EEG Signals Classification by Using an Ensemble TPUnit Neural Networks for the Diagnosis of Epilepsy

Hiroki Yoshimura, Tadaaki Shimizu, Maiya Hori, Yoshio Iwai, and Satoru Kishida

Abstract—**The electroencephalogram (EEG) is necessary for the diagnosis of epilepsy. To make a diagnosis of epilepsy exactly, a full EEG recording for a long stretch of time is needed. The observation for a long record is a big burden for a doctor. To reduce this burden, a computer aid is important. This paper presented classifications of EEG patterns using the ensemble TPunit NNs for the diagnosis of epilepsy. The classification accuracy rates of the proposed classifiers were found to be higher than that of stand alone neural network. In addition, the classification accuracy was higher than previous study. The ensemble of the TPUnit neural networks is highly effective in classification problem.**

Index Terms—**Neural network, epilepsy, EEG, TPUnit.**

I. INTRODUCTION

The electroencephalogram (EEG) is necessary for the diagnosis of epilepsy. The EEG appears as epileptic EEG abnormality by hyper excitability of neurons to cause epileptic seizures. Therefore, if such abnormal EEG appears, it can be diagnosed with epilepsy in many cases. To make a diagnosis of epilepsy exactly, a full EEG recording for a long stretch of time is needed. The observation for a long record is a big burden for a doctor. To reduce this burden, a computer aid is important. Neural networks have been used in medical diagnosis decision support systems because of the belief that they have high prediction ability [1]-[3]. Many report have been shown that combing the predictions of several models often results in a prediction accuracy that is higher than that of the individual models [4]-[6]. From this point of view, a classification of EEG patterns using a neural network has been reported [7]. It presented the use of neural networks to combine the predictions of ensemble neural networks for classification of the EEG signals. The classification accuracy rate by the ensemble neural network was higher than that by the stand-alone neural network.

On the other hand, we have proposed a new neuron unit whose activation function is trigonometric polynomial. We call proposed neuron unit as Trigonometric Polynomial Unit (TPUnit). We have showed that 1) the neural network using TPUnits (TPUnit NN) can learn in a smaller number of hidden units than conventional neural network with sigmoid output and 2) it has high generalization ability in some computational experiments [8]. In this study, we improve the classification capability by using an ensemble TPUnit NN instead of conventional neural network with sigmoid output.

II. COMBINED TPUNIT NEURAL NETWORK

A. TPUnit Neural Network

Fig. 1 shows proposed neural network diagram with TPunits (TPUnit NN). As shown in Fig. 1, the TPUnit NN has three layers. The TPUnit NN has TPUnits on the hidden layer (Fig.2). While a typical neural network, we call it sigmoid neural network (Sigmoid NN), has neuron units with sigmoid activation function. Where, the input layer units have a linear activation function and the output layer unit has sigmoid activation function.
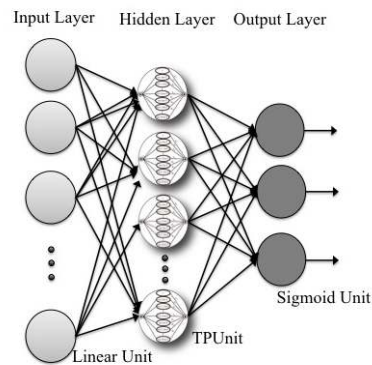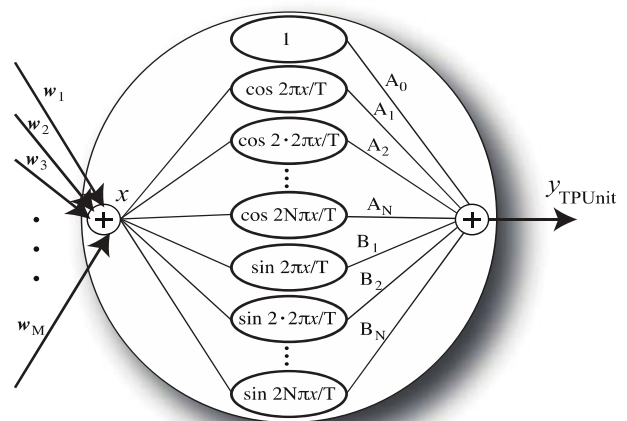

Fig. 1. TPUnit neural network

Fig. 2. Trigonometric polynomial unit (TPUnit)

B. Trigonometric Polynomial Unit

Fig. 2 shows the configuration of TPUnit. The output function of the TPUnit is expressed as follows:

$$y_{TPUnit} = A_0 + \sum_{n=1}^{N}\left(A_n \cos\frac{2n\pi}{T}x + B_n \sin\frac{2n\pi}{T}x \right) \qquad (1)$$

where A_n and B_n are adjustable parameters. N is a number of terms and T is period. N and T are constant. A_n, B_n and weight parameters are modified by back propagation algorithm using $\partial E/\partial A_n$, $\partial E/\partial B_n$.

