Expert Systems With Applications 68 (2017) 1–10

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Windowed electroencephalographic signal classifier based on continuous neural networks with delays in the input

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 24 February 2016
Revised 24 June 2016
Accepted 3 August 2016
Available online 21 September 2016

Keywords:
Time-delay neural networks
Electroencephalographic signals
Pattern classification
Lyapunov stability theory
Epileptic seizures

ABSTRACT

This study reports the design and implementation of a pattern recognition algorithm aimed to classify electroencephalographic (EEG) signals based on a class of dynamic neural networks (NN) described by time delay differential equations (TDNN). This kind of NN introduces the signal windowing process used in different pattern classification methods. The development of the classifier included a new set of learning laws that considered the impact of delayed information on the classifier structure. Both, the training and the validation processes were completely designed and evaluated in this study. The training method for this kind of NN was obtained by applying the Lyapunov theory stability analysis. The accuracy of training process was characterized in terms of the number of delays. A parallel structure (similar to an associative memory) with fixed (obtained after training) weights was used to execute the validation stage. Two methods were considered to validate the pattern classification method: a generalization-regularization and the *k*-fold cross validation processes ($k = 5$). Two different classes were considered: normal EEG and patients with previous confirmed neurological diagnosis. The first one contains the EEG signals from 100 healthy patients while the second contains information of epileptic seizures from the same number of patients. The pattern classification algorithm achieved a correct classification percentage of 92.12% using the information of the entire database. In comparison with similar pattern classification methods that considered the same database, the proposed CNN proved to achieve the same or even better correct classification results without pre-treating the EEG raw signal. This new type of classifier working in continuous time but using the delayed information of the input seems to be a reliable option to develop an accurate classification of windowed EEG signals.

1. Introduction

There are various events that occur and influence the electroencephalogram (EEG) waveform. These events can be usual activities such as listening music (Sawata, Ogata, & Haseyama, 2015), sleeping (Koley & Dey, 2012), smelling particular odors (Byung-Chan et al., 2003), using drugs (Saletu, Anderer, SaletuZyhlarz, Arnold, & Pascual-Marqui, 2002), suffering mental disorders (Manchanda et al., 2014), detecting mood (Yu et al., 2011), among others. EEG signal has been also used to detect different events occurring in certain regions deep in the brain. In particular, epileptic seizures have been studied because they can cause a variety of temporal modifications in perception and behavior. During the epileptic event, EEG signal apparently becomes rhythmic even preceding the first detectable behavioral change (Sierra-Marcos, Scheuer, & Rosseti, 2015). Automatic and reliable detection at the earliest possible moment, can be used in characterizing the epileptic centers as well as detecting if EEG rhythmic signals are truly coming from epileptic events or they are produced by a different physiological or anatomical disorder. Long periods of time and exhaustive signal analysis must be applied to detect epileptic characteristic waveform in EEG recordings (Garces Correa, Orosco, Diez, & Laciar, 2015). Therefore, several attempts have been made to develop automatic detection systems of particular EEG waveform (Ang & Chin, 2012; Hramov, Koronovskii, Makarov, Pavlov, & Sitnikova, 2015).

Diverse studies established certain characteristics in EEG records that can be used to classify them. Due to the fact that EEG signals exhibit complex behavior (Stefanidis, Anogiannakis, evangelou, & Poulus, 2015) with strong non-linear and dynamic properties (Subha, Joseph, Acharya, & MinLim, 2008), several researchers

* Corresponding author.
E-mail addresses: marielalfa@gmail.com (M. Alfaro-Ponce), jamadeo@cic.ipn.mx (A. Argüelles), isaac_chairez@yahoo.com (I. Chairez).

http://dx.doi.org/10.1016/j.eswa.2016.08.020
0957-4174/

have developed algorithms that automatically detect events as a way to avoid manual sorting of EEG signals. Indeed, these algorithms have been developed to reveal time-locked event related modulations of the EEG signal or frequency specific elements. All of these algorithms tried to estimate within each EEG epoch, the dependence of signal amplitude with respect to time and frequency. Therefore, they have used the main characteristics of different signal processing and pattern classification methods that cover a wide range of complexities, from the simple population vector algorithm (Cinar & Sahin, 2013), optimal linear estimator (Ang & Chin, 2012), various versions of Bayesian decoders (Li, Doherty, Lebedev, & Nicolelis, 2011) and different arrangements of the so-called artificial neural networks (NN). Nevertheless, EEG classification techniques have attained limited success when characterizing the brain information, because they must be used in specific applications, and no one has achieved total successful results due to diverse factors (Millan & Mourino, 2003). First of all, the proposed algorithms have to deal with great amounts of data. Besides, the processed information have lots of noise, along with the fact that these methods need to consider the interaction between actual neurons on the cortex yielding to analyze it as an interconnected system (Daly, 2013). Also, the brain outcomes can change dramatically from one individual to another even under the same circumstances and finally most of the proposed algorithms did not quantify all of the information available in the EEG recordings. Indeed, the majority of these techniques applied complicated pretreatments on the signals to obtain better classification results.

Simultaneous time and frequency analysis of EEG signals may produce better results. This kind of analysis has been done using the Wavelet transform technique for example (Faust, Acharya, Adeli, & Adeli, 2015). An alternative method uses the concept of windowing function. This procedure is used to obtain the extraction of both time and frequency characteristics in the EEG signal. In instance, the so-called windowed Fourier transformation has been widely considered to analyze and classify EEG signals (Roshan et al., 2012). The same methodology has been used when the Fourier transformation is substituted by the wavelet transform. A different method uses the so-called short time Fourier transformation (or Gabor) (Kumar, Kanhangad, & Pachori, 2015).

Several factors must be defined when windowing function is used in EEG analysis. For example, when a windowing function is applied to the time-domain signal, the spectral properties of the signal are conditioned a priory. In addition, defining the window characteristics may represent a difficult process. As a matter of fact, this problem continues as a relevant researching problem.

The windowing effect on EEG pattern recognition has been documented in different studies. In particular, this effect has been evaluated when machine learning methods have been used to obtain the EEG signal classification (convolution NN, support vector machines, genetic algorithms) (Hwang, Kim, Choi, & Im, 2013). Just recently, classification algorithms that maintains memory of previous inputs have started to appear. This kind of methods can learn not only by the EEG signal amplitude but also by its variation through time, that is, the temporal dependencies between consecutive samples. In this study, a different approach based on dynamic NN is proposed where the information within the window is considered as delayed versions of the input signal.

This study is aimed to develop an EEG signal classifier based on TDNN. This particular structure of NN was proposed to consider the effect of delayed information in the input signal. This scheme served to consider the usual windowing process that appears when the pattern classifier must consider a period of the signal. Training and validations processes were developed in this study using the so-called Lyapunov stability theory. This solution represents a new and different attempt to use continuous time delay NN as a key tool in the pattern recognition framework. Moreover, the parametric adjustment method used to regulate the weights values in the NN was formally obtained without using heuristic information.

In order to introduce the concept of TDNN, the following section describes briefly their main characteristics.

2. Time-delay dynamic NN

Time-delays are usually considered as sources of instability in dynamic systems, however, for some particular cases, the presence of delays yields to stabilizing effects (Emilia, 2014). Time-delays are also common in biological and chemical systems. A time-delay input signal appears in models of real systems due to different reasons. Usually, delays are forced by the physical nature of the system (Hale, 1977). Transport processes (like in chemical or pneumatic systems) or computational delay (e.g. in digital controllers or communication networks (Kruszewski, Jiang, Fridman, Richard, & Toguyeni, 2012)) are regular sources of delayed input signal.Delays can also be present during an EEG medical test becausethe recorded signal at each electrode is in fact showing the summation of the electrophysiological variance of the brain area closer to the electrode and a lagged version of variances from other subcortical regions (Sargolzaei et al., 2015). Input delay can also be introduced artificially to include the sampling effect in mathematical models (see, for example Fridman, Seuret, and Richard, 2004; Polyakov, 2012). Ignoring the time-delays in biopotential (e.g. EEG) yields to conclusions that do not contain the complete information of the system under study (Mier-y Terán-Romero, Silber, & Hatzimanikatis, 2010).

