Automatic Recognition of Alzheimer's Disease Using Genetic Algorithms and Neural Network[1]

Sunyoung Cho[1], Boyeon Kim[2], Eunhea Park[3], Yunseok Chang[4], Jongwoo Kim[5], Kyungchun Chung[6], Weiwan Whang[5], and Hyuntaek Kim[3]

[1] Basic Science Research Institute, Chungbuk National University, Chungju, Korea, sycho@chungbuk.ac.kr

[2] Department of Electrical & Computer Engineering, Kangwon National University, Chuncheon, Korea

[3] Department of Psychology, Korea University, Seoul, Korea

[4] Department of Computer Engineering, Daejin University, Pocheon, Korea

[5]Department of Oriental Neuropsychiatry, Kyunghee University, Seoul, Korea

[6]Department of Neurology, Kyunghee University, Seoul, Korea

Abstract. We propose an Alzheimer's disease (AD) recognition method combined the genetic algorithms (GA) and the artificial neural network (ANN). Spontaneous EEG and auditory ERP data recorded from a single site in 16 early AD patients and 16 age-matched normal subjects were used. We made a feature pool including 88 spectral, 28 statistical and 2 nonlinear characteristics of EEG and 10 features of ERP. The combined GA/ANN was applied to find the dominant features automatically from the feature pool, and the selected features were used as a network input. The recognition rate of the ANN fed by this input was 81.9% for the untrained data set. These results lead to the conclusion that the combined GA/ANN approach may be useful for an early detection of the AD. This approach could be extended to a reliable classification system using EEG recording that can discriminate between groups.

1 Introduction

A number of quantitative EEG analysis have been used to detect the brain's functional changes in the Alzheimer's disease (AD). Investigators have extracted specific quantitative features from the EEG, which would be characteristics for each stage of this disease. Various spectral and nonlinear analyses were employed and some progress has been established [1-2].

To the spectral nature of the EEG changes in the AD, there is a general agreement that the earliest changes are an increase in theta and a decrease in beta mainly over

[1.] This work was supported by Korean Ministry of Health and Welfare, 00-PJ9-PG1-CO05-0002 and a result of research activities of Advanced Biometric Research Center (ABRC) supported by KOSEF.

P.M.A. Sloot et al. (Eds.): ICCS 2003, LNCS 2658, pp. 695–702, 2003.

parieto-occipital area, followed by a decrease of alpha activity [3-4]. Delta activity increases later in the course of disease [5]. Claus et al.(1998) reported that a slowing spectral EEG could predict the rate of subsequent cognitive and functional decline in the AD, using multiple linear regression analysis [6]. Intra- and inter-hemispheric EEG coherence, which is considered to be a measure of the cortical synchronization and possibly to reflect a functional status of the intracortical communication, was significantly lower in alpha and beta frequency in AD patients [7].

Recent progress in the theory of the nonlinear dynamics has provided new methods for the study of the time-series physiological data. The nonlinear analysis of the EEG data could be a useful tool to differentiate normal and pathologic brain state. Several studies of the EEG in AD patients estimated the correlation dimension (D2) and Lyapunov exponent (L1) [8-9]. They showed significantly lower values of D2 and L1 in AD than age-matched normal subjects, reflecting less complex signal dynamics.

Another useful quantitative electrophysiological assessment for the monitoring of the cortical function is the event-related potential (ERP). Since Goodin et al.(1978) demonstrated the prolonged latency in the P3 component with aging [10], many researchers have studied the ERP components in AD patients but this is still a matter of debate, and the diagnostic sensitivity and specialty of the ERP remain yet to be confirmed [11-12].

In this study, we propose an automatic AD recognition method combined the genetic algorithms (GA) and the artificial neural network (ANN), using the spontaneous EEG and auditory ERP recorded from a single site. The EEG and ERP were analyzed to compute their spectral features as well as statistical and nonlinear features, to make a feature pool. The combined GA/ANN approach was applied to select the dominant features that are most efficient to classify two groups. The selected features were used as a neural network input for training and testing the network.

2 Method

We adopted the artificial neural network as a usual classifier to discriminate the AD patients from the normal subjects, using the computed EEG and ERP features. Applying the ANN as an effective classifier, we have to find the optimum and minimum features as a network input. To solve this problem we used the genetic algorithm to find the dominant input features from a feature pool.

2.1 The Feature Pool

With the electrophysiological data of the AD patients and normal subjects, we made a feature pool that represents their data. The spontaneous EEG data were divided into 30s segments and each segment was analyzed to compute their spectral, statistical and nonlinear characteristics, to generate 118 features. The ERP data to target tone with an averaging epoch 1s including 100ms of the prestimulus baseline, were analyzed to generate 10 features that would describe the characteristic of the patterns.

