Neurocomputing ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A fully complex-valued radial basis function classifier for real-valued classification problems

R. Savitha [a,*], S. Suresh [a], N. Sundararajan [b], H.J. Kim [c]

[a] School of Computer Engineering, Nanyang Technological University, Singapore
[b] Department of Information Science and Engineering, Sri Jayachamarajendra College of Engineering, Mysore, India
[c] Graduate School of International Management and Security, Korea University, Seoul, South Korea

ARTICLE INFO

Keywords:
Classification
Fully complex-valued radial basis function
Complex-valued classifiers
Orthogonal decision boundaries
Acoustic emission signal classification

ABSTRACT

In this paper, we investigate the decision making ability of a fully complex-valued radial basis function (FC-RBF) network in solving real-valued classification problems. The FC-RBF classifier is a single hidden layer fully complex-valued neural network with a nonlinear input layer, a nonlinear hidden layer, and a linear output layer. The neurons in the input layer of the classifier employ the phase encoded transformation to map the input features from the Real domain to the Complex domain. The neurons in the hidden layer employ a fully complex-valued Gaussian-like activation function of the type of hyperbolic secant (sech). The classification ability of the classifier is first studied analytically and it is shown that the decision boundaries of the FC-RBF classifier are orthogonal to each other. Then, the performance of the FC-RBF classifier is studied experimentally using a set of real-valued benchmark problems and also a real-world problem. The study clearly indicates the superior classification ability of the FC-RBF classifier.

1. Introduction

With the evolution of telecommunications [1], adaptive array signal processing [2,3] and medical imaging applications [4] that involve complex-valued signals, there is a need to develop efficient complex-valued neural networks. In addition, recently, T. Nitta has proved that complex-valued neural networks have better computational capability than real-valued neural networks [9] and the orthogonal decision boundaries [10] of the complex-valued neural networks with split-type of activation function provide them an ability to perform classification tasks efficiently. Since complex-valued classifiers are good decision makers, developing complex-valued classifiers to solve real-valued classification problems is gaining interest among researchers.

The Multi-Layer Multi Valued Network (MLMVN) [11,12] using a multi-valued neuron is the first complex-valued classifier available in the literature to solve real-valued classification problems. A multi-valued neuron uses a piecewise continuous activation function to map the complex-valued inputs to C discrete class labels (i.e, C-valued threshold logic, where C is the number of classes). In MLMVN, the normalized real-valued input features (x) are mapped onto a unit circle using $\exp(i2\pi x)$ transformation and the class labels are encoded by the roots of unity in the Complex plane. However, as the input features are mapped to a unit circle, this mapping results in the same complex-valued features for the real-valued features with values 0 and 1 (i.e the transformation is non-unique). In addition, as number of classes (C) increases, the region of sectors per class within the unit circle of output neuron decreases, which influences the classification ability of MLMVN.

Another complex-valued classifier available in the literature is the single layer Phase Encoded Complex-Valued Neural Network (PE-CVNN) [13,14]. In the PE-CVNN, the real-valued input features (x) are phase encoded in $[0,\pi]$, using the transformation '$\exp(i\pi x)$' to obtain the complex-valued input features. However, the activation functions used in the PE-CVNN are similar to those used in the split complex-valued neural networks. Therefore, the gradients used in the parameter update of the PE-CVNN are not fully complex-valued [5] and hence this might result in a significant loss of complex-valued information resulting in misclassification.

The choice of an appropriate activation function and the transformation function used to map the real-valued features to complex-valued features influence the decision making ability of complex-valued neural network. As the aforementioned complex-valued classifiers suffer from either of these issues, there is a need to develop fully complex-valued classifiers to solve real-valued classification problems.

Fully complex-valued neural networks use fully complex-valued activation functions and its complex-valued gradients [5–7] to approximate the functions accurately. The choice of the

* Corresponding author.
E-mail address: savi0001@ntu.edu.sg (R. Savitha).

0925-2312/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2011.05.036

activation function in the Complex domain is limited due the conflict between the Liouville's theorem and the essential properties needed for a fully complex-valued activation function [5]. Therefore, Kim et al. [5] have restricted these essential properties to be analytic and bounded *almost everywhere*. Few fully complex-valued activation functions satisfying these conditions have been developed in the framework of fully complex-valued multi-layer perceptron [5,6] and Radial Basis Function (RBF) networks [7,8].

RBF networks are widely used in the literature to solve classification problems in the real domain due to their localization ability and simple architecture. Similarly, the complex-valued RBF networks can also be used to perform classification tasks. The basic architecture and learning algorithm of complex-valued RBF network were first introduced in [15]. However, the existing complex-valued RBF learning algorithms use the Gaussian function that maps $\mathbb{C}^m \to \Re$ as activation function, resulting in inaccurate phase approximation. To overcome the above limitations, Savitha et al. [7] have used a fully complex-valued radial basis activation function which is a hyperbolic secant function (sech(.)). A Fully Complex-valued Radial Basis Function (FC-RBF) network using this activation function and its learning algorithm was derived in [7]. Recently, a sequential learning algorithm using this fully complex-valued activation function has been developed in [16]. The performance study clearly shows that the FC-RBF network outperforms other complex-valued RBF networks in the literature. As the FC-RBF network outperforms other complex-valued RBF learning algorithms, it is interesting to learn the decision making ability of the FC-RBF classifier.

In this paper, we investigate the decision making ability of the FC-RBF by using it to solve real-valued classification problems in the Complex domain. The classification problem is formulated in the Complex domain and the phase encoded transformation is used as an activation function in the input layer to map the real-valued input features to complex-valued features. Thus, FC-RBF classifier has a nonlinear fully complex-valued activation function in the input/hidden layers and a linear activation function in the output layer. The network parameters are updated using complex-valued gradient descent rule as explained in [7]. This paper provides an analytical proof for existence of orthogonal decision boundaries in FC-RBF classifier. The orthogonal decision boundaries, the nonlinear fully complex-valued activation function and its complex gradient update rules help the FC-RBF classifier to approximate the decision surface more accurately. The performance of the FC-RBF classifier is evaluated using a set of benchmark real-valued classification problems from the University of California, Irvine (UCI) machine learning repository [17] and a real-world acoustic emission signal classification problem [18]. The performance of the FC-RBF classifier is compared with other existing real and complex-valued classifiers in the literature. The results indicate a superior performance of the FC-RBF classifier, especially for the data sets with a high sample imbalance.