A neuron unit with sigmoid function can only divide n-dimensional space linearly. In contrast, a TPUnit has nonlinear separation ability. Because the TPUnit has high spatial separation ability compared with a sigmoid neuron unit, it is expected to reduce the hidden units. We had shown that 1) the TPUnit NN can learn in a smaller number of hidden units than conventional neural network with sigmoid output (Sigmoid NN) and 2) TPUnit NN has high generalization ability in some computational experiments [8].

A back-propagation algorithm is used for modification of weight parameters and adjustable parameters A_n and B_n in order to reduce the following learning error E.

$$E = \frac{1}{2}\sum_{p=1}^{P}\sum_{m=1}^{M}\left(y_m - t_{pm}\right)^2 \qquad (2)$$

where, y_m is output of the output layer units. t_{pm} is target signals. M is a total number of output units. P is a total number of learning patterns.

$\partial E/dw$ is calculated from equation (2). After that, the weight parameters are modified by the following equation:

$$w[l+1] = w[l] - \eta\frac{\partial E}{\partial w} + \gamma\left(w[l] - w[l-1]\right) \qquad (3)$$

where, l is learning times. Third term in equation (3) is the momentum term. η is learning coefficient. γ is momentum coefficient. When the weight parameters from the input layer to the hidden layer are modified, it is noted that $\partial y_{TPUnit}/\partial x$ (eq. (4)) is used in calculating the $\partial E/\partial w$.

$$\frac{\partial y_{TPUnit}}{\partial x} = \sum_{n=1}^{N}\left(\frac{2\pi n}{T}A_n\cos\frac{2\pi n}{T}x - \frac{2\pi n}{T}B_n\sin\frac{2\pi n}{T}x\right) \quad (4)$$

Similarly, adjustable parameters A_n and B_n in the TPUnits are modified by the following equation:

$$A_n[l+1] = A_n[l] + \alpha\frac{\partial E}{\partial A_n} + \beta\left(A_n[l] - A_n[l-1]\right) \quad (5)$$

$$B_n[l+1] = B_n[l] + \alpha\frac{\partial E}{\partial B_n} + \beta\left(B_n[l] - B_n[l-1]\right) \quad (6)$$

$$\frac{\partial E}{\partial A_n} = \frac{\partial E}{\partial y_{TPUnit}}\frac{\partial y_{TPUnit}}{\partial A_n} = \frac{\partial E}{\partial y_{TPUnit}}\sin\frac{2\pi n}{T}x \quad (7)$$

$$\frac{\partial E}{\partial B_n} = \frac{\partial E}{\partial y_{TPUnit}}\frac{\partial y_{TPUnit}}{\partial B_n} = \frac{\partial E}{\partial y_{TPUnit}}\cos\frac{2\pi n}{T}x \quad (8)$$

where, l is learning times. Third term in equation (3) is the momentum term. α is learning coefficient of TPUnit. β is momentum coefficient of the TPUnit.

C. An Ensemble TPUnit Neural Networks

In order to improve the generalization ability, the ensemble learning to integrate the individual responses of multiple learning machines has been proposed [4]-[6]. In the ensemble learning using neural networks, improvement of generalization capability have been made on the basis of variety of neural network that depends on the initial value of weights etc. for the training [7].

