

Elman neural network for the early identification of cognitive impairment in Alzheimer's disease

Francesco Bertè, BSc
Giuseppe Lamponi, PhD
Rocco Salvatore Calabrò, MD, PhD
Placido Bramanti, MD

IRCCS Centro Neurolesi, Bonino-Pulejo, Messina, Italy

Correspondence to: Francesco Bertè
E-mail: francesco.berte84@gmail.com

Abstract

Early detection of dementia can be useful to delay progression of the disease and to raise awareness of the condition. Alterations in temporal and spatial EEG markers have been found in patients with Alzheimer's disease (AD) and mild cognitive impairment (MCI). Herein, we propose an automatic recognition method of cognitive impairment evaluation based on EEG analysis using an artificial neural network (ANN) combined with a genetic algorithm (GA). The EEGs of 43 AD and MCI patients (aged between 62 and 88 years) were recorded, analyzed and correlated with their MMSE scores. Quantitative EEGs were calculated using discrete wavelet transform. The data obtained were analyzed by the means of the combined use of ANN and GA to determine the degree of cognitive impairment. The good recognition rate of ANN fed with these inputs suggests that the combined GA/ANN approach may be useful for early detection of AD and could be a valuable tool to support physicians in clinical practice.

KEY WORDS: artificial neural network, dementia, EEG, genetic algorithm, MMSE.

Introduction

Mild cognitive impairment (MCI, also known as incipient dementia, or isolated memory impairment) is a diagnosis given to individuals who have cognitive impairments beyond what can be expected for their age and level of education, but that do not interfere significantly with their daily activities. It is regarded as the boundary or transitional stage between normal aging and dementia (Petersen et al., 1999). MCI can

present with a variety of symptoms; when memory loss is the predominant symptom it is termed "amnesic MCI" and frequently seen as a risk factor for Alzheimer's disease (AD) (Calabrò et al., 2011). Dementia is a syndrome consisting of a loss, presenting in a number of different combinations, of several separable but overlapping intellectual abilities, such as to interfere with normal daily activities. It may be due to various brain diseases, including neurodegenerative diseases (mainly AD), stroke, trauma and encephalitis.

An estimated 24 million people worldwide have dementia, and the majority of them are thought to have AD. Thus, AD represents a major public health concern and has been identified as a research priority. The cognitive function primarily impaired in AD is short-term memory, followed by long-term memory, attention, language, visuospatial abilities and abstract reasoning; these impairments ultimately lead to a total dependence in activities of daily living (Sunderland et al., 2006). However, not all patients with cognitive impairment progress to dementia; sometimes they can regain some normal function (Jack et al., 2005).

A reliable diagnostic/prognostic tool for cognitive impairment would be extremely useful for screening the elderly population, including subjects with subjective memory complaints; such screening would allow the immediate initiation of appropriate medical and rehabilitation treatments in individuals found to require them. There already exist several tools for evaluating cognitive impairment. These include the well-known Milan Overall Dementia Assessment, Alzheimer Disease Assessment Battery, Cambridge Examination for Mental Disorders of the Elderly and, of course, the Mini-Mental State Examination (MMSE). The MMSE, introduced by Folstein et al. (1975), is a 30-item questionnaire covering mathematics, memory and orientation that is administered to patients to evaluate the severity of cognitive impairment and changes in cognitive abilities. The number of correct answers is converted into a score ranging from a maximum of 30, which indicates normal functioning, to a minimum of zero, which indicates the presence of severe cognitive impairment (Mungas, 1991). However, the test results can be influenced by several conditions, including physical problems such as motor impairment or hearing problems, but also by the age and level of education of the subject investigated (Ganguli et al., 2010).

The MMSE is subject to a “ceiling effect” (when it fails to detect initial cognitive impairment in subjects with a higher level of schooling, who can obtain higher scores), and a “floor effect” (a term that refers to the instrument’s inability to detect changes in established advanced dementia, in subjects with little formal education and those with severe language problems) (Franco-Marina et al., 2010). To overcome these problems, a method based on automated analysis of routine EEG data may be ideal. Indeed, studies in AD patients have shown EEG abnormalities mainly related to alterations in each frequency band. An increase in theta and delta activity and a decrease in alpha and beta activity are commonly observed (Coben et al., 1983, 1985; Giaquinto and Nolfé, 1986; Brenner et al., 1986; Pijnenburg et al., 2004). Furthermore, these abnormalities are correlated with the severity of the disease (Hughes et al., 1989; Kowalski et al., 2001), and a significant correlation has been reported between the generalized frequency value and MMSE scores (Yoshimura et al., 2004). Thus, over the past three decades, the EEG has been found to be a useful tool for detecting, through the extraction of certain features, patients affected by AD and cognitive impairment (Babiloni et al., 2004; Huang et al., 2000; Dierks et al., 1993; Jonkman, 1997; Jeong, 2004; Jung et al., 2007). The aim of this study was to correlate MMSE scores with unique features extracted from EEG traces, in order to develop a tool for computer automated analysis that, combining a genetic algorithm (GA) and artificial neural network (ANN), may prove able to discriminate between cognitively healthy and impaired individuals.