The final feature pool includes as follow;

- 88 power spectral measurements: for example, the maximum power, the frequency at the maximum power, the accumulated and relative power, the mean and variance of the power in δ, θ, α, β, γ band separately
- 28 statistic measurements: for example, the average amplitude, the range between the maximum and minimum amplitude, the ratio between the maximum and mean amplitude, the variance
- 2 chaotic features: the central tendency, the box-counting dimension
- 10 ERP features: for example, the latency and amplitude of the largest peak, the left second peak and right second peak in 300-700ms post-stimulus, the difference of amplitude and latency

2.2 Design of the Chromosomes and the Fitness Function

In the genetic algorithms, the concepts of chromosome are used to encode and manipulate the solution [13]. Each chromosome defines an individual of a population. We set a chromosome as a string consisted of 35 constants, that are representing the feature number in the feature pool and that will be used as a network input.

With this input of a chromosome after learning the ANNs to every training segment, the fitness function gives back a value for the chromosome, which is measuring the performance on the solution. The fitness value of a chromosome was defined as the inverse of the sum of mean square errors of the ANNs, by equation 1 where N is the number of ANNs, m is the number of output nodes of ANNi , do_j is the desired output of output node j and no is the network output of output node j.

$$Fittness = 1/\sum_{i=1}^{N} mean(\sum_{j=1}^{m} (do_j - no_j)^2) \tag{1}$$

2.3 The Genetic Operation

To create a population for a new generation, three basic genetic operations were used: crossover, mutation, and reproduction. One or two chromosomes in a formal generation were selected by the roulette wheel method as the parent chromosomes, with a probability based on its fitness value. In the crossover operation, two offspring chromosomes were produced from two parents by the one-point crossover. In the mutation, only one terminal of a parent chosen randomly would be mutated to generate a new chromosome. The reproduction operation copied a parent chromosome to the next generation.

2.4 The Combined GA/ANN Approach

1. Generate an initial population of the first generation with random proportions of the terminals.
2. Repeat the following steps until the terminating criterion has been satisfied. The evolution would be terminated, once a fitness value reach to 10,000.
   - Evaluate each chromosome in the population. After training the ANN using the features in the chromosome as a network input, it would be assigned a fitness value for each chromosome.
   - Create a population for the next generation by the genetic operations. These operations are applied to the chromosomes in a formal generation with the probabilities based on their fitness.
3. Choose a chromosome that has the maximum fitness value in each generation. Using these chromosomes chosen from several generations, we selected a dominant feature set. We made a histogram showing the number of the selection by these chromosomes for each feature, as shown in the figure 1, which would represent the significance of the feature to fit the solution. We selected the 35 dominant features in order of their significances.
4. Train and test the ANN with these dominant features as a network input.

Table 1 summarized the control parameters related to the execution of our combined GA/ANN approach.

Table 1. Control parameters of GA/ANN approach.

GA	the number of chromosomes in a generation	200
	the maximum number of generations	200
	crossover rate	0.95
	mutation rate	0.05
	reproduction rate	0.001
ANN	ANN model	multi-layered perceptron
	ANN learning rule	backpropagation
	the number of input node	35
	the number of output node	1
	the number of hidden layer	1
	the number of hidden node	13
	learning rate	0.1
	the maximum number of learning iteration	2000

3 Experiments and Results

3.1 Data Acquisition

Subjects. Two groups of the early AD patients and the age-matched normal subjects were studied. Sixteen AD patients were recruited from the oriental neuropsychiatric and neurological sections of Kyunghee University Hospital, aged between 61-82 (72 ± 6.4, mean ± SD). The patients with probable AD were diagnosed using the K-DRS (Korean-dementia rating scale) criteria [14] and their MMSE scores ranged from 15 to 27 with an average score of 19.5. The K-DRS consists of five subcategories including attention, initiation & preservation, construction, computation, and memory. Other medical conditions that are known to produce dementia were excluded following neurological and neuroimaging studies. None of the patients have been previously diagnosed with the psychiatric disorders, such as a depression, attention deficit, or schizophrenia, nor have they any history of significant head trauma, intracranial mass lesion, or any other neurological condition that could be associated with cognitive decline. Sixteen volunteers were the normal subjects aged between 61-78 (70 ± 5.3). They were carefully screened to eliminate individuals with the medical and neuropsychiatric disorder.

EEG recording. The EEG was recorded from an Ag-AgCl electrode placed at P4 based on the 10-20 system, referenced to linked earlobes with recording the EOG. The impedance of the electrodes was kept below 5 kΩ. The EEG was filtered (bandpass 0.5-35 Hz), amplified, digitized (250Hz) and stored on a hard disk for the off-line analysis. Spontaneous EEG with the eyes open was recorded for about 5min, of which artifact-free segments were selected for the analysis. The EEG data was divided into 30s segments, and in each segment 118 features were computed.

