INSTITUTO POLITÉCNICO NACIONAL

CENTRO DE INVESTIGACIÓN EN COMPUTACIÓN

TESIS

Space-time pattern recognition in electrophysiological signals from evoked potentials using dynamic neural networks

Para obtener el grado de

Doctorado en Ciencias de la Computación

P R E S E N T A:

M. en C. Mariel Alfaro Ponce

DIRECTORES DE TESIS

Dr. Amadeo José Argüelles Cruz
Dr. Jorge Isaac Chairez Oria

MÉXICO D.F. JUNIO 2015

SIP-14 BIS

INSTITUTO POLITÉCNICO NACIONAL

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

ACTA DE REVISIÓN DE TESIS

En la Ciudad de México, D.F. siendo las 10:00 horas del día 29 del mes de mayo de 2015 se reunieron los miembros de la Comisión Revisora de la Tesis, designada por el Colegio de Profesores de Estudios de Posgrado e Investigación del:

Centro de Investigación en Computación

para examinar la tesis titulada:

"Space-time pattern recognition in electrophysiological signals from evoked potentials using dynamic neural networks"

Presentada por el alumno:

ALFARO	**PONCE**	**MARIEL**
Apellido paterno	Apellido materno	Nombre(s)

Con registro: B 1 1 0 8 7 8

aspirante de: **DOCTORADO EN CIENCIAS DE LA COMPUTACIÓN**

Después de intercambiar opiniones los miembros de la Comisión manifestaron ***APROBAR LA TESIS***, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA

Directores de tesis

Dr. Amadeo José Argüelles Cruz

Dr. Jorge Isaac Chairez Oria

Dr. Sergio Suárez Guerra

Dr. Oleksiy Pogrebnyak

Dr. Oscar Camacho Nieto

Dr. Cornelio Yáñez Márquez

PRESIDENTE DEL COLEGIO DE PROFESORES

Dr. Luis Alfonso Villa Vargas

INSTITUTO POLITECNICO NACIONAL
CENTRO DE INVESTIGACION EN COMPUTACION
DIRECCION

INSTITUTO POLITÉCNICO NACIONAL

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

CARTA CESIÓN DE DERECHOS

En la Ciudad de México, D.F. el día 1 del mes de Junio del año 2015, la que suscribe **Mariel Alfaro Ponce** alumna del Programa de **Doctorado en Ciencias de la Computación**, con número de registro **B110878**, adscrita al **Centro de Investigación en Computación**, manifiesta que es la autora intelectual del presente trabajo de Tesis bajo la dirección del **Dr. Amadeo José Argüelles Cruz** y el **Dr. Jorge Isaac Chairez Oria** y cede los derechos del trabajo titulado **Space-time pattern recognition in electrophysiologial signals from evoked potentials using dynamic neural networks**, al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso de la autora y/o directores del trabajo. Este puede ser obtenido escribiendo a las siguientes direcciones **alfaromariel@yahoo.com, ichairezo@gmail.com** y **amadeomx@gmail.com**. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

M. en C. Mariel Alfaro Ponce

Nombre y firma de la alumna

Space-time pattern recognition in electrophysiological signals from evoked potentials using dynamic neural networks

por

M. en C. Mariel Alfaro Ponce

Resumen

Durante siglos la humanidad ha tratado de descifrar la manera en que funciona el cerebro, tecnologías como el electroencefalograma (EEG) son capaces de registrar la actividad eléctrica del cerebro [67]. La popularidad del EEG en la investigación científica se debe a la gran cantidad de estudios y aplicaciones que se pueden realizar a través de sus registros, también por la portabilidad y el costo del equipo que es menor a otros que permiten registrar la actividad cerebral. Las técnicas derivadas del EEG incluyen los potenciales evocados (PE), que registra un promedio de la actividad EEG con respecto al tiempo, utilizando la presencia de un estímulo que puede ser de tipo visual, somatosensorial o auditivo [49]. Está claro que la lectura de la respuestas electrofisiológicas del cerebro no representan un reto tecnológico, sin embargo, el problema principal actual respecto a estas señales, centra en la decodificación de la información integrada en estas señales [85]. En la actualidad, existen muchos algoritmos de decodificación que se han propuesto para ello, que van desde estimadores lineales óptimos, hasta diferentes versiones de decodificadores bayesiano, e incluso diferentes topologías de redes neurales artificiales (ANN, por su definición en inglés). Hoy en día, la mayoría de los algoritmos de interpretación automática de EEG se enmarcan en la denominada teoría de reconocimiento de patrones [8].

Talvez la más importante limitación que tienen las técnicas de decodificación del cerebro es, de hecho, la necesidad de implementar tratamientos preliminares para la señal de EEG [56], [64]. Se ha establecido que estos pretratamientos implican una pérdida de información que puede limitar la eficiencia del decodificador.

Con la intensión de limitar los efectos de la perdida de información, se han desarrollado algunos algoritmos que pueden utilizar la señal de EEG sin procesar, no obstante, la velocidad de decodificación de la señal es lenta con incluso bajas tasas de clasificación adecuada.

Las ANNs se han aplicado con éxito en diversas áreas de la disciplina de teoría de computación denomina clasificación de patrones [15], [38], [93], [71]. De hecho, diferentes tipos de NN estáticas se han utilizado para la lograr la correcta decodificación de la señal de EEG [98], [127]. La mayoría estas soluciones se han aplicaciones en diagnóstico de condiciones cerebrales patológicas, tales como; la epilepsia, el autismo, el Alzheimer, la degeneración del cerebro tejido, trastornos del sueño, entre otros. Pero está claro, que este tipo de redes neuronales (NN, por su definición en inglés) desprecia el carácter continuo de la señal electrofisiológica del EEG.

En esta tesis se propone la aplicación de una clase de NN denominadas Redes Neuronales Dinámicas (DNNs, por su definición en inglés), para lograr la decodificación de las señales de

EEG tomando en cuenta la naturaleza continua de estas señales. Las DNNs son conocidas por tener una estructura más compleja para su implementación que las NN estáticas, ademas, son más avanzadas dado que consideran el efecto de la retroalimentación del estado de la red, debido al hecho de que los datos se almacenan y se procesan a través del tiempo. Por otro lado a pesar de que las entradas pueden ser independientes, estás interactuando e influyen entre sí, lo que enriquece la capacidad de clasificación de este tipo de redes.

A lo largo de este trabajo se proponen cuatro topologías de DNNs con sus correspondientes estructuras de ajuste para los pesos. En primer lugar se presenta una red neuronal recurrente (RNN, por su definición en inglés) que analizan una versión discretizada de las señales de EEG con un período de muestreo fijo. Posteriormente, se desarrolla una red neuronal diferencial (DfNN, por su definición en inglés), la cual tiene una estructura continua y cuya ley de ajuste de los pesos obedece una ecuación diferencial ordinaria. Ambos tipos de redes son aplicados para clasificar señales de EEG obtenidas de diferentes bases de datos obtenidas tanto en fuentes publicadas como desarrolladas como parte de este estudio.

En tercer lugar, se desarrolla una red neuronal en cuya estructura se considera el efecto de retardos en el tiempo en la señal de entrada, lo que permite introducir el concepto de ventaneo al algoritmo de clasificación propuesto. Esta red en particular, se evaluó solo con una de las bases de datos mencionada anteriormente. Cabe notar que esta clase de red tiene la ventaja de tomar en cuenta la información previa al momento del análisis de la señal electrofisiológica. Por último se describe una NN con estados complejos que permite el análisis y clasificación de la señal de EEG utilizando la respuesta en frecuencia de esta misma señal.

En el caso particular de la red diferencial, se implementó una versión discretizada en un dispositivo del tipo Field Programmable Gate Arrays (FPGAs), para evaluar su respuesta temporal en un dispositivo dedicado. En este caso se obtuvo una mejora en el tiempo de simulación de 4 min a 23.6s. Posteriormente se implementó una versión de esta misma red en un modelo Very-Large-Scale Integración (VLSI), donde se pudo desarrollar una propuesta de circuito integrado que implementa el funcionamiento de la DfNN. Una última implantación electrónica se desarrolló utilizando un arreglo de circuitos analógicos, para representar la naturaleza continua de la DfNN. Este diseño analógico fue evaluado por simulación y se validó empleando una clase de señales electrofisiológicas similar a la del EEG.

Para todas las NN consideradas en esta tesis se realizó su entrenamiento y validación en software. La eficiencia de clasificación se validó mediante dos métodos: el primero denominado método de generalización – regularización [50] y el segundo denominado un k-fold cross validation. Ambos métodos demuestran una eficiencia de clasificación correcta por encima del 90 %. En el caso de las NN implementadas en hardware el enfoque es diferente dependiendo de la implementación. Las tasas de clasificación correcta de software NN están por encima de otros métodos de clasificación que emplean la misma base de datos y no aplican ningún pre proceso de la señal de EEG. Para el conocimiento de autores que no hay aplicación de la DfNN en hardware haciendo de esto una de las más importantes contribuciones de esta tesis.

Space-time pattern recognition in electrophysiological signals from evoked potentials using dynamic neural networks

by

MSc. Mariel Alfaro Ponce

Abstract

For centuries mankind has sought to understand the brain, technologies such as the electroencephalogram (EEG) are able to record the brain electrical activity [67]. The popularity of EEG in scientific research is due to the large number of studies and implementation of their records , also its portability and the low cost of equipment in comparison than others that allow record brain activity. Derivatives of the EEG technique include evoked potentials (EP), which involves averaging the EEG activity time-locked to the presentation of a stimulus of some sort (visual, somatosensory, or auditory) [49]. It is clear that reading the brain response does not represent a challenge, nonetheless, the main problem is focused in the decodification of the information that is integrated in these signals [85]. Nowadays, there are many decodification algorithms that have been proposed for this task, choices range from the simple population vector algorithm, optimal linear estimator, various versions of Bayesian decoders to different topologies of Artificial Neural Networks (ANNs). Today, most of the EEG automatic interpretation algorithms are framed in the so-called pattern recognition theory [8].

One of the biggest limitation in brain decoding techniques is, in fact, the pretreatments that are applied to EEG signals [56], [64]. It has been established that there is a huge amount of information that has been lost from the pretreatment, and it may affect the decodificator efficiency.

Trying to avoid this information lost, there are some algorithms that work with the raw EEG signal, but their decoding of the signal is slow and with low rates of classification accuracy.

Neural Networks (NNs) have been successfully applied in various areas of the discipline of computational theory known as pattern classification [15], [38], [93], [71]. In fact, different kinds of the well developed Static NN have been used for EEG decoding [98], [127]. Most of these solutions have been applied in the diagnostic of different brain conditions such as; epilepsy, autism, Alzheimer's, degeneration of brain tissue, sleep disorders, among others. But it is clear that this type of NN depreciate the continuous nature of the electrophysiological signal.

This thesis propose the application of Dynamic Neural Networks (DNNs), to achieve the decodification of the EEG signal taking into account the continuous nature of them. DNNs have a more complex structure in their implementation than the static NN, also are more advanced because they consider the effect of retroalimentation on the state of the net, due to the fact that data is stored and processed throughout the time. Although the inputs may be independent, they are actually interacting and influencing each other, enriching the classification capabilities

of this type of NN.

In this work four topologies of DNN were proposed among with their weights adjustment laws. In the first place, a Recurrent Neural Network (RNN) that analyze a discretized version of EEG signals with a fixed sampling period. Next, a Differential Neural Network (DfNN), that have a continuous structure with weights adjustment law set by ordinary differential equations. Both types of NN are implemented for the classification of EEG signals from different databases, obtained from public resources and also from the development of them.

In third place, it was developed a NN in which structure is consider the effect of time-delays in the input signal. As a result the concept of window is introduced to the proposed classification algorithm. It is remarkable that this particular type of DNN have the advantage of taking into account the information previous to the moment of analysis of the electrophysiological signal.

Finally, is described a NN with complex states that allow the analysis and classification of the EEG signal by using the frequency response of the same signal.

In the particular case of the DfNN, it was implemented a discretized version in a Field Programmable Gate Array (FPGA), to evaluate its temporal response in a dedicated device. In this particular case an improvement in the simulation time was achieved from 4min to 23.6s.Then, the same DfNN was implemented in Very-Large-Scale Integration (VLSI), where it was possible to develop a proposal of integrated circuit that implement a full DfNN. A last electronic implementation was developed employing an arrangement of analog circuits, in order to recreate the continuous nature of the DfNN. This analog design was evaluated by simulation and validated by using an electrophysiological signal similar to EEG.

For all the NN considered in this work it was performed the training and validation in software. Their classification accuracy were validated by two techniques the first one is a generalization-regularization method taken from [50] and the other one is a k-fold cross validation method. Achieving a classification accuracy average above 90%. In the case of the NNs implemented in hardware the approach is different according to the implementation. The software NN classification accuracy rates higher from others works that employed the same database and do not apply any preprocess to the EEG signal. To the author knowledge there is no implementation of DfNN in hardware making this one of most important contribution of this thesis.

Contents

1 Preliminaries **20**

1.1 Introduction 20

1.2 Justification 23

1.3 Contribution 25

1.4 Objectives 25

1.4.1 Main 25

1.4.2 Particulars 26

2 Theoretical background **27**

2.1 Other technologies used to acquire brain response data 28

2.2 Electroencephalogram 34

2.2.1 Technological basis of EEG 35

2.2.2 Electrodes 39

2.2.3 Cerebral rhythms observed in the scalp EEG 40

2.2.4 Evoked potentials 41

2.2.5 Event related potentials 43

2.2.6 Analysis and quantification of the EEG 44

2.3 Cerebral plasticity 51

2.4 Pattern recognition in electrophysiological signals 53

2.5 Recurrent and continuous NN in pattern recognition 55

3 General scheme to perform the signal classification **57**

3.0.1 Recurrent neural networks 59

3.0.2 Differential neural networks . . . 62
3.0.3 Time delay neural networks . . . 65
3.0.4 Complex valued neural networks . . . 69
3.1 Neural networks implemented in embedded systems . . . 74
3.2 Off-line adjustment of weights in the classifiers . . . 75
3.2.1 Training . . . 76
3.2.2 On-line training scheme using the continuous version of least mean square method for RNN . . . 76
3.2.3 On-line training scheme using the continuous version of least mean square method for DfNN . . . 77
3.2.4 On-line training scheme using the continuous version of least mean square method for TDNN . . . 78
3.2.5 On-line training scheme using the continuous version of least mean square method for CVNN . . . 80

4 Experimental development **84**
4.1 Database I . . . 86
4.2 Database II . . . 89
4.3 Material and EEG system . . . 92
4.4 Embedded instrumentation of DfNN classifiers . . . 95

5 Results **97**
5.1 Recurrent neural network . . . 97
5.2 Differential neural network . . . 99
5.3 Time delay neural networks . . . 103
5.4 Complex valued differential neural network . . . 104
5.5 Neural networks in real time . . . 107
5.5.1 FPGAs differential neural networks . . . 108
5.5.2 VLSI differential neural networks . . . 110
5.5.3 Analog differential neural networks . . . 111

6 Published results & scholar activities **116**

6.1 Articles published in journal included in the JCR 116

6.2 Articles published as extended manuscript in international conferences 116

6.3 Workshops 117

6.4 Visiting Scholar 117

7 Conclusions & further work **118**

List of Figures

2-1 fMRI cutaway [121]. 29

2-2 Portable TCD system [48]. 30

2-3 Near-Infrared Spectroscopy Device [111]. 31

2-4 PET test; a) Pacient during the test, b) Images obtained from the test (a) [110]. 32

2-5 Basic Diagram of a MEG System [115]. 33

2-6 EEG electrodes 10-20 system: A) Head Sagittal View. B) Head Axial view. Letters refer to brain positions; 0 = occipital, P = parietal, C = central, F = frontal, and T = temporal [112]. 34

2-7 Technical block diagram of an EEG recorder [45]. 37

2-8 Technical block diagram of an EEG recorder [46]. 38

2-9 Some of the most important brain waves as seen during an EEG test [117]. . . . 42

2-10 Visual stimulation produce a response in the cerebral cortex, this response is call VEP [124]. 43

3-1 RNN topology. On the top in black the RNN classifier structure. In purple the W_1 and in blue the W_2 adjustment structures. 60

3-2 Full implementation of the DfNN. On the top the training description and below the classification procedure. 63

3-3 DfNN topology. On the top in black the DfNN classifier structure. In purple the W_1 and in blue the W_2 adjustment structures. 64

3-4 Full implementation of the TDNN. On the top the training description and below the classification procedure. 66

3-5 TDNN topology. On the top in black the TDNN classifier structure. In purple the W_1 and in blue the W_2 adjustment structures. 68
3-6 Function's time domain, shown in red, to the function's frequency domain, shown in blue. Full implementation of the CVNN for the classifier. On the top the training description and below the classification procedure. 70
3-7 CVNN topology. On the top in black the CVNN classifier structure. In purple the W_1 and in blue the W_2 adjustment structures. 72
3-8 Digital and analog NNs implementation. 74

4-1 Experimental development followed, where the different implemented NNs classify the two databases. Generalization (G). Independent Test (IT). 85
4-2 Examples of the signal taken from [92], class 1 signals correspond to extracranial data taken form the neocortical structures. Class 2 signals correspond to extracranial data taken from the hippocampus, class 3 signals correspond to intracranial data taken from neocortical structures, class 4 signals correspond to intracranial data taken from hippocampus and finally class 5 signals correspond to intracranial data taken during seizure episode. 87
4-3 Example of the different trajectories employed for the training of the different NNs topologies, the amplitude of the trajectories could vary according to the network needs. Also the number of trajectories depends in the number of classes that the NN is classifying. 88
4-4 Different patterns (A, B and C) used for the building of the Database II. 89
4-5 Flowchart; Emotiv connection with MATLAB 90
4-6 Volunteer during trial. 91
4-7 Fourteen evoked potentials recorded with the EMOTIV and plotted using MATLAB 92
4-8 Emotiv Research Edition SDK interface window [26]. 93
4-9 Implementation of the DfNN from the PC to the FPGA. Here Matlab and Xilinx are employed to develop the VHDL code that will run in the FPGA. Matlab also help to generate and interface between the FPGA and the PC. 96

5-1 W_1 and W_2 corresponding to the RNN after training with the database 1. In both top images the black line correspond to the weights of the first class, the blue line to the weights of the second class, the red line correspond to the weights of the third class, the magenta line correspond to the weights of the class 4 and finally the green one to the class 5. The images of the bottom are from left to right the trajectories of the RNN Xe and X during the training process for a signal of class 5 and the integral of the LMS error obtained. 98

5-2 W_1 and W_2 corresponding to DfNN after training with the database 1. In both top images the black line correspond to the weights of the first class, the blue line to the weights of the second class, the red line correspond to the weights of the third class, the magenta line correspond to the weights of the class 4 and finally the green one to the class 5. The images of the bottom are from left to right the trajectories of the DfNN Xe and X during the training process for a signal of class 5 and the LMS error obtained. 100

5-3 Integral of the error per class obtained from the DfNN when the input is a signal of class 5, the black line correspond to class 1, the blue line to class2, the red line to class 3, the magenta line to class 4 and finally the green one is class 5. . . 102

5-4 W_1 and W_2 corresponding to DfNN after training with the database 2. In both top images the black line correspond to the weights of the first class, the blue line to the weights of the second class, the red line correspond to the weights of the third class. The images of the bottom are from left to right the trajectories of the DfNN Xe and X during the training process for a signal of class 2 and the LMS error obtained. 103

5-5 Integral of the error per class obtained from the DfNN when the input is a signal of class 2, the black line correspond to class 1, the blue line to class2, the red line to class 3. 104

5-6 W_1 and W_2, W_2 with a delay equal to 1s and W_2 with a delay equal to 2s corresponding to TDNN after training with the database 1. In all the images, the black line correspond to the weights of the first class, the blue line to the weights of the second class, the red line correspond to the weights of the third class, the magenta line correspond to the weights of the class 4 and finally the green one to the class 5. 105

5-7 On the top left the desired trajectory and the TDNN output, on the top right the integral of the LMS error from the TDNN. On the lower part the error of the parallel TDNN when having as an input a class 5 signal. In the lower part the black line belongs to the class1, the blue line the class 2, the red line the class 3, the magenta line to class 4 and finally the green line to the class 5. 106

5-8 W_1 and W_2 corresponding to CVNN after training with the database 2. In both top images the black line correspond to the weights of the first class, the blue line to the weights of the second class, the red line correspond to the weights of the third class. The images of the bottom are from left to right the trajectories of the CVNN Xe and X during the training process the left image is the real part and the right one the imaginary one. 107

5-9 On the top integral of the LMS error from the real and imaginary part of the CVNN. On the lower part the error of the parallel CVNN when having as an input a class 5 signal. In the lower part the black line belongs to the class1, the blue line the class 2, the red line the class 3, the magenta line to class 4 and finally the green line to class 5. 108

5-10 On the top W_1 and W_2 for the hardware implemented DfNN. On the bottom left the desired trajectory on slashed black line and the hardware DfNN output on solid black line. On the bottom right the integral of the LMS error obtained from the left bottom image. 110

5-11 VLSI final design of the DfNN on Cadence. 111

5-12 Activation fucntion of a DfNN builded with Opam's. 112

5-13 W_1 builded with Opam's. 113

5-14 W_2 builded with Opam's. 114

5-15 DfNN builded with Opam's. 114
5-16 On the top left W_1 for the analog NN, on the top right it corresponding W_2, on the bottom left the desired output (slashed line) and the analog NN approximation (solid line). On the bottom left, the LMS error obtained from the signals depicted in the right bottom image. 115

List of Tables

2.1 Comparisson between different classification techniques applied to the EEG Freiburg University database. 55

4.1 Emotiv technical characteristics. 95

5.1 Results from the classification process per class for the RNN. 99
5.2 DfNN results from the 5-fold cross validation process, that determine the classification accuracy. 99
5.3 Results from the classification process per class for the DfNN. 101
5.4 DfNN results from the 5-fold cross validation process, that determine the classification accuracy. 101
5.5 Results from the classification process per class for the DfNN database II. 101
5.6 Results from the classification process per class for the TDNN. 104
5.7 Results from the classification process per class for the CVNN. 109
5.8 DfNN I/O values for software and hardware. 109
5.9 Device utilization. 109

Chapter 1

Preliminaries

1.1 Introduction

Emerging technologies that allow to generation of interfaces between the brain and machines are rising up as a prolific field of research all over the world. The class of devices that allow to decode the electrophysiological brain response have been theorized and developed since the XIX century. Technologies such as the Positron Emission Tomography (PET), functional magnetic resonance imaging or functional MRI (fMRI), Transcraneal Doppler (TCD) and the Electroencephalography (EEG) records have been used as common techniques for measuring the brain activity [67].

Among all of them, the most popular one is the EEG. Indeed, this technique is recognized by physicians and nurseries as a technique that can be used clinically in order to detect aberrant activity in brain response including epilepsy and electrophysiological neuronal disorders. From the point of view of researching field, this technique has been used to show the electrophysiological brain activity in certain psychological states, such as alertness or drowsiness. A really useful alternative technique associated to EEG measurement included evoked potentials (EP), which averages the EEG activity time-locked to the presentation of a stimulus coming up by a predefined sensorial source (visual, somatosensory, or auditory). Event-related potentials (ERPs) refer to averaged EEG responses that are time-locked to more complex algorithms for processing the sensorial stimuli. This technique has been used in cognitive science, cognitive psychology, and psychophysiological research. The popularity of EEG in research is a con-

sequence of its portability and low cost equipment. However, this technique has some relevant limitations. The most important of these inconveniences is its poor spatial resolution specially when working with the scalp EEG [49].

As a result of all recent technological advances, recording the electrophysiological brain response does not represent a challenge. Nonetheless, the main problem when EEG signal is analyzed has focused on decoding the brain response [85]. A lot of automatic algorithms have been proposed to solve the EGG decoding problem.

These algorithms have used the main characteristics of many different methods that cover a wide range of complexities, from the simple population vector algorithm, optimal linear estimator, various versions of Bayesian decoders until different topologies of the so-called artificial neural networks (NN). Nevertheless, until now EEG decoding techniques have attained limited success when the decoded information must be used in specific applications, and no one has achieved total successful results due to diverse factors [78]. First of all, the proposed algorithms have to deal with great amounts of data. Besides, the information that is being processed have lots of noise, along with the fact that these methods need to consider the interaction between neurons yielding to analyze it as an interconnected group or system. Also, the brain responses 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, majority of these techniques applied complicated pretreatments on the signals to obtain better decoding results.

Among others limitations presented by the brain decoding algorithms, pretreatment algorithms have been classified by many authors as the most relevant one [56], [64]. They have established that there is a huge amount of information included in the raw EEG signal that can be lose during the pretreatment algorithm. As a consequence, there exist some algorithms that have tried to work with the raw signal. However, they are not capable of decoding different signals simultaneously. In practice, the standard EEG system is conformed by 21 simultaneously signals. At this moment, most of the EEG automatic interpretation algorithms are framed in the so-called pattern recognition theory [8]. This scientific discipline deals with methods to produce object description and classification but in the case of EEG signals, they demand the application of truly complicated schemes that may provide the classification itself with any

further treatment. Therefore, a new kind of open question appeared here: what is the actual contribution of EEG signal decoding method?

Automatic (machine) recognition, description, classification, and grouping of patterns are important problems in a variety of engineering and scientific disciplines. Even thought there are several electrophysiological signals that have been well described and where finding patterns is considered more and more as an easier task to perform, the case of finding patterns in EEG signals unequivocally is still considered a difficult task due to the complexity of the signal structure in either time or frequency domain. In recent years, NN have been successfully applied in diverse areas of pattern classification [15], [38], [93], [71]. For decoding the EEG signal information, different kinds of the well-developed static NN have been used [98], [127]. Most of them have been applied in finding single specific patterns that may help on the diagnostic of brain conditions such as epilepsy, autism, Alzheimer's, degeneration of brain tissue, sleep disorders, among others. Just some years ago, recurrent algorithms (NN formed with activation function having feedback connections) have started to be used to solve some of the EEG pattern recognition problems. Despite the clear benefits offered by the introduction of NN with feedback, the continuous nature of electrophysiological signal is deprecated.

In order to consider the continuous nature of the EEG signal, Dynamic Neural Networks (DNNs) could be a reliable option. This type of NN is more advanced than static ones, because data are stored and analyzed throughout the time. In DNN, inputs can be independent, but they are interacting with the state of the DNN and influencing each other. Every input is analyzed as function of time as well as the previous value of all DNN inputs; in other words, the network remembers past inputs making the current output an integrated response depending on past inputs and current response of the system. Although all the advantages of the DNN, to the author knowledge there are few applications of this type of NN in pattern recognition theory and none to EEG signal pattern classification.

The lack of applications of the DNNs to pattern recognitions is the result of the weights adjustment algorithm complexity in comparison to the SNNs. There are many software implementations of DNNs, but in order to deal with great mount of data (as in the case of handling EEG information) the computational cost increases dramatically. An affordable way to implement a real time DNN is when its structure can be embedded in hardware devices such as

Field Programmable Gate Arrays (FPGAs). These devices have proven to carry out the adequate resources needed to implement algorithms required to complete the parallel information processing at a very high operation frequency. There are other promising hardware implementations running at transistor level. All these implementations can work even faster than an FPGA and their resources can be seen as unlimited (as it is usually referred, number of transistors can be larger as needed). This may be the key factor to obtain a successfully real time classification and interpretation of EEG signals.

This thesis deals with the continuous nature of the EEG signals by classifying them using a NN that work in a continuous manner rather than the static ones. Two main structures of NNs are proposed: the first one is a DNN that allows to analyze the raw EEG signal in a continuous way and the second one is a Recurrent Neural Network (RNN) that analyze a discretized version of EEG signals with a fixed sampling period. Three topologies of DNN are developed. The first one is a Differential Neural Network (DfNN) that have a single layer and it is applied for classifying two EEG signals obtained from different databases considered in this thesis. The second one implements a Time Delay Neural Network (TDNN) which is tested with the same databases of EEG signals, a TDNN is a kind of DNN that has the advantage of taking in account previous information from the same electrophysiological signal. It is expected that these NN posses a natural advantage reflected in a better pattern classification performance. The third NN is a Complex Valued Neural Network (CVNN) that allows the classification based on the frequency response of the EEG signals. The different NN were applied to test their accuracy, time classification, efficiency, computational cost, and other parameters. Finally, a suitable scheme for performing an implementation scheme of the different DNN from the PC to Very-Large-Scale Integration (VLSI) and FPGA is detailed.

1.2 Justification

The correct decodification and the consequent interpretation of different evoked potentials have many feasible and interesting applications from rehabilitation, systems security, entertainment schemes, among others. These algorithms required to work in the detection of real-time evoked potentials need to make emphasis on their capabilities to classify events in the signals received

from the brain neurons. EEG signals are usually recorded in a complex shape with small amplitudes. This natural and affordable complexity increases as neurons are stimulated. Since the brain is always stimulated by a large variety of factors such as noise, light and other features that are triggered by the different senses, it is difficult to find specific events. The development of algorithms that are capable of classifying the desired events without being disturbed by noises have become a challenge.

Currently, all the aforementioned algorithms require of a training phase in which an individual have to repeat a task always under the same conditions, the people selected for these trials are few, due to the robustness that the algorithms provide to different physiological characteristics from the population under study. Within the developing of signal processing algorithms, it is necessary to take into account the neuronal plasticity factor that is assume to be very particular in each individual tested during the study. So far, several algorithms have shown a degree of precision around 90% (to successfully identify patterns in the EEG signal trace) after long periods of training and most of them have only been developed and tested for one person. The idea of formulating an algorithm that increases the pattern classification performance, with low computational cost and low implementation cost would mean several advantages (it is usual that implementation of these algorithms demand great amount of memory and high processing speeds).

Most of the literature refer that, when working with EEG signals, it is regular to perform a preprocessing technique to improve the signal pattern classification quality. This procedure supplies signals containing lower noise level to the pattern classification method. Even though the preprocessing method makes easier the classification of the signals, most of the information contained in the EEG raw signals got lost. This makes possible to divide this type of algorithms into two phases: a pre-processing section and then a classification procedure. An algorithm that work with the direct raw signals would represent less computational cost. Raw EEG signals contain relevant information that may be lost when it is preprocessed. As a consequence, working with the raw signal may require an algorithm that have the robustness based on the continuous signals gotten from the EEG registers. This thesis presents an alternative to solve this problem using the concept of continuous and recurrent NN.

1.3 Contribution

In this thesis, various DNN and RNN topologies are applied in the pattern classification of EEG raw signals. Extraction and classification of patterns from EEG raw signals is not an easy task and it has been tried by many researching groups that have reported remarkable accuracy in their results (above 90%, when working with clearly different signals). The learning scheme proposed in this work is applied in two stages: the first consisted on the training of the NNs and then, the second tests the efficiency of the proposed algorithms when classifying events coded in continuous raw EEG signals. As a result, three different kinds of continuous NN are tested and their accuracy were compared. This thesis intend to consider the brain responses in a more natural way, it can be stated that the brain response can be represented as a continuous and time-delayed system that is why a DNN and a TDNN are proposed to perform the pattern classification. The third type of NN addressed the problem of identification concerning the frequency response obtained from the EEG information. Then a CVNN was proposed to develop the frequency analysis yielding to a more effective method to solve the classification of patterns in the brain responses. The algorithms presented in this thesis have the capability to classify events in raw EEG signals. Finally, this work presents a suitable way to implement the DNNs in hardware embedded devices by using the technical characteristics offered by either FPGAs, VLSI circuits or operational amplifiers (OPAMs).

1.4 Objectives

1.4.1 Main

Develop computational algorithms to classify patterns in raw EEG signals by both DNNs and RNNs.

1.4.2 Particulars

- Design an algorithm based on RNNs to perform patterns recognition in raw EEG signals.
- Design an algorithm based on DNNs to perform patterns recognition in raw EEG signals based on the application of the so-called Lyapunov stability theory.
- Design an algorithm based on DNNs to perform patterns recognition based on the windowed response of raw EEG signals and time-delay theory.
- Design an algorithm based on DNNs to perform patterns recognition based on the frequency response of raw EEG signals.

Chapter 2

Theoretical background

Making sense of the brain's mind-boggling complexity is not easy. What we do know is that it is the organ that makes us human, giving people the capacity to create art, language, moral judgments and rational thought. It is also responsible for each individual's personality, memories, movements, and how we sense the world. It is one of the body's biggest organs, consisting of around 100 billion nerve cells that not only put together thoughts and highly coordinated physical actions but regulate our unconscious body processes, such as digestion and breathing.

For centuries mankind has sought to understand its operation. By the first century A. D. ancient medical practitioners had provided a general physical description of the brain. Over time, a more detailed description of brain anatomy and physiology was possible. According to the Human Brain Project [17], understanding the human brain is one of the greatest challenges facing 21st century science. If humanity can rise to the challenge, scientists can gain profound insights into what makes us human, develop new treatments for brain disease and build revolutionary new computing technologies.