Materials and methods

Forty-three patients affected by AD (with different levels of impairment) or MCI were enrolled. The subjects (men and women aged between 62 and 88 years) underwent a routine neuropsychological evaluation, i.e. the MMSE, at IRCCS Centro Neurolesi “Bonino-Pulejo” in Messina. Furthermore, five normal controls (sex- and age-matched), without cognitive deficits, were enrolled into the project as the control group.

EEG analysis

All EEGs were recorded and stored on computer for subsequent preprocessing and extraction of distinctive features. EEG data were recorded from 19 electrodes positioned according to the International 10-20 system, using a unipolar montage and a sampling frequency of 1024 samples per second. The EEG was recorded in resting state eyes-closed and eyes-open conditions (10 minutes in continuous recording mode). EEG signals are subject to internal noise from sources such as muscle activity and cardiovascular activity, and external noise from sources such as electrostatic discharge and interferences due to the presence of the electricity distribution network. Therefore, it is nec-

essary to perform pre-filtering operations. To suppress internal noise a cascade of filters was used:

- A high-pass filter to suppress the low-frequency components with a bandwidth (-3dB) of 1.6002 Hz whose transfer function is

$$H(s) = \frac{1}{1 + \frac{1}{s}}$$

and whose magnitude response is shown in figure 1.

- A low-pass filter to suppress the high-frequency components with a bandwidth (-3dB) of 30 Hz whose transfer function is

$$H(s) = \frac{1}{1 + s}$$

and whose magnitude response is shown in figure 2.

To suppress internal noise and delete individual components at 50 Hz, a second-order notch filter, also known as band-stop filter or band-rejection filter, was used with a center frequency of 50 Hz and a Q factor of 35 whose transfer function is

$$H(s) = \frac{s^2 + \omega_n^2}{s^2 + 2\omega_n s + \omega_n^2}$$

and whose magnitude response is shown in figure 3.

The next step was the extraction of the EEG features to evaluate cognitive impairment. In accordance with reported literature (Primavera et al., 1990), we focused on the Alpha rhythm in the T5, T6, O1 and O2 derivations, the Delta rhythm in the F7, O1, and O2 derivations, and the Beta rhythm in the F7 derivation. For the extraction of the aforementioned rhythms the discrete wavelet transform (DWT) was used. A Ψ function is defined a wavelet if

$$\int_{-\infty}^{\infty} \Psi(t) dt = 0$$

The DWT allows the signals to be described in terms of coefficients, which represent the energy content in their specific time-frequency regions. This representation is constructed through decomposition of the signal over a set of functions generated by scaling and translating a wavelet initial function Ψ

$$\Psi_{b,a} = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right)$$

The initial function can be regarded as a band-pass filter. The bandwidth of this filter can be varied by constricting or dilating the wavelet function. The DWT of a signal x can be regarded as the passage of this signal through a series of filters. The sampled signal x is decomposed by passing it through a low-pass filter with impulse response function g resulting in the convolution between the two

$$y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k]g[n-k]$$

and a high-pass filter with impulse response h . Both filters are correlated with each other and constitute a quadrature mirror filter. At the outputs of the filters we obtain a signal sub-sampled by a multiple of 2:

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n - k]$$

$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n - k]$$

Each output signal is characterized by half of the input frequency and twice the frequency resolution. The decomposition can be repeated to increase the frequency resolution and the approximation coefficients decomposed by the high-pass and low-pass filters. As shown in figure 4 (over), at each level, the signal is split into high and low frequencies. Because of the decomposition process the input signal must be a multiple of 2^n where n is the number of levels.

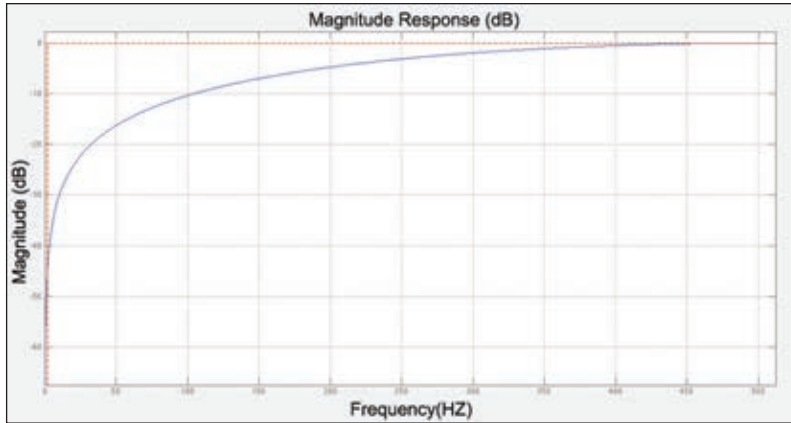


Figure 1 - High-pass filter magnitude response.