Auditory ERP procedure. Event-related potentials were acquired during the auditory oddball task. The stimuli consisted of series of sine-wave tones of 1 kHz and 1.5 kHz, each lasting 300ms. The two tones occurred respectively in 75% and 25% of the trials, in a random sequence. A total of 100 stimuli, including frequent 1 kHz tone and infrequent 1.5 kHz tone were delivered with an inter-stimulus interval of 0.8-1.2s. The subjects were instructed to count internally the number of the 1.5 kHz 'target' tones. The ERP data of the target tone that is averaged with an epoch of 1s including 100ms of the prestimulus baseline, were analyzed to generate 10 features to be included in the feature pool.

3.2 Selection of Dominant Features and Performance of Neural Network

We used 137 EEG segments from 11 AD patients and 10 normal subjects as a training data set. After training, the combined GA/ANN approach found the 35 dominant features including 24 spectral, 8 statistical, 1 nonlinear and 2 ERP features. Figure 1 shows a histogram for selecting the dominant features. It indicates the number of the selection by the 17 chromosomes that have the highest fitness value each in the last 17 generations. We selected the 35 dominant features in order of their heights, which were marked by the rectangular boxes in the figure.

The selected dominant features were applied as a network input to train the ANN again with 137 training EEG segments. After training, the weight values of the ANN were determined to able to test the performance of the network. The 72 EEG segments for the test were from 5 AD patients and 6 normal subjects. Table 2 reports the performance of the network for these untrained data set. For the EEG of the AD patients, the ANN recognized 22 segments out of the 30 test segments. And for the normal EEG, 37 segments out of the 42 test segments were recognized correctly. The 5 segments the network fail to recognize were all from one normal subject.

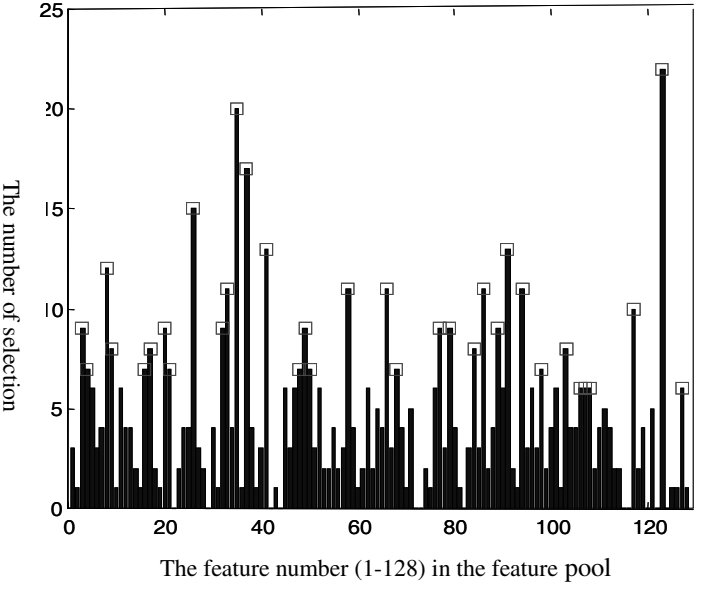

Fig. 1. The histogram for selecting the dominant features. The y axis indicates the number of the selection by the 17 chromosomes that have the highest fitness value each in last 17 generations. The rectangular boxes marked the selected dominant features (1-88: spectral features, 89-116: statistical features, 117-118: nonlinear features, 119-128: ERP features)

Table 2. Network performance for untrained segments

Target / NN output	Segments of AD patients	Segments of normal controls	Performance
AD	22	5	22/30 = 73%
Normal	8	37	37/42 = 88%
Total	30	42	59/72 = 81.9%

4 Discussion

The main goal for the clinical research in the AD is enhancement of the diagnostic accuracy and an earlier diagnosis. It is crucially important for the proper medical treatment and slowing down of the illness progress. We propose a reliable method to recognize the AD, only using one site of EEG recording. Single channel recordings of EEG are extremely easy and convenient to perform, even on the AD patients. If this simple tool has had enough accuracy to differentiate the AD patients from the normal adults, it would be highly helpful to diagnose the disease and to reduce costs. With a single channel EEG, our combined GA/ANN approach could find the dominant feature set and show good performance in determining the AD or normal EEG.