The paper is organized as follows: a detailed description of the FC-RBF network and its learning algorithm is presented in Section 2. This section also includes an analytical proof for the existence of the orthogonal decision boundaries in the FC-RBF network. Section 3 presents the performance evaluation of the FC-RBF classifier in comparison with other complex-valued and best performing real-valued classifiers available in the literature. Finally, Section 4 summarizes the main conclusions from this study.

2. A fully complex-valued radial basis function classifier

In this section, we first present the principles behind solving real-valued classification problem in the Complex domain followed by a detailed description of the FC-RBF classifier. Further, we prove that the FC-RBF classifier has orthogonal decision boundaries which makes it suitable for classification problem.

2.1. Definition of the real-valued classification problem in the complex domain

Let $\{(\mathbf{x}^1,c^1),\ldots,(\mathbf{x}^t,c^t),\ldots,(\mathbf{x}^N,c^N)\}$ be N random observations, where $\mathbf{x}^t=[x_1^t,\ldots,x_m^t]^T \in \Re^m$ is the m-dimensional real-valued input feature vector of the tth observation, and $c^t \in \{1,2,\ldots,C\}$ is its class label.

The coded class label of the tth sample $\mathbf{y}^t=[y_1^t,\ldots,y_C^t]^T \in \mathbb{C}^C$ in the Complex domain is given by

$$y_l^t=\begin{cases}1+i1 & \text{if } c_t=l,\\ -1-i1 & \text{otherwise,}\end{cases} \quad l=1,\ldots,C \tag{1}$$

where $i=\sqrt{-1}$ is a Complex operator.

The training set for a classification problem in the Complex domain can be represented by: $\{(\mathbf{x}^1,\mathbf{y}^1),\ldots,(\mathbf{x}^t,\mathbf{y}^t),\ldots,(\mathbf{x}^N,\mathbf{y}^N)\}$, where $\mathbf{x}^t \in \Re^m$ and $\mathbf{y}^t \in \mathbb{C}^C$. Thus, the classification problem in the Complex domain is defined as estimating that decision function (F) which maps the real-valued input features to the complex-valued coded class labels, i.e., $F:\Re^m \to \mathbb{C}^C$, thereby, predicting the class labels of new, unseen samples with certain accuracy.

2.2. Description of the FC-RBF classifier

The FC-RBF network [7] is a single hidden layer network with m input neurons, K hidden neurons, and C output neurons. The structure of the FC-RBF classifier is presented in Fig. 1. The neurons in the input/hidden layers of the FC-RBF classifier employ nonlinear activation functions while the neurons in the output layer use a linear activation function.

The neurons in the input layer of the FC-RBF classifier utilize a phase encoded transformation [13] to map the real-valued input features to the Complex domain. The complex-valued input features of the tth sample ($\mathbf{z}^t \in \mathbb{C}^m$) thus obtained using this transformation of the real-valued input features ($\mathbf{x}^t \in \Re^m$) is represented as

$$\mathbf{z}^t=\exp(i\pi\mathbf{x}^t) \tag{2}$$

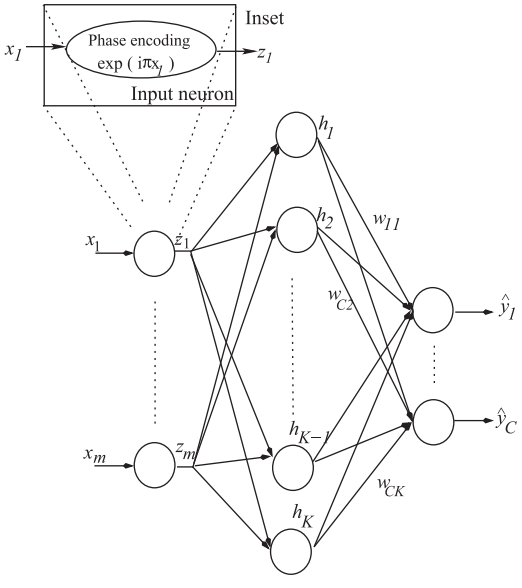

Fig. 1. The structure of a FC-RBF classifier. The inset show the transformation of real-valued input to complex-valued output in the input neuron.

Please cite this article as: R. Savitha, et al., A fully complex-valued radial basis function classifier for real-valued classification problems, Neurocomputing (2011), doi:10.1016/j.neucom.2011.05.036

Since the real-valued input features ($\mathbf{x}^t$) are normalized between [0, 1], the outputs ($\mathbf{z}^t$) are bounded and are distributed along the unit circle in the first and second quadrant of the Complex domain.

The neurons in the hidden layer use a complex-valued activation function of the type of hyperbolic secant. The output ($\mathbf{h}$) of the hidden layer neurons is given by

$$h_k = \mathrm{sech}(\mathbf{u}_k^T(\mathbf{z}^t - \mathbf{v}_k)), \quad k = 1 \cdots K \tag{3}$$

where $\mathbf{u}_k \in \mathbb{C}^m$ is a complex-valued scaling factor of the kth neuron, $\mathbf{v}_k \in \mathbb{C}^m$ is the center of the kth neuron, and $\mathrm{sech}(x) = 2/(e^x + e^{-x})$.

The predicted output ($\widehat{\mathbf{y}}^t$) of the network is given by

$$\widehat{y}_l^t = \sum_{k=1}^{K} w_{lk} h_k, \quad l = 1, \ldots, C \tag{4}$$

where w_{lk} is the complex-valued weight connecting the kth hidden neuron and the lth output neuron.

