Medical Engineering & Physics 24 (2002) 349–360
www.elsevier.com/locate/medengphy

Automatic recognition of alertness and drowsiness from EEG by an artificial neural network

Aleksandra Vuckovic [a,*], Vlada Radivojevic [b], Andrew C.N. Chen [a], Dejan Popovic [a]

[a] Center for Sensory-Motor Interaction, Aalborg University, Aalborg, Denmark
[b] Institute for the Mental Health, Belgrade, Yugoslavia

Received 19 April 2001; received in revised form 27 February 2002; accepted 14 March 2002

Abstract

We present a novel method for classifying alert vs drowsy states from 1 s long sequences of full spectrum EEG recordings in an arbitrary subject. This novel method uses time series of interhemispheric and intrahemispheric cross spectral densities of full spectrum EEG as the input to an artificial neural network (ANN) with two discrete outputs: drowsy and alert. The experimental data were collected from 17 subjects. Two experts in EEG interpretation visually inspected the data and provided the necessary expertise for the training of an ANN. We selected the following three ANNs as potential candidates: (1) the linear network with Widrow-Hoff (WH) algorithm; (2) the non-linear ANN with the Levenberg–Marquardt (LM) rule; and (3) the Learning Vector Quantization (LVQ) neural network. We showed that the LVQ neural network gives the best classification compared with the linear network that uses WH algorithm (the worst), and the non-linear network trained with the LM rule. Classification properties of LVQ were validated using the data recorded in 12 healthy volunteer subjects, yet whose EEG recordings have not been used for the training of the ANN. The statistics were used as a measure of potential applicability of the LVQ: the *t*-distribution showed that matching between the human assessment and the network output was $94.37 \pm 1.95\%$. This result suggests that the automatic recognition algorithm is applicable for distinguishing between alert and drowsy state in recordings that have not been used for the training. © 2002 IPEM. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Alert; Drowsy; EEG; Time series; Cross-spectral density; Neural networks

1. Introduction

The goals of this study were as follows: (1) establishing a method of processing input data from a full spectrum of EEG recordings and (2) selecting an artificial neural network (ANN) that can distinguish between alert and drowsy states in an arbitrary subject by the use of processed EEG signals.

Spontaneous electrical brain activities, i.e. EEG signals, are dynamic, stochastic, non-linear and non-stationary [1–4]. The EEG recordings depend on the location of the electrodes, their impedance and the state of alertness. In addition, the EEG recordings vary substantially between healthy subjects. Extensive expertise is required to visually interpret the EEG recordings in order to isolate and identify characteristic information from a large amount of data. A computerized analysis of the EEG recordings aims to facilitate the time-consuming and difficult visual inspection [5] and automatically extract characteristic features of brain activity.

A computer-assisted EEG classification of drowsiness has been analysed in several studies [6–15]. The classification was based on a spectral analysis of EEG recordings [6,8,9,10] and showed that a limited number of electrodes and spectral analysis of characteristic bands could be used as a classifier. More recently, some studies [9,16] concentrated on detecting the information on drowsiness available from a full EEG spectrum. Principe et al. [12] designed a finite automaton that was capable of categorizing the sleep into seven different stages. McKeown et al. [17] used statistical methods for the analysis of EEG signals and detection of vigilance changes.

Pradhan [18] presented preliminary results for a classification of seizure activities by applying an ANN

* Corresponding author. Tel.: +45-96-35-87-58; fax: +45-98-15-40-08.
E-mail address: av@smi.auc.dk (A. Vuckovic).

1350-4533/02/$22.00
PII: S1350-4533(02)00030-9

based on a learning vector quantization. Kalayci and Ozdamar [19] showed that an ANN performs better if the input and output data can be processed to capture the characteristic features of the signal. They used a wavelet representation for automated detection of the EEG spikes. Wilson and Bracewell [14] developed a method of detecting the alert state by applying a wavelet preprocessing and an ANN, and used a binary output (alert and drowsy states).

The ANN has been proposed as a universal EEG classifier [11,16,20,21,22,23]. In [24] multilayer perceptron (MLP) and Learning Vector Quantization (LVQ) were trained in order to classify six different stages of vigilance from the EEG of infants. Parameters were obtained from EEG epochs of 30 s. The ANN was trained with data from three infants and tested with data from one infant. Yet, no significant differences in scoring rates of MLP and LVQ networks were recognized. Gevins and Smith [11] trained the ANN with Joseph Viglione algorithm to detect transient cognitive impairment in the EEG of nine subjects, and ANN's performance was tested on subjects not included in the training. A perceptron ANN was applied for detecting a left and a right index finger or a right foot from EEG records lasting 1 s [16]. Extended MLP with temporal processing inside the classifier [21] was also proposed for a classification of the EEG record, and performed on three subjects.

ANN have been applied, for example, to EEG recordings in order to generate a link between the conscious activities and environment, i.e. Brain Interface (BI) [16,20,25,26,27,28]. More recently, ANN that apply Bayesian methods are shown to be more robust compared with other techniques because they incorporate measures of confidence in their output for the Levenberg–Marquardt (LM) procedure [29,30]. In addition, standard MLP was improved by using finite impulse response filters (FIR) instead of static weights for a temporal processing of data [21].

The combination of cross-spectral analysis of an EEG with a LVQ in classifying alertness and drowsiness has been proposed as a suitable network to classify data obtained from EEGs [9,20,25].

The following reasons were the basis for improving the methods of automatic detection of changes from alert to drowsy and vice versa states: (1) clinical pre-processing of long-term recordings of wakefulness in order to select sequences of alert and drowsy states for further human inspection [31]; (2) on-line experiments where the timings of a stimulus for cognitive, evoked potential are needed; (3) software for interactive learning [32]; and (4) warning systems for detecting the drowsiness in operator rooms.

The specific design requirement was the applicability of the algorithm to short sequences of EEG recordings; hence, plausible use in real time. The second requirement was to develop a simple algorithm that would work on recordings that have not been used for the training of the same or arbitrary subject.

The study used the following definitions: The *alert state* refers to relaxed wakefulness and operationalized in EEG as the alert wakefulness with the occipital alpha rhythm present. The *drowsy state* is operationalized in both the drowsy EEG (i.e. presence of slow eye movement with the occipital alpha rhythm, a decrease in the amplitude, and/or frequency of the alpha rhythm, low-amplitude activity at the central and posterior EEG channels preceding the Stage 1 sleep) and Stage 1 EEG (non-REM). In our study, two experienced neurologists agreed upon the EEG interpretation of the vigilance status (alert vs drowsy). The agreement was based on a detailed visual inspection of the EEG and (electro-oculogram) EOG recordings, rather than on the behavioral states induced periodically in the manipulated tasks as suggested in [6,7,8,21]. In that way it was possible to obtain an EEG of relaxed wakefulness and drowsiness without interference from transient change in alertness induced by cognitive tasks.