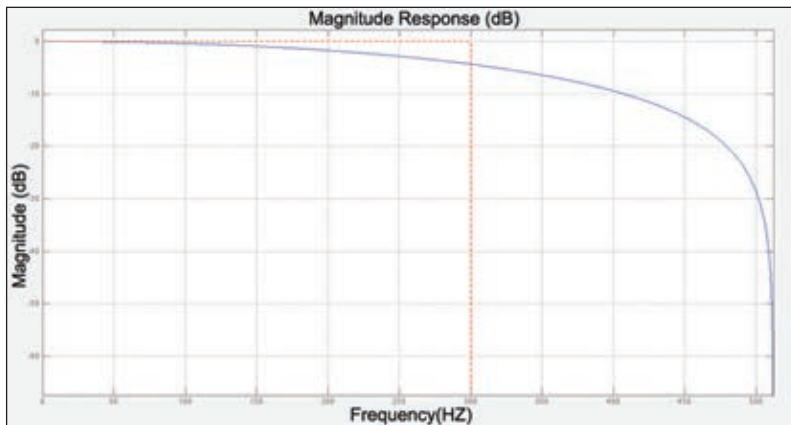


Figure 2 - Low-pass filter magnitude response.

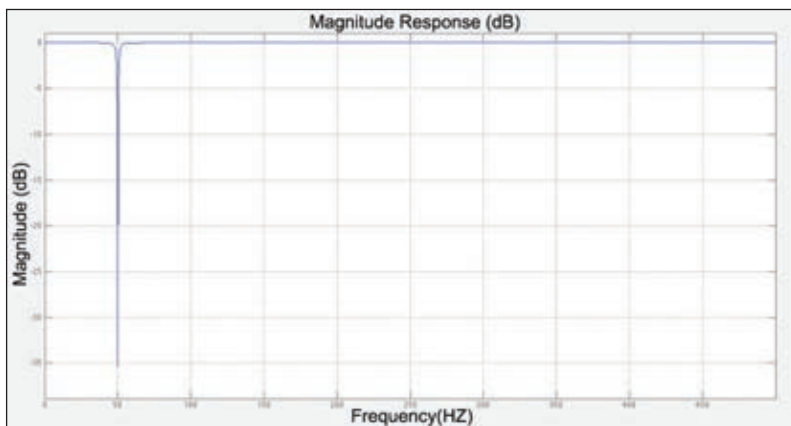


Figure 3 - Notch filter magnitude response.

In our specific case, the sampling frequency of the EEG recordings was 1024 samples per second. Thus, to be able to extract all the rhythms concerned and considering that the Delta rhythm is in the frequency range of 0-4 Hz, we performed a decomposition of the signal over eight levels using the wavelet function Daubechies-8. Once the rhythms had been extracted, the mean frequency (MF) was calculated using the following formula:

$$MF = \frac{\sum_{f=1}^{30} F(f) * f}{\sum_{f=1}^{30} F(f)}$$

and the fractions of total power using the formula

$$PP = \frac{P_x}{T}$$

where P_x is the power of the rhythm concerned and T is the sum of the power in all frequency bands. For the EEG analysis, the Matlab tool was used. To perform the aforementioned filtering operations, finite impulse response digital filters were implemented using the Matlab "Signal Processing Toolbox" software package. Instead, for the extraction operations, a new Matlab script was created. This script analyzes the EEG signal free of artifacts, identifies the sampling frequency, and performs the extraction, processing and saving, in files, of the required features, as described below (Fig. 5).

Artificial neural network and genetic algorithm

Part of the pool of data obtained was used to perform the training of the ANN, whereas the remaining part was used to conduct tests to evaluate the effectiveness of the infrastructure created. For our purposes, it was decided to use the Elman network, a recurrent ANN used in many fields, including analysis of signals such as EEG signals (Palaniappan, 2006; Srinivasan et al., 2005). This choice was based on its good non-linear effect of disturbance elimination and its capacity to identify patterns in a sequence of values. This network consists essentially of an input layer, a hidden layer and an output layer, with feed-forward links

between them. The hidden layer, moreover, is connected not only to the output layer but also to a further layer, called the context layer, with fixed weights of unit values. To create a recurrent connection, the output of the context layer is reported in input to the hidden layer; this allows the neural network to keep a memory of the previous state (Holk Cruse, 2006). The

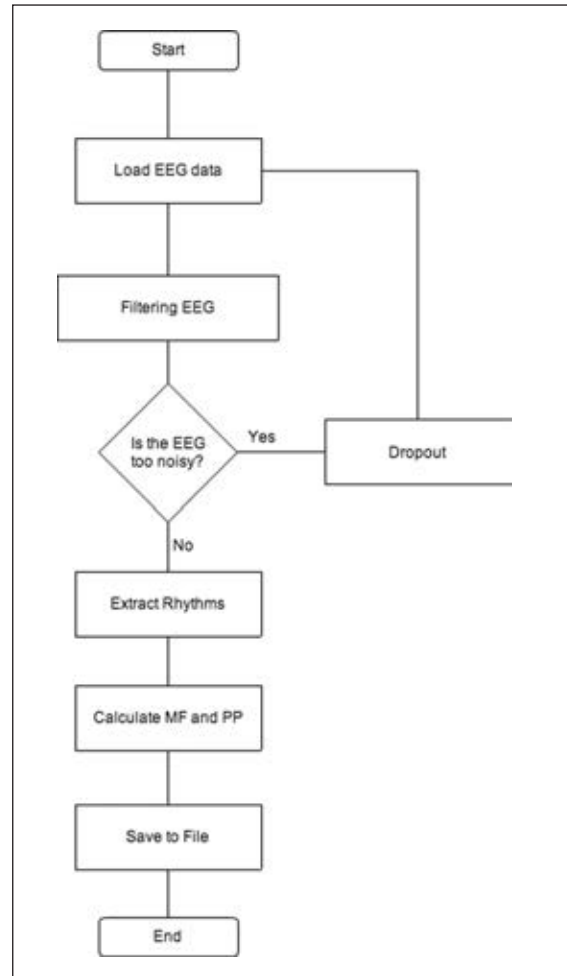


Figure 5 - Flowchart of EEG analysis algorithm.

