Proceedings of the World Congress on Engineering 2011 Vol II
WCE 2011, July 6 - 8, 2011, London, U.K.

Combined Complex-Valued Artificial Neural Network (CCVANN)

Murat Ceylan

***Abstract*— This study presents a new version of complex-valued artificial neural networks (CVANN) for the complex-valued pattern recognition and classification. Proposed new method is called as combined complex-valued artificial neural network (CCVANN) which is a combination of two complex-valued artificial neural networks. To check the validation of proposed method, complex-valued XOR benchmark problem is used. The accuracy of the CCVANN model is more satisfactory as compared to the existing studies in the literature. Moreover the proposed CCVANN models' results have lower recognition error than using a single CVANN model.**

***Index Terms*— Complex-valued artificial neural network, complex-valued XOR problem**

I. Introduction

ANN is a popular approach used especially in difficult and time-consuming engineering problems. The application of ANNs has opened a new area for solving problems not reasonable by other signal processing techniques [1,2]. It is expected that complex-valued artificial neural networks (CVANN) whose parameters (weights, threshold values, inputs and outputs) are all complex numbers, will have applications in fields dealing with complex numbers such as telecommunications, speech recognition, signal and image processing with the Fourier transformation. For complex signal processing problems, many existing neural networks cannot directly be applied. Although for certain applications it is possible to reformulate a complex signal processing problem so that a real-valued network and learning algorithm can be used to solve the problem, it is not always feasible to do so. Moreover it is preferred to preserve the concise formulation and elegant structure of complex signal [3].

The advantage of using complex-valued artificial neural network instead of a real-valued artificial neural network (RVANN) counterpart fed with a pair of real values is well known [4]. In complex-valued neural networks, one of the main problems is the selecting of nodes activation function. In real case, the node activation function is usually chosen to be a continuous, bounded and non-constant function. These conditions on the activation function are very mild and there is no problem in selecting a real function that satisfies these requirements and that is also smooth (derivative exists). In the complex case, any regular analytic function cannot be bounded unless it reduces to a constant. This is known as the Liouville's theorem. In complex case, the main constraints that the activation function should satisfy can be found in literatures [5,6].

In this paper, a novel cascade structure is proposed, called combined complex-valued artificial neural networks (CCVANN). General scheme of CCVANN is implemented in two levels. In the first level, learning of CVANN is realized using the original data set. After that, prediction of the first level and target of the original data are presented to the second level CVANN as inputs. Proposed method is tested by complex-valued XOR benchmark problem. Obtained results are compared with other complex-valued artificial neural network methods in the literature.

Manuscript received June XX, 20XX; revised July XX, 20XX. (Write the date on which you submitted your paper for review.) This work is supported by the Coordinatorship of Selçuk University's Scientific Research Projects.

Murat Ceylan is with the Selcuk University, Electrical-Electronics Engineering Department, Konya, TURKEY (phone: +90-332-2232042; fax: +90-332-2410635; e-mail: mceylan@selcuk.edu.tr).

II. Methods

A. Complex-Valued Artificial Neural Network (CVANN)

Recently, there has been an increased interest in applications of the CVANN to process complex signals [7-9]. In this study, a complex back-propagation (CBP) algorithm has been used for pattern recognition. We will first give the theory of the CBP algorithm as applied to a multi layer CVANN. Figure 1 shows a CVANN model.

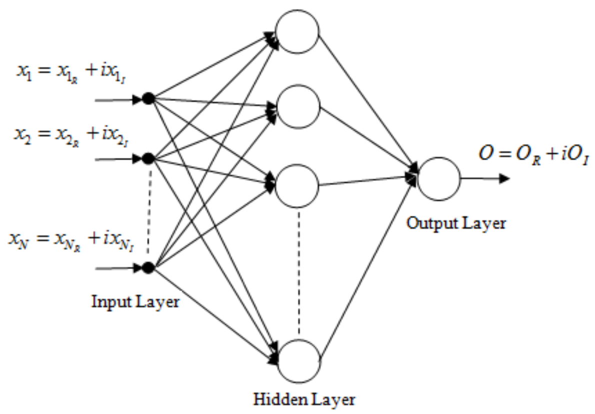

Fig. 1. CVANN model

The input signals, weights, thresholds, and output signals are all complex numbers. The activity Y_n of neuron n is defined as:

$$Y_n = \sum_m W_{nm} X_m + V_n \quad (1)$$

where W_{nm} is the complex-valued (CV) weight connecting neuron n and m, X_m is the CV input signal from neuron m,

