Neural Networks 79 (2016) 88–96

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Pattern recognition for electroencephalographic signals based on continuous neural networks

M. Alfaro-Ponce [a], A. Argüelles [b], I. Chairez [c,*]

[a] Escuela Superior de Tizayuca, Universidad Autonoma del Estado de Hidalgo, Tizayuca, Hidalgo, Mexico
[b] Centro de Investigacion en Computacion, Instituto Politecnico Nacional, Mexico City, Mexico
[c] Unidad Profesional Interdisciplinaria de Biotecnologia, Instituto Politecnico Nacional, Mexico City, Mexico

ARTICLE INFO

Article history:
Received 13 November 2015
Received in revised form 9 March 2016
Accepted 11 March 2016
Available online 6 April 2016

Keywords:
Signal classifier
Continuous neural networks
Electroencephalographic signals
Pattern recognition

ABSTRACT

This study reports the design and implementation of a pattern recognition algorithm to classify electroencephalographic (EEG) signals based on artificial neural networks (NN) described by ordinary differential equations (ODEs). The training method for this kind of continuous NN (CNN) was developed according to the Lyapunov theory stability analysis. A parallel structure with fixed weights was proposed to perform the classification stage. The pattern recognition efficiency was validated by two methods, a generalization–regularization and a k-fold cross validation ($k = 5$). The classifier was applied on two different databases. The first one was made up by signals collected from patients suffering of epilepsy and it is divided in five different classes. The second database was made up by 90 single EEG trials, divided in three classes. Each class corresponds to a different visual evoked potential. The pattern recognition algorithm achieved a maximum correct classification percentage of 97.2% using the information of the entire database. This value was similar to some results previously reported when this database was used for testing pattern classification. However, these results were obtained when only two classes were considered for the testing. The result reported in this study used the whole set of signals (five different classes). In comparison with similar pattern recognition methods that even considered less number of classes, the proposed CNN proved to achieve the same or even better correct classification results.

1. Introduction

Automatic detection and classification of EEG recordings have become important fields of research for the development of brain computer interfaces, security, interactive games and medical diagnosis systems (Bashashati, Fatourechi, Ward, & Birch, 2007; Birbaumer, 2006; Goel et al., 1996; Hwang, Kim, Choi, & Im, 2013). As a result, many algorithms for the EEG classification have been proposed, but most of these algorithms are limited by the fact that they are able to classify just one characteristic of the many that are codified in the EEG signals (Nicolas-Alonso & Gomez-Gil, 2012).

Nevertheless that different methods have been proposed to classify patterns that appear in the EEG signals as, for example, spindles and K-complexes for sleep staging, epileptiform patterns for epilepsy analysis, among others, all of them have in common an automatic system that uses some characteristics from the signals. These characteristics are extracted from the EEG recording.

In general, these pattern recognition algorithms are based on different machine learning techniques such as autoregressive modeling (Ge, Srinivasan, & Krishnan, 2007), Markov chains (Boussemarta & Cummings, 2011), self-organizing maps (Allinson & Yin, 1999), fuzzy c-means clustering techniques (Roy, Charbonnier, & Bonneta, 2014), neural networks (Kannathal, Rajendra, ChooMin, & Suri, 2007) and coefficients of the wavelet transform (Chang, Lin, Wei, Lin, & Chen, 2014), among others. Most of these algorithms implement different pre-treatment algorithms to construct the pattern vector that is to be evaluated in the classifier. This strategy can omit certain characteristics of the information that may be relevant to the characterization of the signal. This omission can be a consequence of the preliminary manipulation of EEG signals before they are evaluated by the classifier (Riaz, Hassan, Rehman, Niazi, & Dremstrup, 2015). Moreover, the continuous nature of the EEG signal is left out in the classifier structure (Akareddy & Kulkarni, 2013).

Although several pattern recognition schemes have been applied over the years on EEG signals, static NN (SNN) (Weng & Khorasani, 1996) based pattern recognition classifiers have

* Corresponding author.
E-mail address: isaac_chairez@yahoo.com (I. Chairez).

http://dx.doi.org/10.1016/j.neunet.2016.03.004
0893-6080/

gained significant prominence with respect to some other alternatives (Gotman & Wang, 1991; Lu, Shin, & Ichikawa, 2004). SNN have been successfully employed to determine complex, nonlinear, multidimensional mathematical relationships between noisy uncertain sets of data with considerable dissimilar natures (Dunne, 2006).

Nowadays, NN based pattern recognition solutions have been frequently applied to classify EEG signals, especially in the domain of function approximation (Alotaiby, Alshebeili, Alshawi, Ahmad, & El-Samie, 2014), pattern recognition (Kasabova, Dhoblea, Nuntalida, & Indiverib, 2013), automated medical diagnostic systems (Amato et al., 2013), decision support systems (Ubeyli, 2009), time series prediction (Coyle, Prasad, & McGinnity, 2005), signal processing (Sanei & Chambers, 2013), image processing. (Bose & Liang, 1996; Hassoun, 1995; Haykin, 1999), wavelets (Chen, 2014) and some others. The success of NN in pattern recognition is a consequence of their capability to approximate nonlinear relationships between the input and output pairs (Cybenko, 1989). Therefore, the method selected to adjust the weights in the NN structure plays a key role on forcing a higher efficiency on the classification task.