In recent studies (Arik, 2000), (Cao, 2000), (Joy, 2000) two kinds of TDNN are recognized, according to how the delay affect their stability. One is referred to as delay independent stability and the other delay dependent stability (Liaoa, Chenb, & Sanchez, 2002). For this paper, we deal with the delay dependent stability case. Nowadays, the results regarding time-delay systems only consider the stability of the TDNN. As a matter of the fact, TDNN has a primary objective to work with sequential data such as the collected in EEG records. This characteristic allows to this particular class of artificial NN to classify patterns or features independent of time-shift. Indeed, the specific movement of an epileptic event in time is a relevant cue to determine the possible source of the event as well as the nature of EEG rhythmic patterns. This pair of characteristics can be used to classify physiological phenomena more than isolated events in the EEG signal. Then, TDNN should have the ability to represent relationships between events in time.

A TDNN may serve as a potential pattern classifier, if it can be represented as a non-parametric identifier system. This requirement arises if we consider that a pattern classifier can be understood as a uncertain even unknown nonlinear system that uses the EEG signal as input (with the corresponding delayed information) while the output is a label or number characterizing the class where the particular windowed EEG waveform belongs. The question regarding how to define the adequate number of delays that must be considered in the TDNN structure is still a matter of research. However, the non-parametric identifier problem has been poorly explored (Ge, Srinivasan, & Krishnan, 2007). In this sense, the remarkable properties of time-delay stability analysis have been wasted for solving the problem of the non-parametric identifier based on Continuous NN or CNN.

3. Pattern classifier problem statement applied on windowed EEG

As mentioned in the previous sections, different classifiers have been applied to categorize EEG. Many of them are based on static NN. Recently, CNN have emerged as powerful tools to extent the

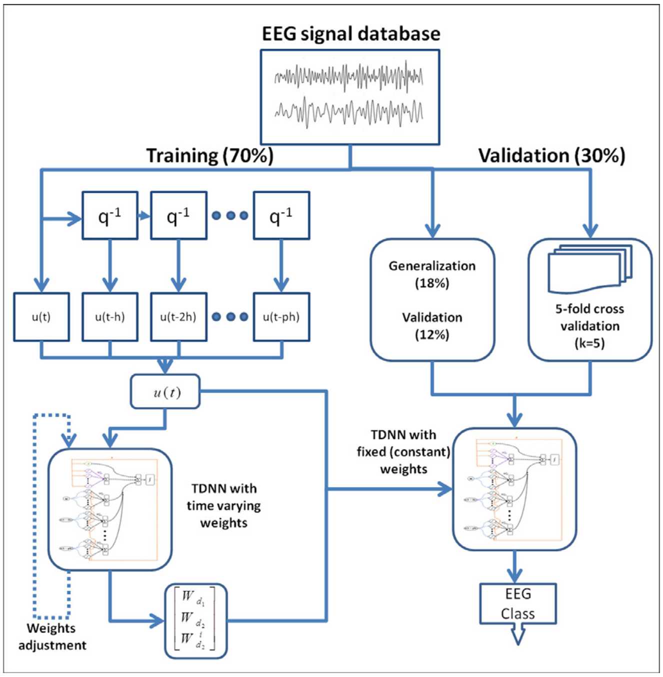

Fig. 1. Flow chart describing the entire process to implement TDNN as EEG signal classifier.

classification capabilities of NN. To the authors' knowledge, TDNN based on CNN have never been used as EEG pattern classifier.

The flow chart showed in Fig. 1 describes all the stages used in this study. The left-hand side describes the training step whit the delayed input produced by the delay operator q^{-1}. The result of this process are the weights fixed at the validation stage. Within the validation stage, two different procedures yielded to the final classification: generalization-validation and *k*-fold cross validation. The final result of the entire process is the specific class for each tested EEG signal.

3.1. TDNN pattern classifier applied on windowed EEG signals

Despite the class of NN used to perform the signal classification, there is a general method that must be applied including the stages of training, validation and testing. The first stage on the EEG signal classification requires to define 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 l, ($u^{j,\, l}$ to the NN), then the output, namely x^l corresponds to the specific class (among the L available classes) where the signal belongs. Then, the *state* x^l corresponds with the concept of target. For this study, this target was represented as a time dependent sigmoid function described by:

$$x^l(t) = \frac{a^l}{1 + e^{-c(t-d)}} \tag{1}$$

where the variable x^l represents the target that belongs to class l ($l = 1, \ldots, L$). The positive constant a^l was modified accordingly to the class where the particular EEG signal belongs. This constant served to modify the amplitude of the sigmoid function and then to characterize each class. The positive constant c was elected in order to regulate the slope of the sigmoid function. Also, the constant d was selected to adjust the transition of sigmoid function within the window. 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 $x^l(t)$ when they both are affected by the same EEG signal. This process seems to be similar to the classical supervised learning. Then, the evaluation of the NN with a percentage of all EGG signals $u^{j,\, l}(t)$ that represents its component number j ($j \in [1, N_l]$, $\sum_{l=1}^{L} N_l = N$, where N is the number of signals of the entire database selected to complete the training process). When the EEG signal $u^{j+1,l}(v)$ is executed, the set of weights produced by this training step is used a part of the NN considered in the next step. When $j = N_l$ then the weights characterizing the EEG signals belonging to class l are obtained.

Therefore, when 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 reused as part of a set of L non-adjustable NN with the same structure to the one used during the training. This part of the process is named the validation stage. Based on the well-known generalization-regularization and the *k*-cross validation methods, a percentage of the whole set of EEG signals $u^{j,\, l}(t)$ is used to evaluate the output of the set of L NN with the corresponding set of $W_{N_l}^{*,l}$. At this part of the validation, all the L NN are evaluated in parallel. The output of each trained NN named $\hat{x}^l$ is compared with the corresponding value x^l. The mean square error $\hat{x}^l - x^l$ is calculated over the period of time corresponding to the length of EEG signal or to the length of the window, that is

$$J^{T,l} = T^{-1} \int_{t=0}^{T} \left(\hat{x}^l(t \mid W_{N_l}^{*,l}) - x^l(t)\right)^2 dt$$

where $\hat{x}^l(t \mid W_{N_l}^{*,l})$ corresponds to the output of the corresponding NN with fixed weights $W_{N_l}^{*,l}$. One must notice that the length of all the testing signals was kept constant. The minimum value (among the L possible results) of this set of mean square (LMS) errors (moving and varying) was the indicator of the class where the EEG signal tested at that moment belongs. The validation state considered that all EEG signals used in this part of the analysis were previously used in the training stage. However, during the testing stage, a set of signals that have never been presented to the classifier was considered.

In summary, the classifier structure proposed for this work was developed according to the following strategy: The first stage was the training process; here the characteristic weights for the CNN were determined for each class. In the next stage, a parallel structure was developed, this structure was based on several CNN with their corresponding fixed weights for each class. This parallel structure uses an EEG signal as input. The signal is evaluated in parallel by the structure and from each CNN in the parallel structure, the LMS error was obtained. Then, the one with the smallest performance index $J^{T,\, l}$ was the one corresponding to the selected class.

The class of stable time-delay system considered is formally described as follows:

$$\begin{aligned} &\frac{d}{dt}x(t) = f(x(t), u(t), u(t-h), \ \ldots, \ u(t-ph)) + \xi(x(t), t) \\ &x_{t_0}(\theta) = x(t_0 + \theta) = \varphi(\theta) \\ &\quad \forall \theta \in [-ph, 0] \ \ p \in \mathbb{Z}^+ \end{aligned} \tag{2}$$

This dynamic representations relates the EEG signals as inputs $(u(t), u(t-h), \ldots, u(t-ph))$ with the corresponding sigmoid signal x. Notice that despite the delay value, the input signal is bounded as $\|u(t-ih)\|^2 \le u^+$, $i = 0, 1, \ldots, p$. The continuous signal $x \in \mathbb{R}$ is the state of the time-delay system (that corresponds to the class where the EEG input signal corresponds) with $|x| < \infty$, $\forall t \ge 0$. One must notice that the source of delay is coming from