The prediction error is

$$\mathbf{e}^t = \mathbf{y}^t - \widehat{\mathbf{y}}^t \tag{5}$$

The mean squared error function is used as minimization criterion and is defined as

$$J = \frac{1}{2}\sum_{t=1}^{N} (\mathbf{e}^{t^H}\mathbf{e}^t) \tag{6}$$

where H represents the Complex Hermitian operator.

FC-RBF network minimizes the mean squared error deviation using a gradient descent based learning algorithm. The gradient descent based update rules are

$$\Delta w_{lk} = \eta_w \overline{\mathbf{h}}_k (y_l^t - \widehat{y}_l^t); \quad l = 1, \ldots, C; \quad k = 1, \ldots, K; \quad t = 1, \ldots N \tag{7}$$

$$\Delta \mathbf{u}_k = \eta_u (y_l^t - \widehat{y}_l^t) \overline{w}_{lk} \overline{f}'(\mathbf{z}^t - \mathbf{v}_k) \overline{(\mathbf{z}^t - \mathbf{v}_k)} \tag{8}$$

$$\Delta \mathbf{v}_k = -\eta_v (y_l^t - \widehat{y}_l^t) \overline{w}_{lk} \overline{f}'(\mathbf{u}_k^T(\mathbf{z}^t - \mathbf{v}_k)) \overline{\mathbf{u}}_k \tag{9}$$

where, η_v, η_u, and η_w are the learning rates for the center, the scaling and the output weight parameters, respectively, and $\overline{z}$ represents the conjugate of a complex-valued signal z. $\overline{f}'$ is the Complex conjugate of the derivative of the sech function.

In the next section, we prove that the FC-RBF network exhibits orthogonal decision boundaries.

2.3. Orthogonality of the decision boundaries

In this section, we show that the decision boundaries of the output neurons are orthogonal. For this purpose, we redefine the following[1]:

Let the response of the kth hidden neuron be $h_k = h_k^R + i h_k^I$. The net input to the kth hidden neuron is represented as $O_k = O_k^R + i O_k^I = \mathbf{v}_k^T(\mathbf{z} - \mathbf{u}_k)$ and the output weights between kth hidden neuron and lth output neuron be $w_{lk} = w_{lk}^R + i w_{lk}^I$. Note that the superscript R and I represent the real and imaginary part of a Complex number, respectively.

The predicted output ($\widehat{y}_l = \widehat{y}_l^R + i\widehat{y}_l^I$) of the lth output neuron is

$$\widehat{y}_l = \widehat{y}_l^R + i\widehat{y}_l^I = \sum_{k=1}^{K} (w_{lk}^R . h_k^R - w_{lk}^I . h_k^I) + i(w_{lk}^R . h_k^I + w_{lk}^I . h_k^R) \tag{10}$$

Using trigonometric/hyperbolic definitions, the sech function can be expanded as

$$h_k^R = \left(\frac{2(\cos(O_k^I)\cosh(O_k^R))}{\cos(2O_k^I) + \cosh(2O_k^R)} \right) \tag{11}$$

$$h_k^I = \left(\frac{2(-i\sin(O_k^I)\sinh(O_k^R))}{\cos(2O_k^I) + \cosh(2O_k^R)} \right) \tag{12}$$

The two decision boundaries formed by real/imaginary parts of the output neuron are

$$\widehat{y}_l^R = \sum_{k=1}^{K} (w_{lk}^R . h_k^R - w_{lk}^I . h_k^I) \rightarrow \mathbf{C}^R \tag{13}$$

and

$$\widehat{y}_l^I = \sum_{k=1}^{K} (w_{lk}^R . h_k^I + w_{lk}^I . h_k^R) \rightarrow \mathbf{C}^I \tag{14}$$

The normal vectors for the above decision boundaries are given by

$$Q^R(\mathbf{O}^R, \mathbf{O}^I) = \left(\frac{\partial \widehat{y}_l^R}{\partial O_1^R} \frac{\partial \widehat{y}_l^R}{\partial O_1^I} \cdots \frac{\partial \widehat{y}_l^R}{\partial O_k^R} \frac{\partial \widehat{y}_l^R}{\partial O_k^I} \cdots \frac{\partial \widehat{y}_l^R}{\partial O_K^R} \frac{\partial \widehat{y}_l^R}{\partial O_K^I} \right) \tag{15}$$

and

$$Q^I(\mathbf{O}^R, \mathbf{O}^I) = \left(\frac{\partial \widehat{y}_l^I}{\partial O_1^R} \frac{\partial \widehat{y}_l^I}{\partial O_1^I} \cdots \frac{\partial \widehat{y}_l^I}{\partial O_k^R} \frac{\partial \widehat{y}_l^I}{\partial O_k^I} \cdots \frac{\partial \widehat{y}_l^I}{\partial O_K^R} \frac{\partial \widehat{y}_l^I}{\partial O_K^I} \right) \tag{16}$$

