

Research on Higher Order Artificial Neuron for E-Information Systems

Manmohan Shukla, B K Tripathi

Abstract— In this paper, a novel model based on Complex Valued Single Neurons which make the composition of MLP and RBf, aggregate their response linearly and non-linearly in two new proposed single neurons respectively is presented. The learning and generalization capabilities of these neurons have been tested over various benchmark and real life problems. Since complex numbers have natural representation of phase and magnitude that locates the complex number uniquely on the plane, these neurons have shown their potent computational power over wide spectrum of geometrical transformations. The artificial neural network based applications in various enterprise information systems are growing interest for comprehensive coverage and understanding of organizational competitiveness. The neurobiologists studies have observed various ways for interaction of synaptic inputs in the dendritic tree which eventually reflect the computation performed by a neuron. The analysis of aggregation linearly or non-linearly has been a subject of interest for NN researchers. Most of the aggregated neuron models are based on 'Multiplication' in an attempt to consider non-linearity in ANN. Prior research work pertaining to implementation of the input derived from the relationship among the second and higher order polynomial as a set of inputs is available. The novel model presented in this paper has the potential to improve the ANN performance by virtue of faster response achieved as a result of aggregated input constituted of MLP and RBf.

Keywords: CVNN, CVBP, QAM, PE, MLP

I. INTRODUCTION

Presently, global competition is demanding the development of an ecosystem that is practically disseminated to get benefitted from the strategic intelligence pertaining to information systems with objective of achieving a supply chain which is integrated globally. There is a rising demand of artificial neural networks based solutions for issues and challenges pertaining to successful applications and effective management. Neural Computing on Complex plain has been emerged an important paradigm in last few years. Researchers interest on it is rising dramatically because of superior learning ability than real valued network and inherent two-dimension motion learning property. This area has shown efficient solution for problems in pattern processing like speech and image processing, telecommunication, computer vision and for various benchmark and real life problems which were intractable or extremely cumbersome when implemented using traditional computing [1]. The structure of neuron cells in human nervous system is extremely complicated in terms of physiological attributes and operations. Details of the biological neuron and their

processing are explained in [2] [3]. The biological researchers in neuroscience and psychology may be interested to understand how computation is organized in the brain in order to explain the capabilities of neurons to solve issues like learning and commonsense reasoning. But computer scientist are interested to study brain in order to investigate the concepts to be used in new computational methodologies [4], [5].

Neuron models are mathematical depiction of genetic neurons which approximate its functional proficiencies. It is very difficult to predict the way of interaction of neurons of different nature, their way of aggregation of signals and forwarding signals to next neurons. In [6],[8] author describes approaches for linear and non-linear interaction of synaptic inputs and information processing in dendritic tree. Looking into dendritic processing of biological neurons we tried little to show how complex behavior can emerge in single neuron. This article describes the interaction of input signals in a neuron unit and single neuron computation.

A series of learning algorithms have been developed for achieving fast convergence in CVNN but not much efforts have been done in the direction of architectural development [1], [9]. So it becomes necessary to find out some new neuron models that could be used as a building blocks for developing powerful networks with extensively high learning and generalization capabilities in view of low processing time, small memory requirement and lesser architectural complexity. In view of architectural diversity of neurons and simultaneously their variations in computational and learning capabilities in human brain we present here two Complex Valued Single Neuron Models based on information aggregation and processing within a neuron or unit. This article refers these new neuron models as "Unit" or "Neuron". Such unit focus on feed forward processing with one output. In its specific architecture (shown in figure 2 and 3), two computational elements named as Processing Elements (PE) interact across by either 'summing element' or 'multiplicative element' to model linear or non-linear interaction of signals in dendritic tree. The major issue discussed in this paper is the illustration and application of single neuron computation and its distributed interaction with global operations in multilayer network. Such powerful computing model will definitely yield the solution for suitable enterprise information systems with the aim of accomplishing flexibility and approachability. The organization of the paper is as follows: The Section II is presented that the architectural structure of complex valued single neuron models and also

Revised Version Manuscript Received on 16 September, 2019.