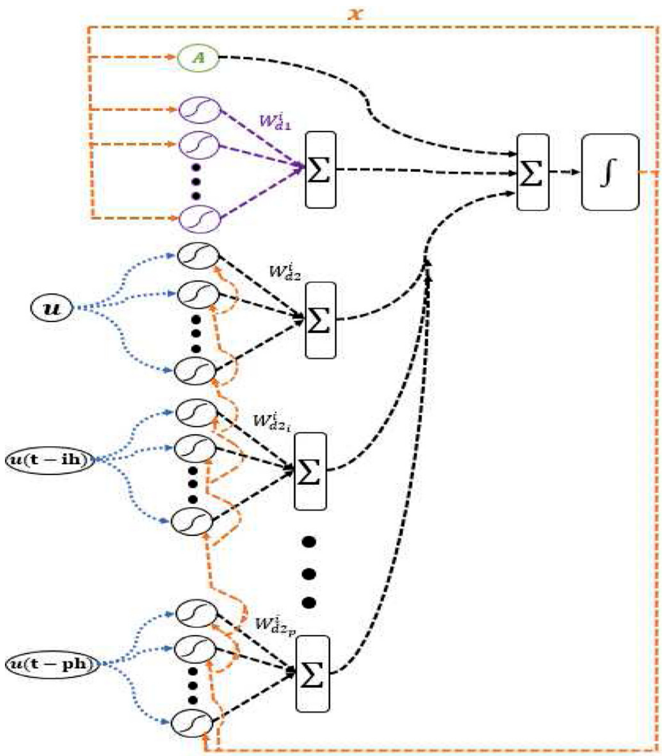

Fig. 2. TDNN structure considered in both stages: training and validation.

the input signal $u \in \mathbb{R}^m$ which now represents the EEG signal at the current time as well as its delayed information. The function f represents the uncertain nonlinear function connecting the state (the class) of the plant with the delayed input signal $u(t-ih)$, $i = 0, \ldots, p$. Indeed, this uncertain function represents the pattern classifier because in its structure, it is hidden the association between the input and the class. The delay value h is known and constant, $h \in \mathbb{R}_+ \forall t \geq 0$.

System uncertainties and perturbations are described by the nonlinear unknown function $\xi(x,t): \mathbb{R}^2 \to \mathbb{R}$ and satisfies

$$|\xi(x,t)|^2 \leq \Upsilon \ \forall t \geq 0 \tag{3}$$

where $\Upsilon \in \mathbb{R}^+$.

3.1.1. Neural network approximation for the time-delay signal classifier

Based on the NN approximation theory, the time delay system presented in (2) can be represented as the following TDNN (Liao, Chen, & Sanchez, 2002; Poznyak, Sánchez, & Yu, 2001):

$$\begin{aligned}\frac{d}{dt}x(t) &= A_d x(t) + \left[W_{d_1}^*\right]^\top \psi_{d_1}(x_d(t)) + \left[W_{d_2}^*\right]^\top \psi_{d_2}(x_d(t))u(t) \\ &\quad + \sum_{i=1}^{p}\left[W_{d_2}^i(t)\right]^\top \psi_{d_2}^i(x_d(t))u(t-ih) \\ &\quad + \tilde{f}(x_d(t), u(t), u(t-h), \ldots, u(t-ph)) + \xi(x(t), t)\end{aligned} \tag{4}$$

The basic structure of the TDNN considered in this study is depicted in Fig. 2. The structure shows the state continuous feedback modulated by weights $W(d_1)^i$ as well as the influence of non-delayed and delayed input information that are regulated by the weights $W_(d_2)$ and $W_{d_2}^i$ respectively. The activation functions used in the TDNN structure are also represented in the same figure.

The approximation based on TDNN tries to obtain a non-parametric approach to determine a feasible description of the EEG signal classifier. The structure of the TDNN, the scalar $A_d \in \mathbb{R}, W_{d_1}^* \in \mathbb{R}^{l_1}, W_{d_2}^* \in \mathbb{R}^{l_2}, W_{d_2}^{i,*} \in \mathbb{R}^{l_2}$ are constant matrices used to approximate the EEG signals. The set of matrices $W_{d_1}^*, W_{d_2}^*, W_{d_2}^i$ are unknown but all their components are bounded $(W_{d_1}^* \Lambda_1 [W_{d_1}^*]^\top \leq V_1^+,\ W_{d_2}^* \Lambda_2 [W_{d_2}^*]^\top \leq V_2^+,\ W_{d_2}^i [W_{d_2}^i]^\top \leq V_2^{+,i}$ with $V_1^+, V_2^+, V_2^{+,i}$ some positive matrices of appropriate dimensions). The modeling error term $\tilde{f}$ is assumed to be bounded as $\left|\tilde{f}(x(t), u(t), u(t-h), \ldots, u(t-ph))\right| \leq \tilde{f}_0$ with $\tilde{f}_0$ a positive scalar.

The vector functions $\psi_{d_1}: \mathbb{R} \to \mathbb{R}^{l_1}$ and $\psi_{d_2}: \mathbb{R} \to \mathbb{R}^{l_2 \times m}, \psi_{d_2}^i: \mathbb{R} \to \mathbb{R}^{l_2 \times m}$ define the set of activation functions used to design the neural network structure. The components of the activation functions were proposed as sigmoid functions, that is

$$\begin{aligned}\psi_{d_1}(x) &= \left.\frac{a_{d_1,j_1}}{1+e^{-\left(c_{d_1,j_1}x\right)}}\right|_{j_1=1,\ldots,l_1} \\ \psi_{d_2}(x) &= \left.\frac{a_{d_2,j_2}}{1+e^{-\left(c_{d_2,j_2}x\right)}}\right|_{j_2=1,\ldots,l_2} \\ \psi_{d_2}^i(x) &= \left.\frac{a_{d_2,j_2^i}^i}{1+e^{-\left(c_{d_2,j_2^i}^i x\right)}}\right|_{j_2^i=1,\ldots,l_2}\end{aligned}$$

3.1.2. Classifier structure

The classifier based on NN is proposed following the classical strategy used to design adaptive parameter identification algorithms (Poznyak et al., 2001); that is, considering a copy of the structural approximation for uncertain nonlinear systems with time-delays defined in (4). Consequently, the approximate classifier based on NN has the following structure.

$$\begin{aligned}\frac{d}{dt}\hat{x}(t) &= A_d\hat{x}(t) + \left[W_{d_1}(t)\right]^\top \psi_{d_1}\left(\hat{x}_d(t)\right) + \left[W_{d_2}(t)\right]^\top \psi_{d_2}\left(\hat{x}_d(t)\right)u(t) \\ &\quad + \sum_{i=1}^{p}\left[W_{d_2}^i(t)\right]^\top \psi_{d_2}^i\left(\hat{x}_d(t)\right)u(t-ih) \\ \hat{x}(t) &= \hat{x}_{t_0}, \quad \forall t \in [-ph, 0] \ \ p \in \mathbb{Z}^+ \quad \hat{x}_{t_0} \in C[-ph, 0]\end{aligned} \tag{5}$$

Where $A_d \in \mathbb{R}, W_{d_1} \in \mathbb{R}^{l_1}, W_{d_2} \in \mathbb{R}^{l_2}, W_{d_2}^i \in \mathbb{R}^{l_2}$. Here $\hat{x}$ defines the identifier state that tries to approximate the corresponding sigmoid signal corresponding to each class. The vectors $W_{d_1}(\cdot)$, $W_{d_2}(\cdot)$ and $W_{d_2}^i(\cdot)$ are adaptive parameters that should be adjusted to approximate the input-output behavior of the TDNN presented in (2). The scalar A_d and functions $\psi_{d_1}(\cdot)$, $\psi_{d_2}(\cdot)$ and $\psi_{d_2}^i(\cdot)$ have the same meaning to the ones introduced in the previous section.

The learning laws for the classifier (5) are defined by

$$\begin{aligned}\frac{d}{dt}W_{d_1}(t) &= -k_{d_1}^{-1} e^{2k_d t} P \Delta_d(t) \psi_{d_1}^\top\left(\hat{x}_d(t)\right) \\ &\quad -\alpha \tilde{W}_{d_1}^{tr}(t) - \left(1 + \Lambda_{Wd_1}\right)\tilde{W}_{d_1}^{tr}(t)\end{aligned} \tag{6}$$

where Λ_{Wd_1} is a small positive scalar while k_{d_1} and k_d are positive gains that shall be adjusted. In the same manner $W_{d_2}^i(t)$ associated to the inputs are described by