The normal vectors of the decision boundaries are orthogonal to each other *iff* their dot product is zero. The dot product is

$$Q^R(\mathbf{O}^R, \mathbf{O}^I) . Q^I(\mathbf{O}^R, \mathbf{O}^I) = \sum_{k=1}^{K} \left(\frac{\partial \widehat{y}_l^R}{\partial O_k^R} \frac{\partial \widehat{y}_l^I}{\partial O_k^R} + \frac{\partial \widehat{y}_l^R}{\partial O_k^I} \frac{\partial \widehat{y}_l^I}{\partial O_k^I} \right) \tag{17}$$

where

$$\frac{\partial \widehat{y}_l^R}{\partial O_k^R} = \frac{\partial \widehat{y}_l^R}{\partial h_k^R} \frac{\partial h_k^R}{\partial O_k^R} + \frac{\partial \widehat{y}_l^R}{\partial h_k^I} \frac{\partial h_k^I}{\partial O_k^R} = w_{lk}^R \frac{\partial h_k^R}{\partial O_k^R} - w_{lk}^I \frac{\partial h_k^I}{\partial O_k^R} \tag{18}$$

$$\frac{\partial \widehat{y}_l^R}{\partial O_k^I} = \frac{\partial \widehat{y}_l^R}{\partial h_k^R} \frac{\partial h_k^R}{\partial O_k^I} + \frac{\partial \widehat{y}_l^R}{\partial h_k^I} \frac{\partial h_k^I}{\partial O_k^I} = w_{lk}^R \frac{\partial h_k^R}{\partial O_k^I} - w_{lk}^I \frac{\partial h_k^I}{\partial O_k^I} \tag{19}$$

$$\frac{\partial \widehat{y}_l^I}{\partial O_k^R} = \frac{\partial \widehat{y}_l^I}{\partial h_k^R} \frac{\partial h_k^R}{\partial O_k^R} + \frac{\partial \widehat{y}_l^I}{\partial h_k^I} \frac{\partial h_k^I}{\partial O_k^R} = w_{lk}^I \frac{\partial h_k^R}{\partial O_k^R} + w_{lk}^R \frac{\partial h_k^I}{\partial O_k^R} \tag{20}$$

$$\frac{\partial \widehat{y}_l^I}{\partial O_k^I} = \frac{\partial \widehat{y}_l^I}{\partial h_k^R} \frac{\partial h_k^R}{\partial O_k^I} + \frac{\partial \widehat{y}_l^I}{\partial h_k^I} \frac{\partial h_k^I}{\partial O_k^I} = w_{lk}^I \frac{\partial h_k^R}{\partial O_k^I} + w_{lk}^R \frac{\partial h_k^I}{\partial O_k^I} \tag{21}$$

Note, $\partial \widehat{y}_l^R / \partial h_k^R = w_{lk}^R$, $\partial \widehat{y}_l^I / \partial h_k^R = w_{lk}^I$, $\partial \widehat{y}_l^R / \partial h_k^I = -w_{lk}^I$ and $\partial \widehat{y}_l^I / \partial h_k^I = w_{lk}^R$.

The derivative of the kth hidden neuron output with respect to its input is given by

$$\frac{\partial h_k^R}{\partial O_k^R} = \frac{\sinh(O_k^R)\cos(3O_k^I) - \sinh(3O_k^R)\cos(O_k^I)}{(\cos(2O_k^I) + \cosh(2O_k^R))^2} \tag{22}$$

$$\frac{\partial h_k^I}{\partial O_k^I} = \frac{\sinh(O_k^R)\cos(3O_k^I) - \sinh(3O_k^R)\cos(O_k^I)}{(\cos(2O_k^I) + \cosh(2O_k^R))^2} \tag{23}$$

$$\frac{\partial h_k^R}{\partial O_k^I} = \frac{\cosh(O_k^R)\sin(3O_k^I) - \cosh(3O_k^R)\sin(O_k^I)}{(\cos(2O_k^I) + \cosh(2O_k^R))^2} \tag{24}$$

$$\frac{\partial h_k^I}{\partial O_k^R} = \frac{\cosh(3O_k^R)\sin(O_k^I) - \cosh(O_k^R)\sin(3O_k^I)}{(\cos(2O_k^I) + \cosh(2O_k^R))^2} \tag{25}$$

From the above equations, it can be observed that

$$\frac{\partial h_k^R}{\partial O_k^R} = \frac{\partial h_k^I}{\partial O_k^I} \tag{26}$$

and

$$\frac{\partial h_k^R}{\partial O_k^I} = -\frac{\partial h_k^I}{\partial O_k^R} \tag{27}$$

[1] For notational convenience, we drop the superscript t.

Please cite this article as: R. Savitha, et al., A fully complex-valued radial basis function classifier for real-valued classification problems, Neurocomputing (2011), doi:10.1016/j.neucom.2011.05.036

This indicates the sech activation function satisfies the Cauchy Riemann equations [31].

Substituting the Eqs. (26) and (27) in Eq. (18)–(21), we get

$$\frac{\partial \hat{y}_l^R}{\partial O_k^R} = w_{lk}^R \frac{\partial h_k^R}{\partial O_k^R} + w_{lk}^I \frac{\partial h_k^R}{\partial O_k^I} \tag{28}$$

$$\frac{\partial \hat{y}_l^R}{\partial O_k^I} = w_{lk}^R \frac{\partial h_k^R}{\partial O_k^I} - w_{lk}^I \frac{\partial h_k^R}{\partial O_k^R} \tag{29}$$

$$\frac{\partial \hat{y}_l^I}{\partial O_k^R} = w_{lk}^I \frac{\partial h_k^R}{\partial O_k^R} - w_{lk}^R \frac{\partial h_k^R}{\partial O_k^I} \tag{30}$$

$$\frac{\partial \hat{y}_l^I}{\partial O_k^I} = w_{lk}^I \frac{\partial h_k^R}{\partial O_k^I} + w_{lk}^R \frac{\partial h_k^R}{\partial O_k^R} \tag{31}$$

