F., Congedo M., Lecuyer A., Lamarche F., and Arnaldi B., 2007. "A review of classification algorithms for EEG-based brain–computer interfaces". *Journal Neural Engineering*, Vol. 4, n.2, p.R1–13.

Mallat, S., 1989. "A theory for multiresolution signal decomposition: The wavelet representation". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, n.7, p.674–693.

Miao, M., Zeng, H., Wang, A., Zhao, C., and Liu, F., 2017. "Discriminative spatial-frequency-temporal feature extraction and classification of motor imagery EEG: An sparse regression and Weighted Naïve Bayesian Classifier-based approach". *Journal of neuroscience methods,* Vol. 278, p.13-24.

Nag, K., and Pal, N. R., 2016. "A multiobjective genetic programming-based ensemble for simultaneous feature selection and classification". *IEEE transactions on cybernetics*, Vol. 46, n.2, 499-510.

Qiu, Z., Allison, B. Z., Jin, J., Zhang, Y., Wang, X., Li, W., and Cichocki, A., 2017. "Optimized motor imagery paradigm based on imagining Chinese characters writing movement". *IEEE Transactions on Neural Systems and Rehabilitation Engineering.*

Scherer, R., Muller, G. R., Neuper, C., Graimann, B., and Pfurtscheller, G., 2004. "An asynchronously controlled EEG-based virtual keyboard: improvement of the spelling rate". *IEEE Transactions on Biomedical Engineering*, Vol. 51, n.6, p.979-984.

Searson, D. P., Leahy, D. E., and Willis, M. J., 2010. "GPTIPS: an open source genetic programming toolbox for multigene symbolic regression". In *Proceedings of the International multiconference of engineers and computer scientists*, Vol. 1, p. 77-80.

Subasi, A., 2007. "EEG signal classification using wavelet feature extraction and a mixture of expert model". *Expert Systems with Applications*, Vol. 32, n. 4, p.1084-1093.

Tan, P., Sa, W., and Yu, L., 2016. "Applying Extreme Learning Machine to classification of EEG BCI". In *Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), 2016 IEEE International Conference on*, p. 228-232.

Tran, B., Xue, B., and Zhang, M., 2016. “Genetic programming for feature construction and selection in classification on high-dimensional data”. *Memetic Computing*, Vol. 8, n.1, p. 3-15.

Werner, G. Á., and Hanka, L., 2016. “Tuning an artificial neural network to increase the efficiency of a fingerprint matching algorithm”. In *Applied Machine Intelligence and Informatics (SAMI), 2016 IEEE 14th International Symposium on*, p. 105-109.

Xu, Q., Zhou, H., Wang, Y., and Huang, J., 2009. “Fuzzy support vector machine for classification of EEG signals using wavelet-based features”. *Medical engineering & physics*, Vol. 31, n.7, p.858-865.

7. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.

Form Presentation: Poster