$$\begin{aligned}\frac{d}{dt}W_{d_2}^i(t) &= -\int_{\tau=t-ih}^{t-(i-1)h} \alpha e^{2k_d(\tau)} \tilde{W}_{d_2}^{tr,i}(t) \bar{\Psi}_d^{+,i}(\tau) d\tau \\ &\quad -\alpha \tilde{W}_{d_2}^{tr,i}(t) + 2e^{2k_d t} P \Delta^\top(t) \psi_{d_2}^i\left(\hat{x}_d(t)\right)u(t-ih)) \\ &\quad -\left(1 + \Lambda_{W_{d_2}}\right)\tilde{W}_{d_2}^{tr,i}(t) - \Omega^i \tilde{W}_{d_2}^{tr,i}(t)(e^{2k_d t_{i+1}} \bar{\Psi}_{d_2}^i(t_{i+1}) \\ &\quad \times \left[\bar{\Psi}_{d_2}^i(t_{i+1})\right]^\top - e^{2k_d t_i} \bar{\Psi}_{d_2}^i(t_i)\left[\bar{\Psi}_{d_2}^i(t_i)\right)\right]^\top \\ \Omega^i &= \left(2k_{d_2} + \int_{\tau=t_{i-1}}^{t_i} e^{2k_d(\tau)} \bar{\Psi}_d^{+,i}(\tau) d\tau\right)^{-1}\end{aligned}$$

$$\tilde{\Psi}_d^{+,i}(\tau) = \tilde{\Psi}_{d_2}^i(\tau)\left[\tilde{\Psi}_{d_2}^i(\tau)\right]^\top, \quad \tilde{\Psi}_{d_2}^i(t) = \psi_{d_2}^i\left(\hat{x}_d(t)\right)u(t),$$
$$\Lambda_{W_{d_2}} \in \mathbb{R}^+, \; k_{d_2} \in \mathbb{R}^+ \quad (t_i = t-(i-1)h) \tag{7}$$

3.2. On-line training scheme using the continuous version of least mean square method for TDNN

The following theorem is used to obtain the evolution of the weights in the TDNN.

Theorem 1. *Let consider the time-delay uncertain system (2) with a known number of fixed delays. Suppose that perturbations and non-modeled system $\xi(x(t), t)$ affecting the classifier of EEG signals fulfills (3). If there exist positive scalars $\Lambda_{d_k} > 0$, $\Lambda_{d_k} \in \mathbb{R}$, $d_k = 1, 2$ and a positive scalar $\alpha > 0$ such that the following Riccati equation $W_d(h, P, Q, R, \alpha) = 0$ with*

$$W_d(h, P, Q, R, \alpha) = 2P\hat{A}_d + P^2R + Q$$
$$\hat{A}_d = A_d + \left(1 + \frac{\alpha}{2}\right)$$
$$R = V_1^+ + V_2^+ + \sum_{i=1}^{p} V_2^{+,i} + \Lambda_1 + \Lambda_2$$
$$Q = \left(\lambda_{\max}\left(\Lambda_1^{-1}\right)h_1 + pu^+\lambda_{\max}\left(\Lambda_2^{-1}\right)h_2\right)$$
$$h_1, h_2 \in \mathbb{R}^+ \tag{8}$$

has at least one positive definite solution $P > 0$, $P \in \mathbb{R}$, then the classification error $\Delta = x - \hat{x}$ converges exponentially to a region characterized by

$$\beta := \frac{\lambda_{\max}\left(\Lambda_1^{-1}\right)}{\lambda_{\max}\left(\Lambda_f\right)}\tilde{f}_0 + \frac{\lambda_{\max}\left(\Lambda_2^{-1}\right)}{\lambda_{\max}\left(\Lambda_\xi\right)}\Upsilon$$

that is

$$\overline{\lim_{t\to\infty}} e^{2kt}\left(\alpha\Delta^\top(t)P\Delta(t) - \beta\right) = 0 \tag{9}$$

Proof. The learning method used to adjust the parameters included in the TDNN proposed in this study can be obtained using the so-called Lyapunov stability analysis. If the full-time derivative was applied to the previous equation, and using (2) and (5), the following delayed differential equation is obtained:

$$\frac{d}{dt}\Delta(t) = A\Delta(t) + [W_1^*]^\top\left(\psi_1(x(t))\text{-}\psi_1\left(\hat{x}(t)\right)\right)$$
$$+\left[\tilde{W}_1(t)\right]^\top\psi_1(\hat{x}(t)) + \left[\tilde{W}_2(t)\right]^\top\psi_2(\hat{x}(t))u(t)$$
$$\sum_{i=1}^{p}\left[W_2^i(t)\right]^\top\tilde{\psi}_2^i\left(\hat{x}(t), x(t)\right)u(t-ih)$$
$$+\tilde{f}(x(t), u(t), u(t-h), \ldots, u(t-ph)) + \xi(x(t))$$

By using the so-called Lyapunov–Krasovskii functional, we constructed the learning algorithms. This kind of functionals has been widely used to prove the existence of an equilibrium point for time delay dynamic systems. For this paper, the candidate of Lyapunov–Krasovskii functional was designed as follows

$$V\left(t, \Delta, \tilde{W}_1, \tilde{W}_2^i\right)$$
$$= e^{2k_d t}P\Delta^2 + k_{d_1}\tilde{W}_1^\top\tilde{W}_1 + k_{d_2}\left[\tilde{W}_2\right]^\top\tilde{W}_2(t)$$
$$+\sum_{i=1}^{p}k_{2,i}\Xi_2^i + \sum_{i=1}^{p}\int_{\tau=t-ih}^{t-(i-1)h}e^{2k_d(\tau)}\left(\left[\tilde{\Psi}_2(\tau)\right]^\top\Xi_2\tilde{\Psi}_2^i(\tau)\right)d\tau$$
$$\Xi_2^i = \left[W_2^i(t)\right]^\top W_2^i(t) \tag{10}$$

The full time derivative with respect to time of this Lyapunov–Krasovskii functional provides the following differential equation

$$\frac{d}{dt}V(t) = 2ke^{2k_d t}P\Delta^2(t) + 2e^{2k_d t}\Delta(t)P\frac{d}{dt}\Delta(t) + 2k_{d_1}\tilde{W}_1^\top(t)\frac{d}{dt}\tilde{W}_1(t)$$
$$+2k_{d_2}\tilde{W}_2^\top(t)\frac{d}{dt}\tilde{W}_2(t) + 2\sum_{i=1}^{p}k_{2,i}d\Xi_2^i(t)$$
$$+\sum_{i=1}^{p}e^{2k_d t_i}\left(\tilde{\Psi}_2^i(t_i)^\top\Xi_2^i(t)\tilde{\Psi}_2^i(t_i)\right)$$
$$-\sum_{i=1}^{p}e^{2k_d t_{i+1}}\left(\tilde{\Psi}_2^i(t_{i+1})\right)^\top\Xi_2^i(t)\tilde{\Psi}_2^i(t_{i+1})$$
$$+2\sum_{i=1}^{p}\int_{\tau=t_{i+1}}^{t_i}e^{2k_d(\tau)}\left(\tilde{\Psi}_2^i(\tau)\right)^\top\tilde{\Psi}_2^i(\tau)d\Xi_2^i(t)d\tau$$
$$d\Xi_2^i(t) := \left[W_2^i(t)\right]^\top\frac{d}{dt}W_2^i(t) \tag{11}$$

The second term in the previous differential equation is analyzed by means of the well-known matrix inequality $XY^\top + YX^\top \leq X\Lambda X^\top + Y\Lambda^{-1}Y^\top$ valid for any $X, Y \in \mathbb{R}^{r\times s}$ and any $0 < \Lambda = \Lambda^\top \in \mathbb{R}^{s\times s}$. Then, using this last inequality a number of times, one has

(a) $$2e^{2k_d t}\Delta P A\Delta = 2e^{2k_d t}PA\Delta^2(t)$$

(b) $$2e^{2k_d t}\Delta(t)P[W_1^*]^\top\left(\psi_1(x(t)) - \psi_1\left(\hat{x}(t)\right)\right)$$
$$\leq e^{2k_d t}\Delta^2(t)(P^2[W_1^*]^\top\Lambda_a W_1^* + \left(\lambda_{\max}\left(\Lambda_a^{-1}\right)h_1\right))$$

(c) $$\sum_{i=1}^{p}2e^{2k_d t}\Delta(t)P\left[W_2^{*,i}\right]^\top\tilde{\psi}_2^i\left(x(t), \hat{x}(t)\right)u(t-ih)$$
$$\leq\sum_{i=1}^{p}2e^{2k_d t}\Delta^2(t)P^2\left[W_2^{*i}\right]^\top\Lambda_b W_2^{*i}$$
$$+\sum_{i=1}^{p}2e^{2k_d t}u^+\lambda_{\max}\left(\Lambda_b^{-1}\right)h_2\Delta^2(t)$$