Using the above results, the dot product of the normal vectors shown in Eqs. (15) and (16) is reduced to

$$Q^R(\mathbf{O}^R,\mathbf{O}^I).Q^I(\mathbf{O}^R,\mathbf{O}^I) = \sum_{k=1}^{K} \left(\frac{\partial \hat{y}_l^R}{\partial O_k^R} \cdot \frac{\partial \hat{y}_l^I}{\partial O_k^R} + \frac{\partial \hat{y}_l^R}{\partial O_k^I} \cdot \frac{\partial \hat{y}_l^I}{\partial O_k^I} \right) \tag{32}$$

$$= \sum_{k=1}^{K} \begin{pmatrix} w_{lk}^R w_{lk}^I \left(\frac{\partial h_k^R}{\partial O_k^R}\right)^2 + (w_{lk}^I)^2 \left(\frac{\partial h_k^R}{\partial O_k^I}\right)\left(\frac{\partial h_k^R}{\partial O_k^R}\right) \\ -(w_{lk}^R)^2 \left(\frac{\partial h_k^R}{\partial O_k^I}\right)\left(\frac{\partial h_k^R}{\partial O_k^I}\right) - w_{lk}^R w_{lk}^I \left(\frac{\partial h_k^R}{\partial O_k^I}\right)^2 \\ + w_{lk}^R w_{lk}^I \left(\frac{\partial h_k^R}{\partial O_k^I}\right)^2 - (w_{lk}^I)^2 \left(\frac{\partial h_k^R}{\partial O_k^I}\right)\left(\frac{\partial h_k^R}{\partial O_k^R}\right) \\ + (w_{lk}^R)^2 \left(\frac{\partial h_k^R}{\partial O_k^I}\right)\left(\frac{\partial h_k^R}{\partial O_k^I}\right) - w_{lk}^R w_{lk}^I \left(\frac{\partial h_k^R}{\partial O_k^R}\right)^2 \end{pmatrix} \tag{33}$$

$$= 0 \tag{34}$$

Thus, the decision boundaries formed by the real and imaginary parts of an output neuron in the FC-RBF network are orthogonal to each other.

Based on the above results, we state the following Lemma:

Lemma 2.1. *The two decision boundaries formed by real/imaginary parts of an output neuron in the FC-RBF network are orthogonal to each other.*

In the next section, we evaluate the performance of the FC-RBF classifier on a set of real-valued benchmark classification problems and a practical classification problem to highlight the advantages of the orthogonal decision boundaries of the FC-RBF classifier.

3. Performance evaluation of the FC-RBF classifier

In this section, we evaluate the performance of the FC-RBF classifier using the real-valued benchmark classification data sets from the UCI machine learning repository [17]. First, we present the results on three multi-category benchmark classification problems, viz., the image segmentation problem, the vehicle classification problem and the glass identification problem. Next, the performance results of the FC-RBF classifier are studied in comparison with other classifiers on four binary classification problems. Finally, we evaluate the performance of the FC-RBF classifier on a practical acoustic emission signal classification problem [18]. All the studies have been conducted in MATLAB 7.10 environment on a desktop PC with Pentium duo processor (2.6 GHz) and 4 GB memory. For all these problems, the number of hidden neurons of the FC-RBF classifier is chosen based on a constructive–destructive procedure presented in [19] and the FC-RBF classifier is trained over 5000 epochs.

Table 1
Description of benchmark data sets selected from [17] for performance study.

Type of data set	Problem	No. of features	No. of classes	No. of samples		I.F.
				Training	Testing	
Multi-Categ.	Image Segmentation	19	7	210	2100	0
	Vehicle Classification	18	4	424	422	0.1
	Glass Identification	9	6	109	105	0.68
Binary	Liver Disorder	6	2	200	145	0.17
	PIMA Data	8	2	400	368	0.225
	Breast Cancer	9	2	300	383	0.26
	Ionosphere	34	2	100	251	0.28

The average (η_a) and over-all (η_o) classification efficiencies are used as the performance measures for comparison:

$$\eta_a = \frac{1}{C} \sum_{l=1}^{C} \frac{q_{ll}}{N_l} \times 100\% \tag{35}$$

$$\eta_o = \frac{\sum_{l=1}^{C} q_{ll}}{\sum_{j=1}^{C} N_j} \times 100\% \tag{36}$$

where q_{ll} is the total number of correctly classified samples in the class c_l, and N_j is the total number of samples belonging to a class j in the data set.

The description of the benchmark problem data sets along with their Imbalance Factor (I. F.) [20] are presented in Table 1. The imbalance factor is defined as

$$(\text{I.F.}) = 1 - \left(\frac{\min_{l=1\cdots C} N_l}{\sum_{j=1}^{C} N_j} \right) * C \tag{37}$$

It can be observed from Table 1 that the image segmentation problem has a well-balanced data set while the remaining data sets are unbalanced data sets with varying values of imbalance factors. Note that glass identification data set has highest imbalance factor of 0.68. The image segmentation, vehicle classification and glass identification problems are multi-category data sets, the remaining are binary data sets.

3.1. Performance evaluation using multi-category benchmark classification problems

In this section, the performance of the FC-RBF classifier is evaluated using three benchmark multi-category real-valued classification problems. The performance results are compared with other complex-valued classifiers, viz., MLMVN [11] and PE-CVNN [14]. To highlight the advantages of the orthogonal decision boundaries, the performances are also compared with a well-known real-valued classifiers, viz., the Minimal Resource Allocation Network (MRAN) [21], the Growing and Pruning Radial Basis Function Network (GAP-RBFN) [22], the Online Sequential Extreme Learning Machine (OS-ELM) [23], the Support Vector Machines (SVM) [24], the Sequential Multi Category RBF (SMC-RBF) [27], the Self-regulatory Resource Allocation Network (SRAN) [25], and the Real Coded Genetic Algorithm Extreme Learning Machines (RCGA-ELM) [26].

The classification performance study of the various classifiers on the well balanced image segmentation problem is presented in

Please cite this article as: R. Savitha, et al., A fully complex-valued radial basis function classifier for real-valued classification problems, Neurocomputing (2011), doi:10.1016/j.neucom.2011.05.036

Table 2
Performance results for the image segmentation problem (balanced data set). The results for MRAN, GAP-RBF, OS-ELM, and SVM are reproduced from Ref. [27].