Bioelectrical modern understanding of brain began with the discovery of the capability of the human brain cortex to be electrically stimulated by G. Fritsch ($1838 - 1927$) and Julius Eduard Hitzig ($1838 - 1907$) in a joint study in 1870. The response of the cortex to electrical stimulation was originally measured with a galvanometer. After several studies, it became obviously that spontaneous electrical phenomena could derive from other kind of non invasive stimulus. Hans Berger is best known as the first person to record the human brain waves in 1924, for which he invented the *EEG*. **Since 1929, he stipulated the possibility of**

read thoughts through the EEG traces [122]**.** Following this idea, in the 70s, researchers developed a primitive control systems based on electrical activity recorded from the brain. Work developed by different research groups, proved that the brain waves could communicate the user intend indeed [125].

2.1 Other technologies used to acquire brain response data

Even though that scalp EEG is the popular choice for developing brain studies, due to its portability, high time resolution and low cost [49]. There are other technologies that allow researching groups to acquire brain information, some of them are mentioned in the following sections:

Functional Magnetic Resonance Imaging. Functional magnetic resonance imaging (Figure $2-1$) is a technique for measuring brain activity. Within the past two decades, fMRI has been developed tremendously, from initial descriptions of changes in blood oxygenation [121]. One of the most popular techniques for fMRI is the one based on the blood oxygenation level dependent (BOLD) contrast, it has become the preferred tool for visualizing neural activity in the human brain. This particular imaging technique is sensitive to changes in intracortical veins, its spatial resolution is determined by the volume of tissue draining to each vein, which is considered to be *the volume of venous unit* [108].

Brain decoding using fMRI includes classification, identification and reconstruction of brain states. It is generally conducted using multi-voxel pattern analysis based on neuroscience evidence that brain functions are mediated by distributed activation patterns. Brain decoding techniques have been successful used in diverse applications such as Brain Computer Interfaces (BCI), patient monitoring, and neurofeedback. However, fMRI is rarely recommended for conventional BCI applications since it is very expensive and has no portability [24].

Transcranial Doppler The use of Doppler ultrasound was first described as early as 1959 for assessing blood velocity in the extracranial vessels [79]. The thickness of the skull bones greatly attenuates the penetration of ultrasonic waves making difficult the noninvasive use of the technique. Ultrasound was therefore limited to surgical procedures, or to be used in children

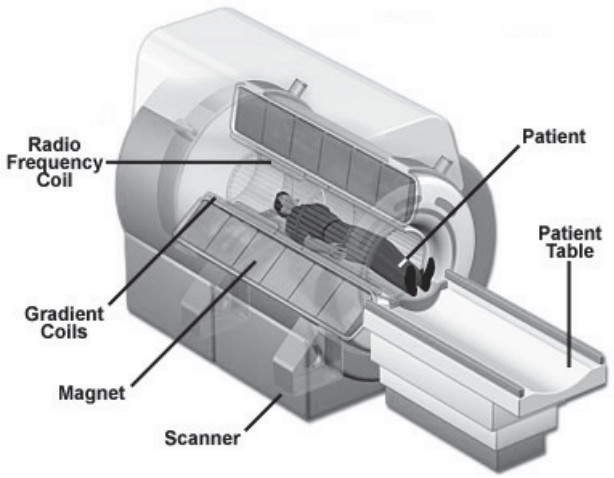

Figure 2-1: fMRI cutaway [121].

with open fontanels. However, [2] demonstrated that the attenuation of sound by bone within the frequency range of 1–2 MHz was far less than conventional frequencies of 3–12 MHz. Indeed, insonation is possible through thinner regions of the skull, termed *acoustic* windows, making it feasible to measure static and dynamic blood velocities within the major cerebral arteries. For the first time, a non-invasive measure of beat-to-beat changes in blood velocity in the vessels of the brain with superior temporal resolution than indicator-dilution techniques was available. However, it is imperative to note that TCD cannot measure CBF *per se*. Rather, TCD measures the velocity of red blood cells within the insonated vessel. Moreover, only the larger basal arteries provide an adequate signal for measurement of cerebral blood velocity with TCD. Because these arteries tend to deliver oxygenated blood to larger regional areas of the brain, TCD gives an index of global, rather than local, stimulus–response relationship. The principles of TCD are the same as extracranial Doppler ultrasound: the Doppler probe emits sound waves that are reflected off moving red blood cells, which are subsequently detected by the transducer. The resultant Doppler shift is proportional to the velocity of the blood [1], [22].

With all the previously mentioned characteristics of Doppler technique, it became factible to think in the development of BCI employing Doppler technologies for the brain scanning. In [83] a investigation of a transcranial Doppler (TCD) sonography was used as the foundation for a new type of non-invasive BCI. TCD can be portable as the one shown in Figure $(2-2)$, light-

Figure 2-2: Portable TCD system [48].

weight, and robust to environmental conditions such as electrical artifacts. It is also relatively inexpensive, particularly in comparison to alternatives such as fMRI. They demonstrated that two types of mental activity can be classified with greater than 80% accuracy on the basis of changes in cerebral blood flow velocity. The accuracy achieved was over 80%, in terms of classification task of word selection and mental rotation. In terms of other studies word selection and mental rotation are well-know and not too difficult to classify.

Near-Infrared Spectroscopy. Near-Infrared Spectroscopy (NIRS). Figure $(2-3)$ show the NIRS device used in this spectroscopic method that employs the near-infrared region of the electromagnetic spectrum (from about 800 nm to 2500 nm). NIRS is a noninvasive optical technique used to measure concentration changes in oxyhemoglobin and deoxyhemoglobin in cerebral vessels based on the characteristic absorption spectra of hemoglobin in the near-infrared range [111], [52].

It has been suggested to be a promising signal acquisition tool for brain research because of its handiness, portability, good spatial resolution, metabolic specificity, and suitability for continuous measurement of brain activity changes with high time resolution [19], [137]. A disadvantage, however, is that NIRS is slow to operate because of the intrinsic latency of the brain hemodynamics [111], in contrast to other techniques. NIRS is based on molecular overtone and combination vibrations. Such transitions are forbidden by the selection rules of quantum mechanics. As a result, the molar absorptivity in the near infrared (IR) region is typically quite

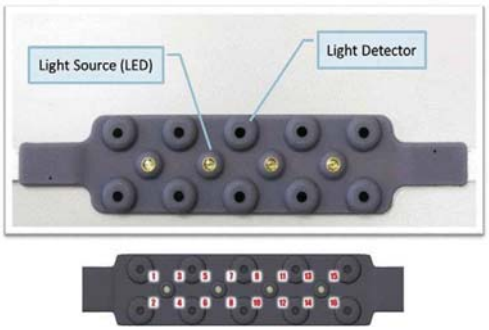

Figure 2-3: Near-Infrared Spectroscopy Device [111].

small. One advantage is that NIRS can typically penetrate much farther into a sample than mid infrared radiation. NIRS is, therefore, not a particularly sensitive technique, but it can be very useful in probing bulk material with little or no sample preparation [84].

Positron Emission Tomography. The potential of positron imaging and the value of eliminating the collimator (is a device that narrows a beam of particles or waves) was recognized by the early developers of nuclear medicine instrumentation, long before the advent of reconstruction algorithms which could allow the generation of transverse sections from data covering a large number of angles. Nuclear medicine images provide a description of metabolic functions in the body, but they contain poor information about patient anatomy and have limited spatial accuracy. Nuclear medicine images tests differ from most other imaging modalities. In that diagnostic tests primarily show the physiological function of the system being investigated as opposed to traditional anatomical imaging such as computer tomography (CT) or magnetic resonance imaging (MRI). Nuclear medicine imaging studies are generally more organ or tissue specific (e.g.: lungs scan, heart scan, bone scan, brain scan, etc. As those shown in Figure $(2-4)$ b) than those in conventional radiology imaging, which focus on a particular section of the body (e.g.: chest X-ray, abdomen / pelvis CT scan, head CT scan, etc.).

Positron Emission Tomography (PET) imaging is not a new concept; it was first proposed in the early 1950s and the first imaging devices were developed by Brownell and by Anger in the 1950s [110]. To generate the images, PET uses an emisor an a detector, this detector records gamma rays. Clinical use of PET is now well-established in clinical oncology and it is therefore

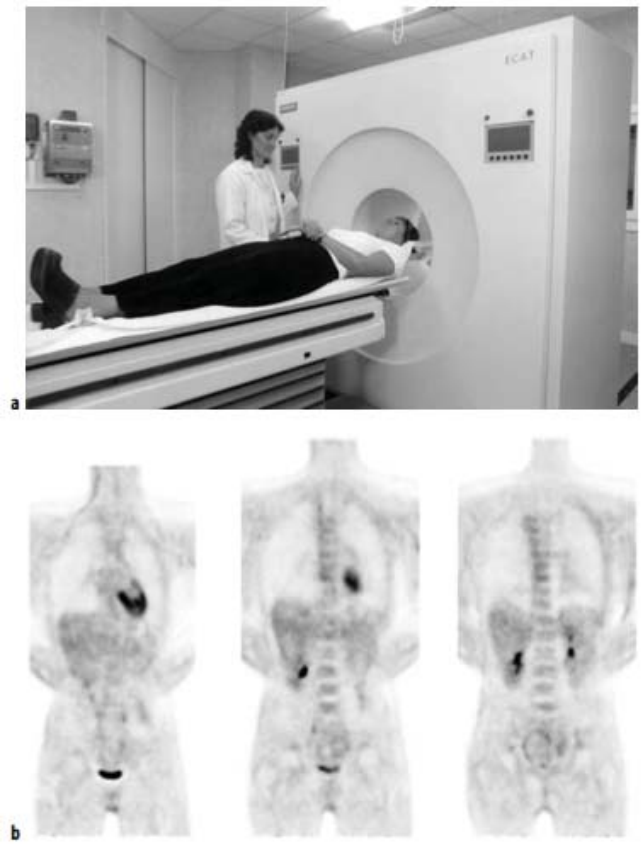

Figure 2-4: PET test; a) Pacient during the test, b) Images obtained from the test (a) [110].

becoming widely available in major hospitals. In addition to its use in research, brain PET also provides diagnostically relevant information mainly in neurodegenerative disorders, focal epilepsy and brain tumours [43].

Magnetoencephalography. Magnetoencephalography MEG technology is relativity young compared with EEG. The responses and phenomena observed in MEG can simultaneously be observed in EEG, so results from years of research in EEG can also be utilized in comparing and interpreting MEG results.

From the functional point of view the main difference between MEG and EEG is: first, MEG if recorded with planar gradiometers shows the response above the source area, while in EEG, the strongest peak is shown further away from the brain source. This is why the results

from MEG are more accurate. Secondly, the MEG is sensitive to tangential currents of large neuronal arrays activated in concert close to surface of the head, while EEG both tangential and radial direction of deep brain areas can be explored [103].

The extracranial magnetic field measured by MEG Figure $(2-5)$ reflects postsynaptic intracellular current flow within the apical dendrites of pyramidal cells oriented parallel to the skull surface. The biomagnetometer is the device used to measure the neuromagnetic signal. It is usually configured as an axial gradiometer, typically consisting of two interconnected induction coils wound in opposition and separated by a few centimeters. The output device for the biomagnetometer is a superconducting quantum interference device (SQUID). The SQUID acts as a high gain, low noise, current to voltage converter that provides the output for the sensor system. The voltage output of the SQUID can be made proportional to the magnetic field sensed [115].

In order to localize the magnetic signals that are changing in space and time, multiple sensors are needed. Sensor coils in a spherical cap are used to surround the head and thereby enable the sampling of large volumes.

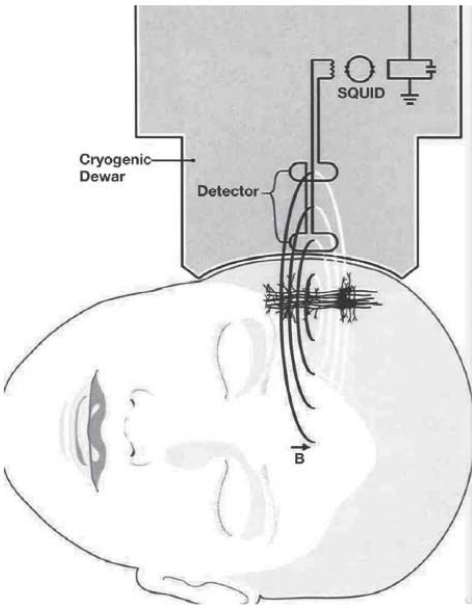

Figure 2-5: Basic Diagram of a MEG System [115].

2.2 Electroencephalogram

The cellular basis of EEG recorded activity has been the topic of intensive studies of extracellular current flow and voltage dependent intrinsic oscillation. The clinical electroencephalographer correlates Central Nervous System (CNS) functions as well as dysfunctions and diseases with certain patterns of the EEG traces on empirical basis [87].

In order to obtain the EEG, the electrodes position is standardized. This standardized arrangement of electrodes over the scalp is known as the International 10/20 system [47] and ensures ample coverage of all parts of the head as shown in Figure $(2-6)$.

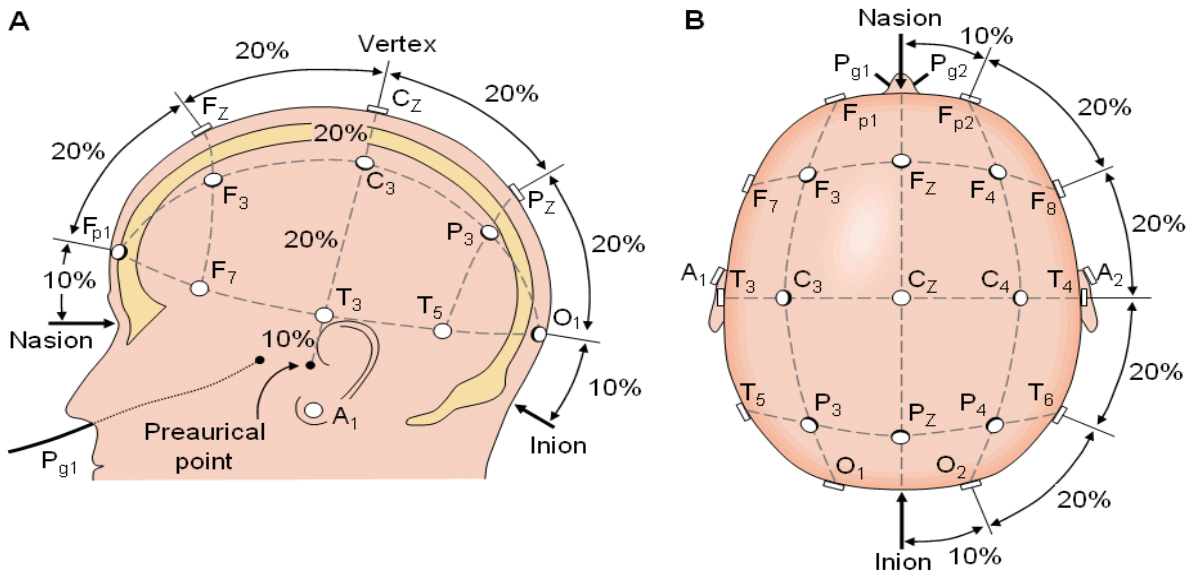

Figure 2-6: EEG electrodes 10-20 system: A) Head Sagittal View. B) Head Axial view. Letters refer to brain positions; 0 = occipital, P = parietal, C = central, F = frontal, and T = temporal [112].

Measuring the EEG requires the setting up of a closed loop passing current from its neuronal source through the various intervening layers of tissue to the electronic amplifier and back again through the head tissue to the neuronal source. Within this circuit, the coupling between skin surface and amplifier input plays a critical role; in contrast to the brain tissue, the skin itself usually exhibits a high resistance and therefore electrodes must be used in combination with a wet electrolyte in order to bridge this junction.

2.2.1 Technological basis of EEG

Technical block diagram of an EEG recorder is shown in Figure $(2-7)$. As one can see in the Figure $(2-7)$ rather than spending separate analog to digital converters (ADC) for each channel, some manufacturers provide just one ADC scanning all channels periodically by means of an analog multiplexer located between the ADC and the low pass filters. During recording, the actual EEG is displayed on the screen. For later evaluation, arbitrary multiple sets of EEG signals can be fetched from the archive and displayed in separate windows. For the adequate design of a EEG recorder device one must take into account:

- Electrodes characteristics
- Analog preamplifier; safety considerations
- Analog to digital conversion velocity and sensibility
- Digital signal processing algorithms
- Signal display on screen and on paper
- Long term data storage and retrieval

EEG Instrumentation Amplifier

There are several layers of tissue over the brain, these tissues make it difficult to read the neural response, so the instrumental amplifier plays an important role in the design of any feasible EEG recorder device. In order to make the EEG signal legible to the human eye, standard low noise operational amplifiers with high input impedance ($> 10\ MW$) are used to amplify the voltage differences between all pairs of electrodes.

The amplifier is split into two modules, with the first stage providing a modest gain of about 10. Before entering the second stage the signals pass a coupling capacitor that removes potential residual high voltage DC potentials that might occur if electrode potentials are not equal over the electrodes involved (which in practice usually cannot be avoided). The overall gain in most EEG systems is on the order of $10,000$ to $20,000$, yielding an EEG amplitude of about 1 volt at the amplifier's output. Due to the DC-blocking capacitor between the two modules, the amplifier has a high pass characteristic with a low frequency cutoff that is defined

by the capacitor. Traditionally, the time constant of this circuit is specified, rather than the lower cut-off frequency. Usual values are $0.03, 0.1, 0.3$ (standard), 1, and 3 s, corresponding to $5, 1.6, 0.5, 0.16$, and 0.05 Hz. Short-time constants facilitate the interpretation of EEG signals when there are large superimposed low frequency components, due either to artifacts or pathological activity. However, if pathological activity at low frequencies needs to be evaluated with high sensitivity, for instance, for brain death diagnosis, larger time constants are required. EEG recorders therefore allow switching between different settings. On the amplifier side, this is accomplished by hardware switches that select between various capacitor values. An alternative (in many available systems, additional) approach is to change the effective time constants by modifying the coefficients of the digital filters during signal post-processing

Besides providing adequate signal transmission, the amplifiers must be designed to match the safety demands specified by the IEC 60601 [18] standard as formulated by the International Electrotechnical Commission, Geneva, Switzerland (1994). The main goal is to rule out the possibility that the current flowing from the amplifier input through the tissue exceeds 100 mV, even in the case of a failure of the electronics. Such protection can be achieved using appropriate resistors between the electrode cable and amplifier's input pins. In addition, modern EEG amplifiers usually provide full electrical insulation (using optical transmitters) of the front-end amplifier from subsequent electronics in order to prevent high voltages entering into the front end. Such decoupling is known as floating input, because there is no stable relation between the absolute signal amplitude within the preamplifier and the ground potential of the subsequent stages.

For special applications DC recording units are available that transmit the input difference signals without any frequency limitations, that is, without any DC-suppressing coupling capacitor. In order to avoid excessively large amplitudes that would exceed the amplifier's dynamic range, it is necessary to provide an individual DC voltage for each channel that is subtracted from the difference signal, thus compensating for the residual DC component resulting from the fluctuating electrode potentials. From time to time, these recorders need to be reseted interactively in order to adapt the voltage of this DC compensation signal. Alternatively, a slow voltage follower may track the fluctuating DC, adapting the compensation signal automatically. However, in a strict sense, this is no longer a pure DC recorder.

Modern EEG systems are designed as referential recorders, meaning that all electrodes are measured with respect to one common reference electrode placed somewhere on the head. Accordingly, this electrode is internally connected to the inverting input pins of all the different amplifier channels. One important characteristic of these difference amplifiers is their common mode rejection (CMR) characteristics. A large CMR is a prerequisite for efficient suppression of noise components present at both input pins. Modern EEG amplifiers achieve a CMR of 80 dB or even better, thereby reducing common noise components by at least $1 : 10,000$.

The intrinsic noise level of modern EEG amplifiers is about 0.5 $mVeff$ at a bandwidth of 100 Hz before amplification Adding the noise originated at the electrodes, a total noise floor up to 0.7 $mVeff$, corresponding to an approximate peak-to-peak level of 2 to 3 mV (Gaussian amplitude distribution), is realistic.

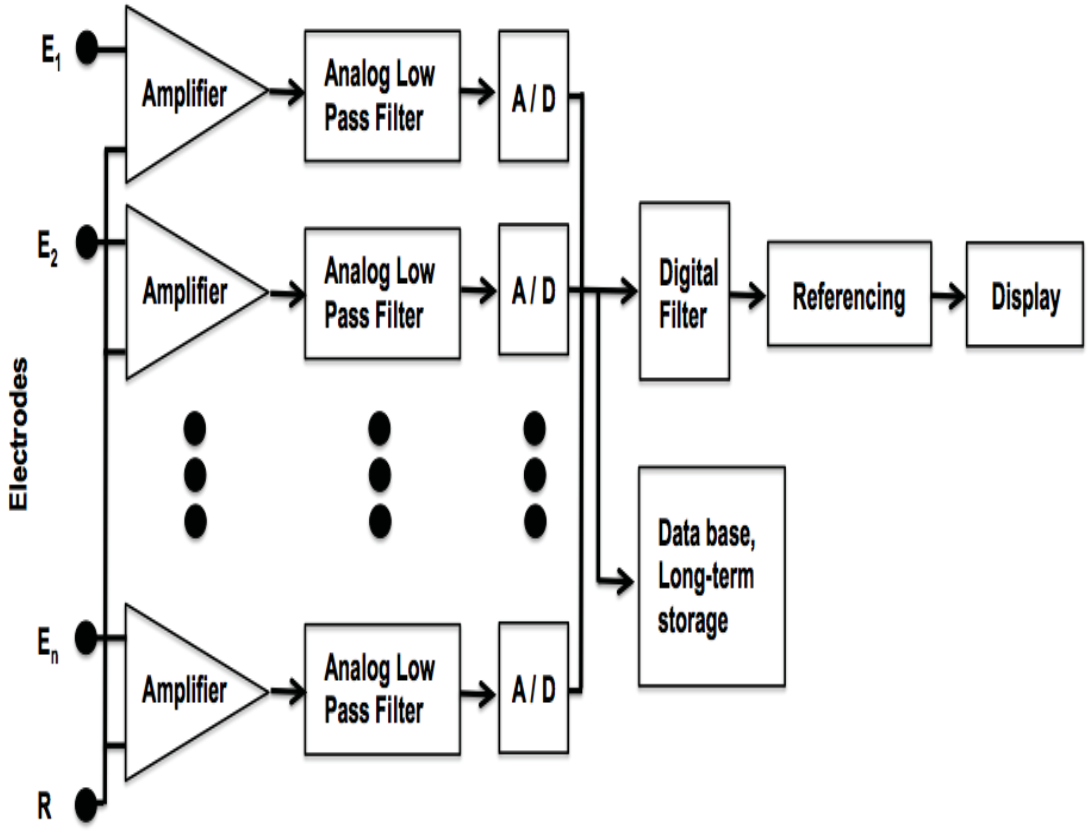

Figure 2-7: Technical block diagram of an EEG recorder [45].

As mentioned before the EP conform a derivative of EEG technique, an EP can be obtained from the EEG record. In Figure $(2-8)$ an EP recorder system is technically described. It is important to remark its similarity to a conventional EEG amplifier. As the one seen in the EEG recorder device, here also, standard low noise operational amplifiers with high input impedance (in most cases$>10\ MW$) are used to amplify the voltage differences between pairs of electrodes (i.e., differential amplifiers are used). The amplifier is split into two modules, with the first stage providing a modest gain of about 10. Before entering the second stage, the signals pass a coupling capacitor that removes potential residual high voltage DC potentials that might occur if electrode potentials are not perfectly equal over the electrodes involved (e.g., due to unclean surfaces). The overall gain in most EP systems is at least $10,000$ yielding a signal amplitude of a few volts at the amplifier's output. Due to the DC blocking capacitor between its two modules, the amplifier has a high pass characteristic with a low frequency cut-off that is defined by the capacitor. If several different capacitors are provided, a switch allows several cut-off frequencies to be selected.

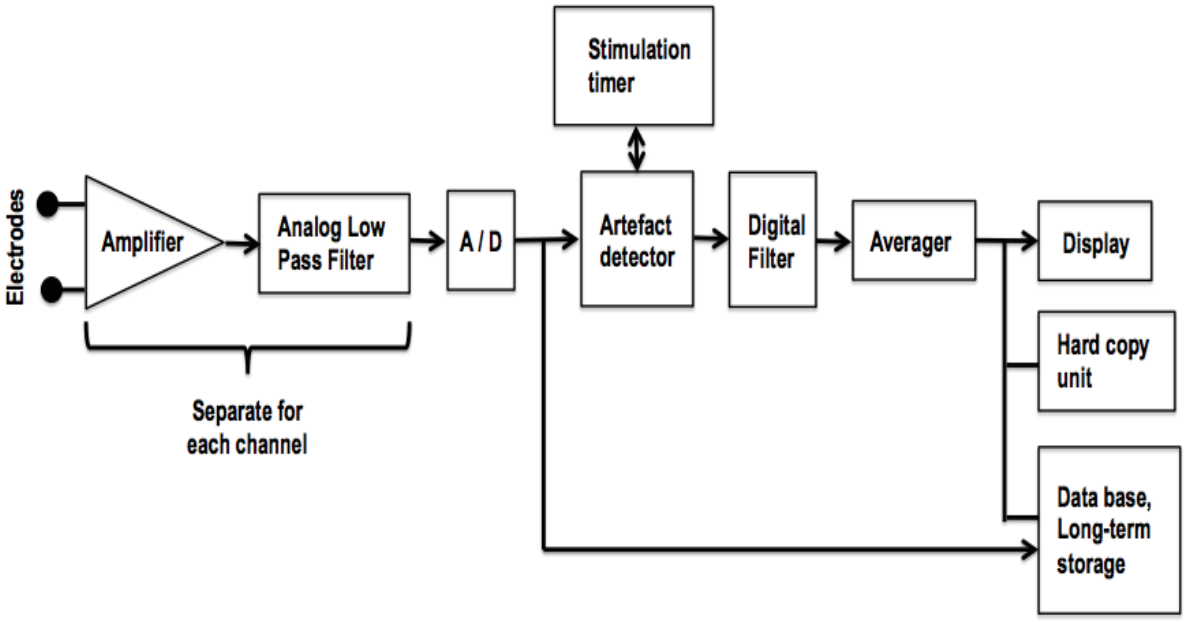

Figure 2-8: Technical block diagram of an EEG recorder [46].

Traditionally as happened with EEG recorders, the time constant of this circuit is specified in addition to the lower cut-off frequency. Typical values are 1, 0.3, 0.1, and 0.03 s, corresponding to 0.16, 0.5, 1.6, and 5 Hz, respectively. Modern digitized EP recorders provide a better flexibility than older ones with respect to these settings. This flexibility is accomplished by providing just one high capacity coupling capacitor in the amplifier and defining the final cut-off frequency using a digital filter (see below). The same approach is used to filter out higher frequencies, according to the demands of the various EP modalities: The amplifier transmits the signal with an upper cut-off frequency meeting the demands of all EP modalities. Then the signal is converted from the analog to the digital domain, with the final bandwidth being defined afterwards using digital filters. Because of the large amplitude range covered by the various EP modalities, as well as the superimposed EEG, it is desirable to have different amplifier gain factors (also known as sensitivity, specified in terms of microvolts per unit).

2.2.2 Electrodes

For recording of brain electrical activities, various types of electrodes are used [107]. The electrode plays an important role in the acquisition of the brain information, The are many types of electrodes, that vary in metal composition, shape and location. Some of them are mentioned next:

- **Reusable disks**. These electrodes can be placed close to the scalp, even in a region with hair because they are small. The electrodes are held by a washable elastic head band. Disks are made of tin, silver, and gold are available.

- **Adhesive Gel Electrodes**. These are the same disposable silver/silver chloride electrodes used to record ECGs and EMGs, and they can be used with the same snap leads used for recording those signals.

- **Subdermal Needles**. These are sterilized, single-use needles that are placed under the skin. Needles are available with permanently attached wire leads, where the whole assembly is discarded, or sockets that are attached to lead wires with matching plugs.

- **EEG Caps with disks**. Different styles of caps are available with different numbers and

types of electrodes. Some caps are available for use with replaceable disks and leads. Gel is injected under each disk through a hole in the back of the disk.

It is important to remark that choosing of electrode kind depends on the location of the electrodes and the behavioral situation in which the recording takes place. A basic property of any type of electrode is that there exists a metal/liquid junction in the electrical connection between tissue and the EEG recording apparatus.

2.2.3 Cerebral rhythms observed in the scalp EEG

The EEG is typically described in terms of rhythmic activity and transients. The rhythms of the EEG are defined as regular recurring waveforms of similar shape and duration as seen in Figure $(2-9)$. Most of the cerebral signal observed in the scalp EEG falls in the range of $1-20Hz$ (activity below or above this range is likely to be artifactual, under standard clinical recording techniques).

There are a plethora of signals, also referred to as components. These signals fall into two major classes: spikes and field potentials [131]. Spikes reflect the action potentials of individual neurons and thus acquired primarily through microelectrodes implanted by invasive techniques. And field potentials, that are measures of combined synaptic, neuronal, and axonal activity of groups of neurons and can be measured by EEG or implanted electrodes.

Delta. Oscillations within the frequency range of 1 to 4 Hz. There are two types of delta activity. One is generated in the cortex, as it survives after thalamectomy (it have not yet been investigated). The other signal is originated in the thalamus, even after decortication, and its cellular mechanisms are quite well understood.

Tetha. This rhythm is usually considered within the frequency range of 4 to 7 Hz. The cellular bases of these waves have been intensively investigated in rodents, but is less evident in other mammals. The normal theta activity,is generally considered as poor or absent in primates.

Alpha. The alpha rhythm occurs in the frequency range of 8 to 13 Hz. Although this is probably one of the most important graphoelements and its description dates back to Berger

(1929) , there is no solid knowledge on its cellular mechanism. Alpha waves are usually described as occurring during awakeness [86]. These signals can be detected in all age groups but are most common in adults.

Beta. Beta rhythms frequency is normally greater than $13Hz$. These have been shown to be localized to the occipital areas and are engaged in the visual discrimination task [117]. They are observed in all age groups.

Gamma. These signals are usually having components within $20-50\ Hz$. Oscillations under $30\ Hz$ are related to the visual system. Also these oscillations display during cross-modal sensory processing (perception that combines two different senses, such as sound and sight).

Mu. Shows rest state of motor neurons. Oscillations components are within $8-13\ Hz$.

In the cortex fast rhythms (Beta and Gamma), both in the EEG and at the level of single neuronal activity have been identified. These are usually called beta of gamma rhythms, depending on the frequency rage. It has been found experimentally that fast frequency rhythms (33-45 Hz) occur in the frontaparietal areas of the brain. A large number of studies have demonstrated the existence of high-frequency rhythms activities in the human EEG under different circumstances, particularly in relation with motor cognitive functions [28].

2.2.4 Evoked potentials

EP are the electrical signals generated by the nervous system in response to sensory stimuli. EP can be extracted from the background EEG if subjects are exposed to repeated brief sensory stimuli. For standard neurological applications, auditory, somatosensory, and visual stimuli are the most important. The corresponding specific EP are called AEP, SEP, and VEP, respectively. In contrast to the more or less random background EEG, EP are to a large extent reproducible when peripheral stimuli are repeatedly presented. This is the reason why they can be measured, despite the fact that their amplitude is much lower than the background EEG [46].