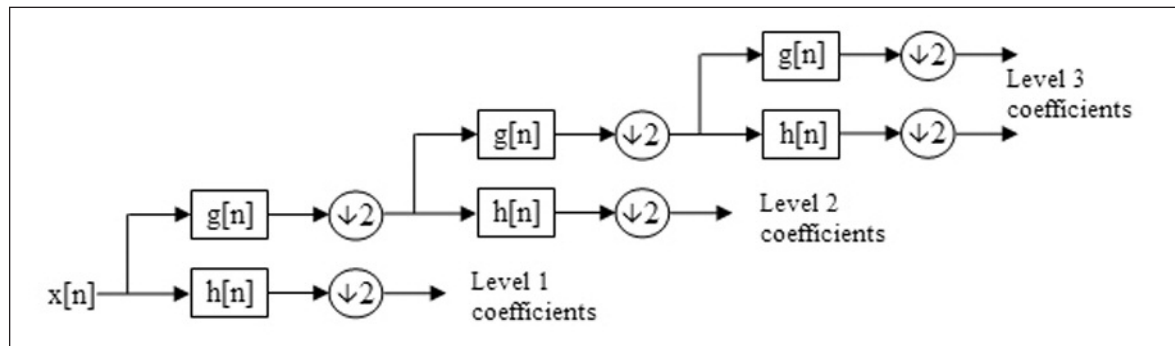


Figure 4 - Block diagram of the decomposition performed by discrete wavelet transform.

ANN was created using the open source framework Encog (E.M.L. Framework), whose structure is shown in figures 6 and 7.

The data to be processed were stored in text files using the comma-separated values format. Each file contains a table of values in which each line of text represents a row of the table while the fields are identified by a special separator character, in our case, a comma “,”. The first line of the file consists of the header, formed from the name of the features extracted from EEG signals and the MMSE “score”. The next

lines of text contain the measures of data previously collected. These measure values, once they have been normalized to the range [0,1], and formatted in a binary representation using tools provided by the Encog software package, are used as input data at the level of the index to be processed by the eight neurons that constitute the input layer of the network. After these operations, the data are processed by the sixty-four neurons that compose the hidden layer. The number of hidden and context units was chosen through experimental tests. After the processing, the

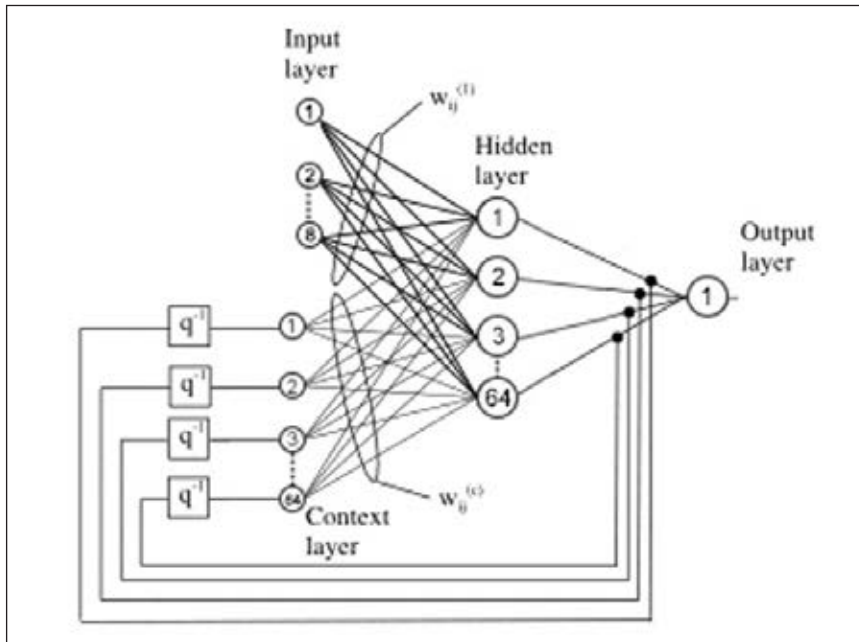


Figure 6 - Block structure of the Elman network (Lundstedt et al., 2002).