Approximation problems can be solved by employing either supervised learning (Nait-Ali, 2009) (where the weights and biases of the SNN are learned in the presence of training data set) or unsupervised learning (where inputs are classified into different clusters in a multidimensional space, in the absence of targets' training data). Today, supervised learning continues as the most preferred option when pattern recognition algorithms are employed. However, most of these algorithms were executed using patterns formed by vectors of characteristics. These vectors were obtained as solution of preliminary signal processing algorithms that usually do not take into account the continuous nature of the EEG signal. Therefore, relevant information can be lost during this preliminary process that can simplify the pattern recognition algorithm.

Regarding the EEG signal pattern classification, several types of SNNs have been proposed (Cheng-Jian & Ming-Hua, 2009). Most of these solutions employed a supervised learning mode which implies high levels of computational effort. Additionally, the complexity and necessity of performing preliminary treatment demand an extra processing effort that may compromise the application of the pattern classifier to obtain an on-line pattern recognition solution (Omerhodzic, Avdakovic, Nuhanovic, & Dizdarevic, 2013).

This study proposes an alternative method to solve the EEG signal pattern recognition problem. A class of CNN (Poznyak, Sanchez, & Wen, 2001) is used to represent the relationship between the EEG signal and its particular pattern class represented by a sigmoid type of function. The CNN concept is defined by the approximation provided by NN to the right hand side of ODEs. The CNN structure preserves the highly parallel structure that characterizes many of the usual pattern recognition forms. By virtue of its parallel distribution, the proposed CNN is tolerant under the presence of faults and external noises, able to generalize the input–output relationships well and capable of solving nonlinear approximation problems (Benvenuto & Piazza, 1992).

The pattern recognition method proposed in this study was applied on the signals contained in the database taken from NA (2012). The classification results obtained in this study were compared to the ones obtained by other researchers (Guo, Rivero, Dorado, Rabuñal, & Pazos, 2010; Guo, Rivero, Seoane, & Pazos, 2009; Nigam, 2004; Tzallas, Tsipouras, & Fotiadis, 2007), who applied signal classifiers based on SNN. This article is organized as follows: Section 2 describes the mathematical structure of the CNN as a classifier and its parallel implementation for the classification task, this section also describes the training and validation process. Section 3 contains the description of the two databases that were employed to test the CNN performance. Section 4 describes the set of two databases used to evaluate the classification performance achieved by the algorithm proposed here. Section 5 details the simulation results obtained from the implementation of the CNN to the two databases. Section 6 closes the article with some conclusions and discussions.

2. CNN EEG pattern classification

There is a general method that must be applied including the stages of training, validation and testing (despite the class of NN used to perform the signal classification). The first stage on the EEG signal classification requires defining a set of targets associated to the specific class of EEG. Therefore, if the EEG signal is considered as the input number j in the class i, $u^{i,j}$ to the NN, then the output, namely x^i corresponds to the specific class where the signal belongs to the L available classes. Then, the state x^i corresponds to the concept of target. For this study, this target was represented as a sigmoid function described by:

$$x^l(v) = \frac{a^l}{1 + e^{-cv}} \tag{1}$$

where the variable x^l represents the target that belongs to class l $(l = 1, \ldots, L)$. The positive constant a^l was modified according to the class where the particular EEG signal belongs. These constants served to modify the amplitude of the sigmoid function and then to characterize each class. The positive constant c was chosen to regulate the slope of the sigmoid function. One may notice that different functions could be selected to define the characteristic of a class but according to the Cybenko's seminal paper (Cybenko, 1989), this selection (sigmoid function) seems to be more natural.

The training process consisted of comparing the output of the NN with the target $x^l(v)$ when they both are affected by the same EEG signal. The training process consisted of executing the evaluation of the NN with a percentage of all EGG signals $u_r^l(v)$ that represents the r signal in the class $l\,(r \in [1, N_l]\,, \sum_{l=1}^{L} N_l = N\,, N$ is the number of signals of the entire database) signal in the class l. Then when the EEG signal $u_{r+1}^l(v)$ is executed, the set of weights produced by this training step $W^{*,l}$ is used a part of NN in this training stage.

Once the whole set of N signals selected to perform the training process has been tested, L different sets of weights $W_{N_l}^{*,l}$ have been produced. If the training process has been correctly executed, the aforementioned weights are recovered as part of a set of L non-adjustable NN with the same structure as the one used during the training. This part of the process is named the validation stage. Based on the well-known generalization–regularization and k-cross validation methods, a percentage of the whole set of EEG signals $u_r^l(v)$ is used to evaluate the output of the set of L NN with the corresponding set of $W_{1,N_l}^{*,l}$ and $W_{2,N_l}^{*,l}$. At this part of the validation, all the L NN are evaluated in parallel. The output of each NN named NN^l is compared with the corresponding value a^l. The mean square error $a^l - x_r^l$ is calculated over the period of time corresponding to the length of the EEG signal, that is

$$J^{T,l} = T^{-1} \int_{t=0}^{T} \left(a^l - x_r^l(t)\right)^2 dt.$$

One must notice that the length of all the testing signals was kept constant. The selection of T should be done according to the nature of signal. This is still a matter of interest and many studies have been proposed during the last 50 years. In this particular case, the window size was selected in agreement to the results presented in Kuncheva and Zliobaite (2009).