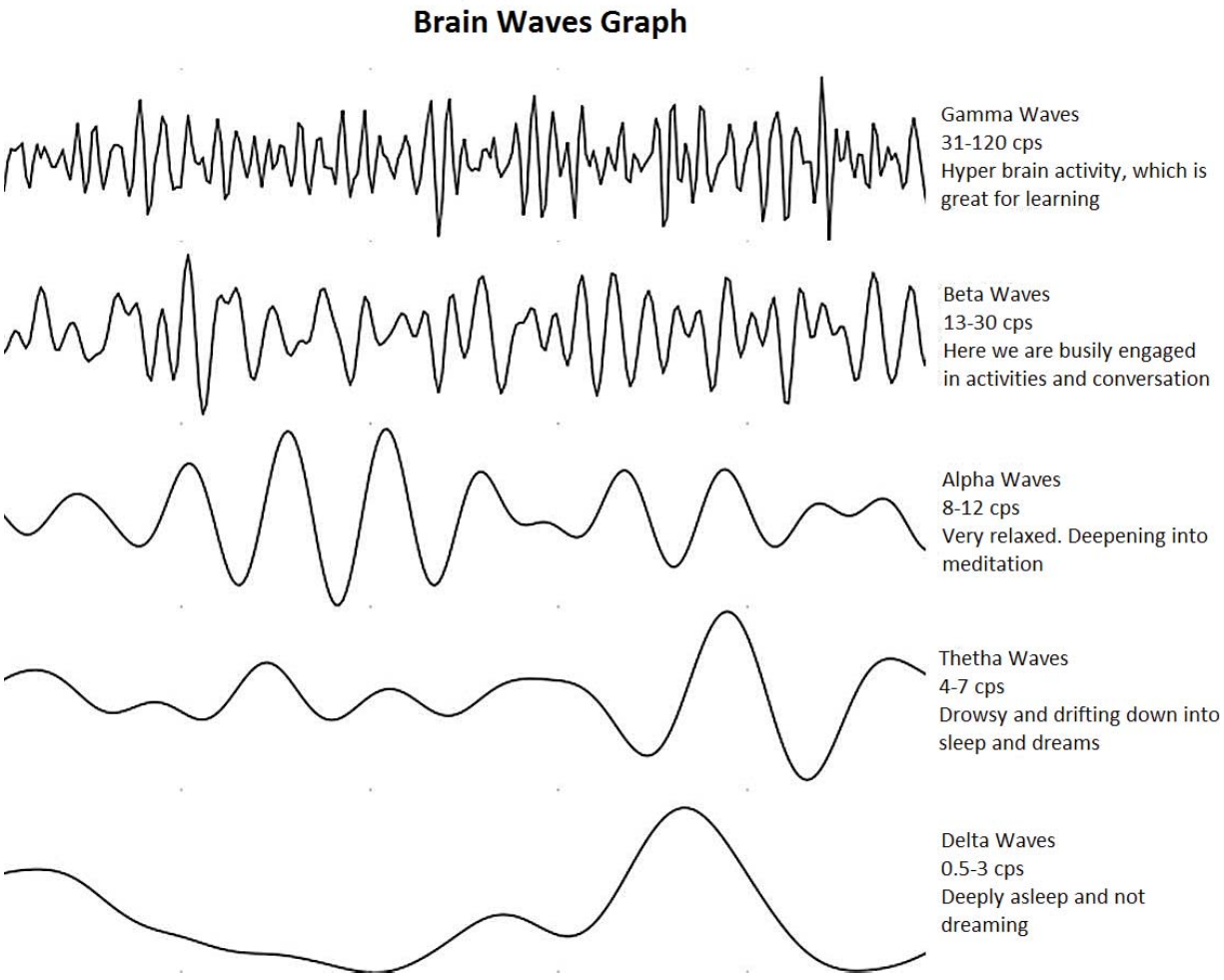

Figure 2-9: Some of the most important brain waves as seen during an EEG test [117].

Visual Evoked Potentials (VEP).

The VEP is an evoked potential produced by sensory stimulation within the visual field [129] as shown in Figure $(2-10)$. In 1934, [3] noticed potential changes of the occipital EEG were observed under stimulation of light. The VEP can be extracted, using signal averaging, from the electroencephalographic activity recorded at the scalp [124]. In order to record VEP, there are two major methods of stimulation.

- Luminance; uniform flash light.

- Pattern; black and white checker board composed by the same number of white and black square checks.

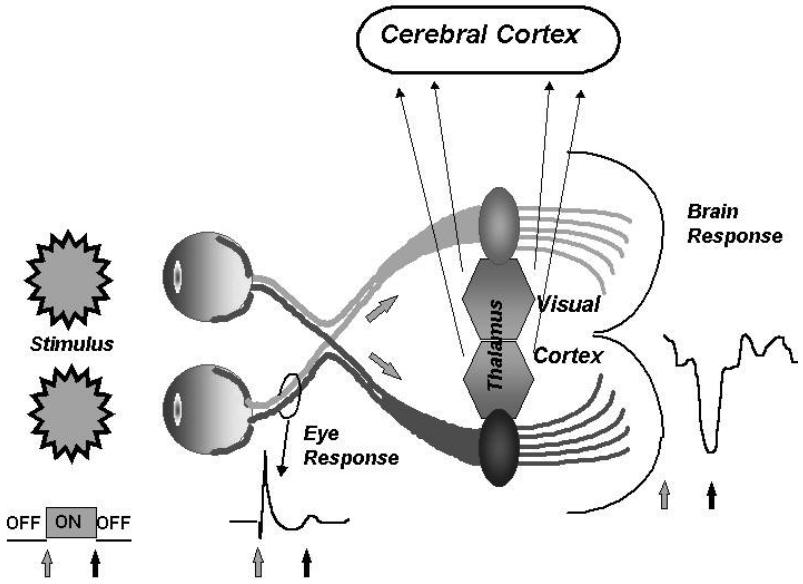

Figure 2-10: Visual stimulation produce a response in the cerebral cortex, this response is call VEP [124].

2.2.5 Event related potentials

Event-related potentials (ERPs) are time-locked responses by the brain that occur at a fixed time after a particular external or internal event. ERPs are measured from EEG signals. From studies, there are two types of ERPs.

- Exogenous ERP components are obligatory responses as consequences of physical stimuli and occur due to processing of the external event but independent of the role of the stimuli in the information processing.

- Endogenous ERP components occur when an internal event is processed. It is dependent on the role of the stimulus in the task and the relationship between the stimulus and the context in which it occurred.

P300. The P300 (P3) wave is an event related potential (ERP) elicited by infrequent, task-relevant stimuli [23]. It is considered to be an endogenous potential as its occurrence links not to the physical attributes of a stimulus but to a personal reaction to the stimulus [113]. This wave was firstly reported in the 60's. The P300 component is measured by assessing its amplitude and latency.

Amplitude (μV) is defined as the difference between the mean pre-stimulus baseline voltage and the largest positive going peak of the ERP waveform within a time window.

Latency (ms) is defined as the time from stimulus onset to the point of maximum positive amplitude within a time window. The P300 scalp distribution is defined as the amplitude change over the midline electrodes (Fz, Cz, Pz), which typically increases in magnitude from the frontal to parietal electrode sites. The origin of this wave have not yet been well defined, it has been suggested that it is originated from task conditions involving working memory, and that conscious awareness may be related to stimulus sequence effects. In addition, P300 amplitude is sensitive to the amount of attentional resources engaged during dual-task performance. As primary task difficulty is increased, P300 amplitude decreases regardless of the modality or the motor requirements of the primary task that must be solved. Passive stimulus processing generally produces smaller P300 amplitudes than active tasks, because stimulus and non-task events engaged attentional resources to reduce amplitude [96].

2.2.6 Analysis and quantification of the EEG

Since the EEG invention, mathematical analysis techniques have played an important part on its interpretation for clinical diagnosis of neural diseases, measuring neural function and for monitoring neurological injury. Nowadays, EEG signals acquisition systems have improved and so the quality of the registered information, also EEG action fields have diversified and most of them require the management of great amounts of data. In order to do a quantitative analysis, it is first necessary to convert the continuous analog EEG signal into a digital form, which is accomplished by an analog to digital converter. Once the signal is represented as individual numbers in time series, it can be manipulated mathematically. The EEG signal can be described as a complicated example of electrophysiological response. This signal originates in the intricate neural system. Traditionally, the spontaneous EEG is characterized as a linear

stochastic process with great similarities to noise [118],[119].

- EEG Preprocessing

The main problem when dealing with the EEG is the signal size, the record of the brain activity signals are small between 10 to 100μV. This small signal could easily get contaminated with various sources of noise and artifacts. Artifacts are undesired electrical potentials which come from sources other than the brain. Some of the artifacts that could affect this signal are; patient movement (such as blinking), muscle artifacts and the heart beat [54]. Noises could come from different sources such as the power line (50/60Hz), mobile phones, TV and radio stations, cardiac pacemakers and medical equipment among others.

Most of the noises and artifacts can be easily removed by filtering, but others are difficult to reject, in order to get rid of artifacts there are many things that can be done like improve the electrodes and the placement of them. A more sophisticated way to remove noise is to apply signal processing methods, such as Independent Component Analysis (ICA) [53].

Linear Methods

- Time Domain Methods

These methods, model time series of EEG signals. In EEG analysis two methods have been classically employed.

Parametric modeling methods presumes that the EEG signals are created with equations, with unknown coefficients to be approximated. It can be divided into Autoregressive model and Sinusoidal model. The first one fits EEG with mathematical model. The second one uses sinusoidal basis functions to represent the signal, the task of such modelling is to find the optimal coefficient of each sinusoidal function. A classical sinusoidal model is the Fourier transform (FT)

$$\begin{cases} x(n) = \dfrac{1}{N}\displaystyle\sum_{k=1}^{N-1}\left(X_r(k)\sin(n\Omega_0 k) + jX_i(k)\cos(n\Omega_0 k)\right) \\ \\ \Omega_0 = \dfrac{2\pi}{N} \end{cases} \tag{2.1}$$

where $X(k)$ is the Fourier coefficient that indicate the strength of the signal frequency [5].

Nonparametric methods study the waveform directly. Factors such amplitude and energy change are important for the analysis of the EEG signal. The amplitude of the signal can vary between subjects, also the position of the electrodes can affect this measure. It is common in this method to apply short time windows to the signal, this window facilitate the energy measurement. The Teager Energy Operator (TEO) is a popular way to measure the energy changes. TEO is defined as

$$\psi(n) = x^2(n) - x(n-1)\,x(n+1) \tag{2.2}$$

TEO depends on the frequency and high frequencies are emphasized [55].

- Frequency Analysis

The basic idea of frequency analysis is to study EEG in several classic non-overlapping frequency bands. The clinical technician interprets the EEG by the features or magnitudes of waves in each frequency band [20]. Spectral analysis has been used for decades as the most important diagnostic tool. Even though the physicians do not calculate the spectrum, they usually focus on some frequency components.

In this case, the most popular method is the Discrete Fourier Transform (DFT) which is defined as [25]

$$x(n) = \frac{1}{N}\sum_{n=1}^{N} X(n)\, e^{\frac{-j2\pi nk}{N}} \tag{2.3}$$

The power spectrum is obtained with

$$\|X(k)\|^2 = P(k) \tag{2.4}$$

DFT assumes that the signal is stationary and slowly varying.

Parametric model-based estimation [82] represent the EEG series with an autoregressive model defined as

$$w(n) = x(n) - a_1 x(n-1) - a_2 x(n-2) - ... - a_p x(n-p) \tag{2.5}$$

taking the Z-transform yields

$$W(Z) = A(Z)\,X(Z) \tag{2.6}$$

where $A(Z) = 1 - \sum a_i z^{-i}$,then

$$X(Z) = A^{-1}(Z)\,W(Z) \tag{2.7}$$

In the practice to approximate estimation of the spectrum $X(\omega)$ of $x(n)$ can be obtained by setting $z = e^{j\omega}$

$$\bar{X}(\omega) = \frac{W(\omega)\,T}{\left|1 - \sum a_i e^{-j\omega T}\right|^2} \tag{2.8}$$

where T is the sampling frequency.

Spectral distance (SD) and cepstral distance (CD) have been deeply studied to detect quantitative changes in EEG signals. A good example SD is the autoregressive based spectral distance (ARSD) measure [63], ARSD is defined with the help of the spectrum obtained through Equation (2.8) by measuring the difference between two autoregressive spectras, $X_t(\omega)$ and $X_r(\omega)$

$$ARSD(X_t, X_r) = \left\{ \frac{1}{L} \sum_{l=0}^{L-1} \left| X_t(j\Omega_l) - X_r(j\Omega_l) \right|^p \right\}^{\frac{1}{p}} \tag{2.9}$$

where $\Omega_L = \pi l/L$ for $l = 0, 1, ...L-1$.

CD measures the differences between two EEG signals, where one is a baseline. It is commonly used to determine a brain injury level.

Domain frequencies, the different peaks that compose a EEG signal are use for its analysis. In order to accomplish this analysis, the equation (2.8) can be written as

$$\bar{X}(\omega) = \frac{W(\omega)\,T}{\left|\prod_{k=1}^{p} (e^{-j\omega T} - P_k)\right|^2} \tag{2.10}$$

where $\{P_k\}$ are the complex poles of $\bar{X}(\omega)$. Therefore the frequencies satisfying $e^{-j\omega_k T} = P_k$, as a result the dominant frequencies can be obtain by

$$F_{do\min ant} = \frac{F_{sampling}}{2\pi} \cdot \omega_k \tag{2.11}$$

CD method has proven to be useful to determine hypoxic-ischemic brain injury [34].

- Time Frequency Analysis

Time frequency methods have been successfully used to analyze the source of the epileptic episodes and electrocorticograms [130]. The most common method uses a short time Fourier transform (STFT) to increase time resolution

$$STFT\left(\omega,t\right) = \int_{-\infty}^{\infty} x\left(\tau\right) g\left(\tau - t\right) e^{-j\omega\tau} d\tau \tag{2.12}$$

where $g(t)$ is the window function. STFT cannot achieve high resolution. A high time and frequency resolution can be obtained through Wigner-Ville distribution (WVD)

$$Wx\left(\omega,t\right) = \int x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j\omega\tau} d\tau \tag{2.13}$$

$Wx\left(\omega,t\right)$ is the FT of the autocorrelation function of signal $x(t)$ with respect to the delay variable [139]. Both methods STFT and WVD are useful when working with high frequency, but when a slow waveform needs to be analyze, there are other techniques that improve results. Such an analysis needs an adaptive time-frequency analysis method like wavelet transform (WT), the WT is defined as

$$Wx\left(a,b\right) = \left|a\right|^{-\frac{1}{2}} \int_{-\infty}^{\infty} x\left(t\right) \varphi\left(\frac{t-b}{a}\right) dt \tag{2.14}$$

where a and b are the scaling and transiting parameters, respectively, and φ is the mother-wavelet function.

Nonlinear Methods

- Information Theory-Based Theory

A series of statistical measures has been developed to evaluate the EEG signals in different domains, including time, frequency, or time-frequency. One measure calculates the information (entropy) of EEG signals in these three domains. Entropy is a method to quantify the order/disorder of a time series. It is calculated from the distribution $\{\rho_i\}$ of one of the signal parameters, such as amplitude, power, or time-frequency representation. By studying the mutual information between different regions on the cortex, we can understand the interdependence of different regions of the brain. Recently, various entropy measures, such as time-dependent entropy [109], wavelet entropy [106], time-frequency complexity [120], and mutual information, have been applied to EEG analysis.

Time-dependent entropy measure, calculate the entropy of the entropy of the EEG time series. The amplitudes of the EEG segment are partitioned into M microstates; the raw sampled signal is denoted as $\{x(k) : k = 1, ..., N\}$. The amplitude range W is therefore divided into M disjointed intervals $\{I_i : i = 1, ..., M\}$ such that [120]

$$w = \bigcup_{i=1}^{M} I_i \tag{2.15}$$

The probability distribution can be obtained by the ratio of the frequency of the samples N_i falling into each bin (I_i) and the total sample number N:

$$p_i = \frac{N_i}{N} \tag{2.16}$$

Then, the entropy can be defined with the amplitude distribution across the M bins:

$$SE = -\sum_{i=1}^{M} p_i \ln(p_i) \tag{2.17}$$

Wavelet entropy and time-frequency complexity. The wavelet entropy evaluates the complexity of the energy distribution in a different frequency band (subbands) [138]. Wavelet entropy

is defined using the Shannon measure of entropy of the energy distribution:

$$\begin{cases} SWE = -\sum_i p_i \log\left(p_i\right) \\ \\ p_j = \dfrac{\sum \left\|c_{j,k}\right\|^2}{\sum_j \sum_k \left\|c_{j,k}\right\|^2} \end{cases} \tag{2.18}$$

where p_i actually is the ratio of the energy in j^{th} scale and the total energy. [66] applied wavelet entropy to the specific frequency bands (Delta, Theta, Alpha, and Beta) of the EEG following hypoxic-ischemic injury, which is also called subband wavelet entropy (SWE).

Mutual information (MI) is a useful measure for studying the dependence or relation between different regions of the brain. Mathematically, the MI between two cortical activity variables X and Y is defined with their joint probability density function, $p(x, y)$, and marginal probability density functions, $p(x)$ and $p(y)$. The $MI(X;Y)$ is the relative entropy between $p(x, y)$ and the product distribution $p(x)p(y)$ [21].

$$MI(X;Y) = \sum_{x \in X} \sum_{y \in Y} p\left(x, y\right) \log \frac{p\left(x, y\right)}{p\left(x\right) p\left(y\right)} \tag{2.19}$$

- High-Order Statistics

The high order statistics (HOS) analysis is a nonlinear method for describing the phase coupling [81]. HOS is suitable for multivariable analysis of measuring the extent of statistical dependence in the time series, mathematically, $B\left(\omega_1, \omega_2\right)$ of a time series is defined as

$$B\left(\omega_1, \omega_2\right) = E\left\{X\left(\omega_1\right) X\left(\omega_2\right) X^*\left(\omega_1 + \omega_2\right)\right\}, \tag{2.20}$$

where $X\left(\omega_i\right)$ is the complex Fourier coefficient spectrum of the EEG and X^* is the complex conjugate. Two frequencies, ω_1 and ω_2, are said to be the phase-coupled when a third component exist at a frequency of $\omega_1 + \omega_2$. To make the differences of EEG comparable, the normalized bispectrum is extracted as bicoherence $bic\left(\omega_1, \omega_2\right)$

$$bic\left(\omega_1, \omega_2\right) = \frac{\left|B\left(\omega_1, \omega_2\right)\right|}{\sqrt{P\left(\omega_1\right) P\left(\omega_2\right) P\left(\omega_1 + \omega_2\right)}} \ ,$$

where $P(\omega_i) = \|X(\omega_i)\|^2$ is the power spectrum at frequency ω_i and $bic(\omega_1, \omega_2)$ varies between 0 to 1 [88]. When $P(\omega_1 + \omega_2)$ is not zero $bic(\omega_1, \omega_2)$ shows the degree of coupling between the frequencies ω_1 and ω_2.

- Chaotic Measures

The most commonly used descriptions are based on chaotic measures, such as; correlation dimension, information dimension, capacity dimension, and multifractal spectrum. The motivation for nonlinear dynamics analysis of the EEG is the high complexity and limited predictability of the neurological signals, which may make them essentially stochastic [68]. Hence, various quantitative measures that help describe nonlinear and chaotic dynamics may be useful in characterizing EEG after trauma or neurological disorders.

Approximate entropy estimation (ApEn) its being successfully used to calculate complexity an irregularities from EEG short data sets, which is useful for real clinical and experimental studies. It is defined with the correlation integer at each point in the embedded space $C_i^m(r)$ [94]. The average logarithm of the correlation integer is obtained by

$$\Phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \log C_i^m(r) \ . \tag{2.21}$$

Then, the ApEn is

$$ApEn(m, r, N)(u) = \Phi^m(r) - \Phi^{m+1}(r), m \succeq 1. \tag{2.22}$$

2.3 Cerebral plasticity

Modern neuroanatomical, electrophysiological and imaging techniques allow scientists to reveal the insights of brain function. As a consequence many algorithms capable of decipher the brain behavior have been proposed. In the development of these algorithms brain plasticity represent an important factor, as a parameter that drastically changes brain response and learning capability according to the age, sex, neurological injury or disease among others [61], [89].

Cortical plasticity is the capability of the cerebral cortex to alter its functional organization as a result of experience. One must recall that plasticity refers to the phenomenon of change, not to the specific underlying mechanisms. Significant progress have been made in understanding what factors drive cortical plasticity in normal and injured brains.

Plasticity is invoked for encoding information during perceptual learning, by internally representing the regularities of the environment [32].

Physiological and anatomical changes are driven by natural sensory stimulation, skill acquisition, peripheral injury, central injury, exogenous growth promoting agents, exogenous neuromodulating drugs and exogenous electrical / magnetic stimulation. These factors that seem to drive cortical plasticity may be especially significant with regard to understanding ways to promote recovery of normal brain response.

Neurorehabilitation face several variables that can contribute to the capacity for functional improvement when treating neurological injuries or diseases [60]. *While developing pattern recognition algorithms for EEG signals, the capacity for functional improvement represents the demanded training time for any algorithm trying to have a real impact on medical issues.* These include patient health status, age, lifestyle, and time after injury in addition to the nature and locus/extent of the brain injury. All of these factors together create a brain that is very different from the *normal* one even within the same injury domain [62]. This leads to further heterogeneity in the way that the residual brain areas adapt to the injury and potentially respond to therapy through neural recovery and compensation.

There are different neural strategies to recover brain normal response that can be identified. These strategies involve neural recovery and/or compensation and take advantage of the inherent functional redundancy within the brain. The strategies are mentioned next:

Restoration Residual brain areas undergo profound neurobiological changes following brain injury or disease resulting of dysfunction within structurally intact brain areas both proximal and distal to the damaged area [90].

Recruitment Recruitment refers to enlisting motor areas that have the capacity to contribute to the lost motor function but may not normally have been making significant contributions to that behavior prior to the injury. These areas are asked to play a larger role in the perform-

ance of the impaired motor behavior compromised because of stroke but are not necessarily acquiring new function (retraining). Within the motor cortex, it can be demonstrated through the expansion of movement representations within areas outside of the original motor map [31].

Retraining In some cases,several areas of motor cortex may be asked to adapt existing function or take on additional functions to support functional improvement. Although this strategy is integrally related to restoration and recruitment in that neural circuits do not simply use their existing functions to contribute to behavior but begin to perform novel or additional functions [7].

There are several researching groups working in improving pattern classifications methods for acquired EEG signals, but the complexity of the brain function, interaction and the fact that it can drastically change its response from one individual to other under similar conditions demands of a robust algorithm and the implementation of it in embedded devices capable of manage big sets of data in real time [122].

2.4 Pattern recognition in electrophysiological signals

Pattern Recognition (PR) plays an important role in medicine, in both treatment and diagnosis of different illnesses. But, before explaining its application over classification of electrophysiological signals, there are some important concepts that need to be explained.

PR is the scientific discipline dealing with methods for object description and classification [75]. A fundamental notion in pattern recognition is the similarity. Similarity, consist in comparing objects that share common valued attributes with a target. The PR task of assigning an object to a class is said to be a classification task.

Electrophysiology is the study of the electrical properties in biological cells and tissues. It involves measurements of voltage change or electric current on a wide variety of scales from single ion channel proteins to whole organs like the heart. In neuroscience, it includes measurements of the electrical activity of neurons, and particularly action potential activity [126].

There is a large number of PR implementations in electrophysiology including cardiac peacemakers, life support devices, health monitoring on mobile devices, among others. Despite of the great advances, most of these systems work with well known electrophysiological signals such

as the heart beat. Electrophysiological signals such as the EEG, are still hard to process and most of the existing processing methods do not achieve high accuracy[116].

Automatic (machine) recognition, description, classification and grouping of patterns are important problems in a variety of engineering and scientific disciplines. In EEG signals, it is difficult to find a pattern due to the complexity of the signals. Mostly of the pattern classification systems use the support of medical specialists which classify a certain number of signals for training the PR systems. Once the system has a well-defined pattern, its recognition/classification may consist of one of the following tasks: 1) supervised classification in which the input pattern is identified as a member of a predefined class, 2) unsupervised classification in which the pattern is assigned to a hitherto unknown class [9].

The recognition problem here is being posed as a classification or categorization task, where the classes are either defined by the system designer (in supervised classification) or learned based on the similarity of patterns (in unsupervised classification). The design of a pattern recognition system essentially involves the following three aspects:

1. Data acquisition and preprocessing.

2. Data representation.

3. Decision making.

The problem domain dictates the choice of sensor, preprocessing technique, representation scheme, and the decision making model. It is generally accepted that a well-defined and sufficiently constrained recognition problem (small intraclass variations and large interclass variations) yields to a compact pattern representation and a simple decision making strategy [70]. Learning from a set of examples (training set) is an important and desired attribute of most pattern recognition systems.

Some examples of different classification techniques applied to EEG signals are shown in the Table(2.1).

Table 2.1: Comparisson between different classification techniques applied to the EEG Freiburg University database.

Researches	Method	Dataset	CfA	Validation method	PA
[114]	Time & frequency domain features-recurrent neural network	Z,S	99.6%	NC	Yes
[58]	Entropy measures-adaptive neuro-fuzzy inference system	Z,S	92.22%	NC	Yes
[57]	Chaotic measures-surrogate data analysis	Z,S	∼90%	NC	Yes
[95]	Fast Fourier transform-decision tree	Z,S	98.72%	k-fold cross (sic)	Yes
[116]	Discrete wavelet transform-mixture of expert model	Z,S	95%	Cross validation (sic)	Yes
[91]	Discrete wavelet transform-approximate entropy (ApEn)	Z,S	96.65%	NC	Yes

*Classification accuracy (CfA)
*Not Considered (NC)
*Preprocessing Apply (PA)
The dataset for all theses papers where taken from [92].

2.5 Recurrent and continuous NN in pattern recognition

The main characteristics of NNs are their ability to learn complex nonlinear input-output relationships, application of sequential training procedures, and self-adaptation to the input data. The increasing popularity of NN models in solving pattern recognition problems has been primarily due to NN seemingly low dependence on domain-specific knowledge and due to the availability of efficient learning algorithms that can be used by practitioners [9]. NNs provide a new set of nonlinear algorithms for feature extraction and classification. In addition, there are several feature extraction and classification algorithms which can also be mapped on

NN architectures for efficient (hardware) implementation.

When analyzing a signal by NNs, the problem can be tackled in two different ways: the first based on the application of the classical static NN (SNN) that do not take into account the continuous nature of the signal or second, based on the application of DNN where the data are stored and elaborated on time. Taking into account the time-varying characteristics of signal, it is expected to obtain more information that allow a simpler classification algorithm with even better results.

Automatic detection and classification of EEG recordings have become an important thrust area of research for the development of BCI, security, games and medical diagnosis systems. Although several automatic diagnostic schemes have been proposed over the years, SNN [128] based pattern recognition classifiers have gained significant prominence with respect to some others alternatives [35], [102], [73]. SNNs have been successfully employed to determine complex, non- linear, multidimensional mathematical relationships between noisy uncertain sets of data with very dissimilar natures.

Along the last couple of decades NNs based solutions have been successfully employed, specially in the domain of function approximation, pattern recognition, automated medical diagnostic systems, decision support systems, time series prediction, signal processing, image processing, etc. [42], [13], [40]. Approximation problems can be solved by employing either supervised learning, where the weights and biases of the SNN are learned in presence of training data set or by a SNN that can be employed with unsupervised learning, where inputs are classified into different clusters in a multidimensional space, in absence of any training data. In the case of EEG classification, several types of SNNs have been proposed. Most of these SNNs have been usually employed in supervisory mode and require pretreatment of the EEG signal.

Chapter 3

General scheme to perform the signal classification

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 u to the NN, then the output namely x corresponds to the specific class where the signal belongs among the L available classes doe the whole set of N EEG signals. This x corresponds with the concept of target. For this thesis, this target was represented as a sigmoid function described by:

$$x^l(v) = \frac{a^l}{1 + e^{-cv}} \tag{3.1}$$

where the variable x^l represents target that belongs to class l $(l = 1, ..., L)$. The positive constant a^l was modified accordingly 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 elected in order to regulate the slope of changing in the sigmoid function. 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 $r - th$ $(r \in [1, N_l]\,,\ \sum_{l=1}^{L} N_l = N)$ 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 $W_{1,r}^{*,l}$ and $W_{2,r}^{*,l}$ is used a part of NN in this training stage.

Therefore, when the whole set of N signals selected to perform the training process has been tested, L different sets of weights $W_{1,N_l}^{*,l}$ and $W_{2,N_l}^{*,l}$ 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 to the one used during the training. This part of the process is named the validation stage. Based on the well-known generalization-regularization method [50], 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 EEG signal, that is

$$J^{T,l} = \begin{cases} \displaystyle\sum_{k=1}^{T} \left(a^l - x_r^l(k)\right)^2 & if \quad \text{NN is a RNN} \\ \\ \displaystyle\int_{t=0}^{T} \left(a^l - x_r^l(t)\right)^2 dt & if \quad \text{NN is a DNN} \end{cases}$$

One must notice that the length of all the testing signals was kept constant. The minimum value of this set of mean square errors 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.

Finally, the testing stage was executed in the same way that the validation stage did. However, the set of EEG signals used in this part of the analysis never was used in either the training or the validation stages. This methodology is one of the major contributions of this thesis. The parallel scheme of NN proposed in the validation stage appears as a contribution to the state of art regarding the signal classification by NN. Figure 3-2 describes in graphical form the signal classification procedure proposed in this thesis.

In the next subsections, an introduction to the different types of NN that have been proposed to deal with the EEG signals classification is given. First a RNN is presented to analyze the

EEG signal with fixed number of discrete measures as inputs, then three types of CNN are propose: DfNN, TDNN and CVNN. Each subsection explain the advantages of using this type of topology in comparison to another possibilities. The corresponding training algorithm for each type of DfNN is presented in a separated section to highlight the differences between all the classifiers considered in this thesis.

3.0.1 Recurrent neural networks

RNN as the one shown in Figure (3-1) is a class of artificial NN where connections between units form a direct discrete time delay structure. This case creates an internal state of the network which allows it to exhibit dynamic temporal behavior. Unlike feedforward NN, RNNs can use their internal memory to process arbitrary sequences of inputs. The basic architecture of this type of NN was developed in the 1980s.

The concept of signal classifier as a recurrent algorithm

The EEG signals can be seen as an input signal described by the sequence $u_r(k)$. In the case of RNN, the target is represented by x_r Therefore, it is possible to represent the relationship between the EEG signal and the class. Therefore, the following RNN structure represents the actual form of the signal classifier

$$x_r\left(k+1\right) = f\left(x_r\left(k\right), k\right) + g\left(x_r\left(k\right)\right) u_r\left(k\right) + \xi\left(x_r\left(k\right), k\right) \tag{3.2}$$

Here the function $f \in \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ (we considered that the class can be characterized even by a vector of n components) represents the brain response obtained after the internal processing enforced by the input signal $u_r \in \mathbb{R}^m$ that represents EEG signal that is filtered by an internal structure in the network represented by $g \in \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$. The term $\xi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ symbolizes the noises and the uncertainties produced by the EEG signals on the classifier structure. Both functions f and g are considered smooth enough. This fact is paramount to propose the approximation of the recurrent structure presented in (3.2) as a RNN.

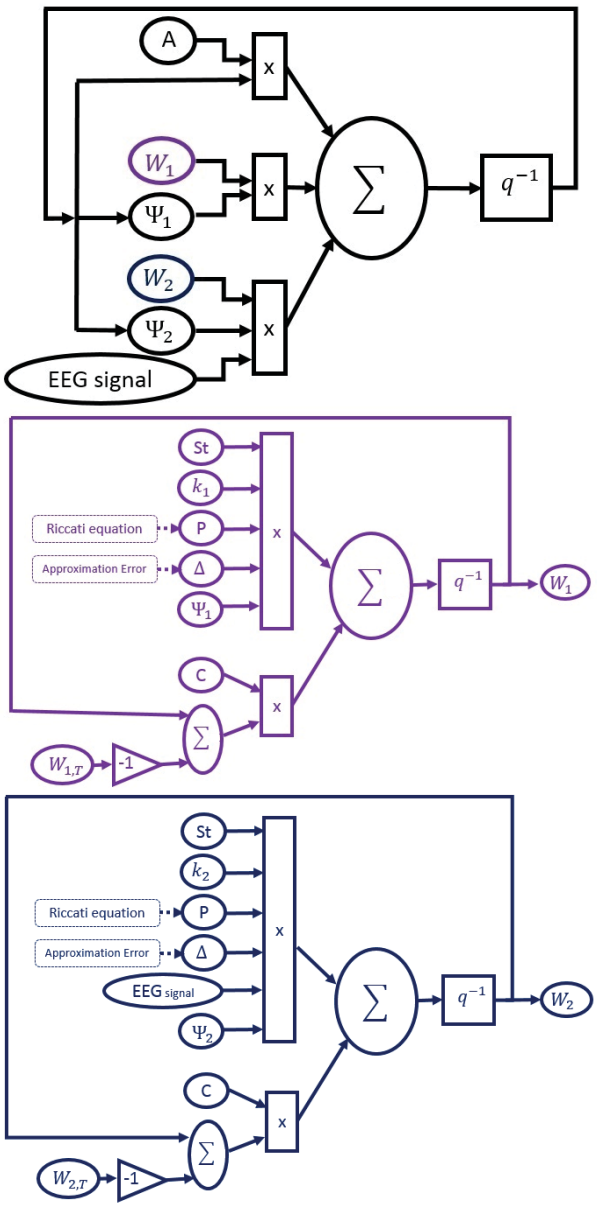

Figure 3-1: RNN topology. On the top in black the RNN classifier structure. In purple the W_1 and in blue the W_2 adjustment structures.