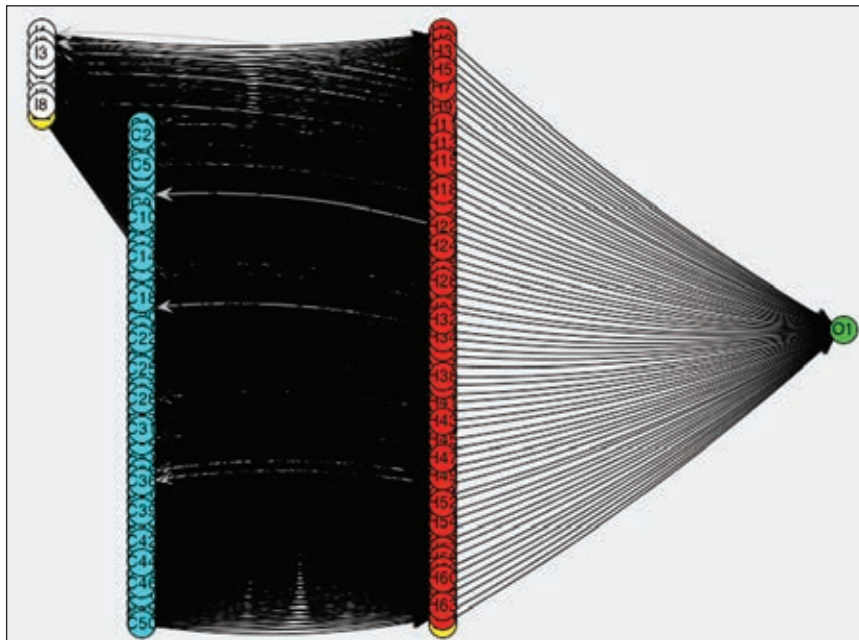


Figure 7 - Real structure of the Elman network.

single neuron in the output layer provides an indication regarding the presence or absence of cognitive impairment in the patients studied. For our purposes, the activation functions used were tan-sigmoid and log-sigmoid for the hidden and the output layers respectively. A GA was used to train the ANN; GA is a research and optimization method that mimics biological mutations and natural evolutions. It is used in different fields including ANNs, and is often used in recursive networks, such as Elman networks. The solution of the problem is encoded in a structure called a chromosome. Initially a certain number of chromosomes are created randomly: these constitute the initial population. A function that evaluates the quality of the solution, called fitness function, is also defined. The algorithm then proceeds to assign a quality indicator to each solution to sort and select the best ones, which constitute the parents of new genes that are then generated by making random changes to a single parent (mutation) or by suitably combining the characteristics of a pair of parents (crossover). Figure 8 summarizes the operations used to train the ANN.

Results

The ANN was trained, trying to minimize the possibility of error. For this reason, the GA was appropriately calibrated to obtain a percentage error of less than one in the classification of the patients used to train. Chromosomes encoded the matrix weights of the network. In order to manipulate shorter chromosomes, each gene on the chromosome was coded using real-valued coding (instead of using binary notation) and represented a weight from the neural network. The starting network weights were initialized with random values. After this, a size population of 100 with randomly assigned values was created. The fitness function was defined as the deviation of the calculated value from the ideal value expressed as a percentage. The process of mating occurs by splitting the parent gene sequence into three splices. These three splices are then used to build the new gene. The mutation process is completely random. Two random genes are chosen, and then swapped as part of the mutation process. Only part of the gene population will be subject to the gene mutation and mating process. After various experiments the best results were obtained by setting the following parameters:

- Learning Rate: 5000
- Mutation Percent: 0.1
- Percent to mate: 0.25

With these settings, after 170 iterations, the error was decreased below the predetermined level (Fig. 9), and the network had balanced the weights to obtain a result that was as near to the ideal one as possible (Fig. 10). After the training phase, tests were conducted to verify the reliability and the goodness of the classification. As shown in table I only thirty-eight EEG recordings out of a total of forty-three were used in the tests; the dropouts showed significant alteration due to artifacts.

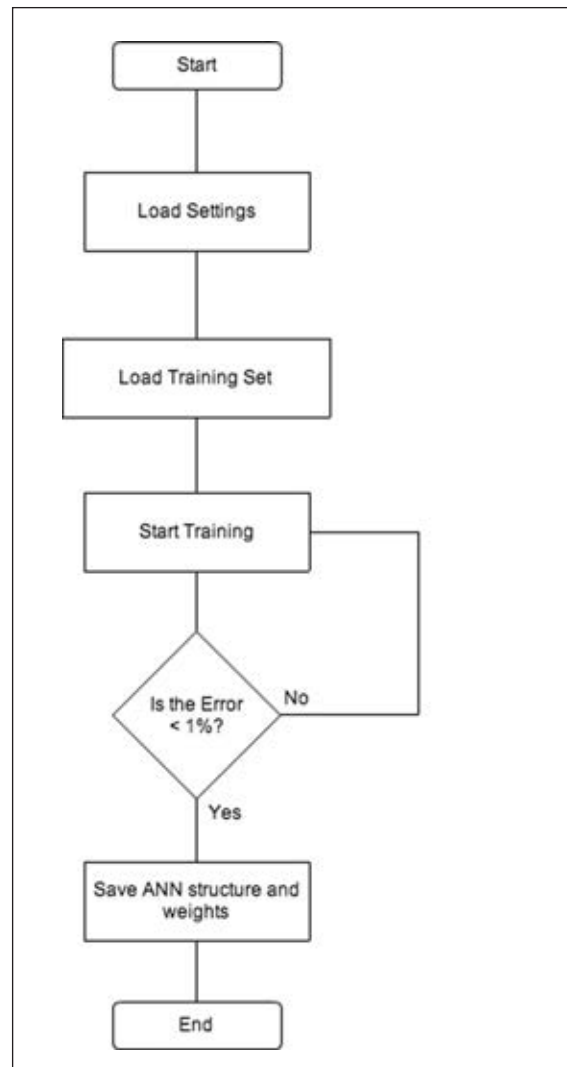


Figure 8 - Flowchart of the training of the artificial neural network.