Considering the dynamics and complexity of the EEG, it is instrumental to use the data that comprise both spatial and temporal information for the training of an ANN. As input to the ANN, we used the parameters determined from the EEG data based on the full spectrum. Based on the extensive literature [8,22,24,25,28] three ANNs were selected for this study: a Learning Vector Quantization (LVQ) [33], a feed-forward network that uses the Levenberg–Marquardt learning (LM) rule [34], and a linear network that uses Widrow–Hoff rules [34].

2. Materials and methods

2.1. Subjects and data preparation

2.1.1. Subjects

Nine males and eight females, age 25–35 yr (mean: 28) of normal intelligence and without mental disorders were included in the study after passing the neurological screening. The subjects were lying in a dark room with their eyes closed. All recordings were performed in accordance with the medical ethical standards after the subjects signed the informed consent approved by the local ethics committee. The subjects were not sleep-deprived, they had no deviations from their usual circadian cycle, and they took no medicine. A neurologist did not allow the subjects to fall asleep (i.e. no further than stage 1 of non-REM sleep) by closely monitoring their stage of vigilance, and by tapping at the desk to awake them when needed.

2.1.2. Data collection

An EEG (MEDELEC 1A97 EEG system, MEDILOG BV, Nieuwkoop, the Netherlands) was used in an electromagnetically shielded room during 30 min recording sessions. Ag/AgCl electrodes (impedances $R < 5$ kΩ) were positioned at 14 locations (F7, F8, T3, T4, T5, T6, F3, F4, C3, C4, P3, P4, O1, O2) following the International 10-20 System with an average reference, and left and right electro-oculogram (EOGL, EOGR). The recordings were band pass filtered between 0.5 and 70 Hz. The EEG recordings were digitized with 12-bit resolution at a sampling rate of 256 Hz per channel (A/D PCI board, Data Translation 2801, Marlboro, MA, USA).

2.1.3. Data preparation

The artifact was removed from the recordings by visual inspection. Two neurologists, with extended experience of interpreting the EEG, evaluated and rated the recordings used for this study. Each of them inspected the EEG/EOG recordings, and then agreed which EEG sequences clearly indicate a drowsy or alert state of a subject. There were sequences where they disagreed: slight diffuse slowing of α without a distinct decrease of amplitude or fragmentation, and very short intervals (< 1 s) which the experts did not classify as drowsiness. The sequences of disagreement of the neurologists were excluded from the evaluation. All data were sequenced to epochs lasting 1 s. We performed experiments 1–4 by using 60 epochs of alertness and 60 epochs of drowsiness (light drowsy and S1) in each subject. The sets were divided into two groups: (1) *the training set* consisted of 20 epochs of drowsiness and 20 epochs of alertness, and 2) *the validation set* consisted of 40 epochs of alertness and 40 epochs of drowsiness. The ANN performance was assessed on both the training and the validation set.

In experiment 5 we used 5 min of raw EEG (300 epochs) selected from a 30 min recording from 17 subjects. The criterion for selection of the 5 min interval was that the recordings had frequent transitions from alertness to drowsiness and vice versa. Also by subjects included in the training set, epochs selected for training were not included in these 300 epochs. These data were used only for testing the ANN performance. The ANN classification was then compared with the classification made by the two experienced neurologists.

This procedure differs from the classification adopted for detecting the drowsiness of operators in the control rooms, where only alertness can be wrongly diagnosed, and all 'problematic' EEG segments have to be classified as drowsiness.

3. Data processing

The work was performed in five experimental paradigms.

3.1. Parameterization of data

We used a moving cross-spectral density csd [Eq.] with signals lasting one epoch (1 epoch consisting of 256 samples=1 s). The interhemispeheric csd was calculated from homologous contralateral electrodes F7–F8, T3–T4, T5–T6, F3–F4, C3–C4, P3–P4, and O1–O2 [Eq. (2)]. The intrahemispheric csd was calculated from electrodes on the left hemisphere F8, T4, T6, F4, C4, P4, and O2. In this case, the csd was calculated for the signal lasting one epoch and the signal from the same electrode delayed for 50% of the epoch [Eq. (3)]. The equations determining the parameters are as follows:

$$\mathrm{csd}(x,y) = \sum_m \gamma_{xy}(m)\mathrm{e}^{-j\varpi m} \quad (1)$$

$$\mathrm{csd}^1 = \mathrm{csd}(x_i,y_i)\ i = 1,2,\ldots,K \quad (2)$$

$$\mathrm{csd}^2 = \mathrm{csd}(x_{1i},x_{2i}) \quad (3)$$

$$P_j = \frac{1}{J}\sum_{j=1}^{j}\sqrt{(\mathrm{Re}(\mathrm{csd}_j)) + (Im(\mathrm{csd}_j))^2}x_{1i} = x_i(n)\ n = 1,2,\ldots,N \quad (4)$$

$$x_{2i} = x_i\left(n + \frac{N}{2}\right) \quad (5)$$

where γ_{xy} is the cross-correlation function for signals x and y, $K = 7$ is the number of electrodes on one hemisphere, $N = 256$ is the number of digitized samples in one epoch, and x_i and y_i are signals from homologous (opposite) electrodes, m is discretized time, ω is frequency. x_{1i}, x_{2i} are signals from the same electrode, but the x_{2i} is delayed for 128 samples from the x_{1i}. P_j are the parameters that were used for the training of the ANN, and J is the number of parameters.

The csd^1 was chosen for this study because the correlation of homologue electrodes decreases in drowsiness. Likewise, the csd^2 detects the transition from alertness to drowsiness and vice versa because of the changes in the synchronization of the intrahemispheric EEG signal [35].

Since the csd does not comprise the needed temporal information (dynamics of the EEG), we also introduced time shifted $\mathrm{csd}_i = \mathrm{csd}(t)$ $\mathrm{csd}_{i+1} = \mathrm{csd}(t + T)$ and $\mathrm{csd}_{i+2} = \mathrm{csd}(t + 2T)$, $T = 1/N = 40$ ms [Eq. (6)], and treated correlations as a time series.

If we denote both csd^1 and csd^2 as csd, then

$$\mathrm{csd}_i = \mathrm{csd}(a_i,b_i),\mathrm{csd}_{i+1} = \mathrm{csd}(a_{i+1},b_{i+1}),\mathrm{csd}_{i+2} = \mathrm{csd}(a_{i+2},b_{i+2}),i = 1,2,\ldots,N \quad (6)$$

where $N = 256$ is the number of samples in a 1 s sequence of EEG recordings per channel, and a and b are EEG signals either from one hemisphere or from homologous hemispheres [Eq. (2) and (3)].

The number of parameters P_j calculated from csd^1 were $7 \times 3 = 21$, and the number of parameters P_j calculated from csd^2 were $7 \times 3 = 21$, resulting in the overall number of 42 parameters that were used as input to ANN for both the training and validation.