ISBN: 978-988-19251-4-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

and V_n is the CV threshold value of neuron n. To obtain the CV output signal, the activity value Y_n is converted into its real and imaginary parts as follows:

$$Y_n = x + iy = z \tag{2}$$

where i denotes $\sqrt{-1}$. Although various output functions of each neuron can be considered, the output function used in this study is defined by the following equation:

$$f_C(z) = f_R(x) + i.f_R(y) \tag{3}$$

where $f_R(u)$=1/(1+exp(-u)) and is called the sigmoid function. For the sake of simplicity, the networks used both in the analyses and experiments will have three layers. We will use W_{ml} for the weight between the input neuron l and the hidden neuron m, V_{nm} for the weight between the hidden neuron m and the output neuron n, θ_m for the threshold of the hidden neuron m, and γ_n for the threshold of the output neuron n. Let I_l, H_m, O_n denote the output values of the input neuron l, the hidden neuron m, and the output neuron n, respectively. Let also U_m and S_n denote the internal potentials of the hidden neuron m and the output neuron n, respectively. U_m, S_n, H_m, and O_n can be defined respectively as $U_m = \sum_l W_{ml} I_l + \theta_m$, $S_n = \sum_m V_{nm} H_m + \gamma_n$, $H_m = f_c(U_m)$, and $O_n = f_c(S_n)$. Let $\delta^n = T_n - O_n$ denote the error between the actual pattern O_n and the target pattern T_n of output neuron n. We will define the square error for the pattern p as $E_p = \left(\frac{1}{2}\right)\sum_{n=1}^{N} \left|T_n - O_n\right|^2$, where N is the number of output neurons.

Next, we define a learning rule for the CBP model described above. We can show that the weights and the thresholds should be modified according to the following equations [7].

$$\Delta V_{nm} = -\varepsilon.\frac{\partial E_p}{\partial \operatorname{Re}[V_{nm}]} - i.\varepsilon\frac{\partial E_p}{\partial \operatorname{Im}[V_{nm}]} \tag{4}$$

$$\Delta \gamma_n = -\varepsilon.\frac{\partial E_p}{\partial \operatorname{Re}[\gamma_n]} - i.\varepsilon\frac{\partial E_p}{\partial \operatorname{Im}[\gamma_n]} \tag{5}$$

$$\Delta W_{ml} = -\varepsilon\frac{\partial E_p}{\partial \operatorname{Re}[W_{ml}]} - i.\varepsilon\frac{\partial E_p}{\partial \operatorname{Im}[W_{ml}]} \tag{6}$$

1.

$$\Delta \theta_m = -\varepsilon\frac{\partial E_p}{\partial \operatorname{Re}[\theta_m]} - i.\varepsilon\frac{\partial E_p}{\partial \operatorname{Im}[\theta_m]} \tag{7}$$

Equations (4)-(7) can be expressed as:

$$\Delta V_{nm} = \overline{H_m}\Delta\gamma_n \tag{8}$$

$$\Delta \gamma_n = \varepsilon\begin{pmatrix}\operatorname{Re}[\delta^n](1-\operatorname{Re}[O_n])\operatorname{Re}[O_n] \\ + i.\operatorname{Im}[\delta^n](1-\operatorname{Im}[O_n])\operatorname{Im}[O_n]\end{pmatrix} \tag{9}$$