If the training set is appropriately balanced in such a way as to have equal numbers of healthy subjects and subjects with cognitive impairment, the network is able to determine with absolute certainty the presence or absence of cognitive disturbance. By contrast, it was seen that an imbalance of the training set increases the percentage error (Fig. 11, over), and decreases the ability to classify the subjects, as reported in table II (over). A tolerable imbalance occurs when the difference between the number of healthy people and the number with deficits does not exceed 15%; conversely, if variation tends to exceed 30% the network has difficulty distinguishing between the various cases. Nevertheless, in the worst cases, the method appeared to have a precision of classification of 70%, with a p value < 0.05. Moreover, taking into consideration the five tests conducted and given a generic EEG pattern, it was found to be possible to distinguish healthy from non-healthy individuals with an accuracy of 85%.

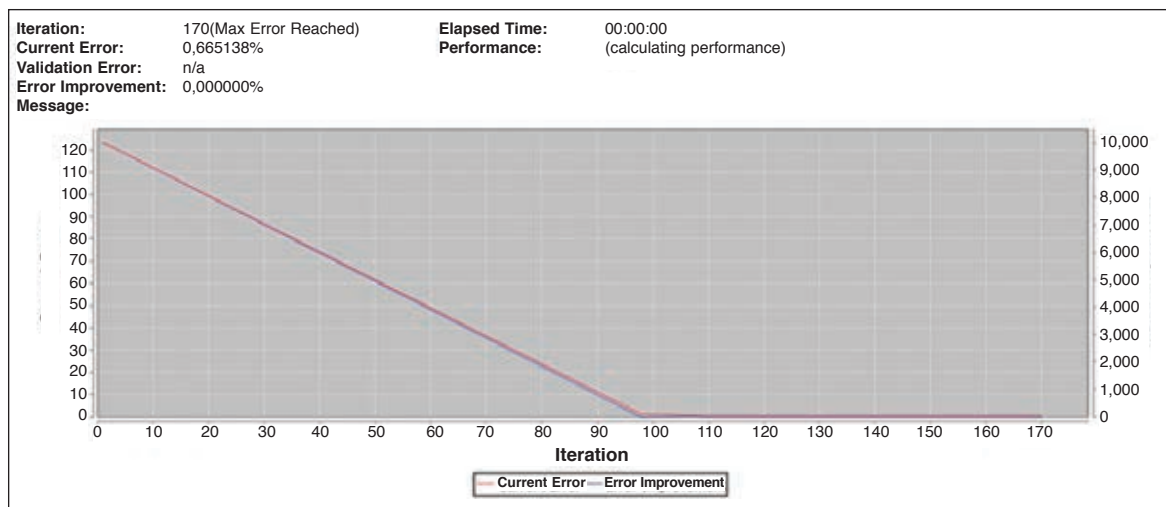


Figure 9 - Error chart showing the number of iterations before the error falls below the threshold (red line) and the improvement in identifying the optimal result (blue line).

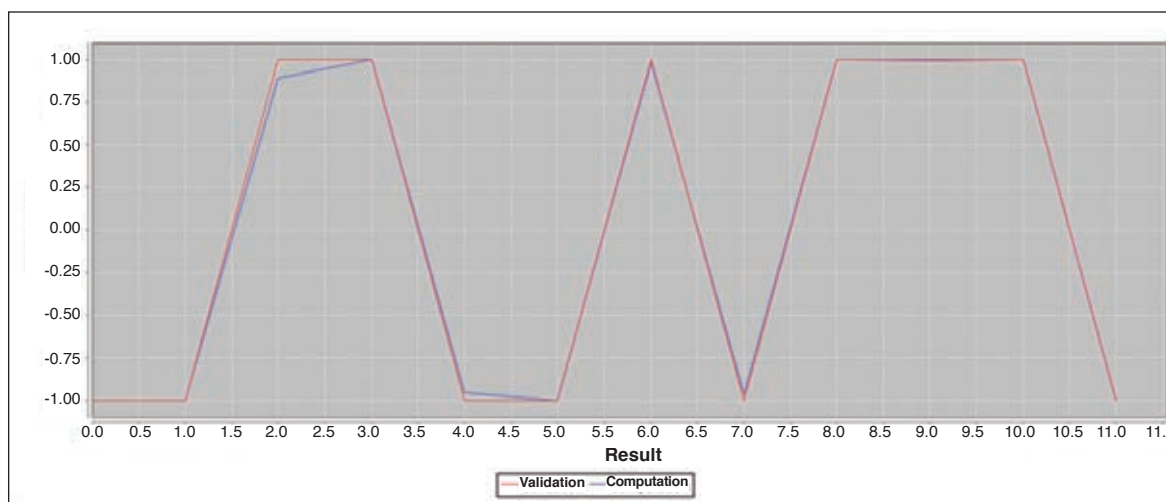


Figure 10 - Validation chart showing how the results obtained by the neural network (blue line) deviated from the ideal results (red line).

Table I - Tests with different training sets.