3.2. Network training

The following three artificial neural networks were used.

3.2.1. Linear network trained by Widrow–Hoff rule [33]

The linear network (Fig. 1) is recommended when the problem is linear or it can be approximated well enough with a linear model. The network can be trained iteratively using Widrow–Hoff rule.

The network used in the analysis had the following parameters: (1) 14 neurons, (2) 100,000 training epochs, and (3) the error goal SSE=0.1. The number of neurons (14) were obtained heuristically by increasing it until the rating of validation set improved sufficiently (SSE decreased less than 1%). There were two alternative stopping rules applied during the training of this network: the number of training epochs and the error goal. The error goal was never exceeded. The number of training epochs were increased (for different a numbers of neurons) until both the training error and the rating of the validation set reached the desired values.

During the training, the value of ‘1’ was given for alertness and ‘0’ for drowsiness at the output of the network. Since the output of linear ANN is continuous, we heuristically chose any output between -0.2 and 0.3 to represent drowsiness, and any output between 0.7 and 1.2 to reflect alertness. If the output did not belonged to the two intervals, this was considered as a failure of the network.

3.2.2. Feed-forward network trained with the Levenberg–Marquardt learning rule [34]

The feed-forward neural network (Fig. 2) had a tanh activation function in the hidden layer (layer 1), and a linear function in the output layer (layer 2). The Levenberg–Marquardt optimization technique, which uses an approximation of Newton’s method rather than a gradient descent technique, was chosen. Networks with biases, sigmoid layers, and linear output layers are known to be capable of approximating any function with an infinite number of discontinuities [36]. Generally, networks trained using the back propagation-based algorithms achieve good generalization. However, they can lead to local rather than global error minima.

The network included in this study had the following parameters: (1) 8 neurons in the hidden layer, (2) 75 training epochs, and (3) the error goal SSE=0.1. The stopping rules were either the error goal or the number of training epochs. The error goal was exceeded within the given number of epochs. These criteria were exceeded even with six neurons. When the ANN with a small number of neurons was used, the network could classify the training set, but failed to classify the validation set. Similar effects were noticed when a small error goal was selected; in this case the network was ‘over-trained’ on the training set, and it was unable to classify the data from the validation set. The output of Levenberg–Marquardt ANN is also continuous; there-

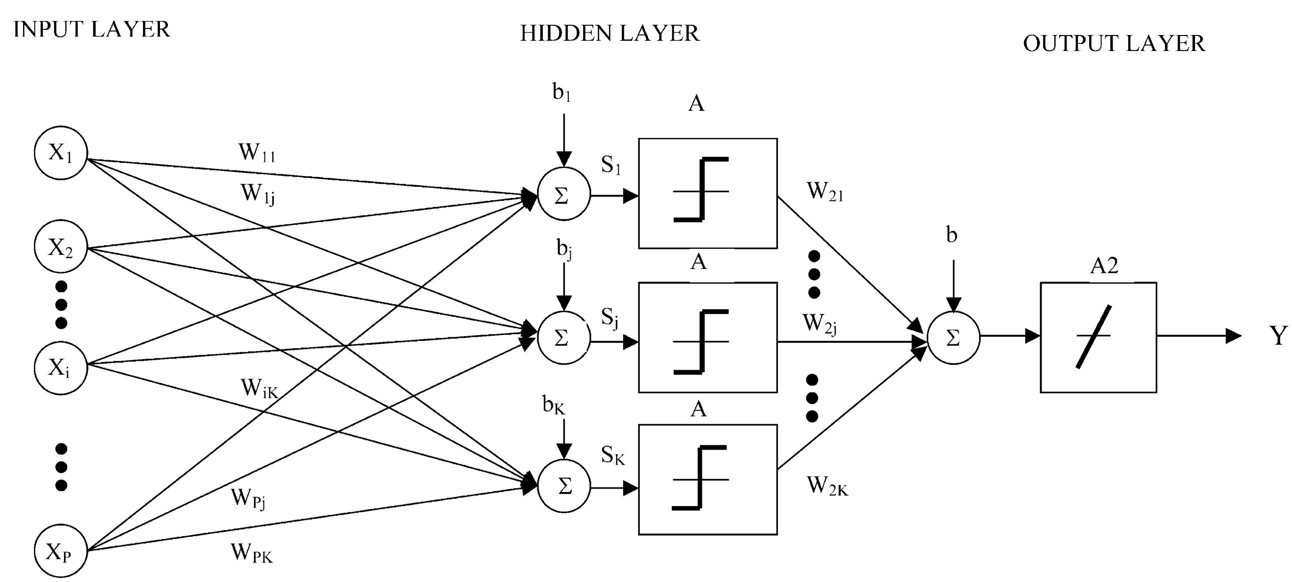

Fig. 1. Feed-forward linear neural network: $\mathbf{X}_i(i = 1\ldots P)$—input corresponds to the number of subject whose data were included in the training of ANN, $\mathbf{W}_{ij}(i = 1\ldots P, j = 1\ldots K)$—weights from input to hidden layer, $b_j(j = 1\ldots K)$—biases of the neurons in the hidden layer, A—activation function in the hidden layer, $\mathbf{W}_{2j}(j = 1\ldots K)$—weighting functions from hidden to output layer, b-bias of the output neuron, $A2$—linear activation function of the output layer, Y—output.