Domain	Algo.	Image segmentation			
		K	Training time[a]	Testing	
				η_o	η_a
Real	MRAN	76	783	86.52	86.52
	GAP-RBF	83	365	87.19	87.19
	OS-ELM	100	21	90.67	90.67
	SVM	96[b]	721	90.62	90.62
	SMC-RBF [27]	43	142	91	91
	RCGA-ELM [26]	50	–	91	91
	SRAN [25]	47	22	92.3	92.3
Comp.	MLMVN	80	1384	83	83
	PE-CVNN[c]	–	–	93.2	
	FC-RBF	**38**	**421**	**92.33**	**92.33**

[a] Time in seconds.
[b] Support vectors.
[c] A single layer network was used in [14]. Also, in [14], a 10-fold validation has been done using 90% of the total samples in training and the remaining 10% for testing in each validation. Training/testing samples used in our work are shown in Table 1.

Table 3
Performance results for benchmark multi-category classification problems with unbalanced data sets. The results for MRAN, GAP-RBF, OS-ELM, and SVM are reproduced from Ref. [27].

Domain	Algo.	Vehicle classification				Glass identification			
		K	Training Time[a]	Testing		K	Training Time[a]	Testing	
				η_o	η_a			η_o	η_a
Real	MRAN	100	520	59.94	59.83	51	520	63.81	70.24
	GAP-RBF	81	452	59.24	58.23	75	410	58.29	72.41
	OS-ELM	300	36	68.95	67.56	60	15	67.62	70.12
	SVM	234[b]	550	68.72	67.99	102	320	64.23	60.01
	SMC-RBF [27]	75	120	74.18	73.52	58	97	78.09	77.96
	RCGA-ELM [26]	75	–	74.2	74.4	60	–	78.1	–
	SRAN [25]	55	113	75.12	76.86	59	28	86.21	80.95
Comp.	MLMVN	90	1396	78	77.25	85	1421	73.24	66.83
	PE-CVNN[c]	–	–	78.7	–	–	–	65.5	–
	FC-RBF	**70**	**678**	**77.01**	**77.46**	**90**	**452**	**83.76**	**80.95**

[a] Time in seconds.
[b] Support vectors.
[c] A single layer network was used in [14]. Also, in [14], a 10-fold validation has been done using 90% of the total samples in training and the remaining 10% for testing in each validation. Training/testing samples used in our work are shown in Table 1.

Table 2. The results for the MLMVN are generated using the toolbox available in the author's web site.[2]

It can be observed from Table 2 that the FC-RBF classifier outperforms other complex-valued classifiers available in the literature. This is because the FC-RBF classifier uses a fully complex-valued activation function at the hidden layer. The performance of the MLMVN classifier is further affected by the transformation used in the input layer to map the real-valued features.

It can also be noted from the table that the performance of the FC-RBF classifier is better than the best performing real-valued classifiers. The better performance of the FC-RBF classifier may be attributed to the presence of orthogonal decision boundaries.

Table 3 presents the performance results of the FC-RBF classifier on the vehicle classification and the glass identification problems. From the table, it can be observed that the performance of the FC-RBF classifier is better than the other complex-valued classifiers (MLMVN and PE-CVNN) available in the literature. The advantage of the FC-RBF classifier in performing the classification on the unbalanced data sets is prominent on these data sets, especially, more so in the glass identification problem with a highly unbalanced data set.

It is to be noted that the FC-RBF classifier outperforms all the real-valued classifiers in the classification of the vehicle classification data set. In the glass identification problem, the performance of the FC-RBF classifier is better than all the real-valued classifiers except that the overall efficiency of the SRAN classifier is a little better than the FC-RBF classifier. However, unlike the FC-RBF classifier that uses all the sample in the learning process, the SRAN classifier uses only the samples with a higher information content to participate in learning. Moreover, the transformation used at the input layer (of the FC-RBF) maps the real-valued input features only to two quadrants of the Complex domain and does not fully exploit the orthogonal decision boundaries of the FC-RBF network.

[2] http://www.eagle.tamut.edu/faculty/igor/Downloads.htm

Please cite this article as: R. Savitha, et al., A fully complex-valued radial basis function classifier for real-valued classification problems, Neurocomputing (2011), doi:10.1016/j.neucom.2011.05.036

3.2. Performance evaluation using binary benchmark classification problems

In this section, we evaluate the performance of the FC-RBF classifier on four benchmark binary classification data sets from the UCI machine learning repository as described in Table 1. The performance of the FC-RBF classifier is studied in comparison with the real-valued Optimization based Extreme Learning Machine (O-ELM) [28], SVM and the SRAN classifiers.

Table 4 presents the performance comparison of the four classifiers on four binary benchmark data sets. From the results, it can be seen that the performance of the FC-RBF classifier is better or similar to that of the best performing results available in the literature for these problems. The performance of the FC-RBF classifier can be still improved with a transformation that maps the real-valued input features to all the four quadrants of the Complex domain.

3.3. Performance evaluation using a real world problem: acoustic emission signal classification

Acoustic emission signals are electrical version of the stress or pressure waves, produced due to the transient energy release caused by irreversible deformation processes in the material [18]. Superficial similarities exist between acoustic emission signals produced by different sources that make their classification difficult. The complexity of classification is further enhanced by the existence of the ambient noise and the pseudo acoustic emission signals.

In this study, we use the FC-RBF classifier to classify the acoustic emission signal data collected in [18]. For further details of the input features and the experimental set up used, one should refer to [30]. Table 5 presents the performance results of the FC-RBF classifier in comparison with the other results for the problem, available in the literature. The results of the FC-RBF classifier is compared against a Fuzzy K-means clustering algorithm [18], the ant colony optimization algorithm [29] and genetic programming [30]. It can be seen that the FC-RBF classifier requires only 10 neurons to achieve an over-all testing efficiency of 96.35%, which is better than the classification efficiency of the other classifiers available in the literature.

Table 4
Performance comparison on benchmark binary classification problems. The results for O-ELM and SVM are reproduced from [28].