Neural network approximation of the recurrent signal classifier

Based on the approximation theory supported on the Stone-Weiwrstrass Lemma, the system described in (3.2) can be represented as a RNN described by

$$x_r(k+1) = A_r x_r(k) + W_{r,1}\sigma_r(x_r(k)) + W_{r,2}\varphi_r(x_r(k))\,u_r(k) + \tilde{f}(x_r(k), u_r(k), k) + \xi(x_r(k), k) \tag{3.3}$$

In the RNN structure, the matrix $A_r \in \mathbb{R}^{n\times n}$ is a Hurwitz matrix in the discrete sense (all its eigenvalues are in side the unitary circle). The set of matrices $W_{r,1} \in \mathbb{R}^{n\times p_1}$ and $W_{r,2} \in \mathbb{R}^{n\times l_2}$ are the weights that best fit the uncertain recurrent target signal to the EEG signal considered as the input. These weights are considered unknown but bounded. The vectorial functions $\sigma_r : \mathbb{R}^n \to \mathbb{R}^{p_1}$ and $\varphi_r : \mathbb{R}^n \to \mathbb{R}^{p_1\times m}$ represent the set of activation sigmoid functions used to construct the RNN structure. Each element of these vector functions satisfy a similar structure to the one presented in (3.1). The unknown function $\tilde{f} : \mathbb{R}^{n+m+1} \to \mathbb{R}^n$ represents the modelling error produced by the finite number of activation functions considered in the NN structure.

One may notice that null knowledge on the actual values of both $W_{r,1}$, $W_{r,2}$ motivates the design of non-parametric identifier that can reproduce the relationship between x_r and u_r

Identifier structure to approximate the recurrent signal classifier

The identifier based on RNN is proposed accordingly the classical strategy used within the adaptive parameter identification framework [99].

$$\hat{x}_r(k+1) = A_r\hat{x}_r(k) + \bar{W}_{r,1}(k)\,\sigma_r(\hat{x}_r(k)) + \bar{W}_{r,2}(k)\varphi_r(\hat{x}_r(k))\,u_r(k) \tag{3.4}$$

The identifier state is defined as $\hat{x}_r \in \mathbb{R}^n$ that must reproduce the target represented by x_r. The matrix $A_r \in \mathbb{R}^{n\times n}$ is used to enforce a degree of stability on the network evolution and the vector functions σ_r, φ_r have components named the activation functions that have been proposed in the approximation given by the RNN. The time varying weights $\bar{W}_{r,1} \in \mathbb{R}^{n\times p_1}$ and $\bar{W}_{r,2} \in \mathbb{R}^{n\times l_2}$ are used to adjust the adaptive identifier in order to reproduce the mentioned relationship between $\hat{x}_r$ and u_r. This is the main tool to construct the general form for the

recurrent numerical approximation. The specific method proposed in this thesis to adjust the weights $\bar{W}_{r,1}$ and $\bar{W}_{r,2}$ are described in the section denoted training of this manuscript.

3.0.2 Differential neural networks

DfNN have emerged as a powerful tool to extent the classification capacities of NN as a class of adaptive systems. The DfNN is a new type of NN described by a set of ordinary differential equations (ODEs). These ODEs may be used to represent the set of signals monitored by the EEG amplifier. To the author's knowledge, DfNN (Figure 3-2) have never been used as pattern classifiers. Generally, the NN require a training procedure, which plays an important part in its successful application as classifier. In instance, in conventional gradient-descendent-type weight adaptation, sensitivity of unknown system are required during the so-called training process [49]. By virtue of its parallel distribution, a NN is generally robust, tolerant under the presence of faults and external noise, able to generalize signal examples and they are capable of solving nonlinear approximation problems [11].

The concept of signal classifier as a differential algorithm

The signal classifier information obtained from the EEG signals collected during trails can be interpreted as an absolute continuous signal namely $x(t)$. Therefore, it is possible to represent the brain classification response as a solution of an ordinary differential equation as follows

$$\frac{d}{dt}x\left(t\right) = f\left(x\left(t\right), u\left(t\right)\right) + \xi\left(x\left(t\right), t\right) \tag{3.5}$$

Here $f\left(\cdot,\cdot\right)$ represents the brain's response obtained after the internal processing enforced by an input $u\left(t\right) \in \mathbb{R}^m$ that may represent the EEG signal. The term $\xi\left(x\left(t\right), t\right)$ symbolizes the noises and the uncertainties produced by the EEG signals measured at the same time when the input is presented as stimulus to the subject.

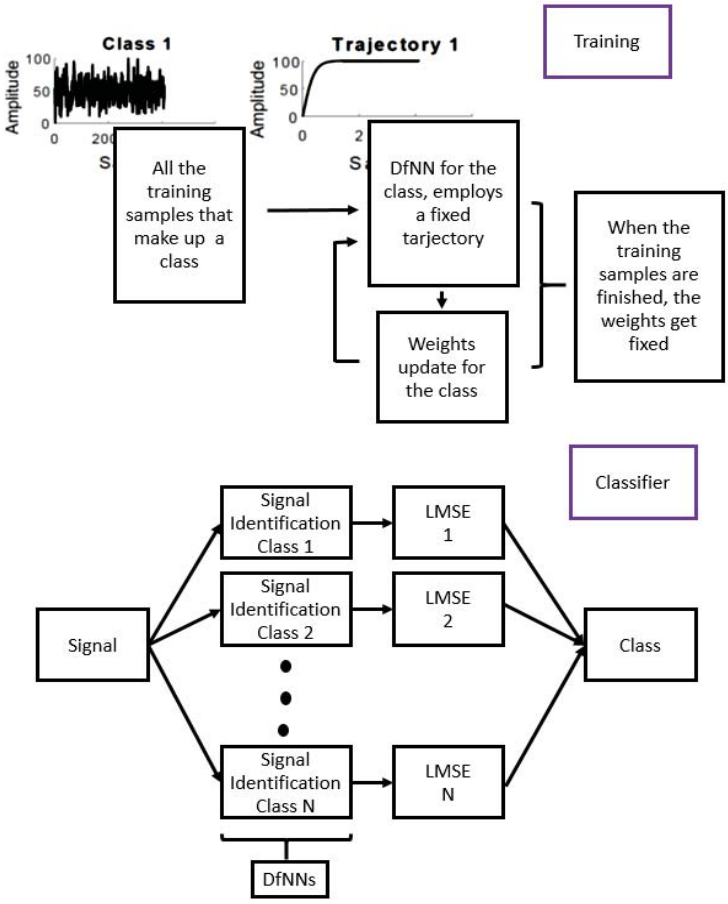

Figure 3-2: Full implementation of the DfNN. On the top the training description and below the classification procedure.

Neural network approximation of the continuous signal classifier

The system can be represented as follows

$$\frac{d}{dt}x\left(t\right) = Ax\left(t\right) + W_1^*\Psi_1\left(x(t)\right) + W_2^*\Psi_2\left(x(t)\right) + \tilde{f}\left(x\left(t\right), u(t), t\right) + \xi\left(x\left(t\right)\right) \tag{3.6}$$

Here the matrix $A \in \mathbb{R}^{n \times n}$ is used to provide a degree of linear approximation for the class of DfNN proposed in this section. The set of matrices $W_1^* \in \mathbb{R}^{n \times p_1}$ and $W_2^* \in \mathbb{R}^{n \times p_2}$ represent the unknown set of weights used to provide the approximation capacity of this kind of CNN. The vector functions $\Psi_1 : \mathbb{R}^n \to \mathbb{R}^{p_1}$, $\Psi_2 : \mathbb{R}^n \to \mathbb{R}^{p_1 \times m}$ have components named the activation

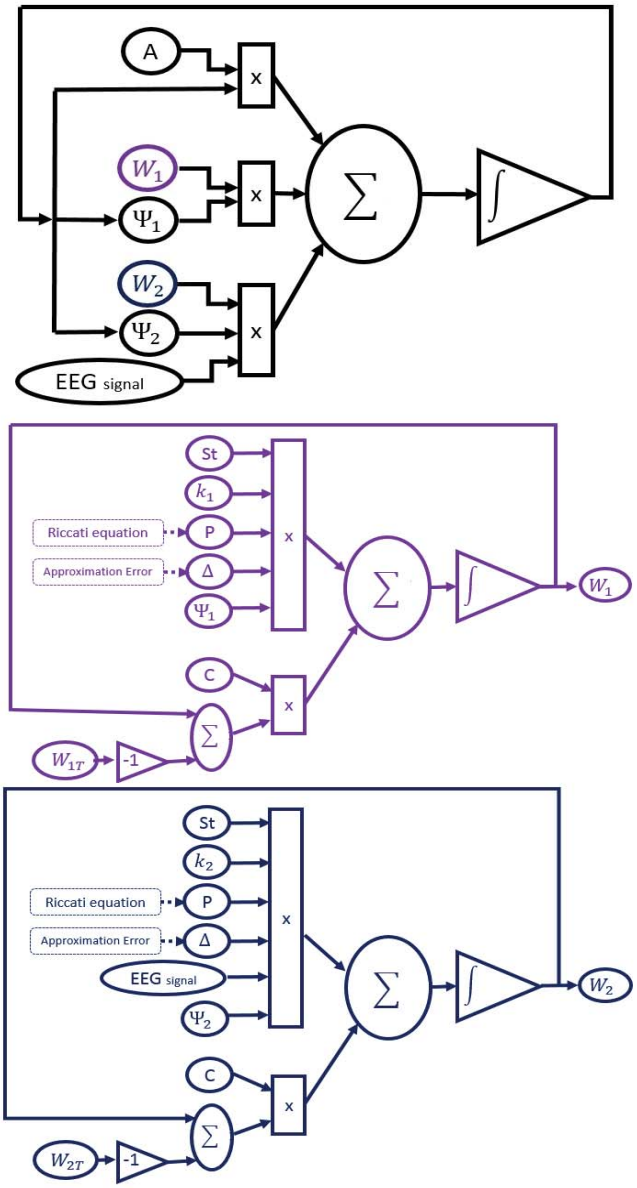

Figure 3-3: DfNN topology. On the top in black the DfNN classifier structure. In purple the W_1 and in blue the W_2 adjustment structures.

functions that have been proposed in the approximation given by the RNN. Despite they are considered unknown, there is an important assumption on they boundeness property. The unknown function $\tilde{f} : \mathbb{R}^{n+m+1} \rightarrow \mathbb{R}^n$ represents the modelling error produced by the finite number of activation functions considered in the NN structure.

Identifier structure to approximate the signal classifier

The classifier based on NN Figure (3-3) is proposed following the classical strategy used in the adaptive parameter identification algorithms [100]. Then

$$\frac{d}{dt}\hat{x}\left(t\right) = A\hat{x}\left(t\right) + W_1\left(t\right)\Psi_1\left(\hat{x}\left(t\right)\right) + W_2(t)\Psi_2\left(\hat{x}(t)\right)u(t) \tag{3.7}$$

Where $W_1 \in \mathbb{R}^{n\times p_1}$ and $W_2 \in \mathbb{R}^{n\times p_2}$ are the weights that best fit the uncertain differential form of the target signal corresponding to the EEG signal considered as the input. The vector labeled $x \in \mathbb{R}^n$ represents the identifier state. The set of weights W_1 and W_2 define the set of adaptive parameters that should be adjusted the eventual classification of the EEG signals in different classes.

3.0.3 Time delay neural networks

Time-delays are common in biological and chemical systems. Some processes in biological and chemical systems [134], [140], such as the gene expression [16], course of an infection [36], biosignal response to an stimulus and feedback control in signal transduction networks involve time-delays [51].

A time delay in input signal appears in models of real systems due to different reasons. Usually its presence is forced by the physical nature of the system [39]. Transport delays (like in chemical or pneumatic systems) or computational delay (e.g. in digital controllers or communication networks [65]) are regular sources of delayed input signal. Input delay can also be introduced *artificially* to include the sampling effect in mathematical models (see, for example, [29], [30] and [97]). Ignoring the time-delays in biological and chemical models yield to conclusions that do not contain the complete information of the system [76].

To obtain a better approach for the class of biological and chemical systems that contain

delays, the topology of the NN should also considered delays. The existence of time-delays on the NN may cause oscillations and instability [136]. By this reason, the stability of TDNN has long been investigated [141], [27].

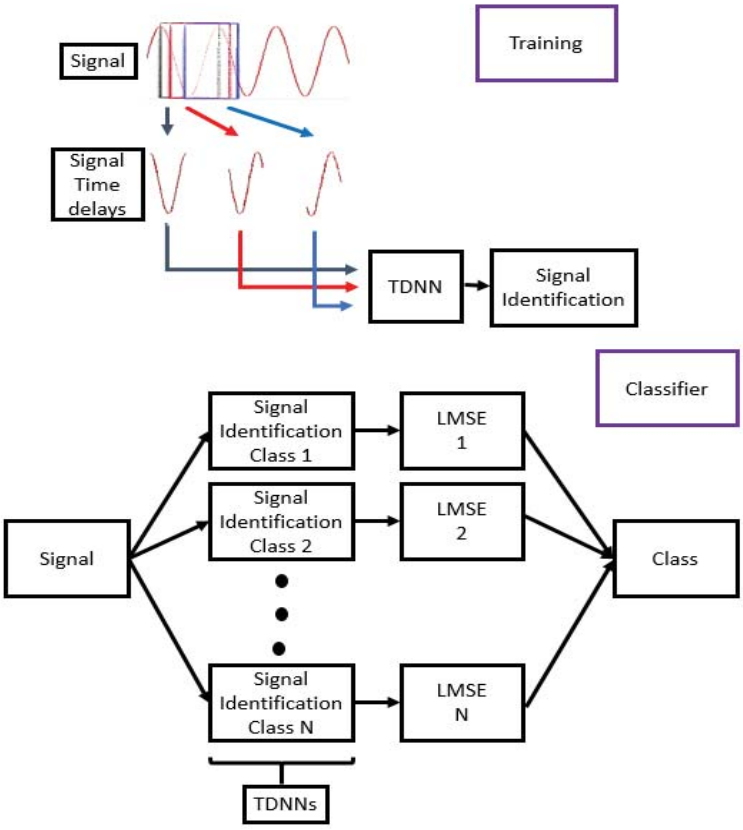

Figure 3-4: Full implementation of the TDNN. On the top the training description and below the classification procedure.

Designing the class of TDNN considered in this study (Figure $3-4$) demanded not only the global stability of the identifier but also the incorporation of some structural properties coming from the uncertain system to be identified. Also, it is often desirable to have a TDNN that converges fast enough to the trajectories of the uncertain system. The exponential stability analysis problem for TDNN can provide such behavior. Indeed, there is a big number of results on this topic [6], [72], [74], [135].

The concept of signal classifier as a time delay differential algorithm

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

$$\begin{gathered}\dot{x}_d(t) = f\left(x_d(t), u_d(t), u_d(t-h), ..., u_d(t-ph)\right) + \xi\left(x(t), t\right) \\ x_{d,t_0}(\theta) = x_d(t_0+\theta) = \varphi(\theta) \\ \forall\theta \in [-ph, 0] \quad p \in \mathbb{Z}^+\end{gathered} \tag{3.8}$$

The continuous signal $x_d(t) \in \mathbb{R}^n$ is the state of the time-delay system with $\|x_d(t)\| < \infty$, $\forall t \geq 0$. One must notice that the source of delay is coming from the input signal $u_d(t) \in \mathbb{R}^m$ which now represents the EEG signal at the current time as well as some of the delayed information. The function f_d represents the uncertain nonlinear function connecting the state of the plant with the delayed input signal $u_d(t-ih)$, $i = 0, ..., p$. The delay value h is known and constant, $h \in \mathbb{R}_+ \forall t \geq 0$.

System uncertainties and noises are described by the nonlinear unknown function $\xi(x(t), t) : \mathbb{R}^{n+1} \to \mathbb{R}^n$ and satisfies

$$\|\xi(x(t), t)\|^2 \leq \Upsilon \tag{3.9}$$

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

Neural network approximation for the time-delay signal classifier

The system can be represented as next

$$\begin{gathered}\dot{x}_d(t) = A_d x_d(t) + W_{d_1}^* \psi_{d_1}\left(x_d(t)\right) + W_{d_2}^* \psi_{d_2}\left(x_d(t)\right) u_d(t) + \\ \sum_{i=1}^{p} W_{d_2}^i(t) \psi_{d_2}^i\left(x_d(t)\right) u_{d_2}(t-ih) + \\ \tilde{f}\left(x_d(t), u_d(t), u_d(t-h), ..., u_d(t-ph)\right) + \xi(x(t))\end{gathered} \tag{3.10}$$

The set of matrices W_d^* are unknown but they are bounded.

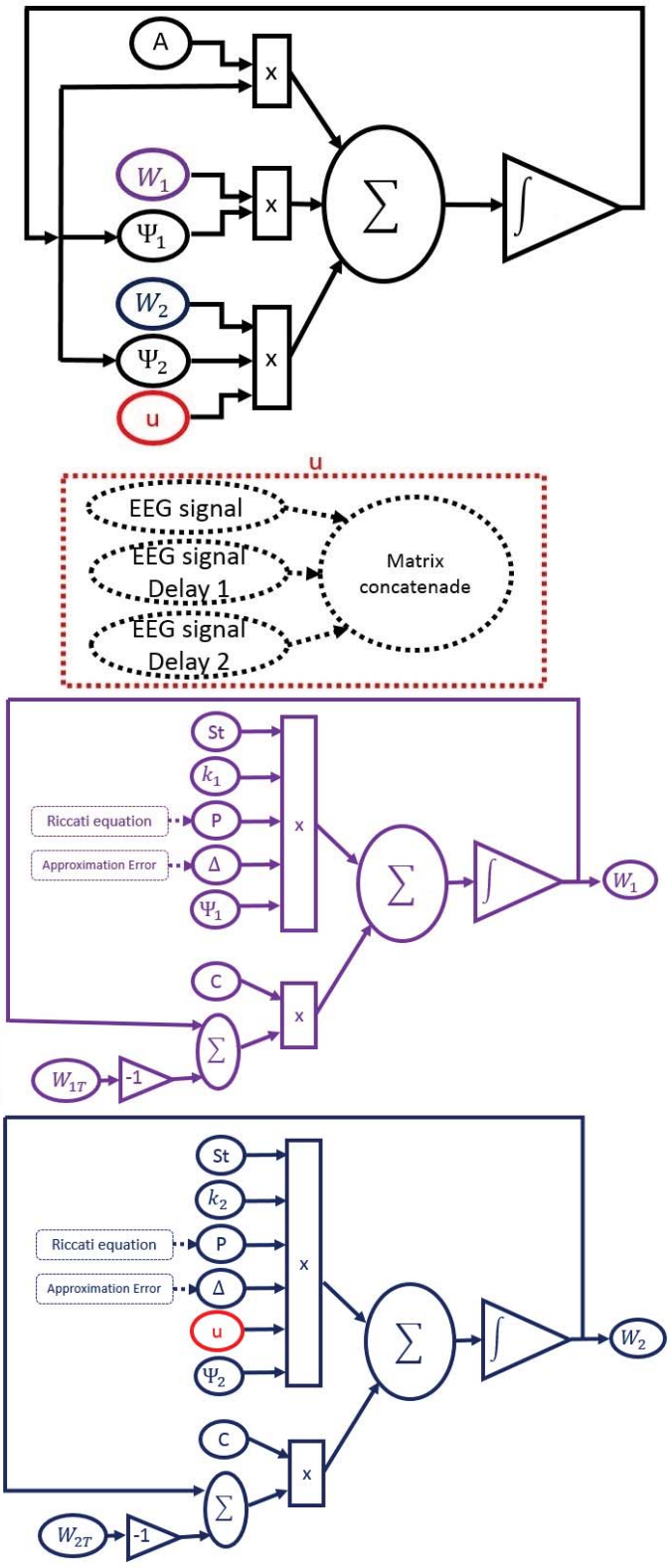

Figure 3-5: TDNN topology. On the top in black the TDNN classifier structure. In purple the W_1 and in blue the W_2 adjustment structures.

Classifier structure

The classifier based on NN Figure (3-5) is proposed following the classical strategy used for proposing adaptive parameter identification algorithms [100]; that is, considering a copy of the structural approximation for uncertain nonlinear systems with time-delays defined in (3.10). Consequently, the classifier based on NN has the following structure.

$$\begin{aligned}\frac{d}{dt}\hat{x}_d(t) = A_d\hat{x}_d(t) + W_{d_1}(t)\,\psi_{d_1}(\hat{x}_d(t)) + W_{d_2}(t)\psi_{d_2}(\hat{x}_d(t))\,u_d(t) + \\ \sum_{i=1}^{p} W_{d_2}^i(t)\,\psi_{d_2}^i(\hat{x}_d(t))\,u_d(t-ih) \\ \hat{x}_d(t) = \hat{x}_{d_{t_0}}, \quad \forall t \in [-ph, 0] \quad p \in \mathbb{Z}^+ \\ \hat{x}_{d_{t_0}} \in C[-ph, 0]\end{aligned} \tag{3.11}$$

Where $A_d \in \mathbb{R}^{n \times n}, W_{d_1} \in \mathbb{R}^{n \times s_a}, W_{d_2} \in \mathbb{R}^{n \times s_b}$ and $W_{d_2}^i \in \mathbb{R}^{n \times s_i}$. Here $\hat{x}_d$ defines the identifier state $W_{d_1}(\cdot)$, $W_{d_2}(\cdot)$ and $W_{d_2}^i(\cdot)$ are adaptive parameters that should be adjusted to approximate the uncertain nonlinear time-delay system (3.8). The matrix A_d and functions $\psi_{d_1}(\cdot)$, $\psi_{d_2}(\cdot)$ and $\psi_{d_2}^i(\cdot)$ have the same meaning to the one introduced in the previous section.

3.0.4 Complex valued neural networks

The complex numbers (CN) concept was not accepted as a valid idea within mathematics field for a long time [123]. Even when their mathematical characteristics and properties were successfully described, its applicability in real life was not understood completely before the middle of 19th century. Actually, during the industrial revolution mathematicians finally recognized the importance of CN not only in pure theoretical problems but also in engineering issues. Today, CN have been actively used in different areas such as physics, circuit theory, fluid analysis, etc.

Once the CN concept was accepted, dynamic systems including states defined within the complex domain were proposed and analyzed. This class of particular systems was used to describe a large set of real plants where waves and related phenomena appear. Among all systems based on CN, those described in terms of the so-called frequency response are particularly in-

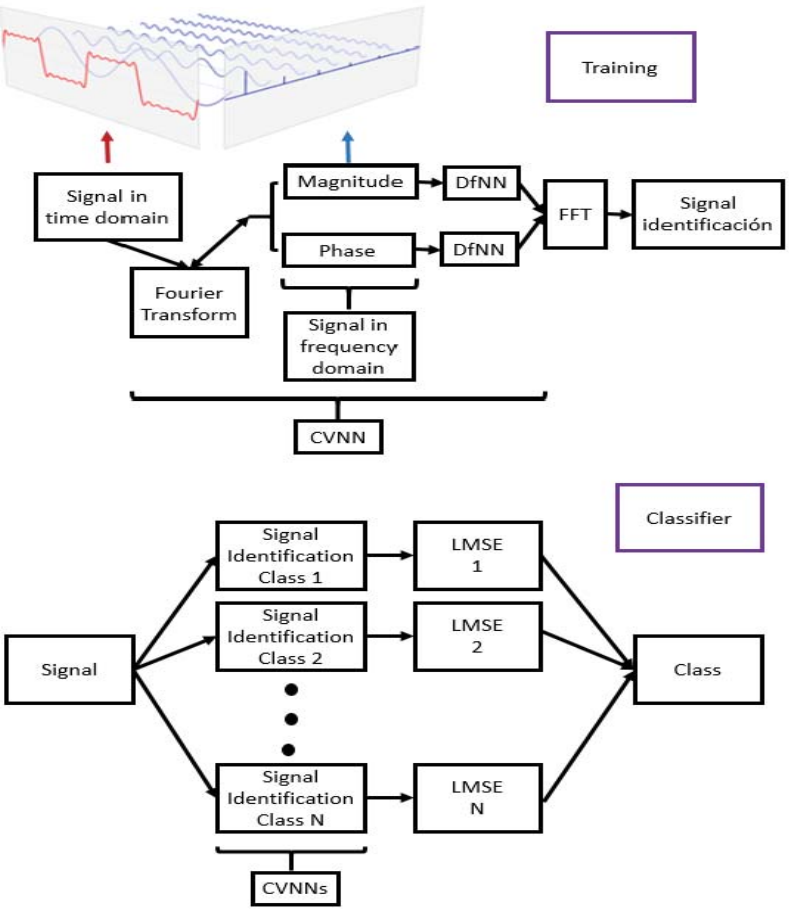

Figure 3-6: Function's time domain, shown in red, to the function's frequency domain, shown in blue. Full implementation of the CVNN for the classifier. On the top the training description and below the classification procedure.

teresting. Collecting all magnitude and phase dynamics through consecutive windows produces a set of nonlinear uncertain dynamics for each frequency described in terms of CN.

The presence of uncertainty in the previous system arises because the formal description of the aforementioned magnitude/phase variation with respect to time is difficult to be obtained accurately. Signal classification based on time variation of continuous signal may seem not be adequate even using CNN. One possible solution to overcome this problem is coming from the application of NN in the complex domain. Indeed, this class of network is the so-called complex valued neural networks (CVNN).

Complex-valued neurons are those whose input and output signals are represented as complex numbers. Validation and recognition of this class of NN was achieved easily because the large number of applications that can be handled by this kind of CNN. Actually, several studies have focused on the engineering usefulness of CVNN [4],[44], [80], [141], [105]. For example, in the human brain, an action potential may have different pulse patterns, and the distance between pulses may be different. This suggests that it is appropriate to introduce CNN representing phase and amplitude into NNs. Pattern recognition in EEG signals using CVNNs as the one propose in Figure $(3-6)$ may become also in active field within the area of so-called BCIs.

The class of complex-valued nonlinear systems

The class of nonlinear dynamics with complex-valued state considered in this part of thesis is characterized by the following complex-valued differential equation

$$\begin{aligned} \dot{x}_c(t) &= f_c(x_c(t), u_c(t)) + \xi(t) \\ x(0) &\text{ fixed and bounded} \\ f_c(\cdot,\cdot) &:= f_{c_r}(\cdot,\cdot) + j f_{c_i}(\cdot,\cdot) \end{aligned} \tag{3.12}$$

where $x_c \in \mathbb{C}^n$, $x_c := [x_{c_1}, ..., x_{c_n}]^\intercal$ defines the system state composed by its corresponding real $x_{c_r} \in \mathbb{R}^n$ and imaginary $x_{c_i} \in \mathbb{R}^n$ parts , that is $x_c = x_{c_r} + j x_{c_i}$. The function $u_c \in \mathbb{R}^m$ represents an exogenous input represented as a vector of complex numbers. Anyway, this signal is assumed to be measurable and bounded. In this thesis, the state is also assumed to be measurable. Measurability of the state is understood in the following sense: the magnitude $|x_{c_i}|$ and phase of each component $arg\{x_{c_i}\}$, $i = \overline{1,n}$ are available by different mechanisms in the whole period of time when the system evolves. This is not a strong assumption because it is usual to obtain such components in real applications where the frequency analysis gives characteristic information for the system beyond the time evolution one. For example, in acoustic engineering, impedance analysis, image and signal processing, etc. The nonlinear CV function $f_c(\cdot,\cdot) : \mathbb{C}^{n+m} \to \mathbb{C}^n$ is composed by two real-valued sections $f_{c_r}(\cdot,\cdot)$ and $f_{c_i}(\cdot,\cdot)$. These nonlinear functions $f_{c_r}(x_c, u_c)$ and $f_{c_i}(x_c, u_c)$ should satisfy a number of conditions such that the complex uncertain nonlinear system given by (3.12) has a solution.

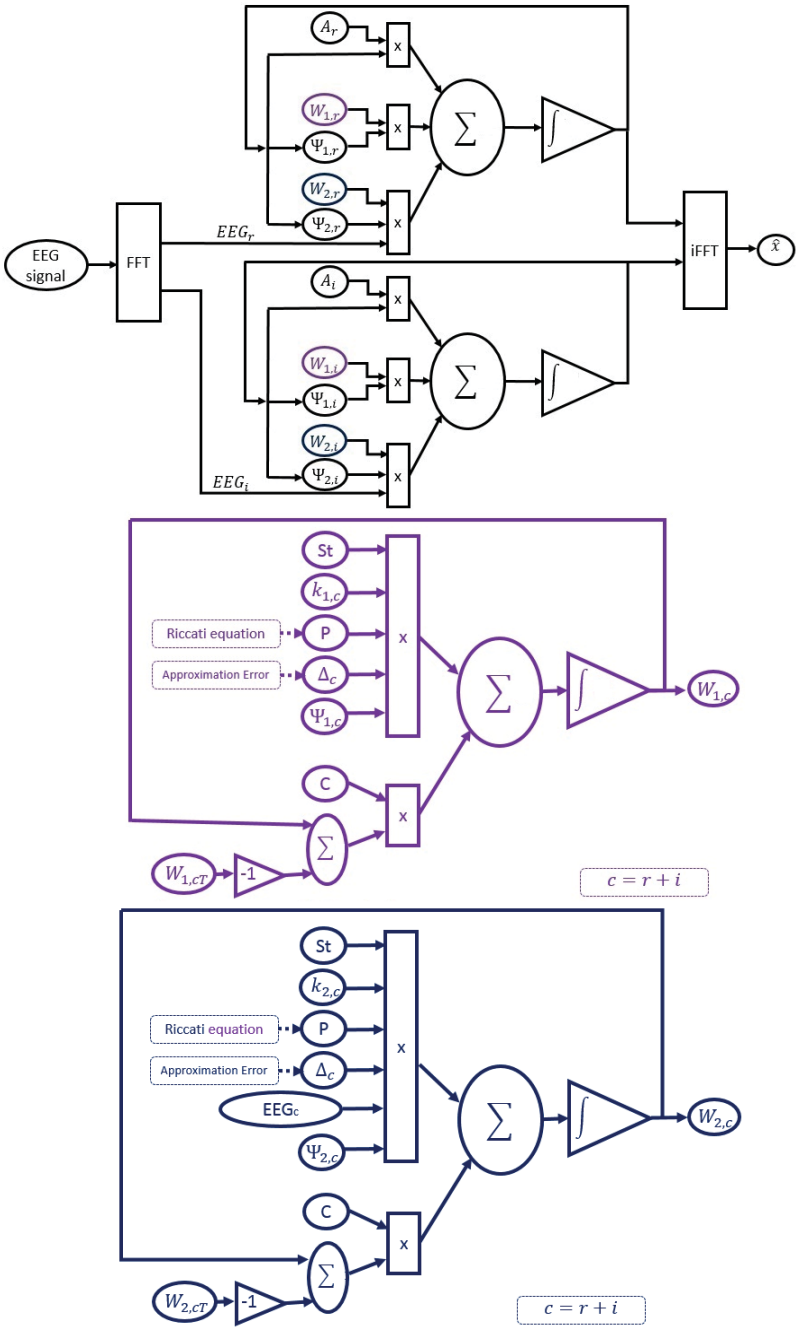

Figure 3-7: CVNN topology. On the top in black the CVNN classifier structure. In purple the W_1 and in blue the W_2 adjustment structures.

Neural network approximation for nonlinear complex-valued systems

The non-parametric mathematical model was obtained using a particular type of nonlinear least mean square algorithm. Using this method demands a very important assumption: the couple of nonlinear structures $f_{c_r}(x_c, u_c)$ and $f_{c_i}(x_c, u_c)$ should admit a numerical reconstruction based on NN. Actually, they should be CVNN. This pair of CVNN are represented by $f_{c_{r,0}}(x_c, u_c)$ and $f_{c_{i,0}}(x_c, u_c)$, respectively. This assumption is paramount to admit the existence of solution for the adaptive modeling problem. A number of possible approximations can be used here. Among others, the classical least mean square based on polynomials, sinusoids functions, wavelet nonlinear functions and NN can be used here. No matter what selection is adopted to obtain the approximation, the following construction for representing the uncertain complex-valued nonlinear system is proposed

$$\begin{aligned} \dot{x}_c(t) = A_c x_c(t) + W_{c_1}^* \psi_{c_1}(x_c(t)) + W_{c_2}^* \psi_{c_2}(x_c(t)) u_c(t) + \\ \eta_c(x_c(t)) + \xi(t) \end{aligned} \tag{3.13}$$

where $A_c = A_{c_r} + jA_{c_i}$, $W_{c_k}^* := W_{c_{r,k}}^* + jW_{c_{i,k}}^*$ $(k = 1, 2)$ and $\eta_c(x_c) := \eta_{c_r}(x_c) + j\eta_{cj}(x_c)$. The set of matrices $W_{c_j}^*$ are unknown but they are bounded.