$$\Delta W_{ml} = \overline{I_l}\Delta\theta_m \tag{10}$$

$$\Delta\theta_m = \varepsilon\begin{bmatrix}(1-\operatorname{Re}[H_m])\operatorname{Re}[H_m] \\ \times\sum_n\begin{pmatrix}\operatorname{Re}[\delta^n](1-\operatorname{Re}[O_n]) \\ \operatorname{Re}[O_n]\operatorname{Re}[V_{nm}] \\ +\operatorname{Im}[\delta^n](1-\operatorname{Im}[O_n]) \\ \operatorname{Im}[O_n]\operatorname{Im}[V_{nm}]\end{pmatrix}\end{bmatrix} - i\varepsilon\begin{bmatrix}(1-\operatorname{Im}[H_m])\operatorname{Im}[H_m] \\ \times\sum_n\begin{pmatrix}\operatorname{Re}[\delta^n](1-\operatorname{Re}[O_n]) \\ \operatorname{Re}[O_n]\operatorname{Im}[V_{nm}] \\ -\operatorname{Im}[\delta^n](1-\operatorname{Im}[O_n]) \\ \operatorname{Im}[O_n]\operatorname{Re}[V_{nm}]\end{pmatrix}\end{bmatrix} \tag{11}$$

where $\bar{z}$ denotes the complex conjugate of a complex number z.

<u>Summary of CBP algorithm:</u>

1. Initialization

Set all the weights and thresholds to small complex random values.

2. Presentation of input and desired (target) outputs

Present the input vector X(1), X(2),….,X(N) and corresponding desired (target) response T(1), T(2),….T(N), one pair at a time, where N is the total number of training patterns.

3. Calculation of actual outputs

Use the formula in Eq.3 to calculate output signals.

4. Adaptation of weights and thresholds

Use the formulas in Eq. (8-11) to calculate adaptated weights and thresholds.

One of the difficulties encountered in applying the CBP algorithm to the complex domain involves the appropriate choice of activation function.

Several researchers developed a set of properties that a complex activation function must satisfy in order to be useful in a multilayer perceptron trained with the back-propagation algorithm [6]. In summary, these properties are as follows:

1- The activation function $\varphi(z)$ should be nonlinear in both Z_Q and Z_I, which denote the real and imaginary parts, respectively, of the argument Z. Otherwise, there is no advantage in having a multilayer perceptron.

2. 2- The function $\varphi(z)$ should be bounded.

3. 3- The partial derivatives of $\varphi(z)$ should exist and be bounded.

4. 4- The function $\varphi(z)$ should not be an entire function, which is defined as a complex function that is analytic everywhere in the complex domain.

5. Complex activation function is a superposition of real and imaginary logarithmic sigmoids, as shown by

$$\varphi(Z) = \frac{C}{1+\exp(-Z_r)} + j\frac{C}{1+\exp(-Z_i)} \tag{12}$$

where Z_r and Z_i are the real and imaginary parts of Z respectively [6].

B. Combined Complex-Valued Artificial Neural Network (CCVANN)

Conventional ANN models for engineering focus on identifying and using a single, neural network model. This approach assumes that a single ANN model can take all the information of input data. The other candidate models are redundant. However, any individual model cannot be a success in the extraction of related information from data. Wolpert [10] proposed the idea of stacked generalization to combine multiple models. Sridhar et al. [11] were applied to stacked generalization for neural network models (SNN) that consist of a combination of the candidate neural networks. In [11], SNNs were limited to using a linear combination of artificial neural networks. Models that are useful in a nonlinear sense are wasted if a linear combination is used. Same authors proposed an information theoretic stacking (ITS) algorithm for combining neural network models in [12]. Proposed algorithm identifies and combines useful models regardless of the nature of their relationship to the network output. In this work, obtained results demonstrate that the SNNs developed using the ITS algorithm can achieve highly improved performance as compared to using a single ANN [12].