We perform the classification of the EEG using the T ensemble TPUnit NNs of two types. The configuration diagrams of classifiers with the ensemble TPUnit NNs are showed in Fig. 3. As shown in Fig. 3(a), the multiple TPUnit NNs are set at the first level. Each TPUnit NNs are learned with the different initial values of weights. At second level, the outputs of the TPUnit NN are processed as simple ensemble. The arithmetic mean of each TPUnit NN outputs corresponding to the classification category is calculated. We call classifier in Fig.3 (a) a Type I. Fig. 3 (b) shows a classifier type II. As with the type I, the multiple TPUnit NNs are set at the first level and are learned by the different initial value. At second level, one TPUnit NN is set. All of the outputs of the TPUnit NNs in the first level are input to the TPUnit NN at the second level.
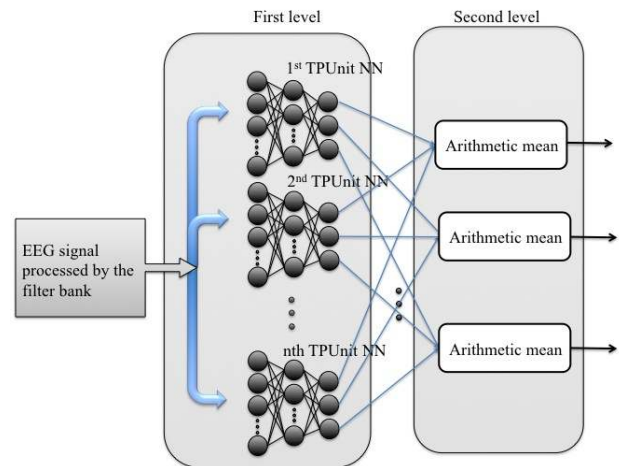

(a) Type I
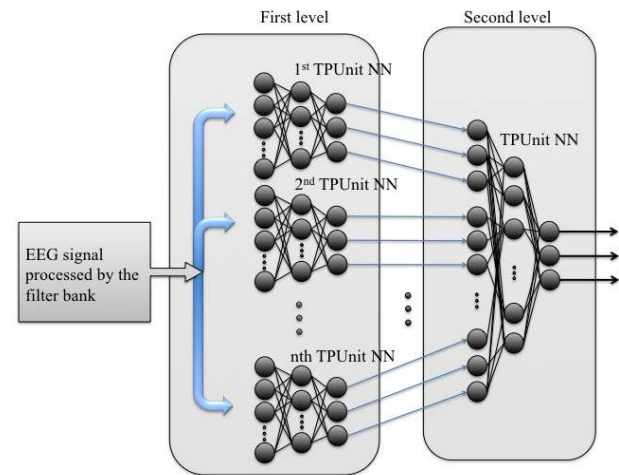
(b) Type II

Fig. 3. The configuration diagrams of classifiers with the ensemble TPUnit NNs

III. EXPERIMENT

In this study, we perform the classification of the following three classes of EEG.

1. EEG signals recorded from healthy volunteers with eyes open

2. EEG signals recorded from epilepsy patients in the epileptogenic zone during a seizure-free interval

3. EEG signals recorded from epilepsy patients during epileptic seizures
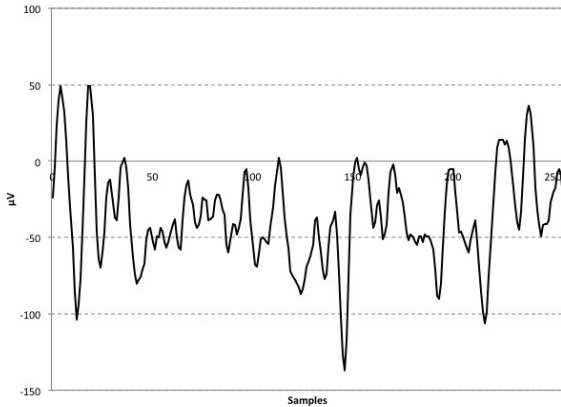
(a) healthy volunteers with eyes open
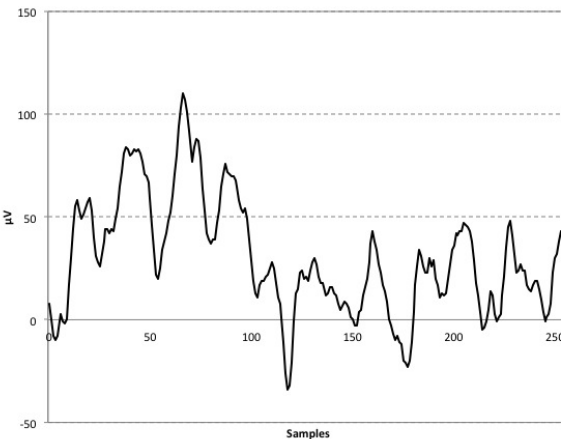
(b) epilepsy patients in the epileptogenic zone during a seizure-free interval
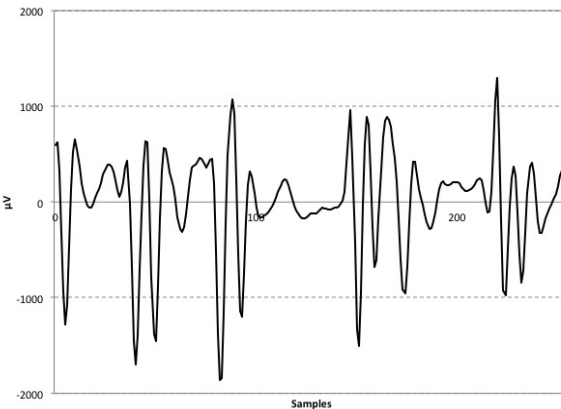
(c) epilepsy patients during epileptic seizures
Fig. 4. An example of a waveform of the EEG time series

The figure X shows an example of a waveform of the EEG time series described above.