Classifier structure

The classifier based on NN Figure (3-7) is proposed following the classical strategy used within the adaptive parameter identification framework [99]. This construction uses a structural copy of the approximation for uncertain system based on NN defined in (3.13). Therefore, the identifier based on NN has the following structure

$$\begin{aligned} \frac{d}{dt}\hat{x}_c(t) = A_c \hat{x}_c(t) + W_{c_1}(t) \psi_{c_1}(\hat{x}_c(t)) + W_{c_2}(t) \psi_{c_2}(\hat{x}_c(t)) u_c(t) \\ \hat{x}_c(0) \text{ fixed and bounded} \end{aligned} \tag{3.14}$$

Where the time varying weights $W_{c_s}(t) := W_{c_{r,s}}(t) + jW_{c_{i,s}}(t)$ $(s = 1, 2)$ are used to adjust the adaptive identifier. This is the main method to construct a general form for this numerical

approximation. The identifier state is defined as $\hat{x}_c(t) \in \mathbb{C}^n$. Matrix A_c and functions $\psi_{c_1}(\cdot)$, $\psi_{c_2}(\cdot)$ are the activation functions.

3.1 Neural networks implemented in embedded systems

Several methods have been proposed over the years to implement NN. These methods can be implemented in personal computers or embedded hardware. Software implementation offers the flexibility and full precision of powerful microprocessors with complex architecture algorithms as well as limited memory conditions. However, they have the disadvantage of being slower and with higher computational costs associated to operative system operations [59].

Hardware implementations are ideal for signal processing which demand high volume adaptive real time processing and learning of large data-sets in reasonable time and need the use of energy-efficient hardware, but it is limited in flexibility and usually offers lower precision, but has the advantage of processing data in parallel (as in FPGA or in VLSI), making possible the implementation of complex algorithms in soft real time.

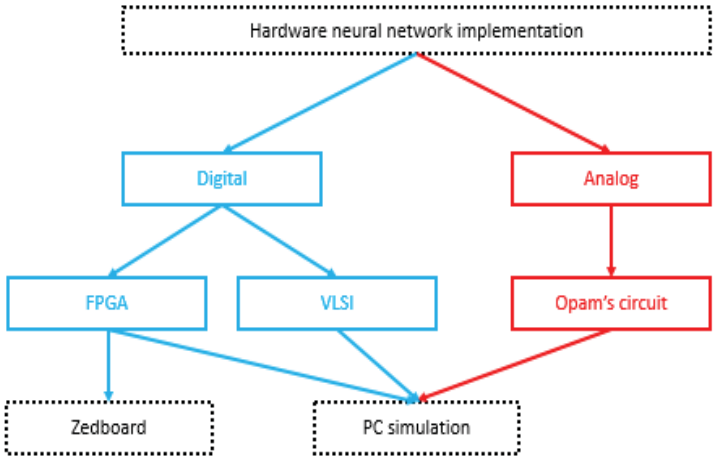

Figure 3-8: Digital and analog NNs implementation.

FPGAs are semiconductor devices that are based around a matrix of configurable logic blocks (CLBs) connected via programmable interconnections. FPGAs can be reprogrammed to desired application or functionality requirements after manufacturing. This feature dis-

tinguishes FPGAs from Application Specific Integrated Circuits (ASICs), which are custom manufactured for specific design tasks. Although one-time programmable (OTP) FPGAs are available, the dominant types are SRAM based which can be reprogrammed as the design evolves [133].

Nowadays, the capacity and performance of FPGAs have increased. Factors such as speed and circuit density make them suitable for signal processing in parallel. Digital signal processing is one of the primary design applications for FPGAs, they have been gaining considerable attraction in high performance applications because of their versatility and speed over traditional processors like microcontrollers [132]. Since 1990 there is a constant development of NNs applications in FPGAs [143], [41], [33] but the great majority of them work in a static way and do not take into account the continuous nature of the signal.

Hardware implementation can be also done by using VLSI and analog circuits. The VLSI is the process of creating an integrated circuit (IC) by combining thousands of transistors into a single chip, an advantage of VLSI over FPGA is that the design is not limited to fixed maximum of memory and speed [10]. It is possible to generate a VLSI design from VHDL code. Since the 1980s there are reports of implementation of NN in VLSI [10], [104].

In the analog implementation of NNs a coding method is used in which signals are represented by currents or voltages. This allows us to think of these systems as operating with real numbers during the NN simulation. Even thought the analog implementation is the fastest one, it also represents a challenge in the characterization of the algorithms in analog circuits.

3.2 Off-line adjustment of weights in the classifiers

For each type of DNN presented on this work, the classifier development follows the same structure. The first part is the training process. By using a part of each database, we are able to set the weights for each class. Next, a validation of this training procedure was developed using a simple strategy of substituting the weights produced by the training and then, testing the on-line learning system. This is a key aspect that reveals the differences between the classical SNN and the DNN.

3.2.1 Training

The identifier structures previously propose demand the knowledge of $W_j^*, \quad j = 1, 2$ that are actually unknown. Therefore, the learning laws cannot be used as they were presented previously. Nevertheless, a training algorithm can be used to overcome this problem. In this thesis, the training follows the method proposed in [14]. The training method will produce a set of values $W_j^{*,i}$ that must be near to W_j^* in some sense. The following procedure will show how the values of $W_j^{*,i}$ are obtained.

1. Collect a few sets X_k^{tra} $(k = 1, ..., r)$ of discrete state's measurements $X_k^{tra} := \{x_k^{tra}(t_s), s$ a positive integer or zero with $t_s - t_{s-1} > 0\}$ or continuous trajectories $X_k^{tra} := \left\{x_k^{tra}(t),\ t \geq 0\right\}$ obtained from the uncertain system that will be identified by the NN. Evidently the information gotten from these sets is affected by measurements noises.

2. Propose a parameter identification algorithm to obtain a suitable value of W^* using the available information collected in the previous step. In this step, just a function of all information sets collected here is used to perform the training. In this case, we have used the 60% of all the sets on the Database I and Database II from each class to perform the training. The remainder sets from each database were used to validate and CNN off-line training.

3. Set the weights W^* on the identifier structure proposed above.

The key step to perform the training is the second one. Several methods have been proposed to solve this issue. Among others, matrix least mean square with several modifications have been used to solve that training algorithm.

3.2.2 On-line training scheme using the continuous version of least mean square method for RNN

In order to probe the evolution of the weights for each class the following theorem is used

$$\dot{W}_{r,i}(k+1) = \left(g_{r,i}^{-1} l - \frac{1}{4}\Psi_{r,i}(k)\,\Psi_{r,i}^{\top}(k)\right)\left(g_{r,i}^{-1} l + \frac{\Psi_{r,i}(k)\,\Psi_{r,i}^{\top}(k)}{4}\right)^{-1} W_{r,i}(k) +$$

$$PA\Delta_r(k)\Psi_{r,i}^{\top}(k)\left(g_{r,i}^{-1} l + \frac{\Psi_{i,r}(k)\,\Psi_{i,r}^{\top}(k)}{4}\right)^{-1}$$

$$\Psi_{r,1}(k) = \sigma(x(k)), \Psi_{r,2}(k) = \varphi(x(k))\,u(k)$$

$$g_{r,i} \in \mathbb{R}^{+}, P = P^{\top} > 0$$

Here the matrices $\dot{W}_{r,i} \in \mathbb{R}^{n\times 1}$ represent the weights values that are evolving and will be fixed in the second stage of the classification process that is called the validation process.

The time varying $\Delta_r \in \mathbb{R}^n$ function is the identification error. Evidently, the accuracy of these values depends on the number of weights adopted to represent the identifier dynamics. The matrix P is the positive definite solutions for the Riccati equations and $\Psi_{r,1}$ and $\Psi_{r,2}$ are the activation functions previously described.

3.2.3 On-line training scheme using the continuous version of least mean square method for DfNN

The nonlinear weight updating (learning) law to adjust the identifier is described by the following matrix differential equations

$$\dot{W}_j(t) = -k_j P\Delta(t)\,\Psi_j^{\top} + 2^{-1} k_j \tilde{W}_j(t)$$

Matrices $\tilde{W}_{1,t} \in \mathbb{R}^{n\times 1}$ and $\tilde{W}_{2,t} \in \mathbb{R}^{n\times s}$ represent the distance between the current values of $W_{1,t}$ and $W_{2,t}$ to fixed fitted values $W_{1,t}^0$ and $W_{2,t}^0$ that is $\tilde{W}_{j,t} = W_{j,t} - W_{j,0}$. The matrices $W_{1,t}^0$ and $W_{2,t}^0$ are fixed after training procedure. The learning laws selected above are used to obtain the so-called best-fitted weights values that will be fixed in the second stage of the classification process that is called the validation process.

The time varying $\Delta_t \in \mathbb{R}^n$ function is the identification error. Evidently, the accuracy of

these values depends on the number of weights adopted to represent the identifier dynamics. The variables $k_j, j = 1, 2$ are the learning rates. The matrix P is the positive definite solutions for the Riccati equations given by

$$Ric(P) := PA + A^{\top}P + PRP + Q$$

$$R = W_1^0 (\Lambda_2)^{-1} \left[W_1^0\right]^{\top} + W_2^0 (\Lambda_4)^{-1} \left[W_2^0\right]^{\top} + \Lambda_1 + \Lambda_3$$

$$Q = \lambda_{\max} ((2\Lambda_2) I_1) I_{nxn} + Q_0$$

Here $\Lambda_k \in \mathbb{R}^{n\times n}, k = 1, 4$ are positive definite matrices too. In fact, they must be selected (over a large set of possible values) just to ensure the existence of the solution for the previous equations.

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

The following theorem is used to prove the evolution of the weights for the TDNN. Let consider the time-delay uncertain system (3.8) with a known number of delays as well as the value of this delay. Suppose that perturbations and non-modeled system $\xi(x(t), t)$ affecting the uncertain system fulfills (3.9). If there exist positive definite matrices $\Lambda_{d_k} > 0$, $\Lambda_{d_k} = \Lambda_{d_k}^{\top}$, $\Lambda_{d_k} \in \mathbb{R}^{n\times n}$,

$k_d = 1, 2$ such that the following matrix Riccati equation $W_d(h, P, Q, R, \alpha) = 0$ with

$$W_d(h, P, Q, R, \alpha) = P\hat{A}_d + \hat{A}_d^\top P + PRP + Q$$

$$\hat{A}_d = A_d + (1 + \frac{\alpha}{2})I$$

$$R = V_1^+ + \sum_{i=1}^{p} V_2^{+,i} + \Lambda_1 + \Lambda_2 \tag{3.15}$$

$$Q = \left(\lambda_{\max}\left(\Lambda_1^{-1}\right) L_1 + pu^+ \lambda_{\max}\left(\Lambda_2^{-1}\right) L_2\right) I_{n\times n},$$

$$\alpha > 0$$

has at least one positive definite solution $P > 0$, $P = P^\top$, $P \in \mathbb{R}^{n\times n}$ used within the learning laws by

$$\dot{W}_{d_1}(t) = -k_{d_1}^{-1} e^{2k_d t} P \Delta_d(t) \psi_{d_1}^\top\left(\hat{x}_d(t)\right) - \alpha \tilde{W}_{d_1}^{tr}(t) - \left(I + \Lambda_{W d_1}\right) \tilde{W}_{d_1}^{tr}(t) \tag{3.16}$$

and $W_{d_2}^i(t)$ associated to the inputs are described by

$$\dot{W}_{d_2}^i(t) = -\int_{\tau=t-ih}^{t-(i-1)h} \alpha e^{2k_d(\tau)} \tilde{W}_{d_2}^{tr.i}(t)\, \bar{\Psi}_d^{+,i}(\tau) d\tau -$$

$$\alpha \tilde{W}_{d_2}^{tr,i}(t) + 2e^{2k_d t} P \Delta^\top(t)\, \psi_{d_2}^i(\hat{x}_d(t))\, u_d(t-ih)) -$$

$$\Omega^i \tilde{W}_{d_2}^{tr,i}(t) \left(e^{2k_d t_{i+1}} \bar{\Psi}_{d_2}^i(t_{i+1}) \left[\bar{\Psi}_{d_2}^i(t_{i+1})\right]^\top -\right.$$

$$\left. e^{2k_d t_i} \bar{\Psi}_{d_2}^i(t_i) \left[\bar{\Psi}_{d_2}^i(t_i)\right)\right]^\top - \left(I + \Lambda_{W_{d_2}}\right) \tilde{W}_{d_2}^{tr,i}(t) \tag{3.17}$$

$$\Omega^i = \left(2k_{d_2} I_{n\times n} + \int_{\tau=t_{i-1}}^{t_i} e^{2k_d(\tau)} \tilde{\Psi}_d^{+,i}(\tau)\, d\tau \right)^{-1}$$

$$\tilde{\Psi}_d^{+,i}(\tau) = \tilde{\Psi}_{d_2}^i(\tau) \left[\tilde{\Psi}_{d_2}^i(\tau)\right]^\top$$

$$\bar{\Psi}_{d_2}^i(t) = \psi_{d_2}^i(\hat{x}_d(t))\, u_d(t)$$

$(t_i = t-(i-1)h)$ governing the dynamic behavior of the time-delay neural network identifier (3.11).

3.2.5 On-line training scheme using the continuous version of least mean square method for CVNN

The discrete version of the algorithm presented in the previous sections can be easily reproduced following a similar procedure. Moreover, nonlinear least mean square method based on a class

of NN also has been introduced in (3.14), the following equation is valid

$$\dot{x}_c(t)\, x_c^{\intercal}(t) = A_c x_c(t) x_c^{\intercal}(t) + W_c^{*}\Psi_c\left(x_c(t), u_c(t)\right) x_c^{\intercal}(t) + \chi_c(t)\, x_c^{\intercal}(t)$$

$$\begin{aligned} W_c^{*} &:= \left[\begin{array}{cc} W_{c_1}^{*} & W_{c_2}^{*} \end{array}\right] \\ \Psi_c\left(x_c(t), u_c(t)\right) &:= \left[\begin{array}{c} \psi_{c_1}\left(x_c(t)\right) \\ \psi_{c_2}\left(x_c(t)\right) u_c(t) \end{array}\right] \\ \chi_c(t) &= \left[\eta_c\left(x_c(t)\right) + \xi(t)\right] \end{aligned} \tag{3.18}$$

The direct integral operation from 0 to T on both sides of the previous equation yields to

$$\begin{aligned} \int_{\tau=0}^{T} \dot{x}_c(\tau)\, x_c^{\intercal}(\tau) d\tau &= \int_{\tau=0}^{T} A_c x_c(\tau) x_c^{\intercal}(\tau) d\tau + \\ W_c^{*} \int_{\tau=0}^{T} \Psi_c\left(x_c(\tau), u_c(\tau)\right) x_c^{\intercal}(\tau) d\tau &+ \int_{\tau=0}^{T} \chi_c(\tau)\, x_c^{\intercal}(\tau) d\tau \end{aligned} \tag{3.19}$$

Because the upper limit in the integral operation described in the last equation is finite, the weights produced after training never will be the true ones. This is one of the reasons to introduce an on-line learning procedure that already has been detailed. The left hand side on the previous equation can be estimated as (using a simple integration by parts)

$$\int_{\tau=0}^{T} \dot{x}_c(\tau)\, x_c^{\intercal}(\tau) d\tau = x_c(T)\, x_c^{\intercal}(T) - x_c(0)\, x_c^{\intercal}(0) - \int_{\tau=0}^{T} \dot{x}_c(\tau)\, x_c^{\intercal}(\tau) d\tau \tag{3.20}$$

Therefore, by the result of the integration by parts, one has

$$\begin{aligned} \frac{1}{2}\left[x(T)\, x^{\intercal}(T) - x(0)\, x^{\intercal}(0)\right] &= A \int_{\tau=0}^{T} x(\tau) x^{\intercal}(\tau) d\tau + \\ W^{*} \int_{\tau=0}^{T} \Psi\left(x(\tau), u(\tau)\right) x^{\intercal}(\tau) d\tau &+ \int_{\tau=0}^{T} \chi(\tau)\, x^{\intercal}(\tau) d\tau \end{aligned} \tag{3.21}$$

The term $\int_{\tau=0}^{T} \chi_c(\tau)\, x_c^{\intercal}(\tau) d\tau$ is not available to solve the previous equation. Here one can obtain an approximation for W_c^*, namely $W_c^{*,id}$ that actually is the solution of the training algorithm. This solution is based on well known results on matrix least mean square algorithms. In this thesis, we omitted to discuss details on the solvability of this problem because there is a lot of information on that ([101] and the reference therein). The formal expression for this term is

$$W_c^{*,id} = \left[\Omega - A_c \int_{\tau=0}^{T} x_c(\tau) x_c^{\intercal}(\tau) d\tau\right] * \left[\int_{\tau=0}^{T} \Psi_c\left(x_c(\tau), u_c(\tau)\right) x_c^{\intercal}(\tau) d\tau\right]^{-1} \tag{3.22}$$

$$\Omega := \frac{1}{2}\left[x_c(T)\, x_c^{\intercal}(T) - x_c(0)\, x_c^{\intercal}(0)\right]$$

When the available information for training process is substantial ($T \to \infty$), the following expression can be also used to obtain the value of $W_c^{*,id}$.

$$W_c^{*,id} = \lim_{T\to\infty} \left[\Omega - A_c \int_{\tau=0}^{T} x_c(\tau) x_c^{\intercal}(\tau) d\tau\right] * \left[\int_{\tau=0}^{T} \Psi_c\left(x_c(\tau), u_c(\tau)\right) x_c^{\intercal}(\tau) d\tau\right]^{-1} \tag{3.23}$$

No matter which one is consulted, the following property has been discovered $\left\|W_c^{*,id} - W_c^*\right\|^2 \leq \Upsilon$ where Υ is bounded and positive constant. Last equation is only valid when $T \to \infty$. However, if T if finite, a small deviation of the real $W_c^{*,id}$ value is obtained. Therefore, instead of $W_c^{*,id}$, a value named $\bar{W}_c^{*,id}$ (obtained with a finite time T in the previous equation) is introduced in the adjustment laws.

Therefore, the learning laws defined previously are transformed into

$$\dot{W}_{c_j}(t) := -k_{c_j} P_c \Delta_c(t) \psi_{c_j}^{\intercal}\left(\hat{x}_c(t)\right) + \alpha_c \bar{W}_{c_j}^{id} \tag{3.24}$$

This change is used to evaluate how the training process can affect the quality of identification based on DNN for the complex-valued uncertain system. The utilization of this process is obligatory because no knowledge on W_c^* is assumed. Therefore, the identifier state is also bounded by the assumption presented in this subsection. Using the conditions established

in this part, changing $W_c^{*,id}$ instead of W_c^* produces a bigger but bounded deviation of the identifier state compared to the trajectories of the uncertain system. Therefore, the quality of the identification process is reduced but the decrement is measurable and bounded.

Chapter 4

Experimental development

This chapter describes the databases that were employed to test the different proposed NNs and how their data were processed. The first database, downloaded from [92] was chosen because it is one of the most referenced databases, there is a lot of literature on classification algorithms that were tested and validated. The second database was self-generated in the facilities of the National Polytechnic Institute the protocol used to produce this database is also described. It is important to remark that the second database is the first step for future migration of the proposed NNs based classification processes from the PC to an embedded system.

Database I is employed to measure the different NN classification accuracy as shown in Figure (4-1), due to the fact that the particular database have been employed for classification with different NN based techniques. The data processing is the next:

First, one of the four NN is selected. This NN is used during the whole process for the training and also for the construction of the parallel NN.

Second, each class that make up database I is divided in three sets in order to perform the validation by employing a generalization-regularization method:

- Set one is compound of 60% of the samples of each class. With this data the training is perform, during the training the weights evolve at the end they concentrate information that characterize the each class.

- Set two is used for the validation and is 30% of each class. The 30% can be make up from samples used for the training and data that have not been known to the NN. At this

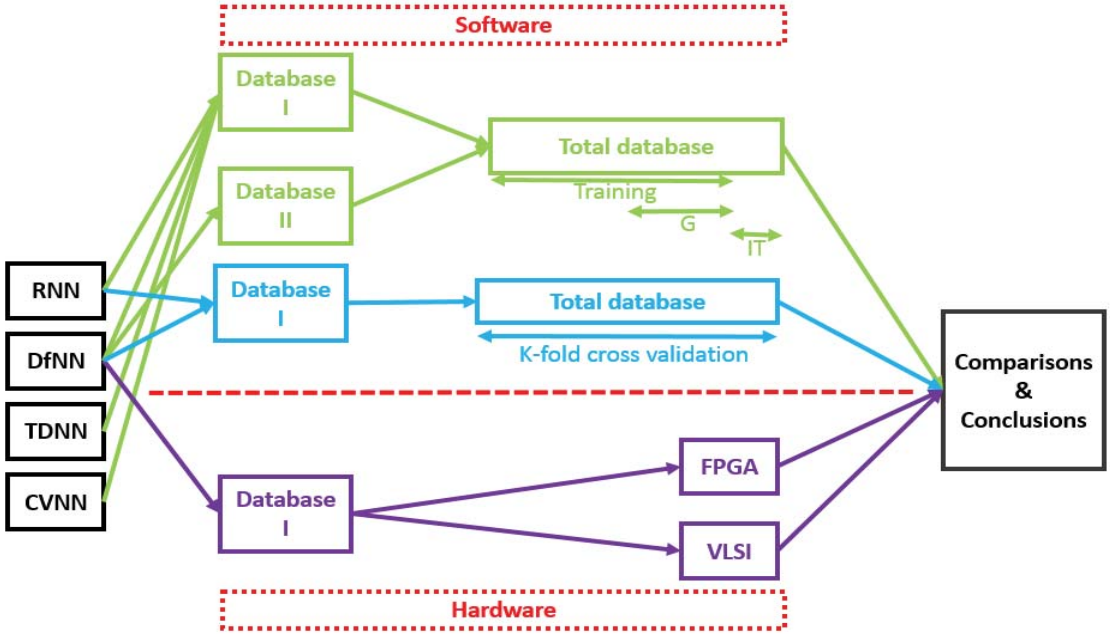

Figure 4-1: Experimental development followed, where the different implemented NNs classify the two databases. Generalization (G). Independent Test (IT).

point the parallel NN is build.

- Set three is the compare or test section, the remaining 10% of each set is employ. The data that make up set three is data have never been presented to the NN.

The next process depicted in Figure (4-1) with a blue color solid line is a 5-fold cross validation method apply to the RNN and DfNN in order to measure their classification accuracy, this method consists in the following steps:

- The database class in divided in 5, 1 of this parts is used to validate the other to develop the weights (training).
- This process is repeat for the 5 classes that make up database I.

4.1 Database I

Database I was downloaded from the *Freiburg University* web-site, a more complete description can be found in their web page and also in several papers that are listed there.

Database I contains invasive EEG recordings of 21 patients suffering from medically intractable focal epilepsy. The data were recorded during an invasive pre-surgical epilepsy monitoring at the Epilepsy Center of the University Hospital of Freiburg, Germany.

The records come from the epileptic focus, 100 samples taken from 11 patients where the EEG record was located in neocortical brain structures, other 100 where from eight patients in the hippocampus. Two patients were analyzed by both methods. It is important to remark that even when the signals correspond to invasive or non invase EEG trails, the signals looks pretty alike, so their classification became difficult even for the trained eye.

In order to obtain a high signal-to-noise ratio, fewer artifacts, and to record directly from focal areas, intracranial grid-, strip-, and depth-electrodes were utilized. The EEG data were acquired using a Neurofile NT digital video EEG system with 128 channels, 256 Hz sampling rate, and a 16 bit analogue-to-digital converter.

Notch or band pass filters have not been applied. So the signal is considered to be raw. Much of the work consulted for this thesis present classification of EEG signals using pre processed signals. For each patients, there are datasets called *ictal* and *interictal*, the former containing files with epileptic seizures and at least 50 min pre-ictal data, the latter containing approximately 24 hours of EEG recordings without seizure activity.

At least 24 h of continuous interictal recordings are available for 13 patients. For the remaining patients interictal invasive EEG data consisting of less than 24 h were joined together, to end up with at least 24 h per patient. For each patient, the recordings of three focal and three extra-focal electrode contacts are available.

Finally, the database is divided in 5 classes. Each class contain 100 samples, the first one extracranial data taken form the neocortical structures Figure $(4-2)$, the second class is formed by 100 samples of the hippocampus EEG signals Figure $(4-2)$, the third set of 100 samples were taken intracranial from neocortical structures Figure $(4-2)$, quarter 100 samples intracranial hippocampus Figure $(4-2)$ and 100 samples taken during a seizure episode Figure

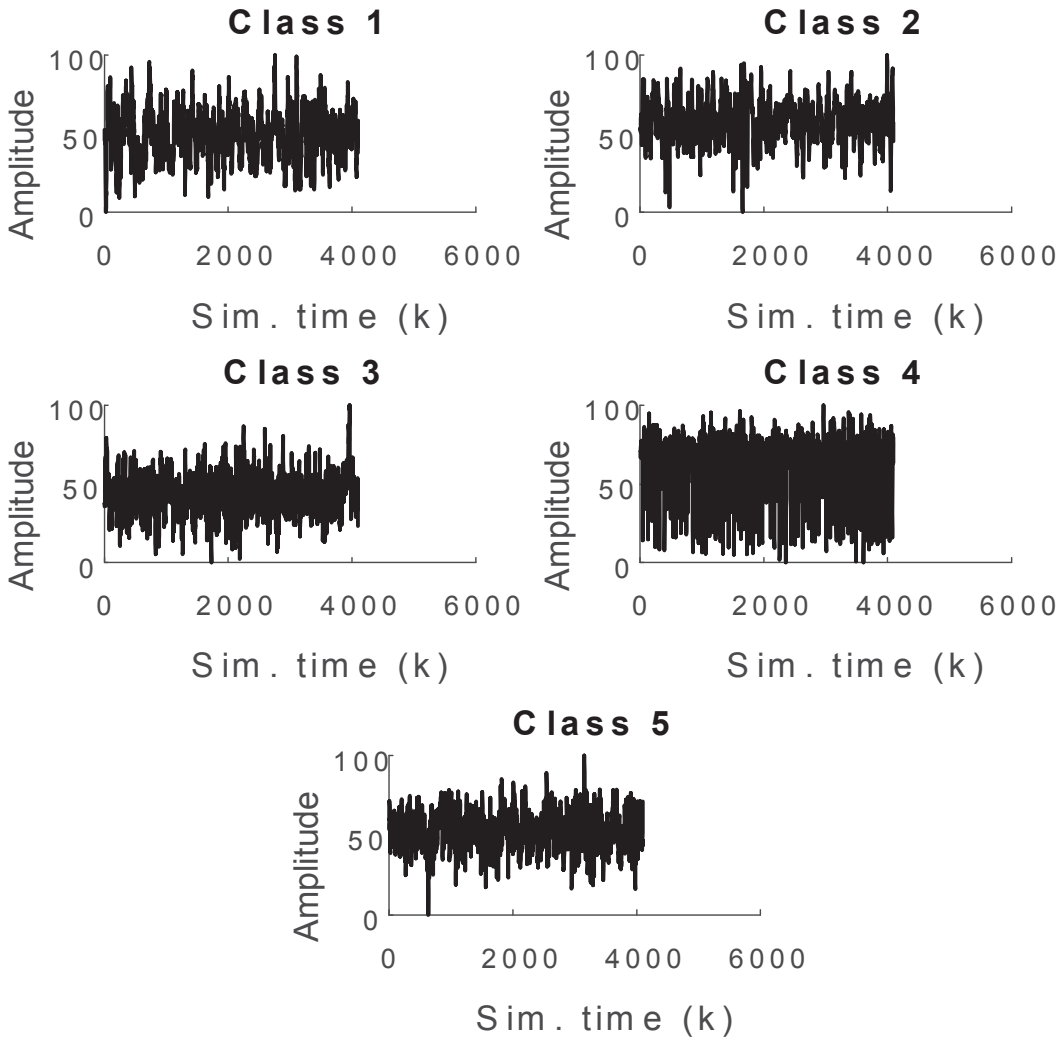

Figure 4-2: Examples of the signal taken from [92], class 1 signals correspond to extracranial data taken form the neocortical structures. Class 2 signals correspond to extracranial data taken from the hippocampus, class 3 signals correspond to intracranial data taken from neocortical structures, class 4 signals correspond to intracranial data taken from hippocampus and finally class 5 signals correspond to intracranial data taken during seizure episode.

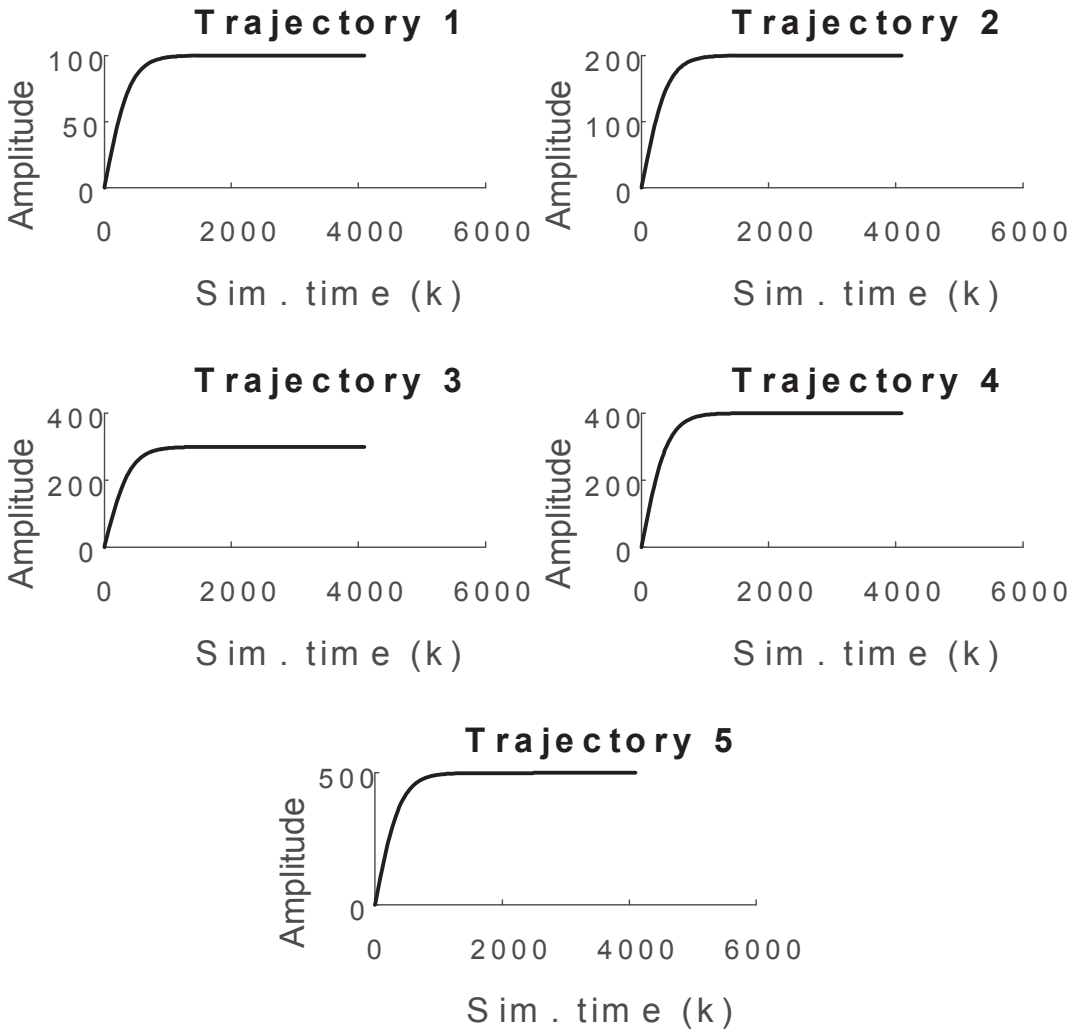

Figure 4-3: Example of the different trajectories employed for the training of the different NNs topologies, the amplitude of the trajectories could vary according to the network needs. Also the number of trajectories depends in the number of classes that the NN is classifying.

$(4-2)$. Unfortunately, the Database I is discontinued and not further available for download since it is superseded by the new European Epilepsy Database. An example of the trajectories employed to train the NNs for this database can be seen in Figure $(4-3)$.

4.2 Database II

The goal in building the Database II is to create a program that acquires data from the EEG device. The information contained in the Database II, represent the recordings of a EEG device produced by the EMOTIV® device produced by the Emotive company. All the presented data are waveforms resulting of VEP generated by visual stimulation.