Manmohan Shukla, PSIT, Kanpur, India

B K Tripathi, H B Technical University, Kanpur, India

reviews the mathematical approaches that are used to understand neuron computing mechanism. This section also investigate the three basic methods of aggregation of complex valued signals in PE. Section III presents update equations for learnable parameters in multi-layered network using CVBP. In section IV, paper explain the learning and generalization ability for composition of transformations. This feature appears only by extending the NN to complex domain because mapping in the complex plane is angle-preserving. Section-V and Appendix ‘B’ is devoted to simulations on benchmark and real life snags and the comparison of performance on different network. This section is also presented that the results of generalization for different compositions of three transformations over geometric figures like line, circle, ellipse, parabola, etc.

II. SINGLE NEURON UNIT

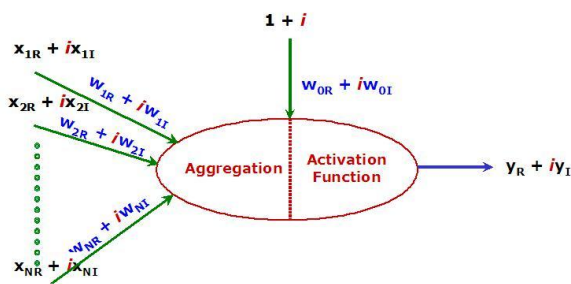


Fig. 1. A general diagram representing Single Neuron computation

The model for information processing is generally considered as neuron model and may also serve as a computational element looking it in the form of an abstraction for the functionality that it requests. For defining information processing in neuron cell, we require knowledge of process of interaction of input signals in neuron and how does conjointly influence the processing as depicted in figure 1. This article explores and models interaction and aggregation of complex valued signals in proposed neuron Unit 1. A unit is composed of following three PEs which are basic computational elements.

A. Summation

This most general PE specifies linear computation on synaptic inputs. In neural modelling, the biological behavior of dendritic linearity may be expressed as:

$$V^S = \sum_{i=0}^n w_{ij} z_i \tag{1}$$

1 The term ‘Unit’ or ‘Neuron’ may be little misleading, but we think that using a different name may be more confusing.

B. Multiplicative

In great variety of neurons in nervous system neurobiologists have popularly found multiplicative way of computations. Michal Schmitt in his study [15] shown the impact of multiplication of neural signals on learning and computation capabilities. Inspired from these concepts we

designed a new complex valued multiplicative processing for aggregation of complex signals. This processing is not simple multiplication of signals but aggregation process outputs the summation of all probable multiplications of weighted inputs, as:

$$V^\pi = \prod_{i=0}^n w_{ij} z_i \tag{2}$$

=Summation of all inputs + summation of multiplication of all grouping of two inputs + summation of multiplication of all grouping of three inputs + - - - - - + products of all inputs.

For example in case of three inputs

$$V^\pi = w_0 z_0 + w_1 z_1 + w_2 z_2 + w_3 z_3 + w_1 z_1 * w_2 z_2 + w_1 z_1 * w_3 z_3 + w_2 z_2 * w_3 z_3 + w_1 z_1 * w_2 z_2 * w_3 z_3 \tag{3}$$

C. Radial Basis Function

The introduction of Rbf networks in NN literature is stimulated from confined reaction in biological neurons. The characteristics like simpler network structure, better approximation capabilities and faster learning capabilities are basis of the proposal of an easy design of complex valued PE which applies radial basis function to determine the activity of the neuron. Where precisely, the norm $\|Z - W\|^2$ is taken and passed through an activation function that decreases as this norm increases. Here Z and W both are n-dimensional vectors. Z represents input weights and W represents synaptic weights.

$$\phi(z) = \exp(-\|Z - W\|^2)$$

where $\|Z - W\|^2 = (Z - W) * (Z - W)^H$

Here superscript H denotes complex conjugate transpose and $\phi: R^+ \rightarrow R$ is Real valued radial basis function.

In this paper we design a powerful model which utilize composite features of both concepts i.e. ‘Rbf’ and ‘MLP’. The characteristic of both ‘Rbf’ and ‘MLP’ processing elements at their output layer combine / aggregate linearly and non-linearly in two models respectively, shown in figure 2,3 and resultant is forwarded through nonlinear activation function. Thus these models enhance the overall performance and computational ability and mimics the biological activities in a single neuron.