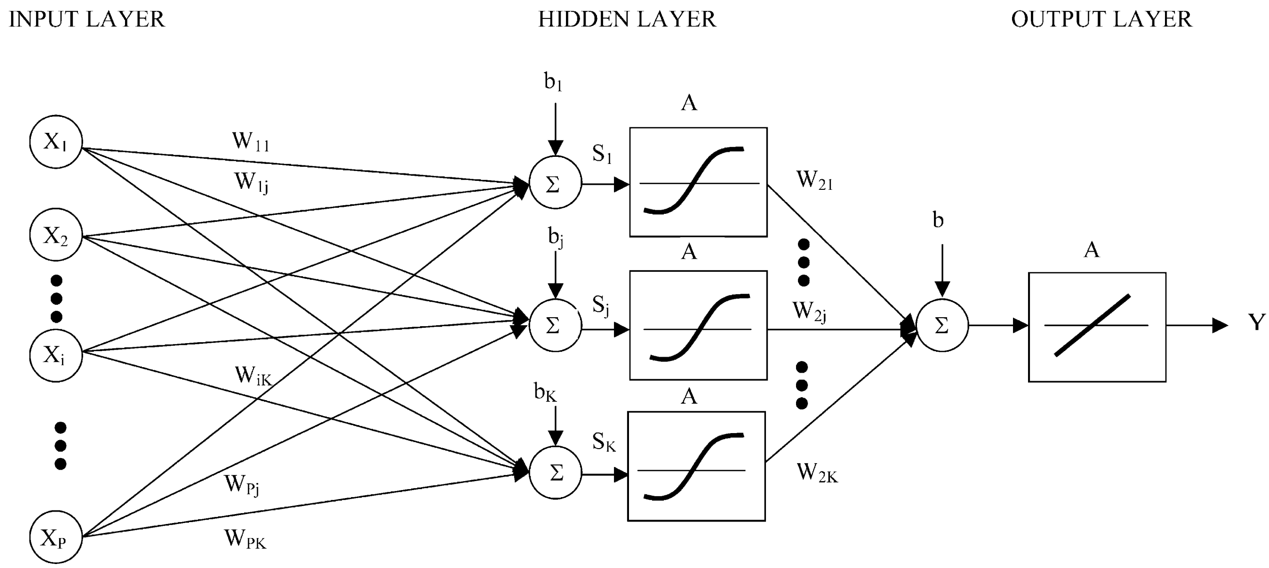

Fig. 2. Feed-forward neural network with sigmoid activation function: $\mathbf{X}_i(i = 1\ldots P)$—input corresponds to the number of subject involved in the training of NN, $\mathbf{W}_{ij}(i = 1\ldots P, j = 1\ldots K)$—weights from input to hidden layer, $b_j(j = 1\ldots K)$—biases of the neurons in the hidden layers, A—sigmoid activation function in the hidden layer, $\mathbf{W}_{2j}(j = 1\ldots K)$—weighting functions from hidden to output layer, b—bias of the output neuron, Y—output.

fore, the output was quantified as described earlier for the linear network.

3.2.3. Learning vector quantization rule [33]

This learning rule is a combined method for training a self-organizing network in a supervised manner. The LVQ learning rule is derived from the Kohonen rule [33]. The network architecture is shown in Fig. 3. The LVQ network consists of two layers. The first layer is a competitive layer, which learns to classify input vectors into one of the subclasses. The competitive neuron whose weight vector **W** forms the closest match with the input vector **X** is classified as output 1. The second linear layer transforms the subclasses of the competitive layer to the output target classes so that each competitive neuron has assigned one target (output). Both the competitive layer and the linear layer have one neuron per class; thus, the competitive layer can learn up to S_K sub-

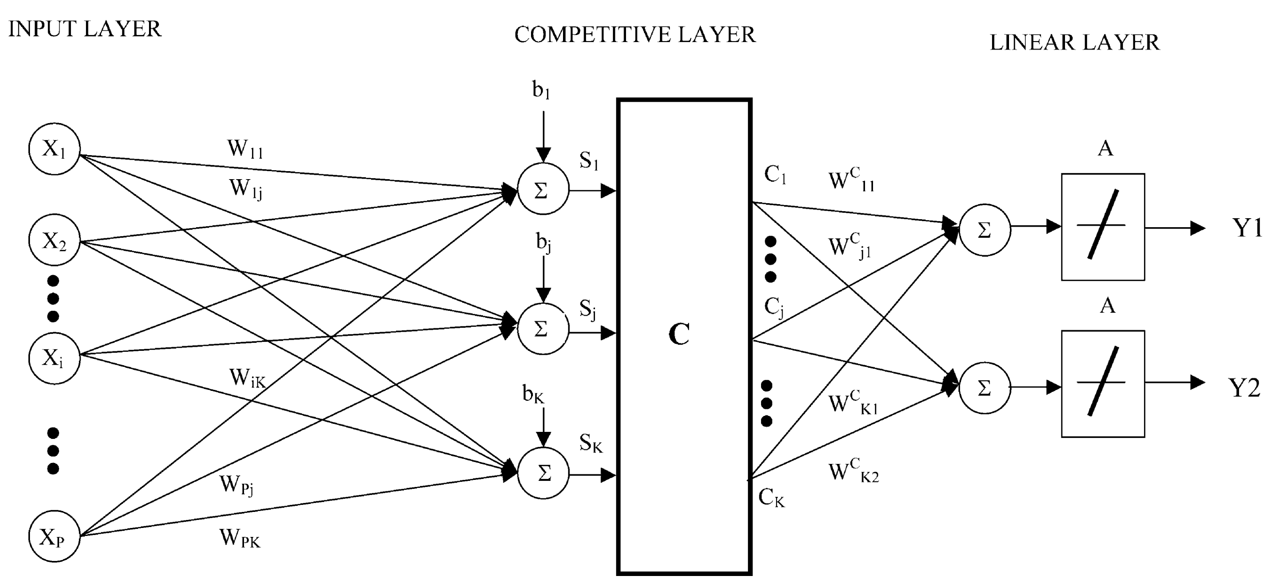

Fig. 3. The structure of neural network learned with learning vector quantization rule: The input and hidden layers are the same as in the competitive NN for unsupervised learning. $\mathbf{X}_i(i = 1\ldots P)$—input corresponds to the number of subjects involved in the training of NN, $\mathbf{W}_{ij}(i = 1\ldots P, j = 1\ldots K)$—weights from input to hidden layer, $b_j(j = 1\ldots K)$—biases, C—competitive layer, W^C_{1j} and $W^C_{2j}(j = 1\ldots K)$—weighting functions from k classes of the competitive layer to neurons 1 and 2 in the output layer, A—linear activation function, Y—two output, one corresponding to drowsiness and the other to alert state. Only one output can have the value 1 at the same time.

classes. This is in turn combined by the linear layer to form the Y_j ($j = 2$ in Fig. 3) target classes. The output neuron Y_j assigned to the winning competitive neuron also has a value of 1 while all the other output neurons have a value of 0. Depending on which neuron had a value of 1, the EEG recording was classified as drowsiness or alertness.

The network used in this study had the following parameters: (1) 12 neurons (classes) in competitive layer, (2) 2 neurons (classes) in output layer, and (3) 10,000 training epochs. The stopping rule was the number of the training epochs. The networks with 8 and 10 classes in the competitive layer were initially tested, but have generated results that were not acceptable. The network with 14 classes did not improve the scoring significantly (due to the fact that error goal was not the stopping criteria, performance of the network was esteemed as a percentage of good matches between ANN and human rating. If an ANN had a percentage of matches greater than 1% more than the other ANN, the improvement in scoring was considered significant). The testing was performed on the validation set. Two neurons in the output layer were used: drowsiness (0) and alertness (1). The LVQ did not need the quantization, and there were no ‘failure’ cases as was the case with the previously described two types of ANNs.

4. Statistical analysis

In all experiments, network scoring was presented as a mean value ± standard deviation SD [37]. These parameters were calculated separately for the training and the validation sets. Two-sided paired *t*-test (scores of ANN for alertness minus scores for drowsiness) was used in order to assess if scoring of an ANN is biased toward drowsiness or alertness. Finally, unpaired two-tail *t*-tests were used to compare data in the training and the validation set for LVQ trained on data of several subjects. In order to estimate the scoring of LVQ over the whole population, two-tailed *t*-distribution $n = 11$ (n degrees of freedom), ($P = 0.95$) was used.