Groups	Training sets	Subjects evaluated	Results %
A	14 → No Dementia 14 → Dementia	5 → No Dementia 5 → Dementia	100%
B	16 → No Dementia 12 → Dementia	3 → No Dementia 7 → Dementia	90%
C	12 → No Dementia 16 → Dementia	7 → No Dementia 3 → Dementia	90%
D	18 → No Dementia 10 → Dementia	1 → No Dementia 9 → Dementia	70%
E	10 → No Dementia 18 → Dementia	9 → No Dementia 1 → Dementia	70%

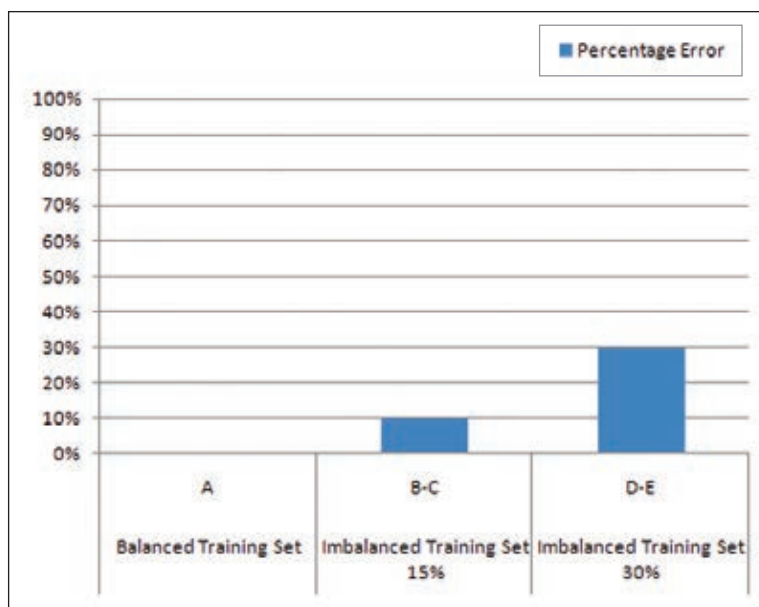


Figure 11 - Graph of percentage error using different training sets to train the artificial neural network.

Table II – Number of errors in different test groups.

Groups	Errors (in 10 tests)
A	0
B	1
C	1
D	3
E	3

Discussion

Diagnosis of cognitive impairment requires a clinical judgment, obtained by clinical observation, as well as blood and neuropsychological tests, and a proper instrumental evaluation. The current challenge for clinical research in this field is to diagnose dementia earlier and with higher accuracy. An earlier diagnosis would allow a more efficient clinical treatment, capable of delaying the progression of the illness. In this work we proposed an automatic system for detecting cognitive deficit that uses ANN combined with signal wavelet decomposition and GA, to analyze continuous EEG recordings during resting state eyes-closed and eyes-open conditions. Recurrent neural networks (like the Elman network used in our work), being capable of detecting linear and non-linear changes in the signal, have been shown to be valuable tools for detecting and analysing EEG features, and thus for demonstrating the presence of dementia. In the work proposed by Petrosian et al. (2001), the use of a recurrent ANN allowed identification of patients affected by AD through analysis of long-term EEG. However, a disadvantage of this specific approach is that it uses a large pool of data in the analysis of the ANN.

The combined use of ANN and GA allowed Kim et al. (2005) to correctly classify normal patients, distinguishing them from those with AD. Indeed, the use of GA

reduced the size of the pool of data needed to identify the minimum set of features required for distinguishing between subjects. Our approach was based on the use of distinctive features extracted from EEG traces and their correlation with the results obtained from the MMSE, in order to start with a restricted data set to be processed by the ANN. The combined use of GA allows further reductions in the number of the features needed for the identification of the pathology without compromising the efficiency of the system. In this regard, the network proved able to recognize healthy subjects, versus non-healthy ones, with good precision ($p < 0.005$). The best results were obtained by configuring the training set in order to have an equal number of features from healthy and non-healthy subjects. Training the network with this set, the system was able to correctly distinguish the five healthy patients from the five patients suffering from dementia with a detection rate of 100%. Nonetheless, the system performance is worse when the training set used is not well-balanced because of the small number of items it contains, which do not allow proper configuration of the weights of the neural network; this is a direct consequence of the small number of patients enrolled. Therefore, more time and further data are needed before the proposed system can be used in common practice and lead to a proper diagnosis. However, the early results are encouraging, and it seems reasonable to conclude that the suggested approach is a valuable tool for identifying patients with dementia and for supporting the clinician in the diagnostic process, by providing a "second opinion".

In conclusion, we proposed a combined ANN/GA approach and wavelet decomposition to analyze continuous EEG for automatic and early detection of patients suffering from dementia. The results obtained by this promising system are positive, but this approach should be extended to a larger patient population in order to properly assess its effectiveness.