An application of combined ANN (CANN) for medical diagnosis was proposed by Hayashi and Setiono [13]. To improve the accuracy of the diagnosis, second level ANN were trained with the outputs of the first level networks. They obtained higher accuracy rate in the second level ANN than the individual ANN in the first level. Ubeyli [14-16] proposed CANN to diagnose and classify tasks for medical data. In this study, diagnosis of internal carotid artery disorders, erythemato-squamous diseases and EEG signals classification was realized using CANN. Different engineering applications of CANN were presented by [17,18]. Subbaraj and Rajasekaran [17] and Dung [18] used CANN instead of single ANN for peak load forecasting and radar target recognition, respectively.

The general CCVANN model which is a combination of two CVANN used in this study is illustrated in Figure 2. The CVANNs were used at the first and second levels for the complex-valued pattern recognition.

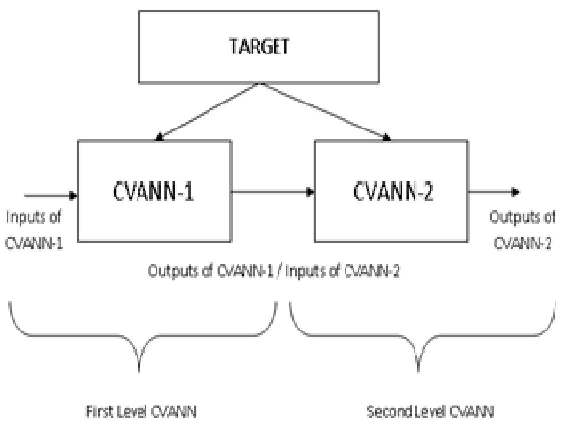

Fig. 2. CCVANN model

III. CALCULATION ERRORS

A. Stopping Criteria

The stopping criteria used for learning of CCVANNs was [8]:

$$\sqrt{\sum_{p} \sum_{n=1}^{N} \left| T_n^{(p)} - O_n^{(p)} \right|^2} = 10^{-1} \tag{13}$$

where $T_n^{(p)}$, $O_n^{(p)} \in C$ denote the desired output value. The actual output value of the neuron n for the pattern p, i.e the left side of (Eq. 13) denotes the error between the desired output pattern and the actual output pattern. N denotes the number of neurons in the output layer.

The training of CCVANNs was stopped when the error goal was achieved. After that, the performances of CCVANNs were tested by presenting test subjects.

B. Numerical Errors

Training and test errors given in tables were conducted according to Eq. (14).

$$\text{Training and Test Error}(\%) = \left(\frac{\sum_{i=1}^{k} \left| t(i) - a(i) \right|}{m * n} \right) * 100 \tag{14}$$

where *t(i)* is desired outputs, *a(i)* is outputs of neural network, *k* is the number of samples in training or test data, *m* is the number of segments in training or test data and *n* is the number of outputs of neural network for training or test procedures [19].

IV. COMPLEX-VALUED XOR PROBLEM

Minsky and Papert [20] clarified the limitations of a single real-valued neuron: in a large number of interesting cases, a single real-valued neuron is incapable of solving the problems. A classic example of this case is the exclusive-or (XOR) problem which has a long history in the study of neural networks, and many other difficult problems involve the XOR as subproblem [21].

In this section, it is proved that the XOR problem can be solved by a CCVANN.

A. Complex-Valued XOR Problem with Four Patterns

The input-output mapping in the XOR problem is shown in Table 1. In order to solve the XOR problem with complex-valued artificial neural networks, the input-output mapping is encoded as shown in Table 2 (similar XOR problem) where the real part of the output can be seen as the XOR of the input's real and input's imaginary part, and the imaginary part of the output is equal to real part of the input [22].

This problem has been simulated with 1-2-1 (one input, two hidden nodes in hidden layer and one output) CCVANN to compare with other methods in the literature (conventional CVANN [21], improved CVANN [22], Complex-Valued Wavelet ANN [23]) for complex-valued

ISBN: 978-988-19251-4-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

XOR problem solving. For all methods, learning rate and maximum iteration number were chosen as 0.5 and 20000, respectively, similarly [23]. Error values in Table 3 were used as stopping criteria (as seen Eq. 13). Results for proposed methods and other methods were presented in Table 3, comparatively. It can be shown in this table, when the error was used as 0.001, the best success rate (nearly 100 %) is obtained using proposed method. For this error value, conventional CVANN and improved CVANN were achieved 93 % and 99 % success rate, respectively. Using proposed method for complex-valued XOR problem were obtained higher success rates than single CVANN for all error values.