5. Results

Five experiments were performed, graduated upon the difficulty of the given task. ANN that showed the poorest result in the previous experiment was eliminated from the next experiment.

5.1. Experiment 1

All three types of ANN were trained to classify the alertness and drowsiness using data from a single subject. The procedure was repeated on EEG recordings of three subjects. Data were extracted from the first data set according to Eq. (1–6). The mean values for the three subjects' data for all types of ANN are shown in Table 1.

The linear network gave the poorest result; hence, it was omitted from the following experiments. The poor classification was somewhat expected due to the nature of the EEG signals and the fact that the network is developed for linear signals. In addition, the total number of weight parameters is very high, compared with the size of the training set; hence, the ability to generalize is relatively low.

5.2. Experiment 2

The LM and LVQ networks were trained using data collected from three subjects in order to test the classification properties of ANNs on mixed data from several subjects. Therefore, the training set consisted of 3x20 epochs of alertness, and 3x20 epochs of drowsiness according to Eq. (1–6). The EEGs included in the training were obtained from subjects 1, 2, and 3 (Fig. 4). Typical appearances of drowsy states with the most obvious α activity were chosen from left to right: (1) anteriorization of α activity (8–13 Hz); (2) slowing and fragmentation of posterior α with occurrence of frontal α and diffuse slowing; and (3) loss of coherence of occipital α, without slowing. To validate this generalization, the data from seven other subjects were evaluated.

The results of experiment 2 are shown in Table 2. The LM network classified the drowsiness reasonably well (83.2 ± 18.3%) but detected the alertness poorly (36.3 ± 32.1%). The paired two-tail *t*-test between alertness and drowsiness [$t(9) = 13.47$, $P < 0.001$] showed that this ANN is very much biased towards the drowsiness. The LVQ network performed much better. It classified the drowsiness (97.6 ± 4.3%) better than the alertness (84.8 ± 13.1%). The results of the paired t test gave [$t(9) = 5.07$, $P < 0.001$] and showed that the LVQ was also biased towards the drowsiness.

Table 1
The percentage of matches between ANN and experienced neurologists for alertness and drowsiness of ANN trained on data of a single subject, training and validation set, mean value for repeated training on data of three different subjects. The notations are Linear for linear ANN, LM for ANN trained with the Levenberg–Marquardt rule, and LVQ for ANN trained with LVQ rule

	Alert training	Drowsy training	Alert validation	Drowsy validation
Linear (%)	75	75	65	65
LM (%)	100	100	100	100
LVQ (%)	100	100	100	100

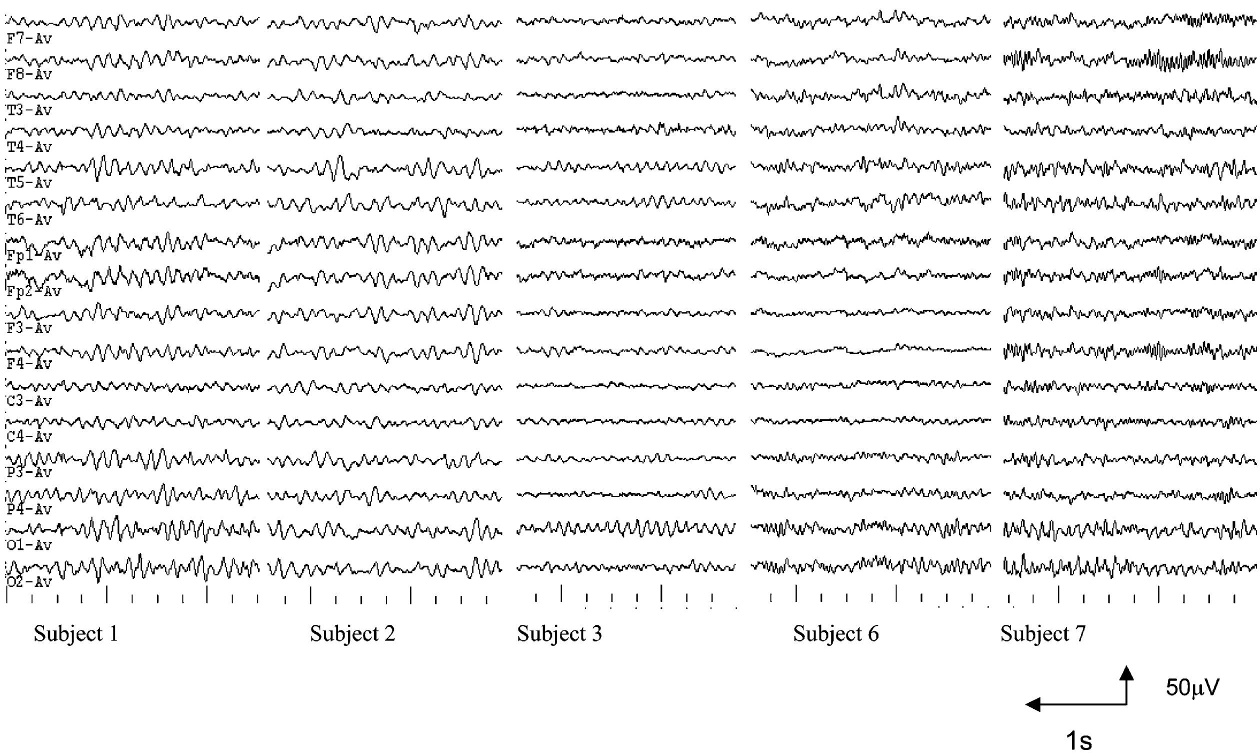

Fig. 4. Five EEG recordings of subjects included in the training of LVQ3.

Table 2
The percentage of good matches of ANN and experienced neurologists for alertness and drowsiness based on mixed data of three subjects, results for non-linear ANN. Data of subjects 1–3 were included in the training set, data of subjects 4–10 are included in the validation set. The abbreviations are: A_t—alert in the training set, D_t—drowsy in the training set, A_v—alert in the validation set, D_v—drowsy in the validation set

Sub. no.	Levenberg–Marquardt (LM) (%)				LVQ (%)			
	A_t	A_v	D_t	D_v	A_t	A_v	D_t	D_v
1	100	74	100	74	100	100	100	87
2	100	89	100	40	96	96	93	93
3	100	38	100	98	90	86	100	100
4		70		100		94		100
5		46		100		87		100
6		3		90		97		98
7		3		85		86		98
8		20		85		70		100
9		10		70		69		100
10		10		90		63		100

5.3. Experiment 3

Since the LVQ network showed better results in the sense of mean values of properly classified EEGs, we continued to use LVQ in further experiments. The training set was used from the same three subjects, and the validation set from the same seven subjects as in the previous experiment. In this experiment we varied the parameters used for training. Three different sets of parameters were used.