Definition 1 HON1 unit, figure 2, is composed of linear summation, radially symmetric function and a combiner. The combiner integrate linearly the weighted output of linear summation and weighted output of radial function. This linear composition is forwarded through non-linear, non-analytic and bounded function. The unit receives one or more inputs from the environment and produces single output to the environment. The outcome of a unit is shown as:

$$O = f(V^S H + \phi(Z)K + w_0 z_0) \tag{5}$$

Definition 2 HON2 unit, figure 3, is composed of linear summation, radially symmetric function and a combiner. The Combiner integrate non-linearly the weighted output of linear summation and weighted output of radial function. This non-linear composition is forwarded through non-linear,



non-analytic and bounded function. The unit receives one or more input from the environment and produces single output to the environment. The outcome of a unit is shown as:

$$O = f\left(\prod(V^s H\phi(Z)K)\right) \quad (6)$$

Definition 3 A multilayer network, formed using units defined in definition 1 and 2, receives input from the environment and outputs to some other conventional neuron or to the environment.

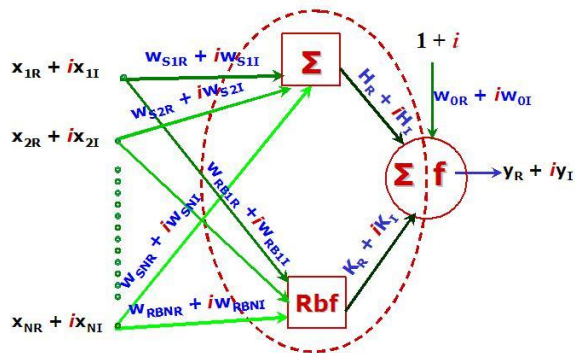


Fig. 2. HON1 Unit : a complex valued single neuron model

III. LEARNING RULE

This section explains complex version of backpropagation algorithm for proposed new models. This error back-propagation learning is concerned with adjusting 2-D motion of complex signals in CVNN.

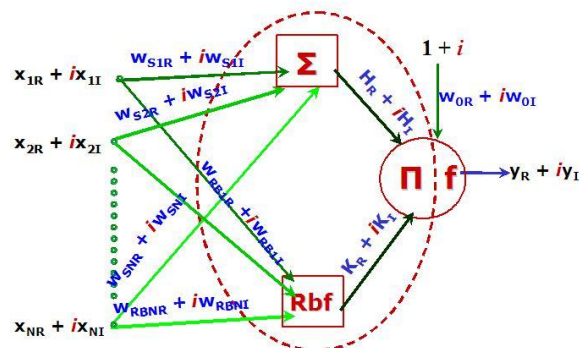


Fig. 3. HON2 Unit : a complex valued single neuron model

This error function is the square of distance amid the outcome of network and required outcome.

$$E = \frac{1}{2} |e|^2 \quad (7)$$

The output O of the neuron can be defined by traditional 2-D augmentation of real activation function for complex variables, as stated in [12]

$$O = f_c(V) = f_R({}^1V) + i * f_R({}^R V) \quad (8)$$

The author in [18] described three motives for espousing function stated in equation (8) as an activation function and proved its universal approximation capability. Numerous non-linear functions like hyperbolic tangent, sigmoidal and Gaussian function are generally used by the neural network.

The selection of function is dependent on the application targeted at by the neural network.

Any continuous nonlinear mapping can be approximated by a neural network of three layers, so we consider here same for discussion. The derivation of weight update equations and details of notations used in this article are presented in Appendix 'A'.

IV. MAPPING ON PLANE

Here we shall introduce and develop the properties of elementary transformations or mappings. In this article we consider scaling, rotation and translation (three basic classes of transformations) and their compositions. The broad arrangement of mapping which is composition of all three

$$Z = \Psi(z) = A_z + B \quad (9)$$

is evidently an extension or reduction thru a factor |A| and rotation through a fixed angle equal to Arg A in counter clock-wise direction, tailed by translation in a direction defined by the Arg B through a distance equal to |B|.

Figure 4 explain the learning and test pattern defined in this discussion. Here line is taken for test pattern but section V presents the generalization behavior of CVNN using proposed new single neuron models on other standard geometric figures.