References

- Babiloni C, Binetti G, Cassetta E et al (2004) Mapping distributed sources of cortical rhythms in mild Alzheimer's disease. A multicentric EEG study. *Neuroimage* 22: 57-67.
- Brenner RP, Ulrich RF, Spiker DG, et al (1986). Computerized EEG spectral analysis in elderly normal, demented and depressed subjects. *Electroencephalogr Clin Neurophysiol* 64, 483-492.
- Calabrò RS, Baglieri A, Digangi G, et al (2011) Psychological tests in neurological diseases: an overview. In: Columbus AM (Ed.) *Advances in Psychology Research*, New York, Nova Publisher 2011.
- Coben LA, Danziger WL, Berg L (1983) Frequency analysis of the resting awake EEG in mild senile dementia of Alzheimer type. *Electroencephalogr Clin Neurophysiol* 55, 372-380.
- Coben LA, Danziger W, Storandt M (1985) A longitudinal EEG study of mild senile dementia of Alzheimer type: changes at 1 year and at 2.5 years. *Electroencephalogr Clin Neurophysiol* 61, 101-112.
- Dierks T, Ihl R, Frolich L et al (1993). Dementia of the Alzheimer type: effects on the spontaneous EEG described by dipole sources. *Psychiatry Res* 50: 151-162.
- E. M. L. Framework, <http://www.heatonresearch.com/encog>.
- Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*, 12 189-198.
- Franco-Marina F, Garcia-Gonzalez JJ, Wagner-Echeagaray F, et al (2010). The Mini-Mental State Examination revisited: ceiling and floor effects after score adjustment for educational level in an aging Mexican population. *Int Psychogeriatr* 22, 72-81.
- Ganguli M, Snitz BE, Lee CW, et al (2010). Age and education effects and norms on a cognitive test battery from a population-based cohort: the Monongahela-Youghiogheny Healthy Aging Team. *Aging Ment Health* 14, 100-107.
- Giaquinto S, Nolfi G (1986) The EEG in the normal elderly: a contribution to the interpretation of aging and dementia. *Electroencephalogr Clin Neurophysiol*, 63, 540-546.
- Huang C, Wahlund L, Dierks T, et al (2000) Discrimination of Alzheimer's disease and mild cognitive impairment by equivalent EEG sources: a cross-sectional and longitudinal study. *Clin Neurophysiol* 111: 1961-1967.
- Hughes JR, Shanmugham S, Wetzel LC, et al (1989) The relationship between EEG changes and cognitive functions in dementia: a study in a VA population. *Clin Electroencephalogr* 20, 77- 85.
- Holk Cruse D (2006), *Neural Networks as Cybernetic Systems*, (2nd and revised edition). bmm615. Brains, Minds and Media, Bielefeld, Germany. (<http://www.brains-minds-media.org/archive/2006#ebook>).
- Jack CR Jr, Shiung MM, Weigand SD, et al (2005) Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnesic MCI. *Neurology* 65, 1227-1231
- Kowalski JW, Gawel M, Pfeffer A et al (2001) The diagnostic value of EEG in Alzheimer disease: correlation with the severity of mental impairment. *J Clin Neurophysiol* 18, 570-575.
- Jeong J (2004) EEG dynamics in patients with Alzheimer's disease. *Clin Neurophysiol* 115, 1490-1505.
- Jonkman EJ (1997) The role of the electroencephalogram in the diagnosis of dementia of the Alzheimer type: an attempt at technology assessment. *Neurophysiol Clin* 27, 211-219.
- Jung H, Lee S, Kim J, et al (2007) Quantitative electroencephalography and low resolution electromagnetic tomography imaging of Alzheimer's disease. *Psychiatry Investigation* 4, 31-37.
- Kim HT, Kim BY, Park EH, et al (2005). Computerized recognition of Alzheimer disease-EEG using genetic algorithms and neural network, *Future Generation Computer Systems* 21, 1124-1130.
- Lundstedt H, Gleisner H, Wintoft P (2002) Operational forecasts of the geomagnetic Dst index. *Geophysical Research Letters* 29, 2181.
- Matlab, <http://www.mathworks.it/>.
- Mungas D (1991) In-office mental status testing: a practical guide. *Geriatrics* 46 54-58
- Palaniappan R (2006) Utilizing gamma band to improve mental task based brain-computer interface design. *IEEE Trans Neural Syst Rehabil Eng* 14, 299-303.
- Petersen RC, Smith GE, Waring SC, et al (1999). Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol* 56, 303-308.
- Petrosian AA, Prokhorov DV, Lajara-Nanson W et al (2001) Recurrent neural network-based approach for early recognition of Alzheimer's disease in EEG. *Clin Neurophysiol* 112:1378-1387.
- Pijnenburg YA, vd Made Y, van Cappellen van Walsum AM, et al (2004). EEG synchronization likelihood in mild cognitive impairment and Alzheimer's disease during a working memory task. *Clin Neurophysiol* 115,1332-1339.
- Primavera A, Novello P, Finocchi C, et al (1990) Correlation between mini-mental state examination and quantitative electroencephalography in senile dementia of Alzheimer type. *Neuropsychobiology* 23:74-78.
- Srinivasan V, Eswaran C, Sriraam N (2005) Artificial neural network based epileptic detection using time-domain and frequency-domain features. *J Med Syst* 29, 647-660.
- Sunderland T, Hampel H, Takeda M., et al (2006) Biomarkers in the diagnosis of Alzheimer's disease: are we ready? *J Geriatr Psychiatry Neurol* 19, 172-179.
- Yoshimura M, Isotani T, Yagyu T et al (2004) Global approach to multichannel electroencephalogram analysis for diagnosis and clinical evaluation in mild Alzheimer's disease. *Neuropsychobiology* 49, 163-166.