The first parameter set consisted only of *interhemispheric csd* from homologues electrodes, Eq. (1,2, 4–6) (only one half of the parameters used in experiment 2). The results are shown in Table 3, and the ANN is denoted as LVQ1. The scoring of ANN for the data in the validation set ($49.0 \pm 35.5\%$ for alertness and $98.3 \pm 2.6\%$ for drowsiness) was comparable to the LM network. A paired two-tail t-test gave [$t(9) = 12.29$, $P < 0.001$], showing significant difference in classifying alertness and drowsiness.

The second data set consisted of *intrahemispheric csd* according to Eq. (1, 4–6),, from both hemispheres. In

Table 3
The percentage of matches between LVQ and experienced neurologists based on different sets of data. The training was performed using data from three subjects. The LVQ1 is trained with interhemispheric csd from homologue electrodes, LVQ2 is trained with intrahemispheric csd from both hemispheres, LVQ3 is trained with intrahemispheric csd from left hemisphere and interhemispheric csd. Abbreviations as in Table 2

Sub. no.	LVQ1 (%)				LVQ2 (%)				LVQ3 (%)			
	Alert		Drowsy		Alert		Drowsy		Alert		Drowsy	
	A_t	A_v	D_t	D_v	A_t	A_v	D_t	D_v	A_t	A_v	D_t	D_v
1	100	100	95	97	95	82	95	95	100	100	100	87
2	90	78	95	92	85	82	95	95	96	96	93	93
3	95	87	100	97	95	89	100	100	90	86	100	100
4		90		100		87		99		94		100
5		27		100		57		100		87		100
6		10		100		65		95		97		98
7		9		100		63		98		86		98
8		35		100		10		100		70		100
9		29		97		10		100		69		100
10		25		100		10		100		63		100

this case, only the information about synchronization change on one single electrode was taken into account. The results are given in Table 3, and this ANN is denoted as LVQ2. The scoring of the LVQ2 (55.5 ± 33.1% for alertness, 98.2 ± 2.3% for drowsiness) was very similar to the previous network, although the training parameters were different. A paired two-tail *t*-test between alertness and drowsiness gave [$t(9) = 11.61$, $P < 0.001$], showing significant difference in classifying the alertness and drowsiness.

For a better comparison, the results of the LVQ network from experiment 2 are shown again in Table 3. The network is denoted as LVQ3 [both *inter* and *intrahemispheric csd*, Eq. (1–6)]. The *P* values for a paired two-tail *t*-test for the three networks show that there is a difference between classifying alertness and drowsiness for all three networks, but the *t*-values for LVQ3 are more than two times smaller compared with the values for the LVQ1 and LVQ2.

5.4. Experiment 4

The data from two more subjects were added in the training set, and LVQ3 was applied. The results for these two subjects were the worst in the previous experiment because their data sets were very different from the data sets used for the training of the LVQ. Thus, the training set consisted of 5x20 epochs of alertness, and 5x20 epochs of drowsiness. We selected data from subjects 6 and 7 shown in Table 3 (Fig. 4). These subjects had more pronounced β activity that diminished and vanished the occipital α (replaced by low amplitude fast β rhythms), moderate fragmentation, and decreased the amplitude of posterior α with increased amplitude fast β rhythms in frontal regions.

The results of experiment 4 are found in Table 4. The network scorings for the data in the validation set were: 95.1 ± 7.9% for alertness, and 93.4 ± 8.1% for drowsiness. The results were much better than those shown in Table 3. The classification was not biased toward drowsiness or alertness. The result of a paired two-tail *t*-test between alertness and drowsiness on validation set was [$t(4) = 0.115$, $P > 0.9$] for data from subjects 1, 2, 3, 6, and 7 (whose data from the training set were included in training of ANNs) and [$t(4) = 2.18$, $P > 0.3$] for data from validation set of subjects whose EEGs were not used for the training set.

We also assessed if there is a statistically significant difference between the scoring for subjects in and out

Table 4
The percentage of matches between experienced neurologists and ANN learned by LVQ, denoted as LVQ3 in Table 3. The network was trained with data from five subjects. The asterisk∗ shows the subjects whose data were only used in the validation set. Training relates to data that have been used for training, and Validation represents the data that have not been used for the training, but only to test the generalization

Sub. no.	Alert training	Drowsy training	Alert validation	Drowsy validation
1	100	100	100	90
2	100	80	100	75
3	100	100	100	100
6	75	100	75	100
7	85	85	88	100
4	—	—	98	98
5	—	—	97	89
8	—	—	97	88
9	—	—	100	94
10	—	—	96	100

of the training set, taking $P = 0.05$ as a statistically significant value. An unpaired two-tail *t*-test gave [$t(4) = 2.8$, $P > 0.02$], indicating a not statistically significant difference.

Fig. 4 shows the five different appearances of EEG records.

5.5. Experiment 5

The EEG data set 2, not previously rated, was passed through the network LVQ3. The results are shown in Table 5. The mis-classification denotes either unsuccessful classification of the network, or disagreement of the two expert neurologists' classification.

The result of a paired two-tail *t*-test on the validation set was [$t(16) = 0.115$, $P > 0.9$] which shows that the classification was not biased toward drowsiness or alertness.

The example of classifying drowsiness from alertness with LVQ3 network trained with data of five subjects is shown in Fig. 5. Channels 7 and 8 show EOG, and typical slow eye movement in drowsiness can be noticed. The bottom trace, denoted as channel 17, shows the network estimation with the time resolution of 1 s. Low signal was used for drowsiness and high signal for alertness. Fig. 5 suggests that the presented algorithm is capable of tracking very fast changes (1–5 s, swap between alertness and drowsiness).

6. Discussion

This study presents a method for classifying a state of vigilance to alert or drowsy states based on an ongoing EEG for an arbitrary healthy subject. The results indicate that for the purpose of classifying alert and drowsy states, the non-linear network with a discrete output (LVQ) performs better compared with the ANN with continuous output (LM). With an ANN having discrete output there is no need to heuristically choose the value of the output, which corresponds to the drowsy and alert states, and there are no undefined values in between.

A spectral analysis of EEG recordings [6,8,9,10] was proven to be a powerful tool for detecting transitions from an alert to a drowsy state when a cognitive task is induced. The spectral analysis used the fact that such an EEG comprises a characteristic rhythm that would disappear when a subject becomes drowsy. We specifically wanted to analyse conditions without cognitive tasks for the reasons listed earlier.