(d) $$2e^{2k_d t}\Delta(t)P\tilde{f}(x(t), u(t), u(t\text{-}h), \ldots, u(t\text{-}ph))$$
$$\leq e^{2k_d t}\left(\Delta^2(t)P^2\Lambda_1 + \frac{\lambda_{\max}\left(\Lambda_1^{-1}\right)}{\lambda_{\max}\left(\Lambda_f\right)}\tilde{f}_0\right)$$

and finally

(e) $$2e^{2k_d t}\Delta(t)P\xi(t) \leq e^{2k_d t}\Delta^2(t)P^2\Lambda_2 + e^{2k_d t}\frac{\lambda_{\max}\left(\Lambda_2^{-1}\right)}{\lambda_{\max}\left(\Lambda_\xi\right)}\Upsilon$$

Bringing all these results together and substitute them in the time derivative of Lyapunov–Krasovskii functional (11), one gets

$$\frac{d}{dt}V(t) \leq e^{2k_d t}2\Delta^2(t)P(A+1) + \Delta^2(t)P^2R + \Delta^2(t)Q$$
$$+2k_{d_1}\tilde{W}_1^\top(t)\frac{d}{dt}\tilde{W}_1(t) + 2e^{2k_d t}\Delta(t)P\left[\tilde{W}_1(t)\right]^\top\psi_1(\hat{x}(t))$$
$$+k_{d_2}\left[\tilde{W}_2(t)\right]^\top\frac{d}{dt}\tilde{W}_2(t) + 2\sum_{i=1}^{p}k_{2,i}d\Xi_2^i(t)$$
$$+2e^{2k_d t}\Delta(t)P\left[\tilde{W}_2(t)\right]^\top\psi_2\left(\hat{x}(t)\right)u(t)$$
$$+\sum_{i=1}^{p}e^{2k_d t_i}\left[\tilde{\Psi}_2^i t_i^\top\right]\Xi_2^i(t)\tilde{\Psi}_2^i(t_i)$$

$$-\sum_{i=1}^{p} e^{2k_d t_{i+1}} \left[\tilde{\Psi}_2^i(t_{i+1})\right]^\top \Xi_2^i(t)\tilde{\Psi}_2^i(t_{i+1})$$
$$+2\sum_{i=1}^{p}\int_{\tau=t_{i+1}}^{t_i} e^{2k_d(\tau)}\left[\tilde{\Psi}_2^i(\tau)\right]^\top d\Xi_2^i(t)\tilde{\Psi}_2^i(\tau)d\tau$$

The assumption given in the theorem statement ensured the negativeness of $W(h, P, Q, R, \alpha)$ with respect to the dependent parameters and using the adjustment laws described in (6) and (7) yields to the following differential inclusion $\dot{V}(t) \leq -\alpha e^{2k_d t}\Delta^2(t)P + e^{2k_d t}\beta$. If we consider the set defined by

$$\Theta(t) = \left\{\Delta(t) \mid \beta < \alpha\Delta^2(t)P\right\}$$

Then, if $\Delta(t) \in \Theta$,

$$\dot{V}(t) \leq 0 \tag{12}$$

It is straightforward to prove that there exist positive definite matrices N_1, $N_{2,\,i}$, $N_{3,\,i}$ and $N_{4,\,i}$ $(i = 1, \ldots, p)$ such that

$$\frac{d^2}{dt^2}V(t) \leq e^{2k_d}\Delta^2 N_1 + \sum_{i=1}^{p} e^{2k_d t_i}\tilde{\psi}_2^\top(t_i)N_{2,i}\tilde{\psi}_2^\top(t_i)$$
$$+\sum_{i=1}^{p} e^{2k_d t_{i+1}}\tilde{\psi}_2^\top(t_{i+1})N_{3,i}\tilde{\psi}_2^\top(t_{i+1})$$
$$+\sum_{i=1}^{p}\int_{\tau=t_{i+1}}^{t_i} e^{2k_d}\tilde{\psi}_2^\top(\tau)N_{4,i}\tilde{\psi}_2^\top(\tau)d\tau$$

In view of (12), the right-hand side of the last differential inclusion is bounded, then by the Barbalat's Lemma, the function $V(t)$ is absolute continuous, and finally one got

$$\overline{\lim_{t\to\infty}} e^{2k_d t}\left(\alpha\Delta^2(t)P - \beta\right) = 0 \tag{13}$$

Firstly, the result presented in (13) implies that $\alpha\Delta^2(t)P - \beta = o\left(e^{2k_d t}\right)$, then $\Delta^2(t)$ approaches its ultimately bound $\beta\alpha^{-1}P^{-1}$ faster than $e^{-2k_d t}$. What is known with this result is that training of TDNN is attained not only with a predefined quality value but also how fast this value is reached. □

4. Database evaluated by the classifier

The database considered in this study was taken from (of Freiburg, 2012). The entire data collection contains 500 EEG recordings divided in 5 different clinical classifications. EEG data were acquired using a Neurofile NT digital video EEG system with 128 channels, 173.61 Hz sampling rate, and a 16 bit analogue-to-digital converter. Notch or band pass filters have not been applied during the recording of the signals. So, each signal is considered to be raw, that is, without any pretreatment. The database is divided in 5 classes. Each class contains 100 samples. The database of EEG signals considered in this study has samples of five different classes but with only one acquisition channel. Therefore, only the information of one single channel already selected was used to perform the classification analysis.

Sets Z and O are signals recorded from the EEG surface with volunteers relaxed in and awaken mode with eyes open and closed respectively. Set N was taken from the hippocampal formation of the opposite hemisphere of the brain, set F was recorded from the epileptogenic zone, while set S only contained seizure activity (Polat & Günes, 2007). For the purpose of this article, only signals included in classes S and Z were considered. Nevertheless, the same pattern classification methodology can be easily extended to include all the signals in the 5 classes previously described as well as more acquisition channels.

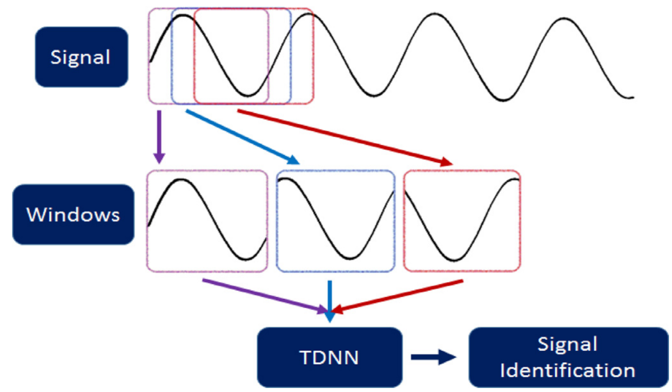

Fig. 3. EEG signal windowing process to fed the TDNN.

5. Numerical evaluations of the TDNN based classifier

The training as well as the validation processes were evaluated with 1,2,..., 15 delays. This part of the study was conducted to evaluate the relation between the classification accuracy with respect to the number of delays. The number of columns for both W_1^i and W_2^i was 1. This selection was also done in agreement of the well-known problem of over-fitting exhibited by different NN. This problem occurs when the training error in each trial is driven to zero or to a very small value. Nevertheless, during the validation process, this error is large enough that yields to obtain low percentages of successful classification results. This condition occurs when the NN memorized the training dataset, but it has not gain the ability to generalize the relationship between delayed inputs and output. Therefore, to reduce as much as possible the number of weights in the TDNN yields to reduce the effect of overfitting on the classification process. Each numerical evaluation in the training process took 7 minutes in average. All the numerical experiments were evaluated in a Dual Xeon E5-2637v3 3.5 GHz, 15 MB cache, 9.60 QPI (Four-Core) 192 GB DDR4-2133 REG ECC (12–16 GB DIMMS) 2 x NVIDIA Quadro K2200 4 GB.

Fig. 3 illustrates how the proposed TDNN works. First, the signal is windowed in same size portions, then those windows are employed as inputs for the TDNN. The windowed information from the signal represent the effect of delayed input. The inputs are organized as follow; present time window, 1 time-delay datum, 2 time-delay datum, up to n time-delay datum (it is expected that increasing the number of delays shall improve the classification capabilities of the TDNN). The purpose of these windowing is to fed the TDNN with more information related to the characteristics of each class.

The sampling frequency of EEG signals used in the classifier study was 173.61 Hz. So, in the case of one delay, the total number of samples was 2 (without delay and one delay) and the time was 5.76 ms. On the other hand, when the number of delay was 15 (the biggest one), the signal time injected into the classifier was 86.40 ms. The complete evaluation of training and testing used the entire signal long (23.6 s). This characteristic was kept constant in all the studies.