Problem	Type	Algorithm	K	Training time (s)	Testing efficiency (η_o)
Liver disorders	Real-valued	SVM	158	0.0972	68.24
		O-ELM	131	0.1734	72.34
		SRAN	91	3.38	66.9
	Complex-valued	**FC-RBF**	**20**	**133**	**74.6**
PIMA data	Real-valued	SVM	209	0.205	76.43
		O-ELM	217	0.2867	77.27
		SRAN	97	12.24	78.53
	Complex-valued	**FC-RBF**	**20**	**130.3**	**78.53**
Breast cancer	Real-valued	SVM	190	0.1118	94.20
		O-ELM	66	0.1423	96.32
		SRAN	7	0.17	96.87
	Complex-valued	**FC-RBF**	**10**	**158.3**	**97.12**
Ionosphere	Real-valued	SVM	30	0.0218	90.18
		O-ELM	32	0.0359	89.48
		SRAN	21	3.7	90.84
	Complex-valued	**FC-RBF**	**10**	**186.2**	**89.48**

Table 5
Performance comparison results for the acoustic emission problem.

Classifier domain	Classifier	Testing	
		η_o	η_a
Real-valued	Genetic Programming	90.58	–
	Ant Colony Optimization	90.51	89.27
	Fuzzy C-means Clustering		93.34
	SRAN	**99.27**	**98.9**
Complex-valued	**FC-RBF**	**96.35**	**95.2**

4. Conclusions

In this paper, we have presented a Fully Complex-valued Radial Basis Function (FC-RBF) classifier for performing real-valued classification tasks. While the neurons in the input layer of the FC-RBF classifier employ the phase encoded transformation, the neurons in the hidden layer use a fully complex-valued activation function of the type of hyperbolic secant. An analytical proof for the existence of orthogonal decision boundaries in FC-RBF has also been presented. The classification performance of the FC-RBF classifier is evaluated using a set of benchmark real-valued multi-category and binary classification problems from the UCI machine learning repository. The FC-RBF classifier is also used to solve a practical acoustic emission signal classification problem. Performance results show the superior classification ability of the FC-RBF classifier. Future works include a study on the performance of the FC-RBF classifier with a self-regulatory learning system to enable selective participation of samples in the training process and identification of a $\Re \rightarrow \mathbb{C}$ transformation that maps the real-valued input features to all the quadrants of the Complex plane, and completely exploits the advantages of the orthogonal decision boundaries.

Acknowledgment

The authors would like to thank the Ministry of Education (MOE), Singapore, for their financial support through the TIER I grant to conduct this study.

References

[1] K. Burse, R.N. Yadav, S.C. Shrivastava, Channel equalization using neural networks: a review, IEEE Trans. Syst. Man Cybern., Part C: Appl. Rev. 40 (3) (2010) 352–357.
[2] J.C. Bregains, F. Ares, Analysis, synthesis, and diagnostics of antenna arrays through complex-valued neural networks, Microwave Opt. Technol. Lett. 48 (8) (2006) 1512–1515.
[3] R. Savitha, S. Vigneshwaran, S. Suresh, N. Sundararajan, Adaptive beamforming using complex-valued radial basis function neural networks, in: Proceedings of the TENCON'09, IEEE Region 10 Annual International Conference, Singapore, November 23–26, 2009, pp. 1–6.
[4] N. Sinha, M. Saranathan, K.R. Ramakrishna, S. Suresh, Parallel magnetic resonance imaging using neural networks, Proceedings of ICIP'07 IEEE International Conference on Image Processing, vol. 3, 2007, pp. 149–152.
[5] T. Kim, T. Adali, Fully-complex multilayer perceptron network for nonlinear signal processing, J. VLSI Signal Process. Syst. Signal Image Video Technol. 32 (1–2) (2002) 29–43.
[6] R. Savitha, S. Suresh, N. Sundararajan, P. Saratchandran, A new learning algorithm with logarithmic performance index for complex-valued neural networks, Neurocomputing 72 (16–18) (2009) 3771–3781.
[7] R. Savitha, S. Suresh, N. Sundararajan, A fully complex-valued radial basis function network and its learning algorithm, Int. J. Neural Syst. 19 (4) (2009) 253–267.

Please cite this article as: R. Savitha, et al., A fully complex-valued radial basis function classifier for real-valued classification problems, Neurocomputing (2011), doi:10.1016/j.neucom.2011.05.036