The data described in Ref.[9], which is publicly available (http://epileptologie-bonn.de/cms/front_content.php?idcat=1 93&lang=3&changelang=3) ,was used. The sampling rate of the data was 173.61 Hz. The spectral bandwidth of the time series is from 0.5 Hz to 40 Hz by filtering.

In this experiment, we used signals processed by the filter bank of 256 sample points for time series of EEG waveform. The Filter bank divided into 12 bands from 0.5 Hz to 40 Hz. Therefore, the number of input units of each the TPUnit NNs in the first level is set to 12. The number of hidden units is set to 15. The number of output units is set to 3 in order to classify the EEG patterns into three classes. The number of the TPUnit NNs in the first level is set to 10. For each TPUnit NN, learning was performed with different initial values (weights etc.). The initial values of weights, An and Bn were set randomly in the range -0.1 to 0.1.

At the second level of the type II, the TPUnit NN has 30 input units (3 output units per TPUnit NN at the first level × 10 TPunit NNs), 15 hidden units and 3 output units (in order to classify into three classes). The initial values of weights,

An and Bn were set randomly in the range -0.1 to 0.1. In the all of the TPUnits, the number of terms is N=7 and the period is T=15.

The number of data patterns used in the learning phase was 3600 patterns (Healthy segment: 1200 patterns, Epileptogenic zone segment: 1200 patterns, Epileptic seizure segment: 1200 patterns). We evaluated classifiers using unlearned data patterns (Healthy segment: 400 patterns, Epileptogenic zone segment: 400 patterns, Epileptic seizure segment: 400 patterns).

IV. RESULT

The confusion matrixes showing the classification results of the type I and the type II are given in Table I. The test performance of the classifiers can be determined by the computation of specificity, sensitivity and total classification accuracy. The specificity, sensitivity and total classification accuracy are defined as:

Specificity: number of true negative decisions/number of actually negative cases;

Sensitivity: number of true positive decisions/number of actually positive cases;

Total classification accuracy:
Number of correct decisions/total number of cases.

TABLE I: CONFUSION MATRIX
TYPE I

Output/desired	Healthy segment	Seizure free epileptogenic zone segment	Epileptic seizure segment
Healthy segment	395	1	19
Seizure free Epileptogenic zone segment	0	398	10
Epileptic seizure segment	5	1	371

TYPE II			
Output/desired	Healthy segment	Seizure free epileptogenic zone segment	Epileptic seizure segment
Healthy segment	387	0	6
Seizure free epileptogenic zone segment	1	393	2
Epileptic seizure segment	12	7	392

A true negative decision occurs when both the classifier and the physician suggested the absence of a positive detection. A true positive decision occurs when the positive detection of the classifier coincided with a positive detection of the physician.

The classification accuracies (specificity, sensitivity, total classification accuracy) on the test sets are presented in Table II. As it is seen from Table II, the type I classified healthy segments, seizure free epileptogenic zone segments, epileptic seizure segments with the accuracy of 98.75%, 99.50%, 92.75%, respectively. The healthy segments, seizure free epileptogenic zone segments, epileptic seizure segments were classified with the accuracy of 97.00%. The type II classified healthy segments, seizure free epileptogenic zone segments, epileptic seizure segments with the accuracy of 96.75%, 98.25%, 98.00%, respectively. The healthy segments, seizure free epileptogenic zone segments, epileptic seizure segments were classified with the accuracy of 97.67%. High classification accuracy was obtained in the type I and the type II. In addition, the classification accuracy was higher than ref. [7] although not directly comparable. In particular, improvement of the total classification accuracy of about 3% was obtained in type II.