5.1. Training process

The number of delays was varied from 1 to 15. We used 12 components for the internal weights W_{d_1} while we used 7 components for each weight associated to the delayed input. Therefore, a total number of weights components given by 12+(p+1)*7 was used where p is the number of delays. Then, the classifier proposed in this part of the work was evaluated with the following parame-

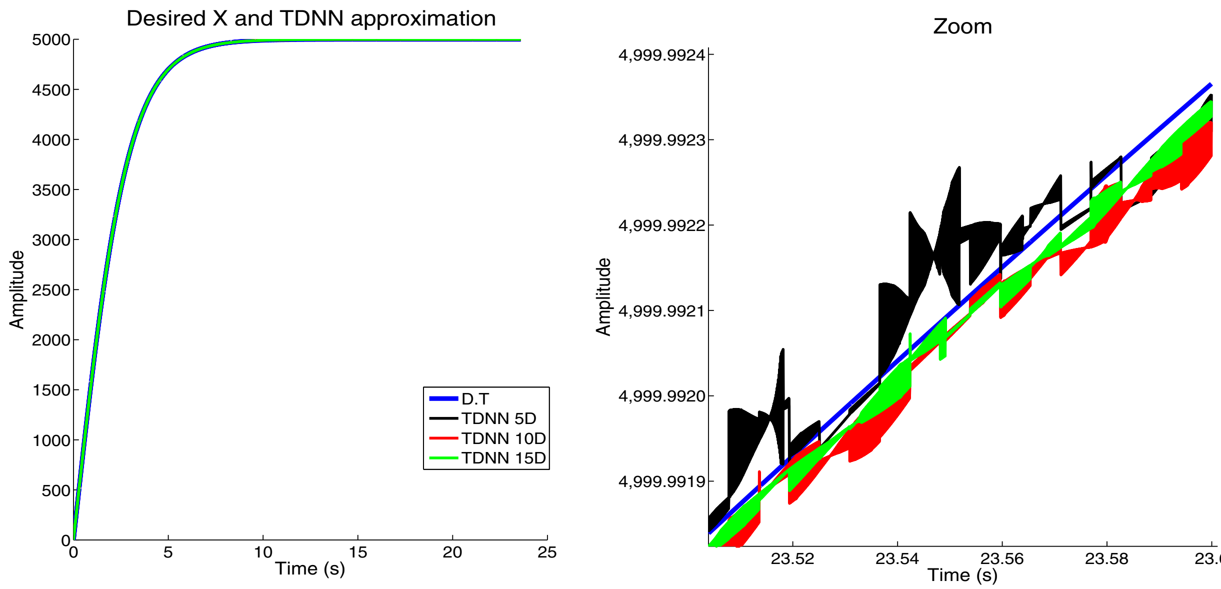

Fig. 4. On the left size of the figure, the desired trajectory (D.T.) in blue line, the TDNN waveform with 5 input delays in black line, TDNN with 10 input delays in red line, TDNN with 15 input delays in green line are shown. On the right size of the figure, an amplified view of the same comparison is depicted. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ters

$$A = -2.6 \qquad W_{d_1}(0) = 2.0 * ones(12) \qquad m = n = 1$$
$$W_{d_2}(0) = 7.5 * ones(7) \qquad W^i_{d_2}(0) = 7.5 * ones(7) \qquad i = 1, \ldots, 7$$

These parameters were obtained after several numerical evaluations. Today, there is not a formal manner to select these parameters in a different form. Notice that the dynamic nature of the TDNN based classifier reduces the necessity of having hidden layers in its structure. However, the inclusion of hidden layers in the classifier structure is still a matter of further investigation. In general, if the number of EEG channels is 1, the number of rows for both W_{d_1} and W_{d_2} can be freely adjusted. Nevertheless, the size of the signal window (number of delays) used in the classifier defines the number of weight $W^i_{d_2}$.

Fig. 4 a shows the approximation performance of the CNN when performing the training for a specific EEG signal taken from the class S. The sigmoid function used to represent the class was

$$x_l(t) = \frac{3}{1 + e^{-2t}} \tag{14}$$

The convergence between both signals is an indirect demonstration of the training efficiency generated by the learning laws proposed for the pattern classifier based on CNN. Fig. 4b shows an amplified view of the tracking provided by the TDNN and the sigmoid function used to define signals included in the class Z. Notice that 4 different waveforms appear in the same plot. These signals correspond to the TDNN trajectories when different number of delays were considered in its structure. Notice that when the number of delays is five, the approximation of the sigmoid function is not as good as in the case when the number of delays increased up to 10 or 15. However, there is no evident enhance when the number of delays is above 10. This condition can introduce an indirect method to define the size of the window needed to obtain an accurate pattern classification quality. Notice that when the number of delays is referred, the length of each window is what is explained. All the windows were evaluated independently and they are not overlapped. Therefore, there is zero samples overlapped.

In order to evaluate the training quality, the least mean square error (LMSE) was calculated with respect to the number of delays considered in the input signal. This error served to evaluate the degree of over-fitting during the training process. Fig. 5 depicts the relationship between the LMSE and the number of delays included in the EEG input signals. LMSE was calculated for both, the signals included in the class S as well as the ones included in the class Z. This value was the result of averaging the LMSE for all the 100 signals included in each class.

Remark 1. Authors have reviewed the existing literature regarding the estimation of window size. There are some results that have exploited the spectral information of the signal or the nature of the electrophysiological signal (Roshan et al., 2012; Sierra-Marcos et al., 2015). These results can be applied to the classifier proposed in this study. Nevertheless, this aspect was not discussed in detail in our article because we were trying to analyze the dependence of training accuracy with respect to the number of delays. This aspect has been much less studied and to the authors' knowledge, and today there is not a formal way to determine the number of delays a-priori if a predefined accuracy error is expected.

5.2. Validation procedure

Validation results for database are divided in two studies: the first uses the well-known training-generalization-validation, the second evaluation used the k-cross fold training with $k = 5$. The results of the classification process obtained for the training-generalization-validation process when the number of delays was 5 are contained in Table 1. This method achieved a 94.88% of correct classification for all the signals included in the selected database. This result is in the same range of those reported in similar studies (using the same database to evaluate different classifiers) where also signals from the same two classes (S and Z) were considered.

From the 5-fold cross validation applied to the database I, a total classification accuracy of 100.0% (Table 2). For this part of the process, the total set of signals in the database was employed. Notice that even when the final classification percentage is high, there is still the necessity of increasing the number of classes that should be included in the study.

Table 3 contains the information of the training-generalization-validation process when the number of delays was increased up to 10. One may notice that classification accuracy also increased from

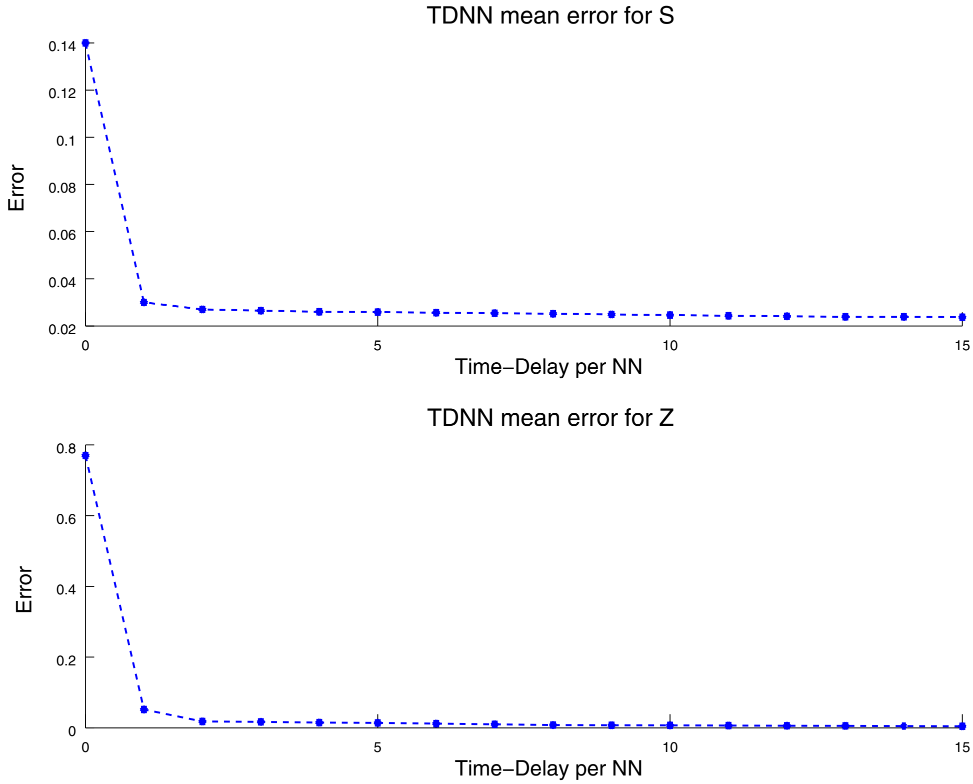

Fig. 5. Effect of number of delays on the training process.