TABLE I
XOR PROBLEM WITH FOUR PATTERNS

INPUTS		OUTPUT
X1	X2	Y
0	0	0
0	1	1
1	0	1
1	1	0

TABLE II
SIMILAR XOR PROBLEM FOR COMPLEX-VALUED PATTERNS

Input Pattern	Output Pattern
0	0
i	1
1	1+i
1+i	i

TABLE III
COMPARISON OF PROPOSED METHOD AND PREVIOUS STUDIES

Origin of Methods	Error = 0.1 / 0.01 / 0.001	
	Success Rate (%)	Iteration Number
[21] Conventional CVANN	95 / 95 / 93	488 / 1878 / 1184
[22] Improved CVANN	100 / 100 / 99	203 / 582 / 3180
[23] Complex-Valued Wavelet ANN	98.58 / 99.77 / 99.97	71 / 665 / 5095
[Proposed] Combined CVANN	**97.55 / 99.75 / 99.99**	**75 / 615 / 5435**

When the error was used as 0.001, outputs of CCVANN was presented in Table 4. It can be shown in Table 4, results of proposed CCVANN algorithms to solving complex-valued XOR problem were converged to target, successfully.

TABLE IV
OUTPUTS OF CCVANN FOR ERROR=0.001

Target	Outputs of CCVANN
0	0.0001 + i0.0001
1	0.9999 + i0.0001
1 + i	0.9999 + i0.9999
i	0.0001 + i0.9999

B. Complex-Valued XOR Problem with Sixteen Patterns

For the second experiment, complex-valued XOR gate patterns of Nitta [8] were used. CCVANN was trained and tested with the sixteen patterns of Table 5. Leave one out cross-validation method [23] was used for obtaining a better network generalization. Obtained training and test errors were averaged.

Optimum network architecture is defined as 2-2-1. Learning rate was chosen as 1.0 in training via experimentation. Maximum iteration number is 1000. Training and test errors were obtained 0.032 % and 0.035 %, respectively. Target and outputs of CCVANN are shown in Table 6.

TABLE V
XOR PROBLEM WITH SIXTEEN PATTERNS

Input 1	Input 2	Output
0	0	1
0	i	i
i	0	0
i	i	1+i
i	1	i
1	1	1+i
1+i	i	i
1+i	1+i	1
0	1	i
0	1+i	0
i	1+i	0
1	0	0
1	i	i
1	1+i	0
1+i	0	0
1+i	1	i

TABLE VI
RESULTS OF CCVANN FOR TABLE V

Target	Outputs of CCVANN
1	0.9988+0.0006i
i	0.0010+0.9991i
0	0.0007+0.0010i
1+i	0.9991+0.9989i
i	0.0010+0.9992i
1+i	0.9994+0.9986i
i	0.0009+0.9991i
1	0.9986+0.0007i
i	0.0010+0.9993i
0	0.0006+0.0011i
0	0.0007+0.0010i
0	0.0005+0.0010i
i	0.0008+0.9993i
0	0.0006+0.0011i
0	0.0005+0.0008i
i	0.0009+0.9991i

V. CONCLUSIONS

In this paper, a combined complex-valued artificial neural network model considering the complex-valued pattern recognition was developed. The following conclusions may be drawn based on the results presented;

1. The results of the CCVANN compared to experimental results are found to be more satisfactory (99.99%). Moreover, the proposed method performs better when compared to existing methods in the literature.

ISBN: 978-988-19251-4-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

2. The proposed CCVANN methods' results have a lower prediction error than using a single CVANN model.
3. Although the performance of the developed CCVANN model is limited to the range of input data used in the training and testing process, the method can easily be applied with additional new set of data.