Our network was able to recognize successfully the EEG of normal subjects except with one subject. Even though her DRS and MMSE scores are in the normal range, the EEG of this subject could be deviant. In case of AD patients, 22 of the 30 segments were recognized correctly, so that the global recognition rate of the network was 73%. Note that it was only the case using the features from the 30s segments of the EEG data, and not with the every whole EEG data of each subject. With each subject, the whole EEG data consisted of 4-8 segments. The network failed to recognize only one or two segments with each AD patient. The remaining segments, that are more numerous, were recognized correctly as AD. Therefore, there should be no flaw whether a subject has AD or not. The spontaneous EEG of AD patients may vary a lot. In fact, the standard deviations in each segment (the 92nd statistic feature of the feature pool) of the AD group were significantly higher than those of the normal group ($F(1,206)=8.162$, $p<.005$). This feature was included in the dominant feature set selected by the GA.

We selected the dominant features by the genetic algorithms and used them as an optimum input of the neural network. We think this procedure enhanced the efficiency of the network. The redundant input from the raw EEG data or the manipulated data could make it rather worse. It was also useful to include the nonlinear characteristics and the ERP features in the feature pool, verified by the experimental results.

It seems reasonable to conclude that a single channel EEG data might be enough to recognize the AD using our combined GA/ANN approach. The suggested approach could be extended to a reliable classification system using EEG recording that can discriminate between groups.

References

1. Stam, C.J., Jelles, B., Achtereekte. H.A.M., Van Birgelen, J.H., Slaets, J.P.J.: Diagnostic Usefulness of Linear and Nonlinear Quantitative EEG Analysis in Alzheimer's Disease. Clin Electroencephalogr, Vol. 27.2. (1996) 69–77
2. Huang,C., Wahlund, L.-O., Dierks, T., Julin P., Winblad, B., Jelic, V.: Discrimination of Alzheimer's Disease and Mild Cognitive Impairment by Equivalent EEG Sources: A Cross-Sectional and Longitudinal Study. Clinical Neurophysiology. Vol. 111. (2000) 1961–1967
3. Prinz, P.N., Vitiello, M.V.: Dominant Occipital (Alpha) Rhythm Frequency in Early Stage Alzheimer's Disease and Depression. Electroencephalogr Clin Neurophysiol, Vol. 73. (1989) 427–432
4. Jelic, V., Shigeta, M., Julin, P., Almkvist, O., Winblad, B., Wahlund, Lo.: Quantitative Electroencephalography Power and Coherence in Alzheimer's Disease and Mild Cognitive Impairment. Dementia, Vol. 7. (1996) 314–323
5. Elmstaêhl, S., Roseân, I., Gullberg, B.: Quantitative EEG in Elderly Patients with Alzheimer's Disease and Healthy Controls. Dementia, Vol. 5. (1994) 119–124
6. Claus, J., Kwa, V., Teunisse, S., Walstra, G., Van Gool, W., Koelman, J., Bour L., Ongerboer De Visser, B.: Slowing on Quantitative Spectral EEG is a Marker for Rate of Subsequent Cognitive and Functional Decline in Early Alzheimer Disease. Alzheimer Dis Assoc Disord, Vol. 12.3. (1998) 167–174
7. Wada, Y., Y. Nanbu, Y. Koshino, N. Yamaguchi, T. Hashi-Moto.: Reduced Interhemispheric EEG Coherence in Alzheimer'S Disease: Analysis During Rest And Photic Stimulation. Alzheimer Dis. Assoc. Disord, Vol. 12. (1998) 175–181
8. Pritchard, W.S., Duke, D.W., Coburn, K.L.: Altered EEG Dynamical Responsivity Associated with Normal Aging and Probable Alzheimer's Disease. Dimentia, Vol. 2. (1991) 102–105
9. Jeong, J., Kim, S.Y., Han S.H., Nonlinear Analysis of Chaotic Dynamics Underlying EEGs in Patients with Alzheimer'S Disease. Electro Encephalogr Clin Neurophysiol, Vol. 106. (1998) 220–228
10. Goodin, D.S., Squires, K.C., Henderson, B., Starr, A.: Age-Related Variations in Evoked Potentials to Auditory Stimuli in Normal Human Subjects. Electroenceph clin Neurophysiol, Vol. 44. (1978) 447–458
11. Patterson, J.V., Michalewski, H.J., Starr A.: Latency Variability of the Compo-Nents of Auditory Event-Related Potentials of Infrequent Stimuli in Aging.Alzheimer-Type Dementia, and Depression. Electroenceph Clin Neuro-Physiol, Vol. 71. (1988) 450–460
12. Ray, P.G., Meador, K.J., Loring, D.W., Murro, A.M., Buccafusco, J.J., Yang, X.H., Zamrini, E.Y., Thompson, W.O., Thompson, E.E.: Effects of Scopolamine on Visual Evoked Potentials in Aging and Dementia. Electroenceph. Clin. Neurophysiol, Vol. 80. (1991) 347–351.
13. Davis, L.: Handbook of Genetic Algorithms. Van Nostrand Reinhold (1991)
14. Choi, J.: Expert Manual in Korean Dementia Rating Scale. (1998)