Table 1
Results from the generalization validation method for 5 delays.

	S	Z
Samples	100	100
Training	100%	100%
Generalization	82%	73%
Independent test	87%	81%

Table 2
Results from the 5-fold cross validation method for 5 delays.

	S	Z
Samples	100	100
1°S. CfA	100.0%	100.0%
2°S. CfA	100.0%	100.0%
3°S. CfA	100.0%	100.0%
4°S. CfA	100.0%	100.0%
5°S. CfA	100.0%	100.0%
Total CfA	100.0%	100.0%

CfA: Classification accuracy, S: Segment.

Table 3
Results from the generalization validation method for 10 delays.

	S	Z
Samples	100	100
Training	100%	100%
Generalization	91%	86%
Independent test	88%	83%

Table 4
Results from the 5-fold cross validation method for 10 delays.

	S	Z
Samples	100	100
1°S. CfA	100.0%	100.0%
2°S. CfA	100.0%	100.0%
3°S. CfA	100.0%	100.0%
4°S. CfA	100.0%	100.0%
5°S. CfA	100.0%	100.0%
Total CfA	100.0%	100.0%

CfA: Classification Accuracy, S: Segment.

82.0% to 93.0% in the case of the signals contained in the class S while a similar increment was observed for the signals included in the class Z (73.0% to 86.0%). These results confirm the impact of increasing the number of delays in the input signal. Notice that a more expensive calculus must be done because the number of weights in the TDNN is also doubled. Nevertheless, the improvement in the pattern classification quality justified this more complex TDNN design.

The results of the 5-fold cross validation applied to the database I showed a total classification accuracy of 100.0% (Table 4). Therefore, this evaluation test only confirmed the results obtained when the number of delays was 5.

Some additional tests were done with a larger number of delays considered in the input signal. Nevertheless, the classification accuracy was not increased at all when the number of delays was larger than 10.

5.3. Classification results for borderline signals

In order to evaluate the classification capacities of the algorithm proposed in this study, some evaluation tests were proposed. A set

Table 5
NPV, PPV, TPR and SPC for 1,5,10 and15 delays.

Delays	*1*		*5*		*10*		*15*	
class	S	Z	S	Z	S	Z	S	Z
NPV	0.9010	0.875	0.92307	0.8888	0.9368	0.9578	0.9473	0.9578
PPV	0.0989	0.125	0.0769	0.1111	0.0631	0.0652	0.0526	0.0421
TPR	0.9213	0.9058	0.9230	0.9302	0.9569	0.9450	0.9574	0.9680
SPC	0.8181	0.7333	0.7777	0.7142	0.8571	0.6666	0.8333	0.6666

Table 6
Comparisson between different classification techniques applied to EEG singel trail signals.

Researches	Dataset	CfA	Validation method
(Srinivasan, Eswaran, & Sriraam, 2005)	Z,S	99.6%	NC
(Kannathala, Rajendra-Acharyab, Limb, & Sadasivana, 2005)	Z,S	92.22%	NC
(Kannathala, Choob, Rajendra-Acharyab, & Sadasivana, 2005)	Z,S	∼ 90%	NC
(Polat & Günes, 2007)	Z,S	98.72%	*k*-fold
(Subasi, Akin, Kiymik, & Erogul, 2005)	Z,S	95%	*k*-fold
(Ocak, 2009)	Z,S	96.65%	NC
This study	Z,S	97.32%	*k*-fold, gen.

of 200 signals was prepared artificially accordingly to the following procedure: consider a first signal S_i from the class *S* and a second one S_j from a class *Z*, the hybrid signal S_{ij} was generated as the convex combination of $S_{ij} = \lambda S_i + (1 - \lambda S_j)$. Then, these signals were tested on the parallel arrangement of trained TDNN. In particular, for this particular analysis, $\lambda = 0.8$ meaning that 80% of a signal belong to a class *S* was combined with a 20% of a single belonging a class *Z*. A second round of analysis considered the analysis of borderline signals using $\lambda = 0.7$. After the total evaluation of classification process, the TDNN based classifier achieved a 95% of correct classification for database I. This result shows the detectability capacity of the classifier proposed in this study.

5.4. Confusion matrix evaluation

In order to detail the classification capacities of the proposed identifier, the analysis of Negative Prediction Value (NPV), Positive Prediction Value (PPV), True Positive Rate (TPR) and Specificity (SPC). These values are calculated according to the following equations (Fawcett, 2005):

$$NPV = \frac{\text{TruePositives}}{(\text{TruePositives} + \text{TrueNegatives})}$$
$$PPV = \frac{\text{TrueNegatives}}{(\text{TrueNegatives} + \text{TruePositives})}$$
$$TPR = \frac{\text{TruePositives}}{(\text{TruePositives} + \text{FalseNegative})}$$
$$SPC = \frac{\text{TrueNegatives}}{(\text{TrueNegatives} + \text{FalsePositive})}$$

Table 5 depicts the results obtained after the evaluation of the parameters included in the confusion matrix. In agreement to the accuracy results as well as the predictive analysis (with all results above 90%), the classifier proposed in this study seems to be a reliable method to classify certain characteristics in EEG signals.

Some examples of different classification techniques applied to EEG signals are shown in the Table 6.

6. Conclusions

In this paper, an EEG signal classifier was developed based on a class of TDNN with delays appearing in the input signal. This characteristic was proposed to take into account the concept of signal windowing. The capability of a the TDNN to be employed as a EEG signal pattern classifier was tested with the information collected in a traditional databases. The training process (based on the method to adjust the weights in the TDNN) was also proposed based on the technique of Lyapunov stability analysis. The training results as well as the validation percentages were reported and evaluated for signals belonging to a couple of classes included in database. In order to evaluate the effectiveness of the classifier proposed in this study, the classification percentages were compared to the results achieved by some other classifiers based on NN that have used similar information. Even though the results reported by others may be higher in their total correct classification accuracy, they are not working with the entire set of signals considered in this study database. Moreover, the method presented here used the raw EEG signal without considering the application of any preliminary signal treatment. Even more, accordingly to the classification results obtained in this study, this kind of classifier can be extended to some other problems where the raw signal can be more informative as is instead of making several steps of pretreatment. In general, the application of this kind of classifier only requires the selection of training time *T* and to perform the exhaustive supervised training. This is an advantage of the classifier structure proposed in this study because non-particular pretreatment should be designed.

Then, the classifier based on continuous delayed neural networks can classify electrophysiological signal on-line by considering the presence of delayed input information. This characteristic can be considered as a novel contribution because today, the classifiers based on neural networks and working on-line uses a different scheme where the delayed information is used as a vector input with the same type of learning laws. This study considers a different option where the learning laws were specially designed to consider the specific impact of each delay on the classifier performance. On the other hand, the classification scheme based on continuous delayed neural networks uses a class of associative memory where a parallel neural network with fixed weights was proposed to obtain the class where the EEG signal belongs.

On our point of view, the more important issue that must be improved in this classifier is the time needed to perform the training process. Because the numerical implementation of integral operation usually takes long periods of time, the training period of time can take several hours. This condition is not longer occurring when the testing phase is evaluated. One additional aspect to consider is the effect of increasing the number of delays. If this number increases, then the simulation time does not increase proportionally. This increment is polynomial.

We believe that the following are potential future research opportunities for the type of classifier considered in this study:

- Finite-time convergence of learning laws in the classifier would provide a higher degree of robustness against noises in the signal.
- Deep learning methods such as extreme learning can be considered as a suitable option to develop new and more powerful variations of the classifier proposed in this study.
- Classifiers based on EEG information as well as visual information (for example) can be used to consider the application of the classifier proposed in this study in the analysis of evoked potentials.
- The analysis of accuracy percentage depending on the type of activation function should be explored in order to determine the best approximation basis of classifier developed in this study.