TABLE II: THE CLASSIFICATION ACCURACIES

Statistical parameters	Type I	Type II	Ref. [7]
Specificity	98.75%	96.75%	96.00%
Sensitivity (seizure free epileptogenic zone segment)	99.50%	98.25%	94.50%
Sensitivity (epileptic seizure segment)	92.75%	98.00%	94.00%
Total classification accuracy	97.00%	97.67%	94.83%

The classification accuracy rates of the type I and the type II were found to be higher than that of stand alone TPUnit NN. This may be attributed to several factors including the estimation of the network parameters and the scattered and mixed nature of the features.

V. CONCLUSION

This paper presented classifications of EEG patterns using the ensemble TPunit NNs for the diagnosis of epilepsy. The type I used the simple ensemble, which is the arithmetic mean of each TPUnit NN outputs corresponding to the classification category is calculated. The type II is the ensemble by the TPUnit NN to combine the predictions of the TPUnit NNs at the first level. From experimental results, high classification accuracy was obtained in the type I and the type II. In addition, the classification accuracy was higher than previous study [7] although not directly comparable.

The classification accuracy rates of the type I and the type II were found to be higher than that of stand alone TPUnit NN. The ensemble of TPUnit NNs is highly effective in classification problem.

Acknowledgment

REFERENCES

[1] D. West and V. West, "Model selection for a medical diagnostic decision support system: A breast cancer detection case," *Artif. Intell. Med*, vol.20, no. 3, pp.183–204, 2000.

[2] D. West and V. West, "Improving diagnostic accuracy using a hierarchical neural network to model decision subtasks," *Int. J. Med. Inform*, vol. 57,no. 1, pp.41–55, 2000.

[3] H. Kordylewski and D. Graupe, "K. Liu, A novel large-memory neural network as an aid in medical diagnosis applications," *IEEE Trans. Inform. Technol. Biomed*, vol. 5, no. 3, pp.202–209, 2001.

[4] M. Taniguchi and V. Tresp, "Combining regularized neural networks," in *Proc, the ICANN'97, in: Lecture Notes in Comput. Sci*, vol. 1327, Springer, Berlin, pp. 349–354, 1997.

[5] E.D. Übeyli and I. Güler, "Improving medical diagnostic accuracy of ultrasound Doppler signals by combining neural network models," Comput. Biol. Med, vol. 35, no. 6, pp.533–554, 2005.

[6] I. Güler and E. D. Übeyli, "ECG beat classifier designed by combined neural network model," *Pattern Recogn*, vol. 38, no.2, pp. 199–208, 2005.

[7] E. D. Übeyli, "Combined neural network model employing wavelet coefficients for EEG signals classification," *Digital Signal Processing*, vol. 19, pp. 297-308, 2009.

[8] H. yoshimura, T. Shimizu, M. Hori, T. Matsumura and Y. Iwai, "Proposal of a Trigonometric Polynomial Unit for Neural Networks," 2*012 International Workshop on Nonlinear Circuits, Communications and Signal Processing, NCSP*, vol. 12, pp.361-364, 2012.

[9] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C.E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev.* E 64, 061907, 2001.

Hiroki Yoshimura graduated from Tottori University in 1993 and completed the M.S. and doctoral programs in 1995 and 1998, respectively. I received my Ph.D. degree from Totori University. I am currently an asisitant professor in the Graduate School of Engineering at Tottori University. I am engaged in studies relating to machine learning.

Tadaaki Shimizu received the Ph.D. degree in engineering from Osaka University in 2002. In 1987, he joined the Department of social systems engineering, Tottori University. He is currently an associate professor in the Graduate School of Engineering at Tottori University.

Maiya Hori received his B.E. degree in systems science from Osaka University in 2005. He received his M.E. and Ph.D. degrees in information science from Nara Institute of Science and Technology in 2007 and 2011, respectively. He is currently a research fellow at Tottori University.

Iwai Yoshio graduated from Osaka University in 1992 and completed the M.S. and doctoral programs in 1994 and 1997, respectively. He was then appointed a research associate at the university, subsequently becoming an associate professor. From 2004 to 2005, he was a visiting researcher at the Cambridge University. He is currently a professor in the Graduate School of Engineering at Tottori University. He is engaged in studies relating to computer vision. He is a member of IEEE, the Information Processing Society, and Japanese Robotics Society. He holds a D.Eng. Degree.

Satoru Kishida graduated from Tottori University and completed the M.S. in 1973 and 1977, respectively. He received his Ph.D. degree from Tohoku University. He is currently a professor in the Graduate School of Engineering at Tottori University and a director of the Tottori University Electronic Research Center (TEDREC). He is engaged in studies relating to neural network.