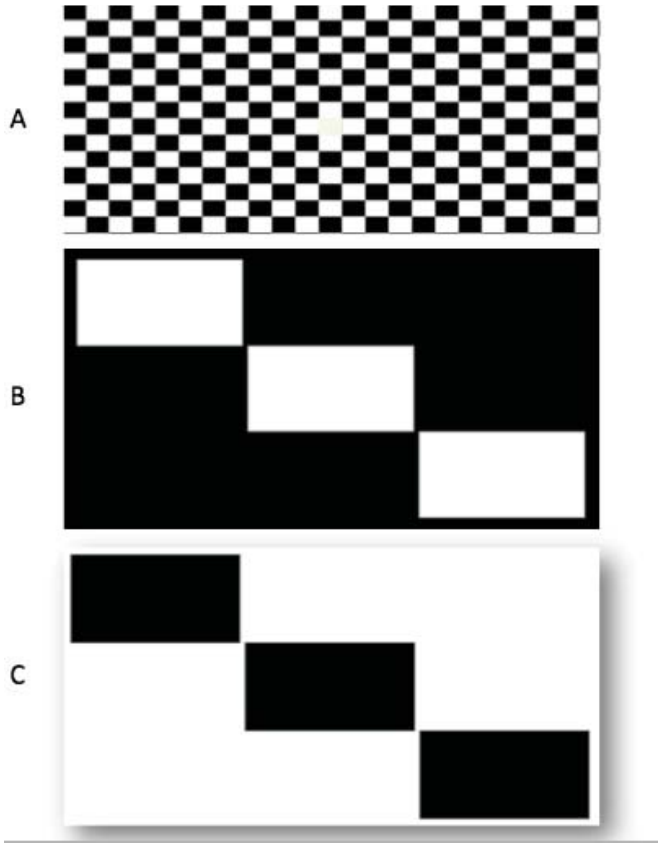

Figure 4-4: Different patterns (A, B and C) used for the building of the Database II.

The type of stimuli used to generate Database II is known in neurosciences as single graphics stimuli, where simple shapes such as rectangles, squares, or arrows, among others are rendered

on a computer screen and appear from and disappear into a background/ whiteground at a specified rate [142]. The stimulation rate is reported as the number of full cycles per second, normally simply referred to as the frequency of the stimulus. All repetitive visual stimuli have various properties such as frequency, color, and contrast. Both the type and properties of stimuli affect the elicited VEP response.

The patterns used for Database II are shown in Figure $(4-4)$, A was taken from regular VEP studies, B and C patterns were proposed according to well-known visual patterns used in neurological studies. Each pattern showed in Figure $(4-4)$ is part of a 40 seconds long video. During this time the pattern blinks with a certain rhythm that is the same for all the stimulation sequences. Each sequence was applied three times to the same volunteer, a break of 1 to 2 minutes was taken between trials so the eyes of the volunteers could rest. The hardest thing to implement was the EEG recording system, because the stimulation and the recording have to start and finish at the same time for each trial.

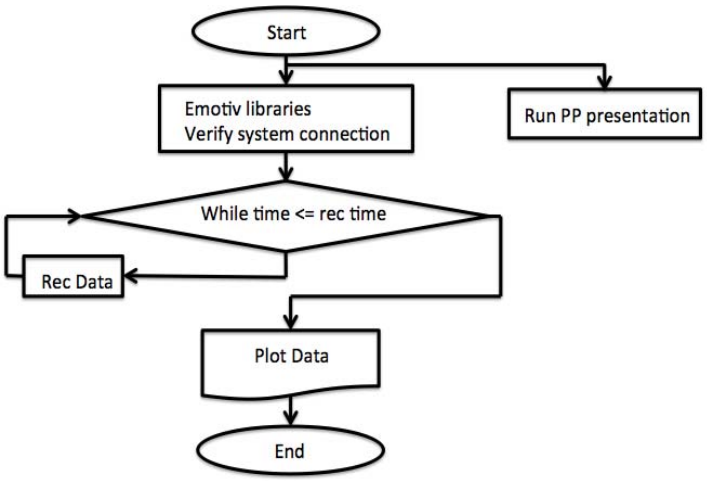

Figure 4-5: Flowchart; Emotiv connection with MATLAB

Even though, the EMOTIV® provides an interface that allow the recording of the VEP data, it was not suitable to perform the visual stimulation tests. The main problem with the EMOTIV software is the fact that you are not able to control the process remotely, so it was not possible to start the stimulation and the recording of the EEG signals at the same time. In order to fulfill this requirement, a program controlled the recording and the stimulation (Figure

Figure 4-6: Volunteer during trial.

$4-5$).

The program has the ability to control other Windows environment applications, including the ones that reproduce the sequence of stimulation patterns. Parallel to this, the libraries that control the EMOTIV drivers are loaded, giving the user the advantage of controlling parameters such as; recording time, start of stimulation, quality of the EEG electrode contact and the file format of signals recorded.

Once the software was proven to work properly, the trials conditions were detailed. The EEG signals are susceptible to the environment and factors such as noise, light and even the mood of the volunteer may be reflected on them. The trials took place in a confined room as shown in Figure (4-6), avoiding noises, with low light level, taking in account a break time during the different stimulation patterns, asking the volunteers to be as focused as possible on the information coming up on the screen. During the trials, it was important to provide a relaxing environment for volunteers.

Then a second step of the program take the file produced during the EEG evaluation and evaluated in automatic run within any of the proposed NNs. It is important to remark that

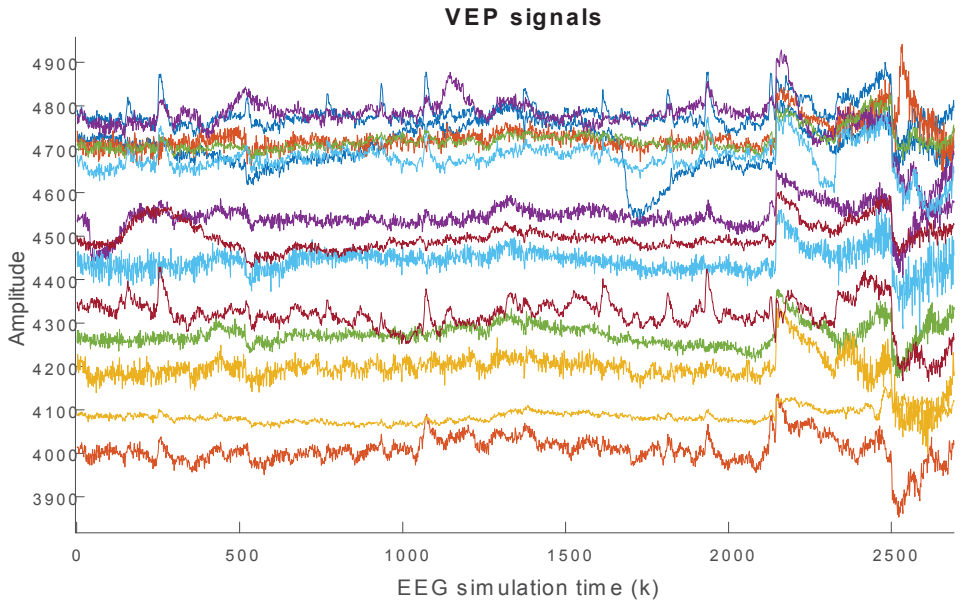

Figure 4-7: Fourteen evoked potentials recorded with the EMOTIV and plotted using MATLAB

the information is not submitted to any kind of prelimnary treatment during this process, so the EEG signals obtained by the EMOTIV are raw (Figure 4-7).

Database II is divided in 3 classes. Each one correspond to one of the visual patterns used to produce the visual stimulus. Each volunteer produces a set of 5 trails. A total of 6 volunteers helped in developing this database. The volunteers were healthy adults between 25 and 35 years old. Finally, a total of 90 records integrated the database.

4.3 Material and EEG system

As mentioned in the previous section the device used to acquire the information of the Database II was an EMOTIV EEG Neuroheadset. This is a high resolution, multi-channel, portable system which has been designed for practical research applications [26]. Some of the EMOTIV EEG Neuroheadset characteristics are mentioned next.

The EMOTIV also provides a research software development kit (Figure (4-8) that allow the user to obtain real-time display of the EMOTIV headset data stream, including EEG, contact quality, FFT, gyro (if fitted – custom option), wireless packet acquisition / loss display, marker

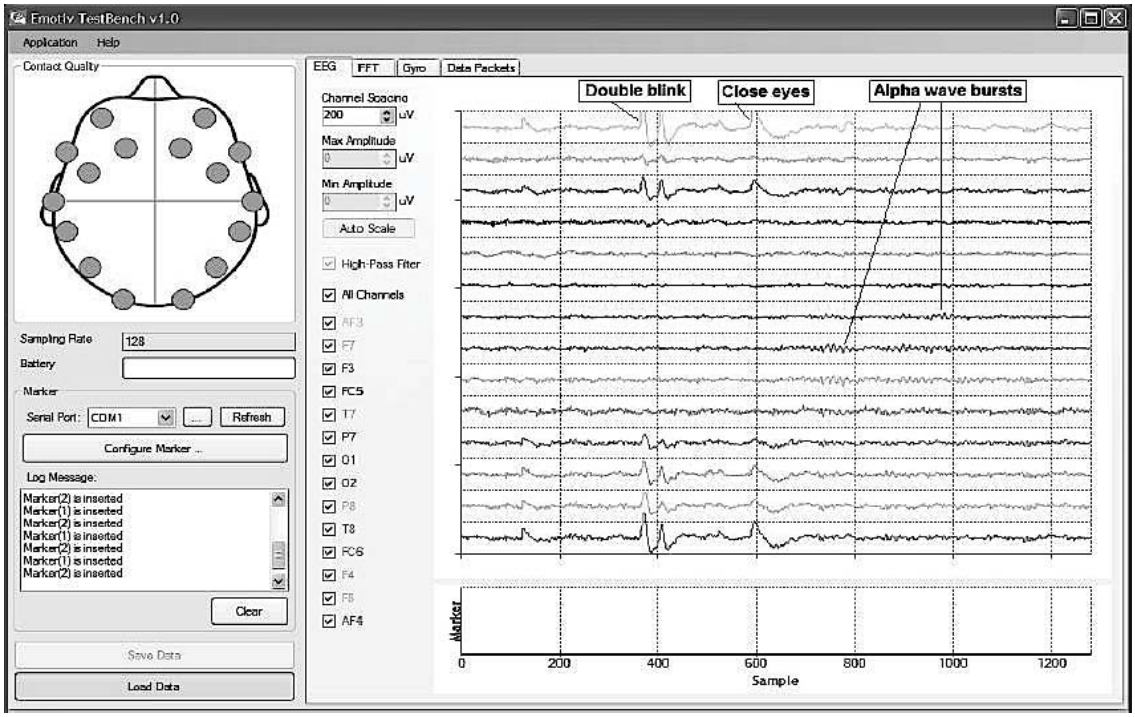

Figure 4-8: Emotiv Research Edition SDK interface window [26].

events, headset battery level, etc. Some characteristics include: recording and replaying EEG files in binary EEGLAB format, command line file converter that produce .csv format files, define and insert timed markers into the data stream, including on-screen buttons and defining serial port events, markers are stored in EEG data files, marker definitions can be saved and reloaded, markers are displayed in real time and playback modes.

EEG display

- 5 second rolling time window (chart recorder mode)
- All or some selected channels can be displayed
- Automatic or manual scaling (individual channel display mode)
- Adjustable channel offset (multi-channel display mode)
- Synchronized marker window

FFT display

- Selected channel only
- All or selected channels can be displayed
- Adjustable sampling window size (in samples)
- Adjustable update rate (in samples)
- dB mode – power or amplitude calculations
- dB scale
- FFT window methods: Hanning, Hamming, Hann, Blackman, Rectangle
- Predefined and custom sub-band histogram display – Delta, Theta, Alpha, Beta, custom bands

Gyro display

- 5 second rolling time window (chart recorder mode)
- X and Y deflection

Data Packet display

- 5 second rolling graph of Packet Counter output
- Packet loss – integrated count of missing data packets
- Verify data integrity for wireless transmission link

Data Recording and Playback

- Fully adjustable slider, play/pause/exit controls
- Subject and record ID, date, start time recorded in file naming convention

Most of the EMOTIV® technical characteristics are mentioned in the Table (4.1).

Table 4.1: Emotiv technical characteristics.

Technical characteristics	Description
Number of channels	14 (plus CMS/DRL references, P3/P4 locations)
Channel names (International 10-20 locations)	AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4
Sampling method	Sequential sampling. Single ADC
Sampling rate	128 SPS (2048 Hz internal)
Resolution	14 bits 1 LSB = $0.51\mu V$ (16 bit ADC, 2 bits instrumental noise floor discarded)
Bandwidth	$0.2 - 45$ Hz, digital notch filters at 50 Hz and 60Hz
Filtering	Built in digital 5th order Sinc filter
Dynamic range (input referred)	8400 μV (pp)
Coupling mode	AC coupled
Connectivity	Proprietary wireless, 2.4 GHz band
Power	LiPoly
Battery life (typical)	12 hours

4.4 Embedded instrumentation of DfNN classifiers

The original DfNN was employed for the generation of the hardware NNs presented in this work. There are two digital implementation and one analog. For the digital development first the VHDL code that describe the DfNN algorithm was firstly programmed in a FPGA and after verifying that it work properly, the same code was used to develop a VLSI design. For the analog section, operational amplifiers were used to recreate the DfNN algorithm, continuous voltages are employed as constants.

Nowadays, FPGAs have increase their speed and memory. This conditions allow them to process great amount of data in parallel making them ideal for NN implementation. For the FPGA-NN a Zedboard development board was chosen. This board was selected because of it hardware characteristic and because of the easy implementation of AXI-4 interfaces between Matlab and the board. Using Matlab 2014 and Xilinx 14.4 the VHDL code that describe the DfNN was generated. Then the VHDL code was programmed in the FPGA. After Simulink generates an interface that was directly placed on the board. From Simulink, the signals were

sent, they were processed by the DfNN inside the FPGA and the DfNN output was sent to the Simulink interface where signals were display.

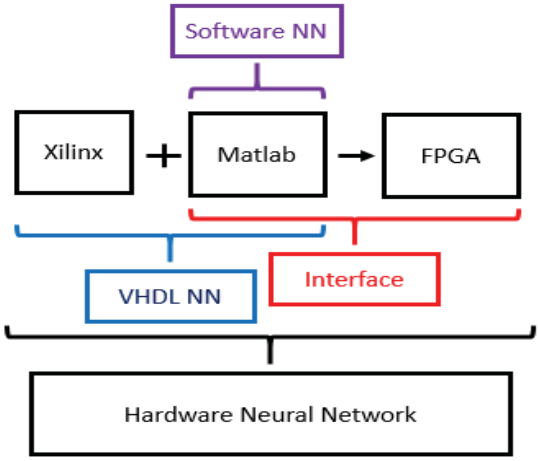

Figure 4-9: Implementation of the DfNN from the PC to the FPGA. Here Matlab and Xilinx are employed to develop the VHDL code that will run in the FPGA. Matlab also help to generate and interface between the FPGA and the PC.

For the VLSI-NN the basic structure of the VHDL code was employed, the parts of the code that implemented the AXI-4 interface was dismiss. Cadence system design tools the VLSI design allow the verification of the VLSI design also to move the position of the transistors in order to avoid holes or left space in the design (make it as compact as possible). The software also have tools that allow the user to see the VLSI design digital response.

Chapter 5

Results

Four NNs were tested in this thesis. The first one was a RNN that make a non-continuous classification of EEG signals. Then, the classification problem is addressed based on three different continuous NN. The brain response to an stimulus is by nature continuous and it can be infer that the more natural way to address its classification is by a NN that can learn in continuous time like a CNN. The consulted literature [12], [77], [69] made reference to the brain response as a plant with time delays, so a TDNN for classification was also tested. Finally, a CVNN was used for classification, by making a frequency analysis of the brain response. All the changes presented on the signal are more evident, this is reflected in the quality of information that is used for the learning laws of the CVNN.

5.1 Recurrent neural network

The RNN used the database I and two validation method (generalization-regularization and 5-fold cross validation) for measuring its classification accuracy. In the first method, 100 samples from each one of the 5 classes were taken. Considering these 100 samples per class, 60 samples were used for the training process, 30 for the generalization and the 10 for the independent test.

Figure (5-1) depicts the weights obtained after the training from the 60 samples for each class. It is clear that weights W_1 are separated accordingly to the class of signals used to train the RNN. The separation of the weights gotten at this point is important because W_1 characterized the information from each class of the EEG signal. In the bottom, right hand-side

in the image, the type of output that was used for the training and the RNN approximation is also depicted. On the left, the mean square error obtained after the training procedure demonstrates the RNN approximation capabilities.

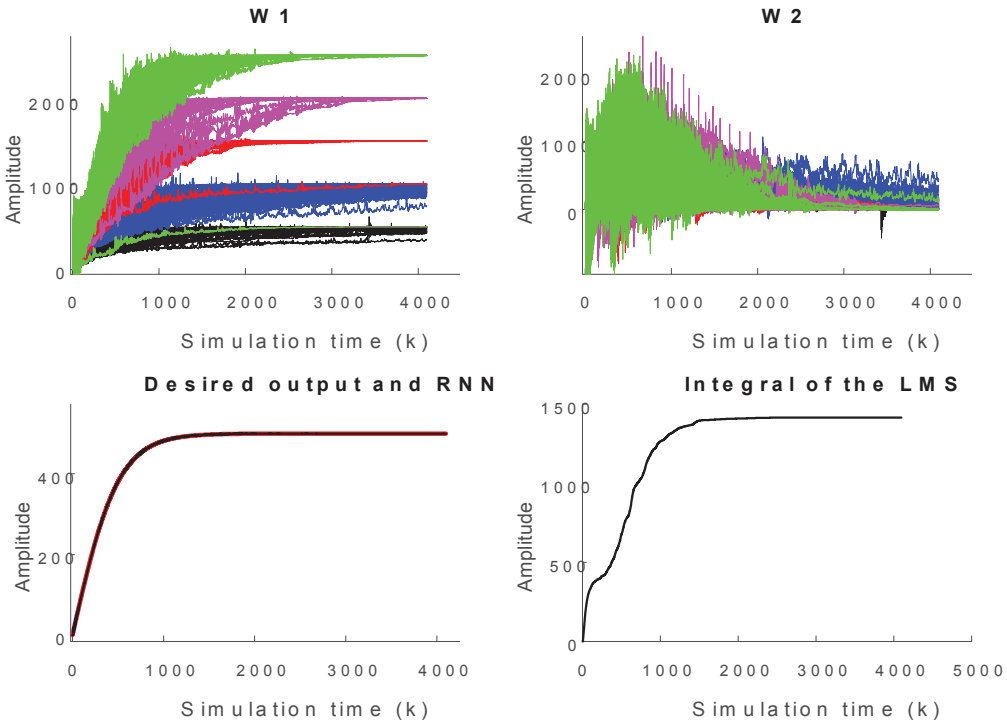

Figure 5-1: W_1 and W_2 corresponding to the RNN after training with the database 1. In both top images the black line correspond to the weights of the first class, the blue line to the weights of the second class, the red line correspond to the weights of the third class, the magenta line correspond to the weights of the class 4 and finally the green one to the class 5. The images of the bottom are from left to right the trajectories of the RNN Xe and X during the training process for a signal of class 5 and the integral of the LMS error obtained.

Table (5.1) contains the details on the results gotten from the generalization-regularization process in each class. A classification accuracy percentage of 94.21% was achieved.

Table (5.2) contains the results achieved from the 5-fold cross validation method applied to database I, a total classification accuracy of 97.48% was achieved.

Table 5.1: Results from the classification process per class for the RNN.

	C1	**C2**	**C3**	**C4**	**C5**
Samples	100	100	100	100	100
Training	100%	100%	100%	100%	100%
Generalization	93.33%	93.33%	90%	100%	96.6%
Independent Test	70%	100%	100%	70%	100%
CfA	87.77%	97.77%	96.66%	90%	98.86%

*CfA Correct Accuracy, according to the number of samples per section.

Table 5.2: DfNN results from the 5-fold cross validation process, that determine the classification accuracy.

	C1	**C2**	**C3**	**C4**	**C5**
Samples	100	100	100	100	100
1°S. CfA	94%	100%	90%	100%	100%
2°S. CfA	100%	94%	100%	92%	100%
3°S. CfA	100%	100%	100%	95%	100%
4°S. CfA	92%	100%	100%	100%	94%
5°S. CfA	100%	100%	92%	94%	100%
Total CfA	97.2%	98.8%	96.4%	96.2%	98.8%

*CfA Classification Accuracy.
*S. Segment.

5.2 Differential neural network

Employing the database downloaded from [92], the full DfNN validations scheme was performed using two validation methods methods. For the first method the training process employed a universe of 60 signals of the complete set of 100 that make up each class. As mentioned before, during the training process, we assigned a desired output value per class that was used to enforce the identifier trajectory. This value was unique for each class, specifically for this database, each class trajectory increased in 50 units of amplitude. Interestingly, after the training the same tendency can be seen in W_2. Figure $(5-2)$ shows how the weights evolved after the training process.

Once the training was finished, the weights were fixed by using the resulting W_1 and W_2 values for each class. The DfNN classifier works in the following way; the weights are fixed for each class and a parallel DfNN is built. For database I, the parallel DfNN is compound of 5 parallel DfNN. Each DfNN was designed to find exclusively one type of class as the one shown in Figure (5-3). To determine if a signal belongs to a class, the least mean square error of the

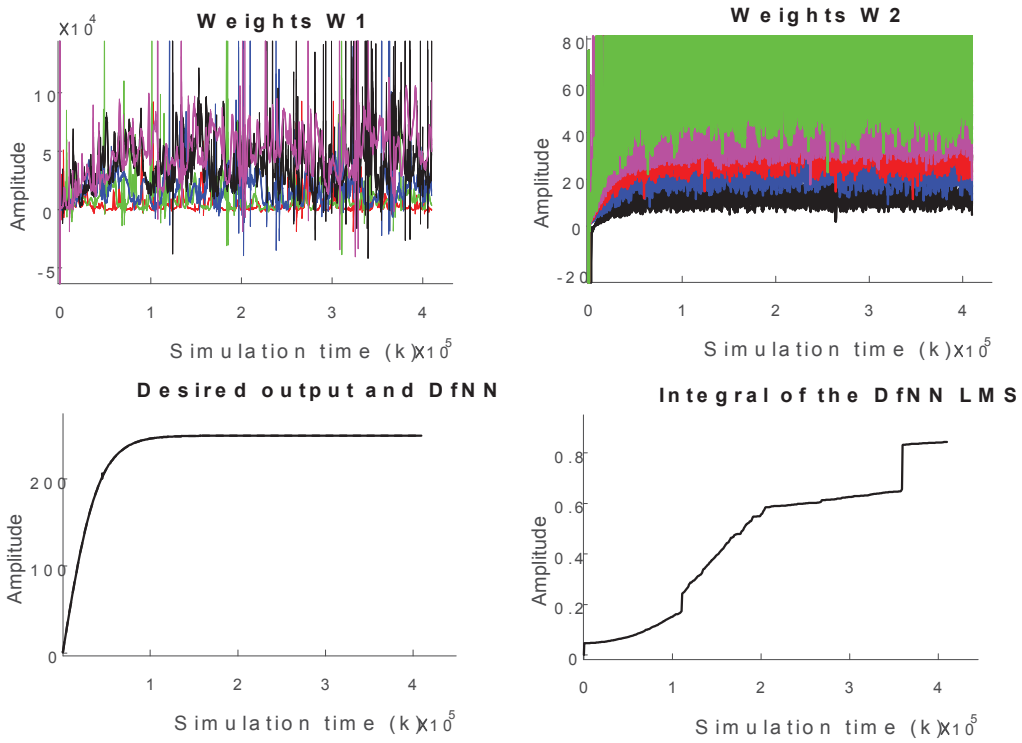

Figure 5-2: W_1 and W_2 corresponding to DfNN after training with the database 1. In both top images the black line correspond to the weights of the first class, the blue line to the weights of the second class, the red line correspond to the weights of the third class, the magenta line correspond to the weights of the class 4 and finally the green one to the class 5. The images of the bottom are from left to right the trajectories of the DfNN Xe and X during the training process for a signal of class 5 and the LMS error obtained.

5 neurons output was used.

The validation methods applied are generalization-regularization and a 5-fold cross validation. In the first one a generalization-regularization validation process, here only 30% of the signals that were used during the training part where employed for the generalization, and the remaining 10% is used for the independent test and validation of the DfNN. The results of this validation method are depicted in Table (5.3) . A classification accuracy of 94.88% was achieved from the parallel DfNN. All the results presented in this thesis work with the raw EEG signals.

The second validation process employed to measure the accuracy of this NN was a 5-fold cross validation method. The results of this process can be consulted in Table(5.4) .This method

Table 5.3: Results from the classification process per class for the DfNN.

	C1	**C2**	**C3**	**C4**	**C5**
Samples	100	100	100	100	100
Training	100%	100%	100%	100%	100%
Generalization	100%	53.33%	100%	100%	100%
Independent Test	100%	70%	100%	100%	100%
CfA	100%	74.44%	100%	100%	100%

*CfA Correct Accuracy, according to the number of samples per section.

Table 5.4: DfNN results from the 5-fold cross validation process, that determine the classification accuracy.

	C1	**C2**	**C3**	**C4**	**C5**
Samples	100	100	100	100	100
1°S. CfA	100%	100%	90%	85%	100%
2°S. CfA	100%	100%	95%	80%	100%
3°S. CfA	100%	100%	95%	90%	100%
4°S. CfA	100%	100%	100%	100%	100%
5°S. CfA	100%	100%	100%	100%	100%
Total CfA	100%	100%	96%	91%	100%

*CfA Classification Accuracy.
*S. Segment.

demonstrated a total classification accuracy of 97.4%.

The DfNN was also tested with the database II. The weights obtained after the training performed for the generalization - regularization process are depicted in Figure (5-4). Also a desired trajectory employed for the validation process and the DfNN obtained output.

Table 5.5: Results from the classification process per class for the DfNN database II.

	C1	**C2**	**C3**
Samples	30	30	30
Training	100%	100%	100%
Generalization	100%	100%	88.88%
Independent Test	100%	66.66%	66.66%
CfA	100%	88.88%	84.66%

*CfA Correct Accuracy, according to the number of samples per section.

The next step was to test the DfNN with the database II. The same three stages described above for the generalization-regularization validation method were also implemented in this case. Therefore, training, validation and testing procedures were executed with the same distribution of EEG signals. The results achieved by the DfNN in this part are interesting due to the

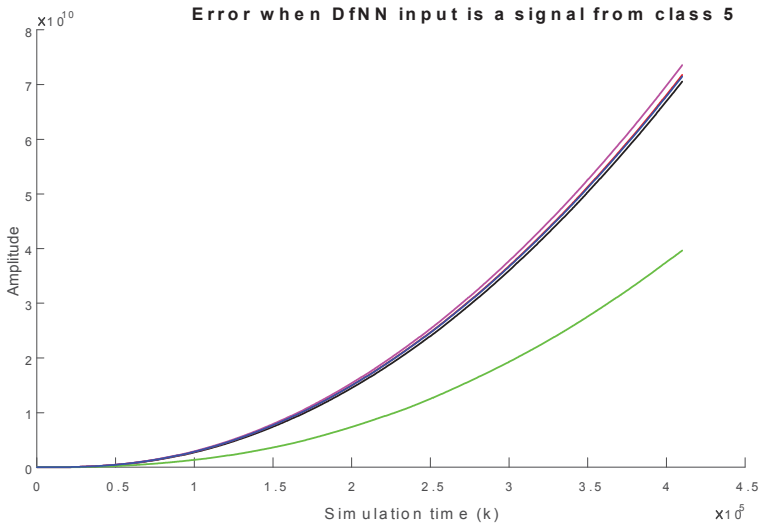

Figure 5-3: Integral of the error per class obtained from the DfNN when the input is a signal of class 5, the black line correspond to class 1, the blue line to class2, the red line to class 3, the magenta line to class 4 and finally the green one is class 5.

fact that the signals that make up this database II were obtained from an EMOTIV® device. Even though the signals quality was not as good as the one observed in the first database, the classification task was possible with remarkable results. The classification efficiency of DfNN for this database can be consulted in Table (5.5).

A problem when dealing with database II is the quality of the signal, due to the technology employed for the acquisition of the EEG signals their quality could be considered as poor. Nonetheless, well separated weights could be obtained for the three classes that make up this database as show in Figure $(5-4)$. This separation in weights values yields to a total classification accuracy of 91.24%.

Finally Figure $(5-5)$ describes the integral of the mean square error for a signal of class 2 when running in the parallel DfNN. It is clear that the DfNN is working correctly when the error of class 2 is the smaller of the three depicted in the above mentioned Figure.

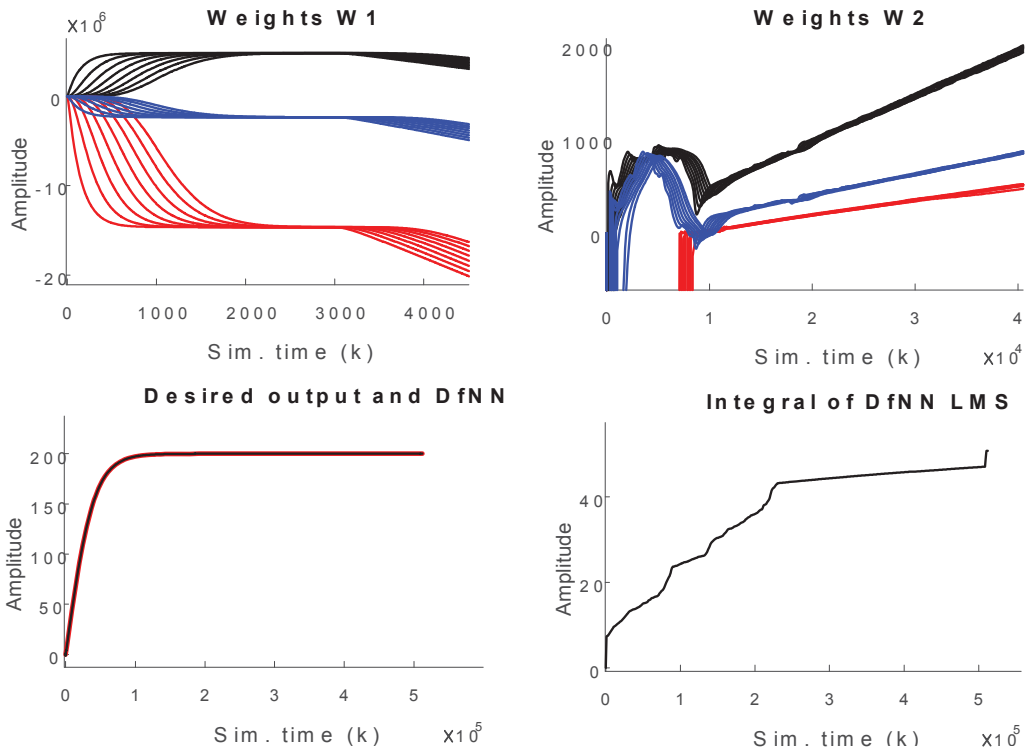

Figure 5-4: W_1 and W_2 corresponding to DfNN after training with the database 2. In both top images the black line correspond to the weights of the first class, the blue line to the weights of the second class, the red line correspond to the weights of the third class. The images of the bottom are from left to right the trajectories of the DfNN Xe and X during the training process for a signal of class 2 and the LMS error obtained.

5.3 Time delay neural networks

In order to test the classification capabilities of the proposed TDNN, 50 EEG signals of each class were taken from the database I. The classification procedure was executed following the generalization-regularization presented for the DfNN. The complete validation scheme was evaluated. In this particular case, 10 EEG signals from each class make up the validation set. The results achieved by this NN can be consulted in Table (5.6) .

The weights for W_1 and W_2 with its two delays obtained after the training process can be seen in Figure (5-6), the separation of the weights can be seen in W_2, W_2 with a delay equal to 1s and W_2 with a delay equal to 2s. Also the delay can be seen in the beginning of each W_2.

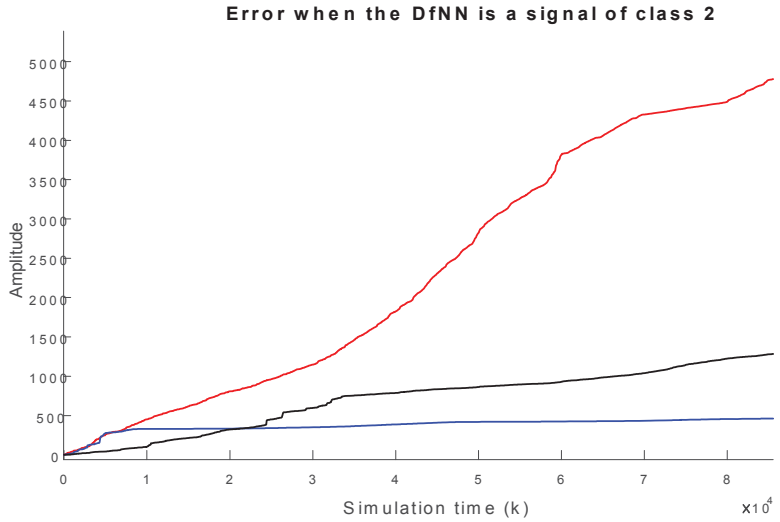

Figure 5-5: Integral of the error per class obtained from the DfNN when the input is a signal of class 2, the black line correspond to class 1, the blue line to class2, the red line to class 3.