Acknowledgment

Thanks to the Instituto Politecnico Nacional for supporting the project SIP-20150287 and Thanks to the project supported by CONACYT- CB 2013-01 221867.

References

Ang, K., & Chin, Z. (2012). New classification techniques for electroencephalogram (EEG) signals and a real-time EEG control of a robot. *Pattern Recognition, 45*(6), 2137–2144.
Arik, S. (2000). Stability analysis of delayed neural networks. *IEEE Transactions on Circuits Systems, 47*, 1089–1092.
Byung-Chan, M., Seung-Hyun, J., In-Hyeng, K., Lee, D., JinKyu, K., SangTae, L., & Kazuyoshi, S. (2003). Analysis of mutual information content for EEG responses to odor stimulation for subjects classified by occupation. *Chemical Senses, 28*(9), 741–749.
Cao, J. (2000). Periodic oscillation and exponential stability of delayed CNNs. *Physics Letters A, 270*, 157–163.
Cinar, E., & Sahin, F. (2013). New classification techniques for electroencephalogram (EEG) signals and a real-time EEG control of a robot. *Neural Computing and Applications, 22*(1), 29–39.
Cybenko, G. (1989). Approximation by superpositions of sigmoidal function. *Mathematics of Control Signals and Systems, 1989*(2), 303–314.
Daly, J. (2013). *Introduction to neural engineering for motor rehabilitation* (pp. 463–475). John Wiley & Sons, Inc.
Emilia, F. (2014). Tutorial on Lyapunov-based methods for time-delay systems. *European Journal of Control, 20*(2014), 271–283.
Faust, O., Acharya, H., Adeli, R., & Adeli, A. (2015). Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis. *Seizure, 26*, 56–64.
Fawcett, T. (2005). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(2006), 861–874.
Fridman, E., Seuret, A., & Richard, J.-P. (2004). Robust sampled-data stabilization of linear systems: an input delay approach. *Automatica, 40*(8), 1441–1446.
Garces Correa, A., Orosco, L., Diez, P., & Laciar, E. (2015). Automatic detection of epileptic seizures in long-term EEG records. *Computers in Biology and Medicine, 57*(1), 66–73.
Ge, D., Srinivasan, N., & Krishnan, S. M. (2007). *Advances in cardiac signal processing* (pp. 209–224). Berlin, Heidelberg: Springer-Verlag.
Hale, J. K. (1977). *Theory of functional differential equations*. Springer Verlag Gmbh.
Hramov, A., Koronovskii, A., Makarov, V., Pavlov, A., & Sitnikova, E. (2015). Wavelets in neuroscience. In *Springer Series in Synergetics* (pp. 253–312). Berlin, Heidelberg: Springer.
Hwang, H., Kim, S., Choi, S., & Im, C. (2013). EEg-based brain-computer interfaces: A thorough literature survey. *International Journal of Human-Computer Interfaces, 29*(13), 814–826.
Joy, M. (2000). Results concerning the absolute stability of delayed neural networks. *Neural Networks, 13*, 613–616.
Kannathala, M., Choob, N., Rajendra-Acharyab, U., & Sadasivana, P. (2005). Entropies for detection of epilepsy in EEG. *Computer Methods and Programs in Biomedicine, 2005*(80), 187–194.
Kannathala, N., Rajendra-Acharyab, U., Limb, C., & Sadasivana, P. (2005). Characterization of EEG–a comparative study. *Computer Methods and Programs in Biomedicine, 2005*(80), 17–23.
Koley, B., & Dey, D. (2012). An ensemble system for automatic sleep stage classification using single channel EEG signal. *Computer in Biology and Medicine, 42*(12), 1186–1195.
Kruszewski, A., Jiang, W. J., Fridman, E., Richard, J.-P., & Toguyeni, A. (2012). A switched system approach to exponential stabilization through communication network. *IEEE Transactions on Control Systems Technology, 20*(4), 887–900.
Kumar, S., Kanhangad, V., & Pachori, R. (2015). Classification of seizure and seizure-free EEG signals using local binary patterns. *Biomedical Signal Processing and Control, 15*, 33–40.
Li, Z., Doherty, J., Lebedev, M., & Nicolelis, M. a. (2011). Adaptive decoding for brain–machine interfaces through Bayesian parameter updates. *Neural Computation, 23*(12), 3162–3204.
Liao, X., Chen, G., & Sanchez, E. (2002). Delay dependent exponential stability analysis of delayed neural networks: An LMI approach. *Neural Networks, 15*, 855–866.
Liaoa, X., Chenb, G., & Sanchez, E. (2002). Delay dependent exponential stability analysis of delayed neural networks: an LMI approach. *Neural Networks, 15*, 855–866.
Manchanda, R., Norman, R., A. , M., Harricharan, R., Northcott, S., & Richard, J. (2014). Electroencephalographic abnormalities and 5-year outcome in first-episode psychosis. *Canadian Journal of Phychiatry, 59*(5), 285–288.
Millan, J., & Mourino, J. (2003). Asynchronous BCI and local neural classifiers: An overview of the adaptive brain interface project. *IEEE Transactions on Neural Systems and Rehabilitation Engineerings, 11*(2), 159–161.
Ocak, H. (2009). Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Systems with Applications, 36*, 2027–2036.
Polat, K., & Günes, S. (2007). Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. *Applied Mathematics and Computation, 87*, 1017–1026.
Polyakov, A. (2012). Minimization of disturbances effects in time delay predictor-based sliding mode control systems. *Journal of The Franklin Institute, 394*(4), 1380–1396.
Poznyak, A., Sánchez, E., & Yu, W. (2001). *Differential neural networks for robust nonlinear control (identification, state estimation and trajectory tracking)*. World Scientific.
Roshan, J., Acharya, R., Hong, J., Petznick, A., Yanti, R., Kuang, C., & Tong, L. (2012). Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals. *International Journal of Neural Systems, 22*(06), 16.
Saletu, B., Anderer, P., SaletuZyhlarz, O., Arnold, G. M., & Pascual-Marqui, R. (2002). Classification and evaluation of the pharmacodynamics of psychotropic drugs by single-lead pharmaco-EEG, EEG mapping and tomography (loreta). *Methods and Findings in Experimental and Clinical Pharmacology, 24*(C), 97–120.
Sargolzaei, S., Cabrerizo, M., Sargolzaei, A., Noei, S., Eddin, A., Rajaei, H., et al. (2015). A probabilistic approach for pediatric epilepsy diagnosis using brain functional connectivity networks. *BMC Bioinformatics, 16*(7), 11.
Sawata, R., Ogata, T., & Haseyama, M. (2015). Human-centered favorite music estimation: EEG-based extraction of audio features reflecting individual preference. In *2015 IEEE international conference on digital signal processing (DSP)* (pp. 818–822). IEEE.
Sierra-Marcos, A., Scheuer, M., & Rosseti, A. (2015). Seizure detection with automated EEG analysis: A validation study focusing on periodic patterns. *Clinical Neurophysiology, 126*(3), 456–462.
S. P. P. F. U. of Freiburg (2012). *EEG database*. online. http://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/eeg-database.
Srinivasan, V., Eswaran, C., & Sriraam, N. (2005). Artificial neural network based epileptic detection using time-domain and frequency-domain features. *Journal of Medical Systems, 29*(6), 647–660.
Stefanidis, V., Anogiannakis, G., evangelou, A., & Poulus, M. (2015). *Optimization, control, and applications in the information age: (Vol. 130* (pp. 349–357)). Springer International Publishing.
Subasi, A., Akin, M., Kiymik, K., & Erogul, O. (2005). Automatic recognition of vigilance state by using a wavelet-based artificial neural network. *Neural Computing and Applications, 14*, 45–55.
Subha, P., Joseph, P., Acharya, R., & MinLim, C. (2008). EEG signal analysis: A survey. *Journal of Medical Systems, 34:*(2010), 195–212.
Mier-y Terán-Romero, L., Silber, M., & Hatzimanikatis, V. (2010). The origins of time-delay in template biopolymerization processes. *PLoS Computational Biology, 6*(4), e1000726.
Yu, X., Fumoto, M., Nakatani, T., Sekiyama, Y., Kikuchi, H., Seki, Y., … Arita, I. (2011). Activation of the anterior prefrontal cortex and serotonergic system is associated with improvements in mood and EEG changes induced by zen meditation practice in novices. *International Journal of Psychophysiology, 80*(2), 103–111.