[8] R. Savitha, S. Suresh, N. Sundararajan, A self-regulated learning in fully complex-valued radial basis function networks, in: Proceedings of International Joint Conference on Neural Networks (IJCNN 2010) (doi: 10.1109/IJCNN.2010.5596781), 2010.
[9] T. Nitta, Solving the XOR problem and the detection of symmetry using a single complex-valued neuron, Neural Networks 16 (8) (2003) 1101–1105.
[10] T. Nitta, On the inherent property of the decision boundary in complex-valued neural networks, Neurocomputing 50 (1) (2003) 291–303.
[11] I. Aizenberg, C. Moraga, Multilayer feedforward neural network based on multi-valued neurons (MLMVN) and a backpropagation learning algorithm, Soft Comput. 11 (2) (2007) 169–193.
[12] I. Aizenberg, D.V. Paliy, J.M. Zurada, J.T. Astola, Blur identification by multilayer neural network based on multivalued neurons, IEEE Trans. Neural Networks 19 (5) (2008) 883–898.
[13] M.F. Amin, K. Murase, Single-layered complex-valued neural network for real-valued classification problems, Neurocomputing 72 (4–6) (2009) 945–955.
[14] M.F. Amin, M.M. Islam, K. Murase, Ensemble of single-layered complex-valued neural networks for classification tasks, Neurocomputing 72 (10–12) (2009) 2227–2234.
[15] S. Chen, S. McLaughlin, B. Mulgrew, Complex-valued radial basis function network, Part I: network architecture and learning algorithms, EURASIP Signal Process. J. 35 (1) (1994) 19–31.
[16] R. Savitha, S. Suresh, N. Sundararajan, A sequential learning algorithm for complex-valued self-regulating resource allocation network-CSRAN, IEEE Trans. Neural Networks 22 (7) (2011) 1061–1072.
[17] C. Blake, C. Merz, UCI Repository of Machine Learning Databases, Department of Information and Computer Sciences, University of California, Irvine, 1998. 〈URL: http://archive.ics.uci.edu/ml/〉.
[18] S.N. Omkar, S. Suresh, T.R. Raghavendra, V. Mani, Acoustic emission signal classification using fuzzy C-means clustering, Proceedings of the ICONIP'02 nineth International Conference on Neural Information Processing, vol. 4, 2002, pp. 1827–1831.
[19] S. Suresh, S.N. Omkar, V. Mani, T.N. Guru Prakash, Lift coefficient prediction at high angle of attack using recurrent neural network, Aerosp. Sci. Technol. 7 (8) (2003) 595–602.
[20] S. Suresh, N. Sundararajan, P. Saratchandran, Risk-sensitive loss functions for sparse multi-category classification problems, Inf. Sci. 178 (12) (2008) 2621–2638.
[21] Y. Lu, N. Sundararajan, P. Saratchandran, A sequential learning scheme for function approximation using minimal radial basis function neural networks, Neural Comput. 9 (2) (1997) 461–478.
[22] G.-B. Huang, P. Saratchandran, N. Sundararajan, An efficient sequential learning algorithm for growing and pruning RBF (GAP-RBF) networks, IEEE Trans. Syst. Man Cybern. B 34 (6) (2004) 2284–2292.
[23] N.-Y. Liang, G.-B. Huang, P. Saratchandran, N. Sundararajan, A fast and accurate on-line sequential learning algorithm for feedforward networks, IEEE Trans. Neural Networks 17 (6) (2006) 1411–1423.
[24] N. Cristianini, J.S. Taylor, An Introduction to Support Vector Machines, Cambridge University Press, Cambridge, UK, 2000.
[25] S. Suresh, K. Dong, H.J. Kim, A sequential learning algorithm for self-adaptive resource allocation network classifier, Neurocomputing 73 (16–18) (2010) 3012–3019.
[26] S. Suresh, R. Venkatesh Babu, H.J. Kim, No-reference image quality assessment using modified extreme learning machine classifier, Appl. Soft Comput. 9 (2) (2009) 541–552.
[27] S. Suresh, N. Sundararajan, P. Saratchandran, A sequential multi-category classifier using radial basis function networks, Neurocomputing 71 (7–9) (2008) 1345–1358.
[28] G.B. Huang, X. Ding, H. Zhou, Optimization method based extreme learning machine for classification, Neurocomputing 71 (1–3) (2010) 155–163.
[29] S.N. Omkar, U.R. Karanth, Rule extraction for classification of acoustic emission signals using ant colony optimisation, Eng. Appl. Artif. Intell. 21 (8) (2008) 1381–1388.
[30] S. Suresh, S.N. Omkar, V. Mani, C. Menaka, Classification of acoustic emission signal using genetic programming, J. Aerosp. Sci. Technol. 56 (1) (2004) 26–41.
[31] R. Remmert, Theory of Complex Functions, Springer Verlag, 1991.

Ramasamy Savitha received the BE degree in electrical and electronics engineering from Manonmaniam Sundaranar University in 2000, and ME degree in Control and Instrumentation Engineering from College of Engineering Guindy, Anna University, India in 2002. She received her PhD from Nanyang Technological University, Singapore, in 2011. Currently, she works as a research associate at the School of Computer Engineering, Nanyang Technological University, Singapore. Her research interests are neural networks and control.

Sundaram Suresh received the BE degree in electrical and electronics engineering from Bharathiyar University in 1999, and ME (2001) and PhD (2005) degrees in aerospace engineering from Indian Institute of Science, India. He was post-doctoral researcher in School of Electrical Engineering, Nanyang Technological University from 2005 to 2007. From 2007 to 2008, he was in INRIA-Sophia Antipolis, France as ERCIM research fellow. He was in Korea University for a short period as a visiting faculty in Industrial Engineering. From Jan 2009 to Dec 2009, he was in Indian Institute of Technology, Delhi, as an Assistant Professor in Department of Electrical Engineering. Currently, he is working as an Assistant Professor in School of Computer Engineering, Nanyang Technological University, Singapore since 2010. His research interest includes flight control, unmanned aerial vehicle design, machine learning, optimization and computer vision.

Narasimhan Sundararajan received the BE in Electrical Engineering with First Class Honors from the University of Madras in 1966, MTech from the Indian Institute of Technology, Madras in 1968 and PhD in Electrical Engineering from the University of Illinois, Urbana-Champaign in 1971. From 1971, he worked in various capacities in the Indian Space Research Organization (ISRO). He was also a NRC Research Associate at NASA—Ames in 1974 and a Senior Research Associate at NASA Langley in 1981–1986. He was working as a Professor in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore since Feb 1991 and retired in June 2010. Currently, he is a visiting research Professor at Sri Jayachamarajendra College of Engineering, Mysore, India. His research interests are in the areas of aerospace control, neural networks and he has published more than 200 papers and four books. He has served as an associate editor for a number of journals and has been a program committee member in a number of international conferences. Sundararajan is a Fellow of IEEE, an Associate Fellow of AIAA and also a Fellow of the Institution of Engineers, (IES) Singapore.

Hyoung Joong Kim received his BS, MS, and PhD degrees from Seoul National University, Seoul, Korea, in 1978, 1986, 1989, respectively. He joined the faculty of the Department of Control and Instrumentation Engineering, Kangwon National University, Korea, in 1989. He is currently a Professor of the Graduate School of Information Management and Security, Korea University, Korea since 2006. His research interests include parallel and distributed computing, multimedia computing, and multimedia security.

Please cite this article as: R. Savitha, et al., A fully complex-valued radial basis function classifier for real-valued classification problems, Neurocomputing (2011), doi:10.1016/j.neucom.2011.05.036