REFERENCES

[1] A.S., Miller, B.H. Blott and T.K. Hames, "Review of neural network applications in medical imaging and signal processing", *Medical and Biological Engineering and Computing*, 30 449-464. 1992.
[2] W.G. Baxt, "Use of an artificial neural network for data analysis in clinical decision making, the diagnosis of acute coronary occlusion", *Neural Computing*, 2, 480-489, 1990.
[3] C. Li, X. Liao and J Yu, "Complex-valued wavelet network", *Journal of Computer and System Sciences*, 67, 623-632, 2003.
[4] N. Benvenuto, M. Marchesi, F. Piazza, and A. Uncini, "A comparison between real and complex-valued neural networks in communication applications", *Proceedings of the International Conference on Artificial Neural Networks* IEEE, Spain, 1991.
[5] N. Georgin and C. Koutsougeras, "Complex domain backpropagation", *IEEE Transactions on Circuits And Systems,* 39, 330-334, 1992.
[6] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 2002.
[7] T. Nitta, "A back-propagation algorithm for complex numbered neural networks", *Proceedings International Joint Conference on Neural Networks,* IEEE, Nagoya, 1649-1652, 1993.
[8] T. Nitta, "An extension of the back-propagation algorithm to complex numbers", *Neural Network*, 10, 1391-1415, 1997.
[9] T. Nitta, "An analysis of the fundamental structure of complex-valued neurons", *Neural Processing Letters*, 12, 239-246, 2000.
[10] D. H. Wolpert,. Stacked generalization, *Neural Networks,* Vol. 5 (2), 241-259, 1992.
[11] V. Sridhar, C Seagrave and B. Barlett, "Process modeling using stacked neural networks", *American Institute of Chemical Engineering Journal*, Vol. 42 (9), 2529-2539, 1996.
[12] V. Sridhar, B. Barlett and C. Seagrave, "An information theoretic approach for combining neural network process models", *Neural Networks,* Vol. 12, 915-926, 1999.
[13] Y Hayashi and R. Setiono," Combining neural network predictions for medical diagnosis", *Computers in biology and medicine*, Volume 32 (4), 237-246, 2002.
[14] E. D. Übeyli, "Combining neural network model for automated diagnostic systems", *Journal of Medical Systems*, Vol. 30, 483-488, 2006.
[15] E. D. Übeyli, "Combined neural networks for diagnosis of erythemato-squamous diseases", *Expert Systems with Applications*, Vol. 36, 5107-5112, 2009.
[16] E. D. Übeyli, Combined neural network model employing wavelet coefficients for EEG signals classification, *Digital Signal Processing*, Vol. 19, 297-308.
[17] P. Subbaraj and V. Rajasekaran, "Evolutionary techniques based combined artificial neural networks for peak load forecasting", World Academy of Science, Eng., Tech., Vol. 45, 680-686, 2008.
[18] P. T. Dung, "Combined neural Networks for radar target recognition from radar range profiles", *International Conference on Advanced Technologies for Communications*, 353-355, 2008.
[19] Y. Özbay, R. Ceylan and B. Karlik, "A fuzzy clustering neural network architecture for classification of ECG arrhythmias", *Computers in Biology and Medicine*, 36, 376-388, 2006.
[20] M. L. Minsky and S. A. Papert,, *Perceptrons*, MIT Pres, Cambridge, 1969.
[21] T. Nitta, "Solving the XOR problem and the detection of symmetry using a single complex-valued neuron", *Neural Networks*, 16, 1101-1105, 2003.
[22] X. Chen, Z. Tang C. Variappan, S. Li and T. Okada, "A modified error backpropagation algorithm for complex-value neural networks", *International Journal of Neural Systems*, 15, 435-443, 2005.
[23] Y. Özbay, S. Kara, F. Latifoğlu, R Ceylan and M. Ceylan, "Complex-valued wavelet artificial neural network for Doppler signal classifying", *Artificial Intelligence in Medicine*, 40, 143-156, 2007.