Table 5.6: Results from the classification process per class for the TDNN.

	C1	**C2**	**C3**	**C4**	**C5**
Samples	50	50	50	50	50
Training	100%	100%	100%	100%	100%
Generalization	100%	100%	86.66%	100%	93.33%
Independent Test	80%	80%	100%	100%	100%
CfA	93.33%	93.33%	95.55%	100%	97.77%

*CfA Classification Accuracy, according to the number of samples per section.

To perform the training and validation procedures, the signals used to evaluate the TDNN classification properties were artificially fixed to 1000c where c is the number of class. This process is regular in signal processing algorithms in order to enforce an enchaced classification process as shown in Figure (5-7).

The final classification accuracy percentage achieved for this NN was 95.99%.

5.4 Complex valued differential neural network

The CVNN employed the half of samples included in database I. A set of 30 signals of each class was used for the training section, 15 signals of each class for the training and 5 signals of

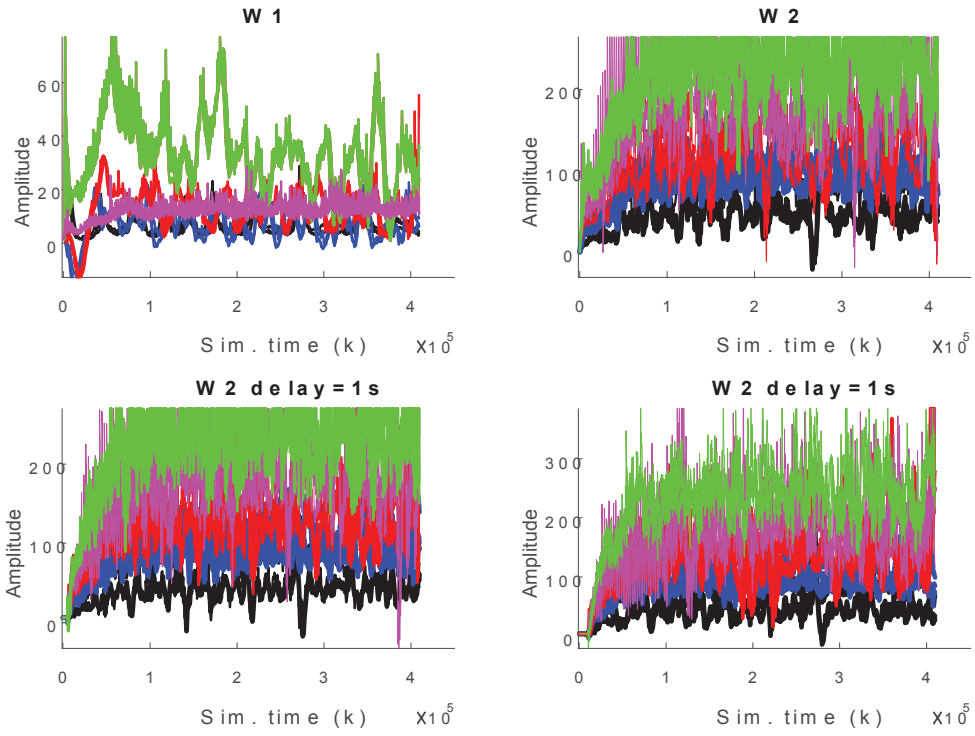

Figure 5-6: W_1 and W_2, W_2 with a delay equal to 1s and W_2 with a delay equal to 2s corresponding to TDNN after training with the database 1. In all the images, the black line correspond to the weights of the first class, the blue line to the weights of the second class, the red line correspond to the weights of the third class, the magenta line correspond to the weights of the class 4 and finally the green one to the class 5.

each class for the testing part. This NN proved to have the most complex topology from all the NN presented in this work. This complexity was reflected in the processing time needed to complete the evaluation of each stage in the classification method proposed in this thesis. Each sample took around 5 times more than the ones obtained with another NNs.

The weights W_1 and W_2 for this NN can be seen on the top of Figure (5-8). It is clear the separation of the weights in W_1. On the bottom left-hand side of the image, the comparison between desired trajectory for the real part and the CVNN real part approximation is detailed. On the right-hand side, the imaginary part of EEG signal transformed by FFT and the corresponding CVNN imaginary part approximation is also described.

For each part (real and imaginary), the integral LMS error was obtained as shown on the

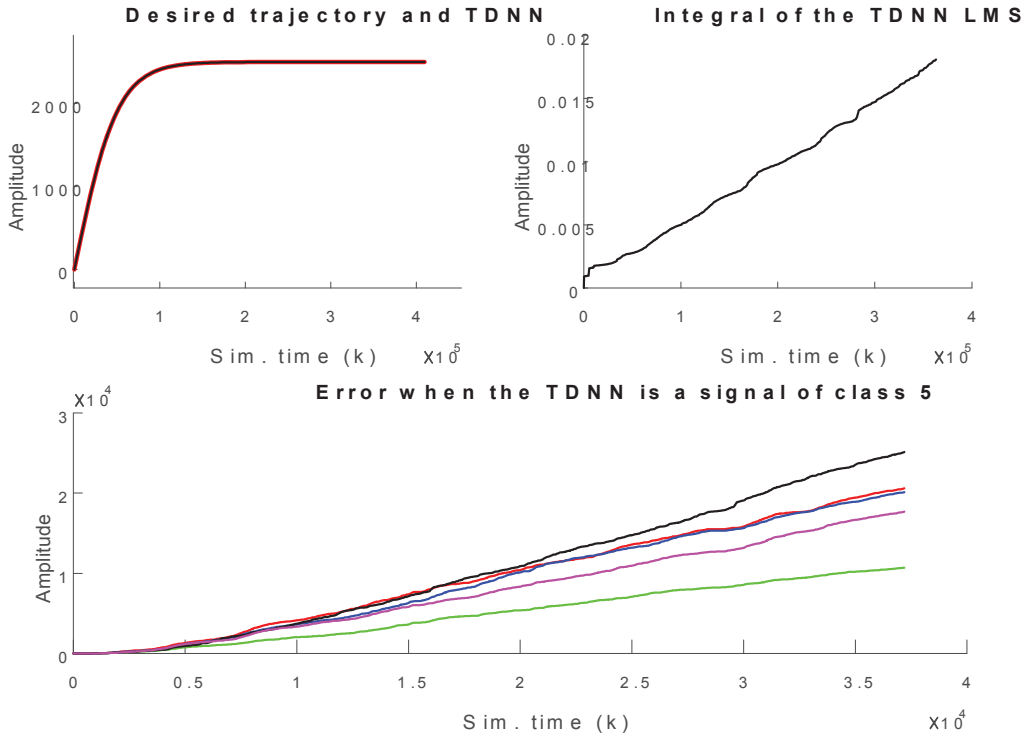

Figure 5-7: On the top left the desired trajectory and the TDNN output, on the top right the integral of the LMS error from the TDNN. On the lower part the error of the parallel TDNN when having as an input a class 5 signal. In the lower part the black line belongs to the class1, the blue line the class 2, the red line the class 3, the magenta line to class 4 and finally the green line to the class 5.

top of Figure (5-9). On the bottom of figure, the response of parallel CVNN is depicted. In this case, it is used a signal of class 5 as an input of the parallel NN. It can be seen that CVNN is classifying correctly this signal because the green line showed in figure that corresponds to the evaluation error is the one with less value. This NN achieved a total correct classification percentage of 92.43%.

The results achieved by CVNN can be consulted in Table (5.7). The table presents the scores obtained by the CVNN on each section of the generalization-regularization validation method.

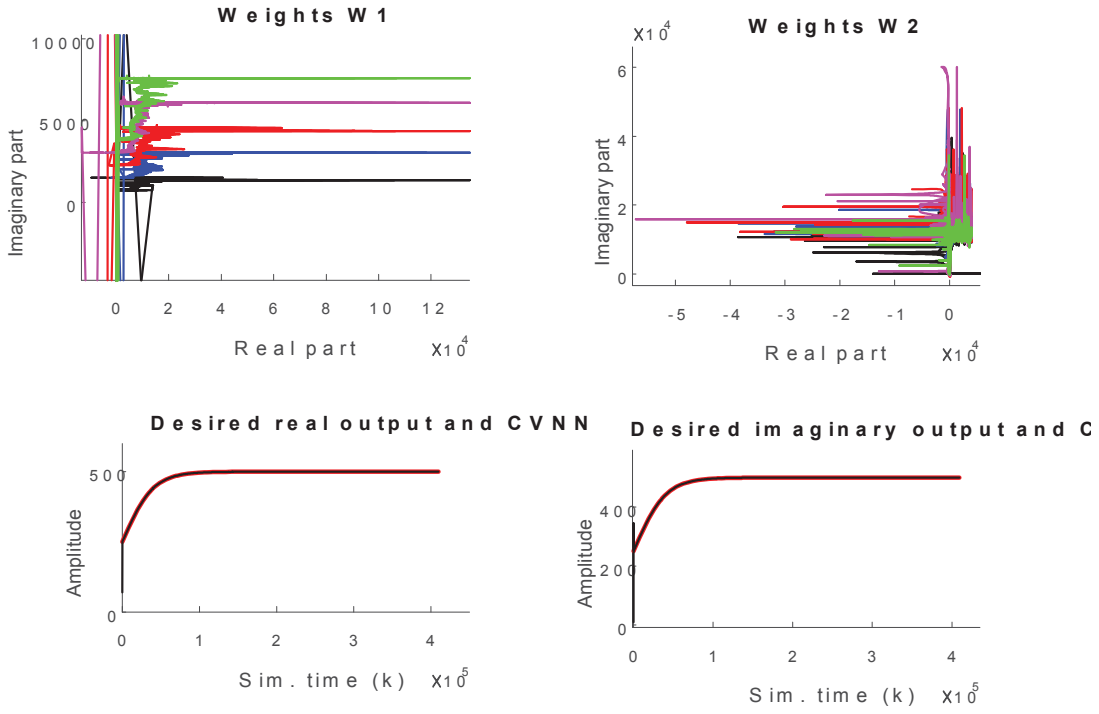

Figure 5-8: W_1 and W_2 corresponding to CVNN after training with the database 2. In both top images the black line correspond to the weights of the first class, the blue line to the weights of the second class, the red line correspond to the weights of the third class. The images of the bottom are from left to right the trajectories of the CVNN Xe and X during the training process the left image is the real part and the right one the imaginary one.

5.5 Neural networks in real time

This section present a summary of the results achieved during a visiting scholar period that took place at Johns Hopkins University from July 2014 to April 2015. At the Computational Sensory-Motor Systems Lab under the supervision of Dr. Ralph Etienne - Cummings. The purpose of this visiting was to develop a hardware approximation of the NNs that have been previously study in this thesis.

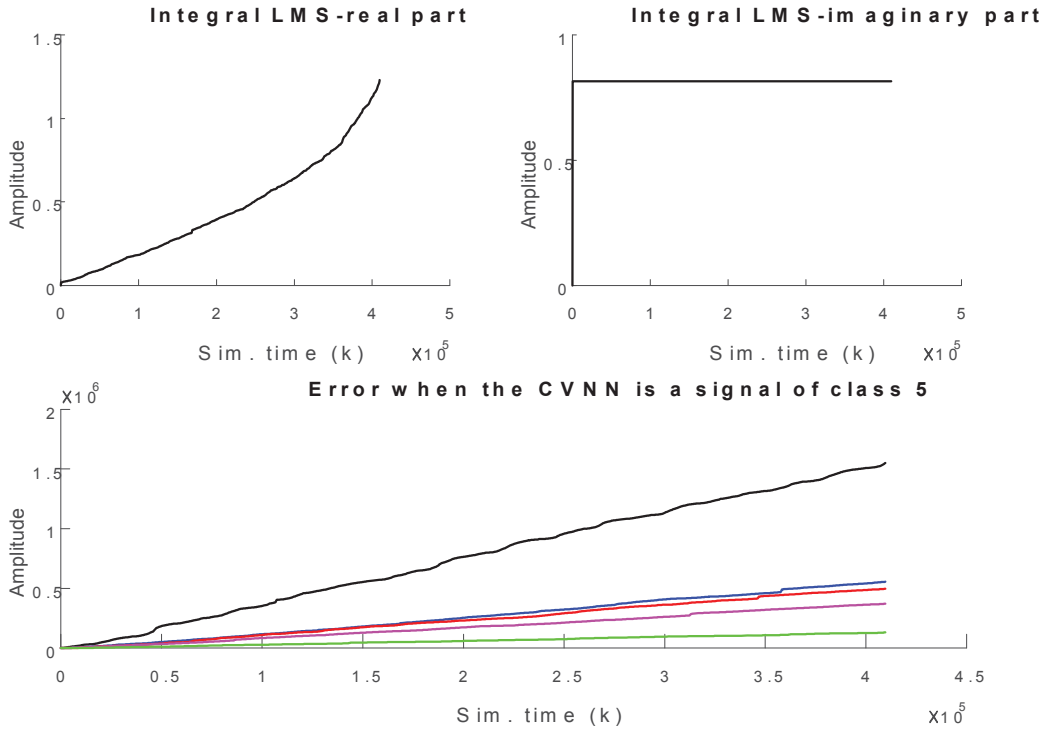

Figure 5-9: On the top integral of the LMS error from the real and imaginary part of the CVNN. On the lower part the error of the parallel CVNN when having as an input a class 5 signal. In the lower part the black line belongs to the class1, the blue line the class 2, the red line the class 3, the magenta line to class 4 and finally the green line to class 5.

5.5.1 FPGAs differential neural networks

The first approach to implement an embedded NN used the particular topology of the DfNN and described in the embedded FPGA system. Fixed point representation of both the identifier state as well as the identification error were developed. Table (5.8) shows the values employed to implement the low precision representation of the identifier.

The DfNN identifier was implemented on a Zedboard development board. This board is based on the Zynq^{TM}-7000 AP SoC xc7z020-CLG484-1 processor. Memory resources of this development board are $512MB$ DDR3, $256Mb$ Quad-SPI Flash, $4GB$ external SD card. Communication systems of this board are USB-JTAG programming, 10/100/1000 Ethernet, USB OTG 2.0 and USB-UART.

Table 5.7: Results from the classification process per class for the CVNN.

	C1	**C2**	**C3**	**C4**	**C5**
Samples	50	50	50	50	50
Training	100%	100%	100%	100%	100%
Generalization	86.66%	86.66%	86.66%	100%	86.66%
Independent Test	100%	80%	80%	80%	100%
CfA	95.55%	88.88%	88.88%	93.33%	95.55%

*CfA Classification Accuracy, according to the number of samples per section.

Table 5.8: DfNN I/O values for software and hardware.

I/O	**Floating P**	**Fixed P**
Sample	double	<±,16.11>
W1	double	<±,16.8>
W2	double	<±,16.8>
X	double	<±,16.12>
Xe	double	<±,16.12>
Delta	double	<±,16.12>

The programming language to implement the 16 bits fixed point representation of DfNN was VHDL. The corresponding code was generated using the Matlab HDL Coder 3.5 and ISE Xilinx 14.4. An auxiliary AXI-4 interface was used to perform the Ethernet communication between the computer and the development board. This method was implemented to follow the time evolution on-line of the DfNN based classifier. Table (5.9) shows the number of total registers used to implement the identifier.

Figure (5-10) shows the performance of the DfNN when working on the Zedboard, even thought working with the low resolution provided by the 16 bits the DfNN is able to estimate the desired output. The obtained weights W_1 and W_2 for this DfNN can be observed in Figure (5-10).

Table 5.9: Device utilization.

Type of register	**Used**	**Total**
Slice Registers	471	106400
Slice LUTS	50284	53200
Slice LUT-Flip Flop pairs	50291	53200

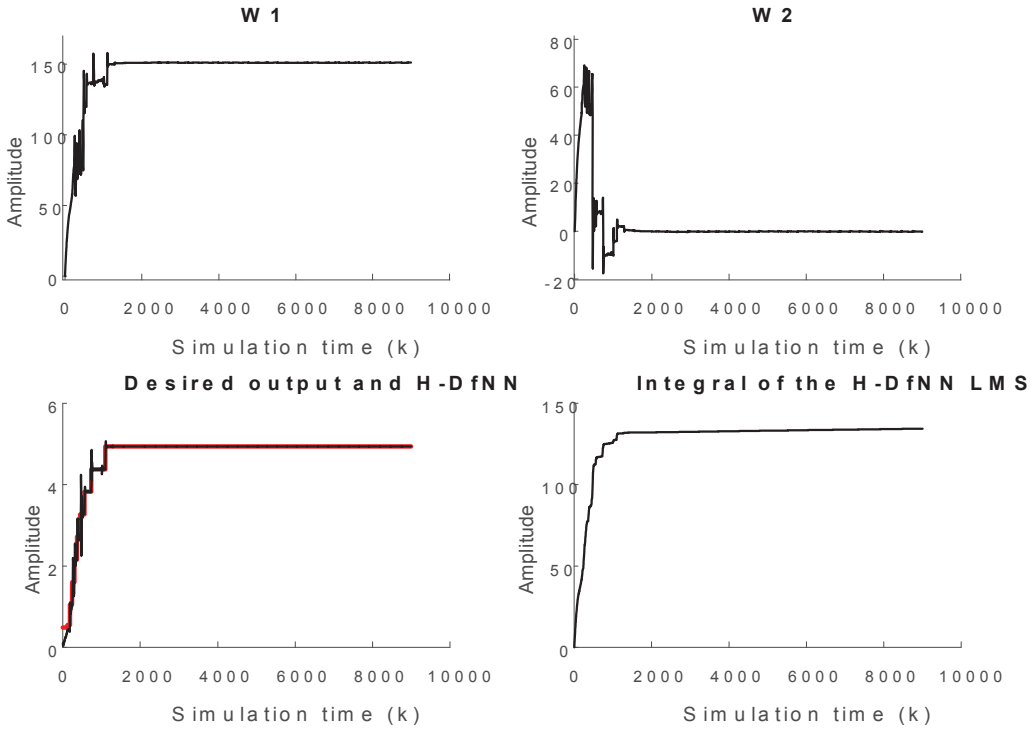

Figure 5-10: On the top W_1 and W_2 for the hardware implemented DfNN. On the bottom left the desired trajectory on slashed black line and the hardware DfNN output on solid black line. On the bottom right the integral of the LMS error obtained from the left bottom image.

5.5.2 VLSI differential neural networks

For the designing of the VLSI circuit, the VHDL code generated for the DfNN was firstly tested in order to check time issues in Cadence by generating a test-bench file and then synthesis in the design compiler called Sinopsys. Then, the corresponding compiled file was imported to Cadence Encounter to perform the Place and Routing procedures of this design. As a result, the DfNN VLSI design occupied an estimated area of 4.5x4.5 mm (9 tinychips). Figure $(5-11)$ depicts the full design of the DfNN. The final design was sent to MOSIS, a private company that produces final integrated devices. The IBM fabrication process was chosen for the design, the estimated time for fabrication is about 2 months according to the evaluation performed by MOSIS.

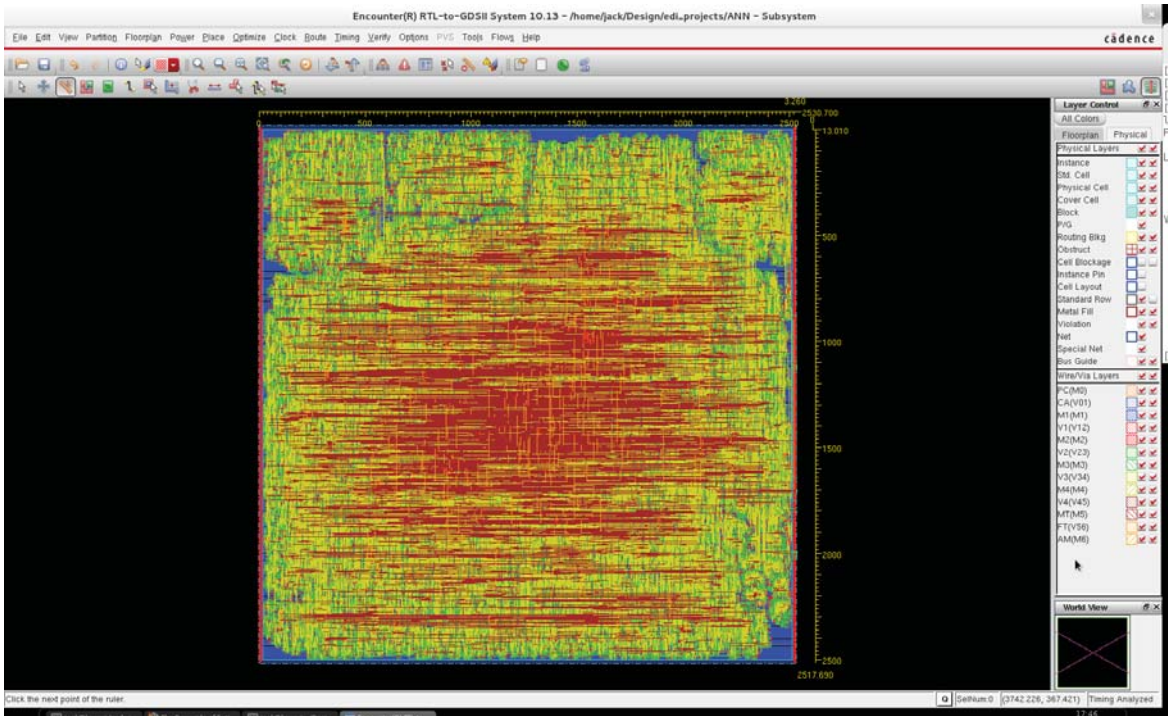

Figure 5-11: VLSI final design of the DfNN on Cadence.

5.5.3 Analog differential neural networks

This section describes the implementation of the DfNN identifier for a EKG signal in analog circuits. The EKG signal was artificially generated with a microcontroler and later converted to analog by using a digital to analog converter IC. In order to replicate the NN algorithms with analog circuits, the segments that make up the NN (activation functions, weights, constants, NN structure) were build in parts.

The first part to built were the NN activation functions (σ, Ψ). The equations S_1, S_2 and S_3 are the three slopes that make up the sigmoid like signal, this sigmoid generator design was taken from [37], for simulation purpose the values employed for σ were $R_1 = 1M\Omega$, $R_2 = 7M\Omega$, $R_3 = 1.5M\Omega$ and $R_F = 200\Omega$,and in the case of Ψ were $R_1 = 1M\Omega$, $R_2 = 7M\Omega$, $R_3 = 1.5M\Omega$ and $R_F = 196\Omega$. The circuit shown in Figure (5-12) is the sigmoid generator circuit with operational amplifiers.

$$S_1 = \frac{R_F}{2R_1} + \frac{R_F}{R_3}$$

$$S_2 = \frac{R_F}{2R_1} + \frac{R_F}{2R_2} + \frac{R_F}{R_3} \tag{5.1}$$

$$S_3 = \frac{R_F}{2R_2} + \frac{R_F}{R_3}$$

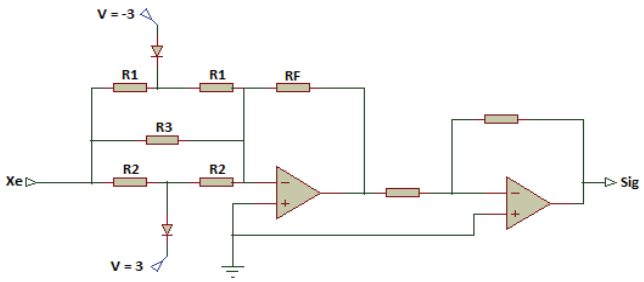

Figure 5-12: Activation fucntion of a DfNN builded with Opam's.

The circuit that describes W_1 as the following equation

$$\dot{W}_1(t) = -k_1 P \Delta(t) \sigma^\top + 2^{-1} k_1 \tilde{W}_1(t)$$

can be seen in Figure (5-13), in this circuit DC signals of value $ST = 3v, K_1 = 1.2v$ and $P = 5v$, the operational amplifier section of the circuit is the integrator with values of $R_1 = 8.5K\Omega$, $R_2 = 10K\Omega$, $C_1 = 22\mu f$ and finally a high input impedance buffer with values of $R_3 = 1K\Omega$ and $R_4 = 1K\Omega$.

The W_2 is described as

$$\dot{W}_2(t) = -k_2 P \Delta(t) u(t) \Psi^\top + 2^{-1} k_2 \tilde{W}_2(t)$$

where ST and P have the same value as in W_1 and K_2 is a DC signal of valued $1.8v$, u is the EKG signal. The next part of the circuit is the integrator with $R_1 = 7.5K\Omega$, $R_2 = 10K\Omega$, and $C_1 = 22\mu f$, the last section is the high input impedance buffer with values of $R_3 = 1K\Omega$ and $R_4 = 1K\Omega$. The circuit is depicted in Figure (5-14).

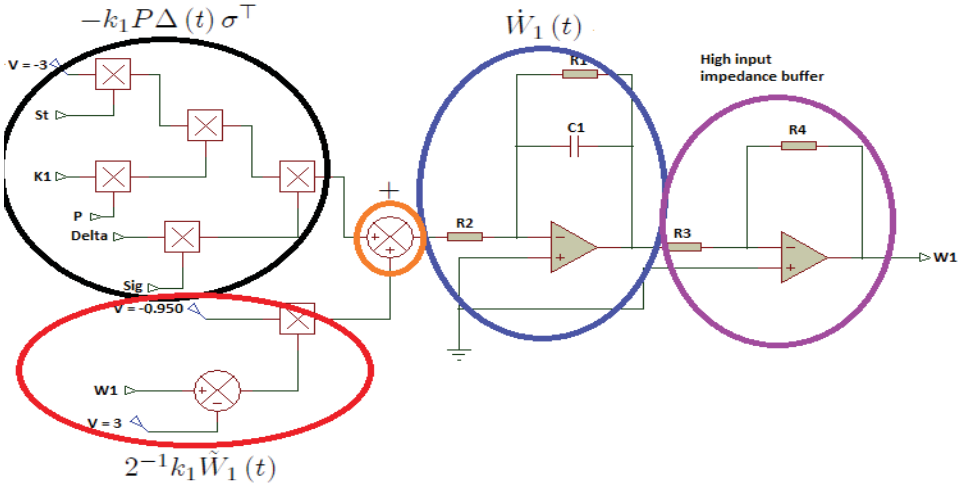

Figure 5-13: W_1 builded with Opam's.

The identifier based on NN employed is described by

$$\frac{d}{dt}\hat{x}(t) = A\hat{x}(t) + W_1(t)\,\sigma_1(\hat{x}(t)) + W_2(t)\Psi_2(\hat{x}(t))\,u(t) \tag{5.2}$$

The circuit that describe the equation (5.2) is shown in Figure (5-15), where $A = -1.2v$,the integrator values are $R_1 = 8K\Omega$, $R_2 = 10K\Omega$, and $C_1 = 33\mu f$, finally a high input impedance buffer with values of $R_3 = 1K\Omega$ and $R_4 = 1K\Omega$.

On the top left section of Figure (5-16) the resulting W_1 for this NN is depicted on the top right side is the resulting W_2. On the bottom left is an analog EKG simulated signal and the hardware NN approximation, a full beat takes around 1s, so it is clear the time that takes the HNN to approximate to the desired output, this time can be seen two in W_1 and W_2. On the left is the LMS error from the desired output and HNN approximation.

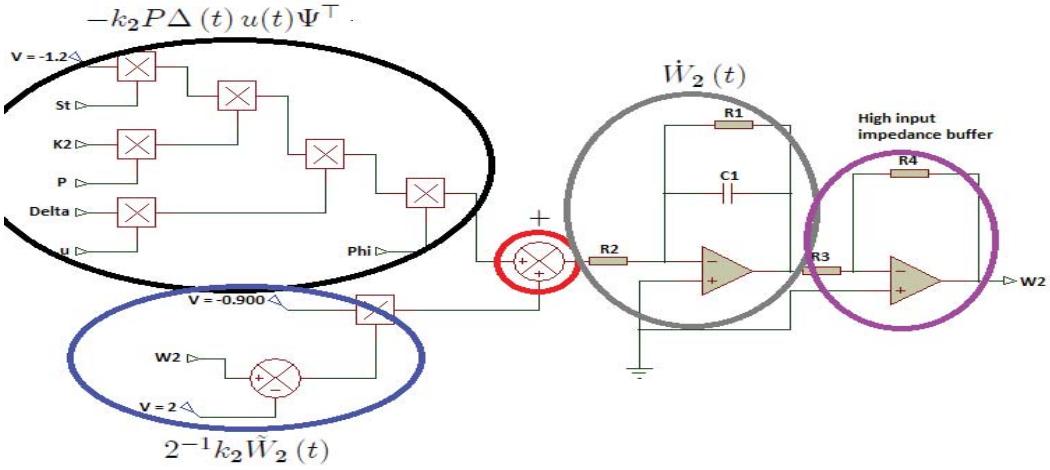

Figure 5-14: W_2 builded with Opam's.

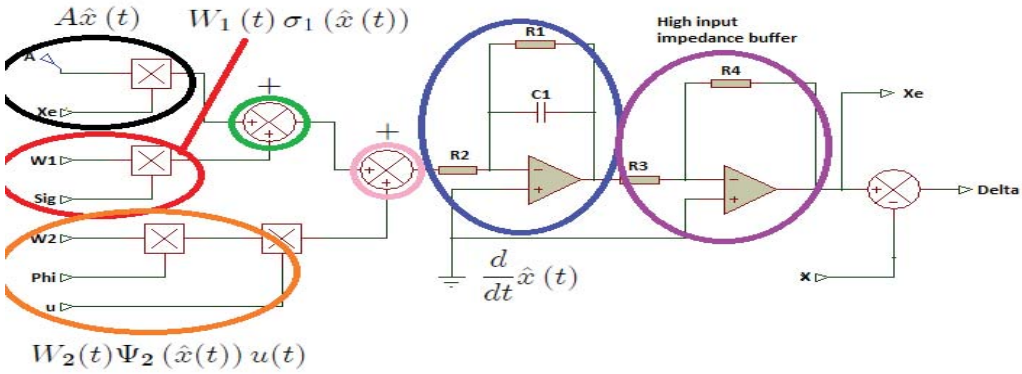

Figure 5-15: DfNN builded with Opam's.

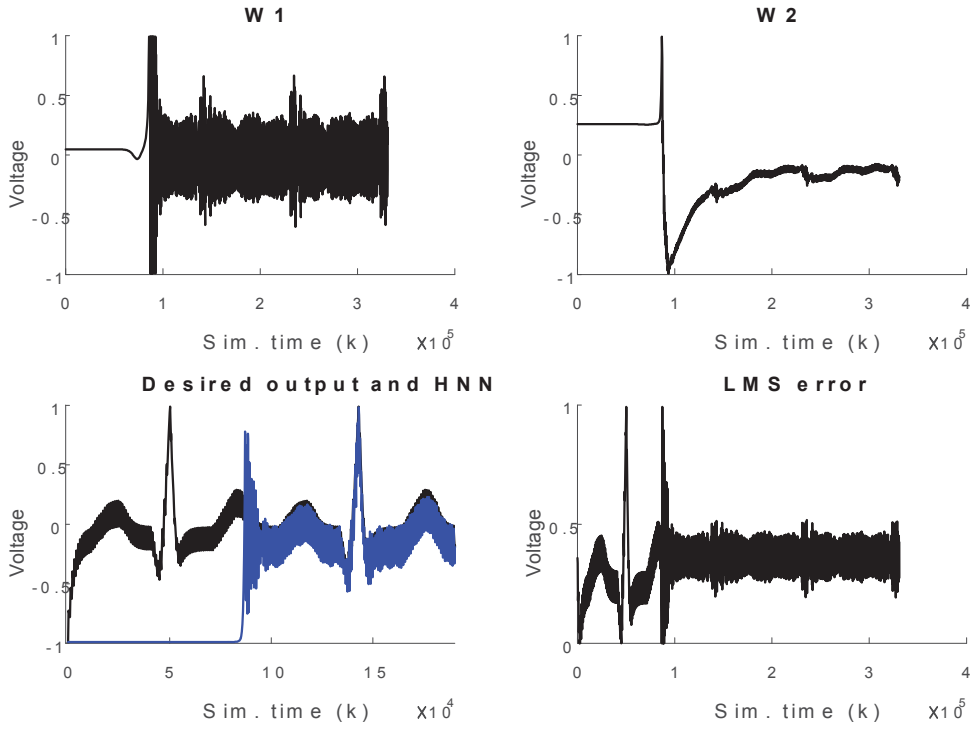

Figure 5-16: On the top left W_1 for the analog NN, on the top right it corresponding W_2, on the bottom left the desired output (slashed line) and the analog NN approximation (solid line). On the bottom left, the LMS error obtained from the signals depicted in the right bottom image.