The use of a spectral analysis for detecting transitions from the alert to the drowsy state from recordings that do not include a specific cognitive activity is a difficult task. The statistical classification of EEG signals [17]

Table 5
The percentage of good matches of LVQ3 (experiment 5) and experienced neurologists for data obtained from raw EEG from which only artifacts were removed (by visual inspection), and which were not previously classified, lasting 5 min. Drowsy denotes which percentage of wrongly classified EEG belongs to drowsiness. Alert denotes which percentage of wrongly classified EEG belongs to alertness. The last row shows the percentage of drowsy state in the 300 s row EEG recording. The subjects 1, 2, 3, 6, and 7 have been included in the training of the ANN. The averaged performance is 94.37 ± 2.60%

Subject no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Percent	92	86	100	92	94	96	96	92	90	98	95	92	95	97	95	98	94
Drowsy	5	6	0	4	4	2	2	3	5	2	1	6	2	3	2	1	4
Alert	3	8	0	4	2	4	4	5	5	0	4	2	3	0	3	1	2
Perc. of drowsiness in the whole EEG	60	45	42	45	47	60	48	57	52	45	47	53	47	62	60	55	45

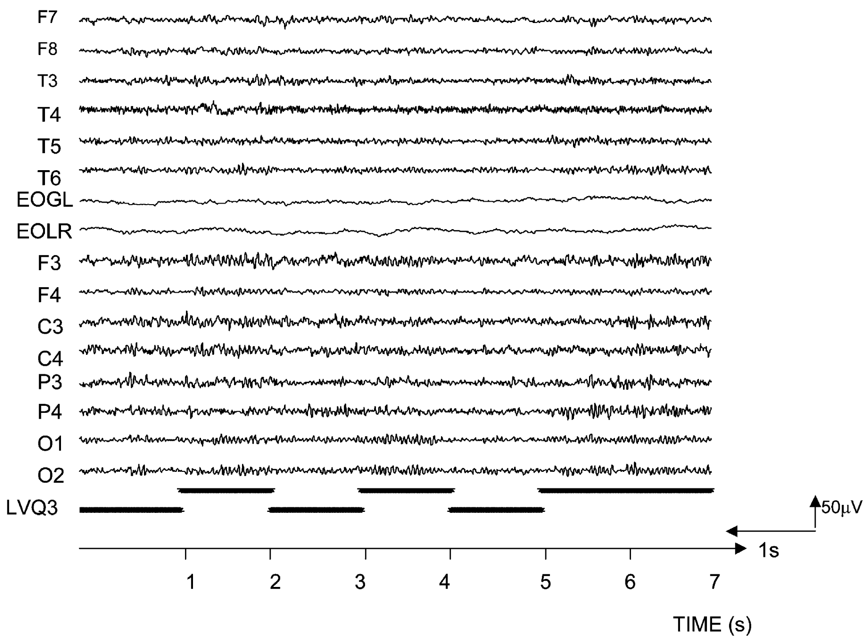

Fig. 5. Example of scoring with LVQ3 on raw EEG signal. Channels 7 and 8 are EOG (left and right), and channel 17 represents the estimation of drowsiness (low level) and alertness (high level) obtained by the neural network.

could be an effective method for classification and detection of changes in vigilance, though it was not used for distinguishing between the alert and drowsy states. These studies were an indication for the pre-processing of the EEG recordings. The initial parameters for all data sets in this study were based both on intrahemispheric and interhemispheric csd. To show that both csds are necessary, but not sufficient, we trained the LVQ network with interhemispheric csd (LVQ1) only, or with intrahemispheric csd (LVQ2) only, and both networks showed inferior results compared with the LVQ3 network trained on both interhemispheric and intrahemispheric csd (Table 3). In addition, the study used three consecutive values of csd and treated them as time series. Treating EEG as time series was an important element of data preprocessing, which had an important impact on generalization properties of the ANN, both LVQ and LM.

The ANN raised high hopes concerning promising results with respect to analysis, classification, pattern recognition, and functional monitoring [22]. This has been especially true for areas such as signal processing, including the analysis and interpretation of EEG recordings during cognitive load, rest, or sleep. The general conclusion from the studies briefly described in the Introduction was that the ANN output could detect the changes; thus, they could be used as universal EEG classifiers [11,16,20,21,22,23]. Although some studies [24] found no difference in the classification of an ANN with a discrete and a continuous output, we found the LVQ favorable to the given problem and the given training set. The combination of cross-spectral analysis of EEG with LVQ in classifying alertness and drowsiness was shown previously to be a suitable algorithm for classifying events from raw EEG signals [9,20,25], but for specific conscious tasks.

In the first tested ANN, a linear network was used primarily because of its structure and a simple learning algorithm. This type of ANN failed to classify data efficiently (Table 1) because the data set used in this study was either too complex or the amount of data was not enough for a given ANN. Both LVQ and LM were tested on mixed data from several subjects. However, when the number of different appearances of the EEG record was not sufficient, the classification of both ANNs was biased to drowsiness (experiments 2 and 3). It was independent on the type of ANN, and on the parameters of a training set. The LVQ network showed better results compared with the LM when tested on mixed data of three subjects for both the training and validation sets (Table 2). Therefore, this network was shown to be a candidate for the efficient classifier as already suggested in the literature [14,25].

The results of this study are similar to those presented by Wilson and Bracewell [14]; the difference is that they used a wavelet pre-processing method. The use of wavelet neural networks is not always a straightforward procedure and requires substantial computing power. The goal of this study was to develop a simple algorithm, which could also be applied in real-time. In [29], the principal component analysis was used for data pre-processing for training the feed-forward ANN. EEG recording was used from only two channels that decreased the amount of processed data. An ANN was efficiently used for the prediction of an ongoing drowsiness.

It should also be mentioned that classifying of drowsiness was successfully achieved (10 from 12 subjects) with some other tool such as a type of expert system [13].

Although much more sophisticated ANN and methods for signal analysis (such as wavelets, principal component analysis) exist according to the results obtained (high percent of good classification, unbiased rating), we considered our choice adequate for a given problem.

The main difference between this and the previous systems for detecting drowsiness is that it is applicable to an arbitrary healthy subject. The efficiency of the novel technique can best be explained by using the result of experiment 5. LVQ was trained on five subjects, yet the performance was tested using recordings from 17 subjects. The *t*-distribution was applied to data from 12 subjects (Table 5) not included in the training (with variance estimates) for subjects 4, 5, 8, 9, 10–17 in Fig. 4. This distribution demonstrates a match of mean ($t(11)*s/(n^{1/2}) = 94.37 \pm 1.95\%$ would be obtained between LVQ and expert human assessment. A paired two-tail *t*-test indicates that there is no difference in classifying alert and drowsy state.

This paper clearly demonstrates that the novel method is applicable for distinguishing between alert and drowsy states of arbitrary humans. It also shows that the classifier works even better if the data of a given subject were used for the training. The qualities of the method are that it is simple to apply, and it does not require high computation power. The method can be used as a standalone tool, but it can be implemented as a building block of a brain–computer interface for computer-assisted EEG diagnostics.