Chapter 6

Published results & scholar activities

6.1 Articles published in journal included in the JCR

- M. Alfaro-Ponce, A. Argüelles and I. Chairez, "Adaptive Identifier for Uncertain Complex Nonlinear Systems Based on Continuous Neural Networks", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 3, pp. 483 - 494, 2013.

- M. Alfaro-Ponce, A. Argüelles and I. Chairez, "Continuous neural identifier for uncertain nonlinear systems with time delays in the input signal", *Neural Networks,* vol. 60, pp.53 -63, 2014.

- M. Alfaro-Ponce, I. Salgado, A. Argüelles and I. Chairez, "Adaptive identifier for uncertain complex-valued discrete-time nonlinear systems based on recurrent neural networks", *Neural Processing Letters,* Published online February 2015.

6.2 Articles published as extended manuscript in international conferences

- Alfaro M., Argüelles A.,Yañez C. and Chairez I., *Continuous Neural Networks for Electroencephalography Waveform Classification*, ANDESCON 2012, Cuenca, Ecuador, November 2012.

- Alfaro M., Argüelles A. and Chairez I., *Continuous neural identifier for uncertain nonlinear systems with time delays in the input signal*, IJCNN 2013, Dallas, USA, August 2013.

6.3 Workshops

- 2014 Telluride Workshop of Neuromorphic Cognition Engineering, 29 June 2014 to 19 July 2014.Telluride, Colorado, USA.

6.4 Visiting Scholar

- Johns Hopkins University, Computational Sensory-Motor Systems Lab., from 1 June 2014 to 28 April 2015, MD., USA.

Chapter 7

Conclusions & further work

In this thesis four topologies of NNs in software were addressed; RNN, DfNN, TDNN and a CVNN. All these different topologies of NNs showed classification capabilities over 91% when working with the raw EEG signal.

The first NN topology to be approach was the one based on RNNs to perform patterns recognition in raw EEG signals, the proposed algorithm classification accuracy was measured by two validation techniques employing the database I. The generalization-regularization method achieve a total classification accuracy of 94.21% on the other hand the 5-fold cross validation obtained a total classification accuracy of 97.48%.

The second topology to be addressed was a DfNN also the adjusting laws for the weight were develop to perform pattern recognition in raw EEG signals. For this algorithm two databases were employed. With the database I two validation process were applied, the generalization-regularization method accomplish a total classification accuracy of 94.89%, and the second validation was 5-fold cross validation with a total classification accuracy of 97.4%. Next, database II was validated only by the generalization-regularization method scoring a total classification accuracy of 91.24%.

Also a NN to perform patterns recognition based on the windowed response of raw EEG signals and time-delay theory was developed. This NN was only validated with half of the total EEG signals that make up database I and with the generalization-regularization method. From this NN a total classification accuracy of 95.99 was achieved.

The final NN topology software to be approach on this thesis was a DNN to perform patterns

recognition based on the frequency response of raw EEG signals. Database I was employed to validate this topology, achieving a total classification accuracy of 91.43%. In Table (2.1) different classification methods where applied to the database I employing preprocessing techniques to raw EEG signals achieving results between 90 to 99% of classification accuracy. The difference between these results and those presented in this theses is that they only work with two classes of the whole database and the results contained in this work were achieve from the 5 classes when working with the raw EEG signal.

The further work for the software NN will be to test the TDNN and CVNN with the full database I, this to verify if the correct classification scores keeps over 90%. In the particular case of database II, the idea is to achieve a 100 EEG signals per class in order to make it more reliable for the NN training and test. Until now the EEG signal that is employed as a NN input is just one, in database I is already described that is a just one electrode EEG, but for database II there are 14 possible electrodes however it is only use the one corresponding to the VEPs part of the brain. The next step would be to generate the four NN topologies that take into account the 14 EEG channels as an input.

All the software NNs that are deal in this works take a lot of software resources around 60% of the computer ram and the time to run a single EEG signal from database I that are 23.6 s signals, the RNN timing was about 5:25, the DfNN took around 10 min, the TDNN 30-40 min and the CVNN between 2-3 hours. All these times for the training which demands more computer resources, then when the weights are fixed the times get reduced.

Even whit this time reduction it is not possible to say that the NNs are able to classify the signal as soon at they end. There are two ways to deal whit this, first would be to work the software NN in a computer with more resources., the second to implement the hardware NN. The hardware implementation results that are present in this thesis are not applicable to the classification task, even thought the NN were implemented in hardware and demonstrated to work in soft real time.

The only topology that was put on hardware was the DfNN, but it was not tested with any of the two EEG databases. For the FPGA and the VLSI implementation, it was only to test the NN as a identifier, employing a EEG or EKG signal as an input and a desired output trajectory. In the FPGA the system show to work and being able to reproduce the desired

output trajectory also the interface that was generated allow the comparison between the same software NN and the hardware NN, it is remarkable the fact that even with low resolutions that have the FPGA signals the performance from both NN was almost the same, on the other hand the VLSI was only test in simulation due to time of fabrication of the IC.

For the analog implementation of the NN it did not work with EEG signals instead it work with an EKG signal, due to the limitation in time this section was just a simulation, in order demonstrated that the analog NN work correctly as a identifier for this signal. The work to do will be to implement the full classifier for EEG first in the FPGA, but this is still limited due to the size of the FPGA. Next the VLSI design of the full classifier scheme for EEG, finally to design the analog classifier in transistor level for its circuit fabrication. It would be ideal to also implement the other topologies in order to measure their correct classification.

References

[1] R. Aaslid and K. Lindegaard, *Transcranial Doppler Sonography.* Springer, 1986, ch. Cerebral Hemodynamics.

[2] R. Aaslid, T. Markwalder, and H. Nornes, "Noninvasive transcranial doppler ultra- sound recording of flow velocity in basal cerebral arteries," *Journal of Neurosurgy*, vol. 57, pp. 769–774, 1982.

[3] E. Adrian and B. Matthews, "The interpretation of potential waves in the cortex." *Journal of Physiology*, vol. 81(4), pp. 440–471, 1934.

[4] I. Aizenberg, *Complex-Valued Neural Networks with Multi-Valued Neurons*, I. Aizenberg, Ed. Springer Berlin Heidelberg, 2011, vol. 353.

[5] H. Al-Nashash, Y. Al-Assaf, J. Paul, and N. Thakor, "Eeg signal modeling using adaptive markov process amplitude," *IEEE Transaction on Biomedical Engeeniering*, vol. 51, no. 5, pp. 744–751, 2004.

[6] S. Arik, "An analysis of exponential stability of delayed neural networks with time varying delays," *Neural Networks*, vol. 17, pp. 1027–1031, 2004.

[7] S. Barbay, E. Plautz, K. Friel, S. Frost, N. Dancause, and A. Stowe, "Behavioral and neurophysiological effects of delayed training following a small ischemic infarct in primary motor cortex of squirrel monkeys," *Experimental Brain Research*, vol. 169(1), pp. 106–116, 2006.

[8] A. Bashashati, M. Fatourechi, R. Ward, and G. Birch, "A survey of signal processing algorithms in brainŰcomputer interfaces based on electrical brain signals," *Journal of Neural Engineering*, vol. 4, no. 2, pp. R32–R57, June 2007.

[9] J. Basu, D. Bhattacharyya, and T. Kim, "Use of artificial neural network in pattern recognition," *International Journal of Software Engineering and Its Applications ts Applications*, vol. 4, no. 2, pp. 23–34, Abr 2010.

[10] J. Bechennec, C. Chanussot, V. Neri, and D. Etiemble, *VLSI for Artificial Intelligence and Neural Networks*, J. Delgado-Frias and M. Moore, Eds. Springer, 1991.

[11] N. Benvenuto and F. Piazza, "On the complex backpropagation algorithm," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 967–969, 1992.

[12] F. Bocharova, G. Rihan, "Numerical modelling in biosciences using delay differential equations," *Journal of Computational and Applied Mathematics*, vol. 125, pp. 183–199, 2000.

[13] N. Bose and P. Liang, *Neural Network Fundamentals with Graphs, Algorithms and Applications*, N. Bose and P. Liang, Eds. Tata McGraw-Hill, 1996.

[14] I. Chairez, "Wavelet differential neural network," *IEEE Transactions on Neural Networks*, vol. 20, pp. 1439–1449, 2009.

[15] A. Chatterjee, A. Nait-Ali, and P. Siarry, *Advanced Biosignal Processing.* Springer-Verlag Berlin Heidelberg, 2009, ch. Chapter 8 Neural Network Approaches for EEG Classification, pp. 165–182.

[16] T.-H. Chueh and H. Lu, "Inference of biological pathway from gene expression profiles by time delay boolean networks," *PLoS ONE*, vol. 7(8), p. e42095, 2012.

[17] E. Commission, "Human brain project," https://www.humanbrainproject.eu/.

[18] I. E. Commission, "The medical device developer's guide to iec 60601," 1994.

[19] S. Coyle, T. Ward, and C. Markham, "Brain-computer interface using a simplified functional near-infrared spectroscopy system," *J Neural Eng.*, vol. 4, pp. 219–226, 2007.

[20] N. Crone, L. Hao, J. Hart, D. Boatman, and R. Lesser, "Electrocorticographic gamma activity during word production in spoken and sign language," *Neurology*, vol. 57, pp. 2045–2053, 2001.

[21] G. Darbellay and I. Vajda, "Estimation of the information by an adaptive partitioning of the observation space," *IEEE Transaction on Informatic Theory*, vol. 45, pp. 1315–1321, 1999.

[22] L. Dewitt and L. Wechsler, "Transcranial doppler," *Stroke*, vol. 19, pp. 915–921, 1988.

[23] E. Donchin, K. Spencer, and R. Wijesinghe, "The mental prosthesis: Assessing the speed of a p300-based brain-computer interface," *IEEE Transaction on Rehabilitation Engineering*, vol. 8 (2), pp. 174–179, 2000.

[24] L. Dongha, P. Bumhee, J. Changwon, and H.-J. P., "Decoding brain states using functional magnetic resonance imagine," *Biomed Eng Lett*, vol. 1, pp. 82–88, 2011.

[25] P. Duhamel and M. Vetterli, "Fast fourier transforms: a tutorial review and a state of the art," *Signal Process*, vol. 19, pp. 259–299, 1990.

[26] Emotiv, "Eeg specifications," EMOTIV, Tech. Rep., 2012.

[27] O. Faydasicok and S. Arik, "An approach stability criterion fo dynamical neural networks with multiple time delays," *Neurocomputing*, vol. 99, pp. 290–297, 2013.

[28] W. Freeman, B. Burke, and M. Holmes, "A periodic phase resetting in scalp eeg of beta-gamma oscillations by state transitions at alpha-tetha rates," *Human Brain Mapping*, vol. 19, pp. 248–272, 2003.

[29] E. Fridman, "A refined input delay approach to sampled-data control," *Automatica*, vol. 46, no. 2, pp. 421–427, 2010.

[30] E. Fridman, A. Seuret, and J.-P. Richard, "Robust sampled-data stabilization of linear systems: an input delay approach," *Automatica*, vol. 40, no. 8, pp. 1441–1446, 2004.

[31] S. Frost, S. Barbay, K. Friel, E. Plautz, and R. Nudo, "Reorganization of remote cortical regions after ischemic brain injury: A potential substrate for stroke recovery," *Journal of Neurophysiology*, vol. 89(6), pp. 3205–3214, 2003.

[32] C. Gilbert, "Adult visual cortical plasticity," *Neuron*, vol. 75, no. 2, pp. 250–264, 2012.

[33] B. Girau, *FPGA IImplementation of Neural Networks*, A. Omondi and J. Rajapakse, Eds. Springer, 2006.

[34] V. Goel, A. Brambrink, A. Baykal, R. Koehler, D. Hanley, and N. Thakor, "Dominant frequency analysis of eeg reveals brainŠs response during injury and recovery," *IEEE Transaction on Biomedical Engeeniering*, vol. 43, pp. 1083–1092, 1996.

[35] J. Gotman and L. Wang, "State-dependent spike detection: Concepts and preliminary results," *Electroencephalography and Clinical Neurophysiology*, vol. 79, pp. 11–19, 1991.

[36] S. Gourley, Y. Kuang, and J. Nagy, "Dynamics of a delay differential equation model of hepatits b virus infection," *Journal of biological Dynamics*, vol. 2(2), pp. 140–153, 2008.

[37] G. Graeme and G. Tobey, *Operational Amplifiers Design and Operation*, G. Graeme, G. Tobey, and P. Laurence, Eds. Mc-Graw Hill, 1971.

[38] F. Güler, E. Ubeyli, and I. Güler, "Recurrent neural networks employing lyapunov exponents for eeg signals classification," *Expert Systems with Applications*, vol. 29, pp. 506–514, 2005.

[39] J. K. Hale, *Theory of functional differential equations*. Springer Verlag Gmbh, 1977.

[40] M. Hassoun, *Fundamentals of Artificial Neural Networks*, M. Hassoun, Ed. MIT Press, 1995.

[41] S. Hauck, T. Fry, M. Hosler, and J. Kao, "The chimaera reconfigurable functional unit," in *Proceedings of the IEEE Symposium on Field-Programmable Custom Computing Machines (FCCMŠ97)*, April 1997, pp. 87–96.

[42] S. Haykin, *Neural NetworksŮA Comprehensive Foundation*, N. Jersey, Ed. Prentice - Hall, 1999.

[43] K. Herholz, S. Carter, and M. Jones, "Positron emission tomography imaging in dementia," *Bristish Journal of Radiology*, vol. 80, no. 2, pp. 160–167, February 2014.

[44] D. Hernandez, S. Ledesma, R. Martinez, G. Avia, and G. Canedo, "Using complex neural networks on noise cancelling," in *MICAI '08. Seventh Mexican International Conference on Artificial Intelligence*, 2008, pp. 138–142.

[45] H. Hinrichs, *Biomedical Technology and Devices Handbook*. CRC Press, 2004, ch. 6 Electroencephalography, pp. 153–173.

[46] ——, *Biomedical Technology and Devices Handbook*. CRC Press, 2004, ch. 5 Evoked Potentials, pp. 136–152.

[47] R. Homan, J. Herman, and P. Purdy, "Cerebral location of international 10-20 system electrode placement," *Electroencephalogr Clin Neurophysiol*, vol. 66, pp. 376–382, 1987.

[48] http://www.chinamedevice.com/Suppliers/5180/TranscranialDopplerwithRoboticProbeEMS9UA455353.html, "Transcranial doppler with robotic probe (ems-9ua)," Webpage (http://www.chinamedevice.com/Suppliers/5180/TranscranialDopplerwithRoboticProbeEMS9UA-455353.html), December 2012.

[49] H. Hwang, S. Kim, S. Choi, and I. CH., "Eeg-based brain-computer interfaces: A thorough literature survey," *International Journal of Brain Computer Interaction*, vol. 29, no. 12, pp. 814–826, Sep 2013.

[50] M. Inc., "Improve neural network generalization and avoid overfitting," Software help, May 2015.

[51] L. Jeong-Woo and O. Jun-Hoo, "Time delay control of non linear systems with neural networks modeling," *Mechatronics*, vol. 7 no.7, pp. 613–640, 1997.

[52] F. Jobsis, "Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters," *Science*, vol. 198, pp. 1264–1267, 1977.

[53] T. Jung, S. Makeig, C. Humphries, T. Lee, and M. McKeown, "Removing electroencephalographic artifacts by blind source separation," *Physchophysiology*, vol. 37, pp. 163–178, 2000.

[54] T. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and t. Sejnowski, "Removal of eye activity artiartifacts from visual event-related potentials in normal and clinical subjects," *Clinical Neurophysiology*, vol. 111, no. 10, pp. 1745–1758, 2000.

[55] J. Kaiser, "On a simple algorithm to calculate the ŚenergyŠ of a signal," in *Proc. Int. Conf. Acoustics*, 1990, pp. 381–84.

[56] Y. Kamitani and F. Tong, "Decoding the visual and subjective contents of the human brain," *Nature Neuroscience*, vol. 8, pp. 679–685, 2005.

[57] M. Kannathala, N.and Choob, U. Rajendra-Acharyab, and P. Sadasivana, "Entropies for detection of epilepsy in eeg," *Computer Methods and Programs in Biomedicine*, vol. 2005, no. 80, pp. 187–194, June 2005.

[58] N. Kannathala, U. Rajendra-Acharyab, C. Limb, and P. Sadasivana, "Characterization of eeg-a comparative study," *Computer Methods and Programs in Biomedicine*, vol. 2005, no. 80, pp. 17–23, June 2005.

[59] A. Khazaeea and A. Ebrahimzadeha, "Heart arrhythmia detection using support vector machines," *Intelligent Automation and Soft Computing*, vol. 19, no. 1, pp. 1–9, Mar 2013.

[60] J. Kleim, "Neural plasticity and neurorehabilitation: Teaching the new brain old tricks," *Journal of Communication Disorders*, vol. 44, pp. 521–528, 2011.

[61] M. Knikou, "Neural control of locomotion and training - induced plasticity after spinal and cerebral lesions," *Clinical Neurophysiology*, vol. 121, pp. 1655–1668, 2010.

[62] B. Kolb, A. Muhammad, and R. Gibb, "Searching for factors underlying cerebral plasticity in the normal and injured brain," *Journal of Communication Disorders*, vol. 44, pp. 503–514, 2011.

[63] X. Kong, A. Brambrink, D. Hanley, and N. Thakor, "Quantification of injury-related eeg signal changes using distance measures," *IEEE Transaction on Biomedical Engeeniering*, vol. 43, pp. 189–197, 1999.

[64] S. Koyama, S. Chase, A. Whitford, A. Velliste, A. Schwartz, and R. Kass, "Comparison of brainŰcomputer interface decoding algorithms in open-loop and closed-loop control," *Journal of Computational Neuroscience*, vol. 29, no. 1-2, pp. 73–87, Nov 2009.

[65] A. Kruszewski, W. J. Jiang, E. Fridman, J.-P. Richard, and A. Toguyeni, "A switched system approach to exponential stabilization through communication network," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 4, pp. 887–900, 2012.

[66] V. Latora, M. Baranger, A. Rapisarda, and C. Tsallis, "The rate of entopy increase at the edge of chaos," *Physical Letters A*, vol. 273, pp. 97–103, 2000.

[67] N. Lazar, B. Luna, J. Sweeney, and W. Eddy, "Combining brains: A survey of methods for statistical pooling of information," *NeuroImage*, vol. 16, no. 2, pp. 538–550, June 2002.

[68] Y. Lee, Y. Zhu, Y. Xu, M. Shen, H. Zhang, and N. Thakor, "Detection of nonlinearity in the eeg of schizophrenic patients," *Clinical Neurophysiology*, vol. 112, pp. 1288–1294, 2001.

[69] X. Liaoa, G. Chenb, and E. Sanchez, "Delay dependent exponential stability analysis of delayed neural networks: an lmi approach," *Neural Networks*, vol. 15, pp. 855–866, 2002.

[70] H. Lin, H. Wensheng, Z. Xiaolin, and C. P., "Recognition of ecg patterns using artificial neural network," in *Sixth International Conference on Intelligent Systems Design and Applications*, vol. 02, 2006.

[71] C. Lina and M. Hsieh, "Classification of mental task from eeg data using neural networks based on particle swarm optimization," *Neurocomputing*, vol. 72(4-6), pp. 1121–1130, 2009.

[72] Z. Liu and L. Liao, "Existence and global exponential stability of periodic solution of cellular neural networks with time varying delays," *Journal of Mathematical Analysis and Applications*, vol. 290, pp. 247–262, 2004.

[73] L. Lu, J. Shin, and M. Ichikawa, "Massively parallel classification of single- trial eeg signals using a min-max modular neural network," *IEEE Transactions on Biomedical Engenieering*, vol. 24, pp. 551–558, 2004.

[74] C. Marcus and R. Westervelt, "Stability of analogy neural networks with delay," *Physics Reviews A*, vol. 39, pp. 347–359, 1989.

[75] J. MarquesDeSá, *Pattern recognition: concepts, methods and applications: with 197 figures*, J. MarquesDeSá, Ed. Springer, 2001.

[76] L. Mier-y Terán-Romero, M. Silber, and V. Hatzimanikatis, "The origins of time-delay in template biopolymerization processes," *PLoS Computational Biology*, vol. 6(4), p. e1000726, 2010.

[77] L. MieryTeran, M. Silber, and V. Hatzimanikatis, "The origins of time-delay in template biopolymerization processes," *PLoS Comput Biol*, vol. 6(4), p. e1000726, 2010.

[78] J. Millan and J. Mourino, "Asynchronous bci and local neural classifiers: An overview of the adaptive brain interface project," *IEEE Transactions on Neural Systems and Rehabilitation Engineerings*, vol. june, pp. 159–161, 2003.

[79] M. Miyazaki and K. Kato, "Measurement of cerebral blood flow by ultrasonic doppler tech- nique; hemodynamic comparison of right and left carotid artery in patients with hemiplegia," *Japanese Circulation Journal*, vol. 29, pp. 383–386, 1965.

[80] A. Murari, D. Mazon, N. Martin, G. Vagliasindi, and M. Gelfusa, "Exploratory data analysis techniques to determine the dimensionality of complex nonlinear phenomena: The l-to-h transition at jet as a case study," *IEEE Transactions on Plasma Science*, vol. 40, no. 5, pp. 1386–1394, 2012.

[81] J. Muthuswamy, D. Sherman, and N. Thakor, "Higher-order spectral analysis of burst patterns in eeg," *IEEE Transactions on Biomedical Engenieering*, vol. 46, pp. 92–99, 1999.

[82] T. N. Muthuswamy J, "Spectral analysis methods for neurological signals," *Journal of Neuroscience Methods*, vol. 83, pp. 1–14, 1998.

[83] A. Myrden, A. Kushki, E. Sejdic, A. Guerguerian, and T. Chau, "A brain-computer interface based on bilateral transcranial doppler ultrasound," *PLOS ONE*, vol. 6(9), p. e24170, September 2011.

[84] T. Nagaoka, K. Sakatani, T. Awano, N. Yokose, T. Hoshino, Y. Murata, Y. Katayama, A. Ishikawa, and H. Eda, *Oxygen Transport to Tissue XXXI, Advances in Experimental Medicine and Biology*. Springer ScienceþBusiness Media, 2010, ch. Development of a New Rehabilitation System Based on a Brain-Computer Interface Using Near-Infrared Spectroscopy.

[85] A. Nait-Ali and P. Karasinski, *Advanced Biosignal Processing*. Springer-Verlag Berlin Heidelberg, 2009, ch. 1 Biosignals: Acquisition and General Properties, pp. 1–13.

[86] E. Niedermeyer, *Electroencephalophy*. Baltimore- Munich, 1987, ch. The normal EEG of the waking adult, pp. 99–117.

[87] E. Niedermeyer and F. LopesDaSilva, *electroencephalography: basic principles, clinical applications and related fields*, M. Sydor, Ed. Lippincott Williams & Wikins, 2005.

[88] C. Nikias and J. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Processing Magazine*, vol. 10, pp. 10–37, 1993.

[89] R. Nudo, "Neural bases of recovery after brain injury," *Journal of Communication Disorders*, vol. 44, pp. 515–520, 2011.

[90] R. Nudo, G. Milliken, W. Jenkins, and M. Merzenich, "Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys," *Journal of Neuroscience*, vol. 16(2), pp. 785–807, 1996.

[91] H. Ocak, "Automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy," *Expert Systems with Applications*, vol. 36, pp. 2027–2036, 2009.

[92] S. P. P. F. U. of Freiburg, "Eeg database," online, may 2012, http://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/eeg-database.

[93] B. Park, W. Pedrycz, and S. Oh, "Polynomial-based radial basis function neural networks (p-rbf nns) and their application to pattern classification," *Applied Intelligence*, vol. 32, pp. 27–46, 2010.

[94] S. Pincus, "Approximate entropy (apen) as a complexity measure," *Chaos*, vol. 5, pp. 110–117, 1995.

[95] K. Polat and S. Günes, "Classification of epileptiform eeg using a hybrid system based on decision tree classifier and fast fourier transform," *Applied Mathematics and Computation*, vol. 87, p. 1017Ű1026, 2007.

[96] J. Polich, "Updating p300: An integrative theory of p3a and p3b," *Clinical Neurophysiology*, vol. 118, pp. 2128–2148, 2007.

[97] A. Polyakov, "Minimization of disturbances effects in time delay predictor-based sliding mode control systems," *Journal of The Franklin Institute*, vol. 394, no. 4, pp. 1380–1396, 2012.

[98] M. Poulos, M. Rangoussi, N. Alexandris, and A. Evangelou, "On the use of eeg features towards person identification via neural networks," *Informatics for Health and Social Care*, vol. 26(1), pp. 35–48, 2001.

[99] A. Poznyak, E. Sanchez, and W. Yu, *Differential Neural Networks for Robust Nonlinear Control (Identification, State Estimation an Trajectory Tracking.* World Scietific, 2001.

[100] A. Poznyak, E. Sánchez, and W. Yu, *Differential Neural Networks for Robust Nonlinear Control (Identification, state Estimation and trajectory Tracking).* Worl Scientific, 2001.

[101] A. S. Poznyak and J. Medel, "Matrix forgetting factor," *Int. Journal of Systems Science*, vol. 30, no. 2, pp. 165–174, 1999.

[102] N. Pradhan, P. Sadasivan, and G. Arunodaya, "Detection of seizure activity in eeg by an artificial neural network: A preliminary study," *Computing Biomedical Research*, vol. 29, pp. 303–313, 1996.

[103] H. Preissl, *Magnetoencephalography*, H. Preissl, Ed. Academic Press, 2005.

[104] U. Ramacher, *VLSI Design of Neural Networks*, U. Ramacher and U. Rückert, Eds. Springer, 1991.

[105] M. Rawat, K. Rawat, and F. M. Ghannouchi, "Adaptive digital predistortion of wireless power amplifiers/transmitters using dynamic real-valued focused time-delay line neural networks," *IEEE transaction on microwave theory and techniques*, vol. 58, pp. 95–104, 2010.

[106] A. Renyi, *Probability Theory*, ., Ed. Amsterdam: North Holland: North-Holland Publishing Company, 1970.

[107] C. Sciences, "Eeg electrodes," iWorx, Tech. Rep., 2006.

[108] K. Seong-Gi, J. Tao, and M. F., *fMRI - Basics and Clinical Applications*. Springer - Verlag, 2010, ch. 3. Spatial Resolution of fMRI Techniques, pp. 15 – 21.

[109] C. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[110] P. Sharp and A. Welch, *Practical Nuclear Medicine*. Oxford University Press, 2005, ch. 3 Positron Emission Tomography, pp. 35–49.

[111] R. Sitaram, H. Zhang, and C. Guan, "Temporal classification of multichannel near- infrared spectroscopy signals of motor imagery for developing a brain-computer interface." *Neuroimage*, vol. 34, pp. 1416–1427, 2007.

[112] S. Slobounov, *Injuries in Athletics: Causes and Consequences*. Springer, 2008, ch. 22 EEG & Neurofeedback in Rehabilitation, pp. 493 – 513.

[113] K. Spencer, J. Dien, and E. Donchin, "Spatiotemporal analysis of the late erp responses to deviant stimuli," *Psychophysiology*, vol. 38(2), pp. 343–358, 2001.

[114] V. Srinivasan, C. Eswaran, and N. Sriraam, "Artificial neural network based epileptic detection using time-domain and frequency-domain features," *Journal of Medical Systems*, vol. 29, no. 6, pp. 647–660, December 2005.

[115] M. Steinmetz, J. Luders, and E. Benzel, *Current Clinical Oncology: High-Grade Gliomas: Diagnosis and Treatment*. Humana Press Inc., 2010, ch. 10 Magnetoencephalography.

[116] A. Subasi, M. Akin, K. Kiymik, and O. Erogul, "Automatic recognition of vigilance state by using a wavelet-based artificial neural network," *Neural Comput. and Applications*, vol. 14, pp. 45–55, 2005.

[117] W. Tatum, A. Husai, S. Benbadis, and P. Kaplan, *Normal EEG*, Kaplan, Ed. Medical Punlishing, 2008.

[118] N. V. Thakor and S. Tong, "Advances in quantitative electroencephalogram analysis methods," *Biomedical Engineering*, vol. 6, pp. 453–495, August 2004.

[119] F. Theis and M.-B. A., *Biomedical Signal Analysis: Methods and Applications*, F. Theis, Ed. MIT Press, 2010.

[120] C. Tsallis, "Possible generalization of boltzmann-gibbs statistics," *Journal of Statistical Physics*, vol. 52, pp. 479–487, 1988.

[121] S. Ulmer, *fMRI- Basic and Clinical Applications*. Springer - Verlag, 2010, ch. 1. Introduction, pp. 5–13.

[122] A. Vallabhaneni, T. Wang, and B. He, *Brain Machine Interfaces*. Springer, US, 2005, ch. 3, pp. 85–121.

[123] B. L. van der Waerden, *A History of Algebra*. NY, USA: Springer Verlag, 1985.

[124] J. Vernon, M. Bach, C. Barber, M. Brigell, M. Marmor, A. Tormene, and G. . Holder, "Visual evoked potentials standard," *Documenta Ophthalmologica*, vol. 108, pp. 118–123, 2004.

[125] J. Vidal, "Real-time detection of brain events in eeg," *Proc. IEEE*, vol. 65, pp. 633–664, 1997.

[126] A. Vuckovic, V. Radivojevic, A. Chen, and D. Popovic, "Automatic recognition of alertness and drowsiness from eeg by an artificial neural network," *Medical Engineering and Physics*, vol. 24, pp. 349–360, 2002.

[127] S. Walder, H. Preiss, E. Demandt, C. Braun, N. Birbaumer, A. Aertsen, and C. Mehring, "Hand movement direction decoded from meg and eeg," *The Journal of Neuroscience*, vol. 28(4), pp. 1000–1008, 2008.

[128] W. Weng and K. Khorasani, "An adaptive structure neural network with application to eeg automatic seizure detection," *Neural Networks*, vol. 9, pp. 1223–1240, 1996.

[129] C. Westall, *Optometry: Science, Techniques and Clinical Management.* Elsevier Health Sciences, 2009, ch. Visual Development, pp. 37 – 50.

[130] W. Williams, H. Zaveri, and J. Sackellares, "Time-frequency analysis of electrophysiology signals in epilepsy," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, pp. 133–143, 1995.

[131] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, pp. 767–791, 2002.

[132] Xilinx, "Dsp as fpgas," Web, May 2015, http://www.xilinx.com/training/dsp/dsp-digital-signal-processing.htm.

[133] ——, "Field programmable gate array," Web, May 2015, http://www.xilinx.com/training/fpga/fpga-field-programmable-gate-array.htm.

[134] M. Xu, F. Wu, and H. Leung, "A biologically motivated signal transmission approach based on stochastic delay differential equations," *Chaos*, vol. 19, p. 033135, 2009.

[135] S. Xu, T. Chen, and J. Lam, "Robust h filtering for uncertain markovian jump systems with mode dependent time delays," *IEEETransactions on Automatic Control*, vol. 48, pp. 900–907, 2003.

[136] S. Xua, J. Lamb, D. Hoc, and Y. Zoua, "Delay dependent exponential stability for a class of neural networks with time delays," *Journal of Computational and Applied Mathematics*, vol. 183, pp. 16–28, 2005.

[137] S. Yoo, T. Fairneny, and N. Chen, "Brain computer interface using fmri:spatial navigation by thoughts," *Neuroreport*, vol. 15, pp. 1591–1595, 2004.

[138] J. Yordanova, V. Kolev, O. Rosso, M. Schurmann, and O. Sakowitz, "Wavelet entropy analysis of eventrelated potentials indicates modalityindependent theta dominance," *Journal on Neurosciense Methods*, vol. 117, pp. 99–109, 2002.

[139] H. Zaveri, W. Williams, L. Iasemidis, and J. Sackellares, "Time-frequency representation of electrocorticograms in temporal lobe epilepsy," *IEEE Transaction on Biomedical Engeeniering*, vol. 39, pp. 502–509, 1992.

[140] X. Zhang, M. Wu, J. She, and Y. He, "Delay dependent stabilization of linear systems with time varying state and input delays," *Automatica*, vol. 41, pp. 1405–1412, 2005.

[141] H. Zhao, X. Zeng, Z. He, W. Jin, and T. Li, "Complex-valued pipelined decision feedback recurrentneural network for non-linear channel equalisation," *IET Communications*, vol. 6, no. 9, pp. 1082–1096, 2012.

[142] D. Zhu, J. Bieger, G. Garcia-Molina, and R. Aarts, "A survey of stimulation methods used in ssvep-based bcis," *Computational Intelligence and Neuroscience*, vol. 2010, pp. 1–12, 2010.

[143] J. Zhu and P. Sutton, "Fpga implementation of neural networks - a survey of a decade of progress," in *13th International Conference on Field-Programmable Logic and Applications (FPL 2003)*. Springer-Verlag, 2003, pp. 1062–1066.