Acknowledgements

The work on this project was partly supported by the Ministry for Science and Technology of Serbia, Belgrade. We would like to acknowledge Mr Predrag Sukovic for providing us with some of the EEG recordings used for this analysis. Our sincere thanks to Prof. Dr Zarko Martinovic for second opinions in the visual analysis and scoring of our EEG records. We thank all subjects for participating in the experiments.

References

[1] Nunez P. Neocortical dynamics and human EEG rhythms. Oxford University Press: New York, 1995.
[2] Radivojevic V, Timotijevic D, Rajkovic M, Car M. Deterministic chaos in EEG signal. Consciousness—scientific challenge of the 21st century. Belgrade: Cigoja Press, 1996.
[3] Galka A. Topics in non-linear time series analysis: with implications for EEG analysis. Advanced series in non-linear dynamics. World Science Publishers, 2000.
[4] Santamaria J, Chiappa KH. The EEG of drowsiness in normal adults. J. Clin. Neurophysiol. 1987;4(4):327–82.
[5] Escourrou P, Luriau S, Rehel M, Nedelcoux H, Lanoe JL. Needs and costs of sleep monitoring. Stud Hlth Technol Inform 2000;78:69–85.
[6] Jung T-P, Makieg S. Estimating level of alertness from EEG. Proceedings of the 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1994:1103–4.
[7] Makeig S, Jung TP. Tonic, phasic and transient EEG correlates of auditory awareness in drowsiness. Cogn Brain Res 1996;4:15–25.
[8] Makeig S, Jung TP, Sejnowski TJ. Using feedforward neural networks to monitor alertness from changes in EEG correlation and coherence. In: Advances in neural information processing systems. Cambridge, MA: MIT Press; 1996. p. 931–7.
[9] Jung TP, Makeig S, Stensmo M. Sejnowski. T.J. Estimating alertness from the EEG power spectrum. IEEE Trans Biomed Engng 1997;44:60–9.
[10] Doghramji K, Doghramji K, Miller Merrill M, Sangal SB. A normative study of the maintenance of wakefulness test (MWT). Electroenceph Clin Neurophysiol 1997;103:554–62.
[11] Gevins A, Smith ME. Detecting transient cognitive impairment with EEG pattern recognition methods. Aviat Space Environ Med 1999;70(10):1018–24.
[12] Principe JC, Gala SK, Chang TG. Sleep staging automation based on the theory of evidence. IEEE Trans Biomed Engng 1989;36:503–9.
[13] Jansen B, Dawant BM. Knowledge-based approach to sleep EEG analysis- a feasibility study. IEEE Trans Biomed Engng 1989;36:510–8.
[14] Wilson BJ, Bracewell TD. Alertness monitor using neural networks for EEG analysis. Proc Neur Net Signal Process X (ISPS) 2000;2:814–20.
[15] Huang RS, Kuo CJ, Tsai LL, Chen OTC. EEG pattern recognition-arousal states detection and classification. In: Proc IEEE Intern Conf Neural Networks 1996; 2: 641–6.
[16] Peters BO, Pfurtscheller G, Flyvbjerg H. Automatic differentiation of multichannel EEG signals. IEEE Trans Biomed Engng 2001;48:111–6.
[17] McKeown MJ, Humphries C, Achermann P, Borbely AA, Sejnowski TJ. A new method for detecting state changes in EEG: exploratory application to sleep data. J Sleep Res. 1997;7:48–56.
[18] Pradhan N, Sadasivan PK, Arunodaya GR. Detection of seizure activity in EEG by an artificial neural network: a preliminary study. Comput Biomed Res 1996;29(4):303–13.
[19] Kalayci T, Ozdamar O. Wavelet preprocessing for automated neural network detection of EEG spikes. IEEE Eng Med Biol Mag 1995;16:0–166.
[20] Pfurtscheller G, Flotzinger D, Mohl W, Peltoranta M. Prediction of the side of hand movements from single-trial multi-channel EEG data using neural networks. Electroenceph Clin Neurophysiol 1992;82(4):313–5.
[21] Haselsteiner E. Pfurtscheller G. Using time-dependent neural networks for EEG classification. IEEE Trans Rehab Eng 2000;8:457–63.
[22] Dorffner G. Can neural networks improve signal processing? A critical assessment from the ANNDEE project. Proceedings of Measurement ‘97. 1997
[23] Dorffner G, Rappelsberger P, Flexer A. Using self-organising feature maps to classify EEG coherence maps. Heidelberg: Springer-Verlag, 1993.
[24] Pfurtscheller G, Flotzinger D. Maruschik K, Sleep classification in infants based on artificial neural networks. Biomed Tech (Berlin) 1992;37(6):122–30.
[25] Flotzinger D, Kalcher J, Pfurtscheller G. EEG classification by learning vector quantization. Biomed Tech (Berlin) 1992;37(12):303–9.

[26] Pregenzer M, Pfurtscheller G, Flotzinger D. Selection of electrode positions for an EEG-based brain computer interface (BCI). Biomed Tech (Berlin) 1994;39(10):264–9.

[27] Pfurtscheller G, Kalcher J, Neuper C, Flotzinger D, Pregenzer M. On-line EEG classification during externally-paced hand movements using a neural network-based classifier. Electroenceph Clin Neurophysiol 1996;99(5):416–25.

[28] Chen JJ, Tsai JJ, Sheu CH, Lin HY, Yeh JG. Application of dipole modeling in localization of mesio-temporal epileptogenic focus. Proc Natl Sci Counc Repub China B 1997;21(2):61–70.

[29] Bishop CM. Neural networks and their applications. Rev Sci Instrum 1994;65(6):1803–32.

[30] MacKay DJC. Probable networks and plausible predictions - a review of practical bayesian methods for supervised neural networks. Network 1995;6:469–505.

[31] Maus A, Enderson J. Misuse of computer-generated results. Med Biol Eng Comput 1979;17:126–9.

[32] Akay M, Marsic I, Medl A, Bu GA. System for medical consultation and education using multimodal human/machine communication. IEEE Trans Inf Tech Biomed 1998;2:282–91.

[33] Kohonen T. The self-organising map. In: Neural Networks. New York: IEEE Press; 1992. p. 74–90.

[34] Wasserman P. Neural computing: theory and practice. New York: ANZA Research, 1989.

[35] Wada Y, Nanbu Y, Koshino Y, Shimada Y, Hashimoto T. Inter- and intrahemispheric EEG coherence during light drowsiness. Clin Electroenceph 1996;27:84–8.

[36] Hornik K. Multilayer networks are universal aproximators. Neur Net 1989;2:359–66.

[37] Armitage P, Berry G. Statistical methods in medical research. Cambridge: Blackwell Science, 1995.