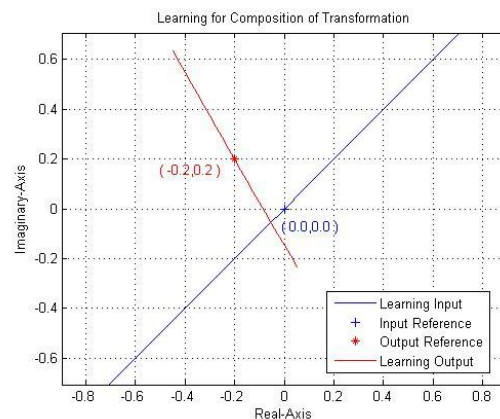


Fig. 4. Learning patterns for mathematical analysis, mapping shows scaling by factor α , angle of rotation τ and displacement by distance β in direction of σ

V. EVALUATION OF NEW STRUCTURE FOR DIFFERENT APPLICATIONS & RESULTS

To assess the métier and efficiency of NN in solving different types of problems arising from various fields, researchers have identified some Benchmark Problems. The tradeoff between training time and mapping accuracy lead to iterative adjustment of learning parameter using simulation.

A. Classification Problem

The experiments were run ten times and averaged values of epochs for convergence have been tabulated in Appendix 'B'. We used different learning rate at different class of neurons or at different layer, a ingenious select of learning rate gives good results.



C-XOR Problem: The mapping of XOR is famous non-linearly separable problem. The comparison of training and testing performance of 2-2-1 network with new neurons and 2-4-1 network taken in [17] is shown in Table I and in Figure 5.

Results shows improved training and testing enactment in novel arrangement while learnable parameters are same.

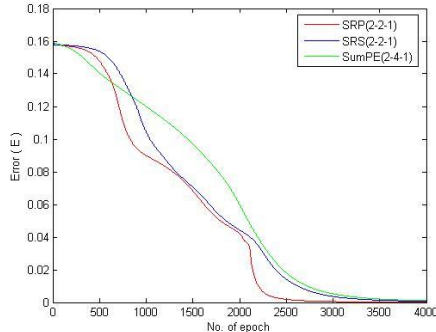


Fig. 5. Learning comparison of C-XOR with three networks

The Detection of Symmetry: Set input x to real part of input and 0 to imaginary part of input. Set y to real part of output and add one more bit y_1 to represent imaginary part of output. When input is symmetric or input is asymmetric with x_3 is 1, y_1 is 1 else

Prearranged input-output mapping is rephrased in Table II of Appendix ‘B’.

In experiment we find that only one proposed neuron is capable to solve detection of symmetry problem, without any theoretical presumptions. In [19], author have solved this problem with two ‘three-layered’ CVNN, 3-1-1 and 3-2-1, using two approaches of CVBP (conventional and improved) and also repeated same experiment with 3-2-2 and 3-4-2 real valued network. In table III we have presented the comparative performance of our model with their work. We find that HON2 model in 3-1 network give best performance among all with least complexity. In view of potent computational power of multiplicative unit described in section-II we have also considered a 3-1-1 network, with one product unit at hidden layer, it perform better than 3-1-network with improved complex-BP and is comparable to 3-2-1 network with improved complex-BP of [19].

The 4-bit Parity Problem: Using divide and conquer approach a 4-bit parity problem may be thought as combination of two 2-bit parity problem. The two bits of 2-bit parity sub problem are encoded as one complex number, first bit as real part and second bit as imaginary part. Similarly, the output of two sub problems has been encoded as a complex number expressing output of problem.

The parity of 4-bits :

Even if parity of two sub problems are simultaneously even or odd

Odd if parity of two sub problems are different

Where output $1 + j*1$ and $0 + j*0$ are interpreted as 1 or even parity and $1 + j*0$ and $0 + j*1$ are interpreted as 0 or odd parity. In this scheme, 4-bit parity input-output mapping on real plane has been encoded as two input and one output mapping on complex plane as presented in Table IV, which reduces the number of connections and parameters

tremendously in CVNN.

Tesauro and Janssens in [13] studied parity problem for a variety of scales using real back-propagation. They used an N-2N-1 network, for N ranging from 2 to 8. In our study, we solved this problem with 2-2-1 network where hidden layer consists of two proposed neurons and one summing neuron at output layer. The experiments have been performed for both models with target error $E = 0.001$. Whole experiment repeated 20 times and on average testing results give only 5% misclassification. Comparison on structure, parameter and connections are presented in Table V.

B. Channel Equalization Problem

Figure 6 shows the 4-QAM training set channel output distribution. Both HON1 and HON2 equalizer has 3-1-1

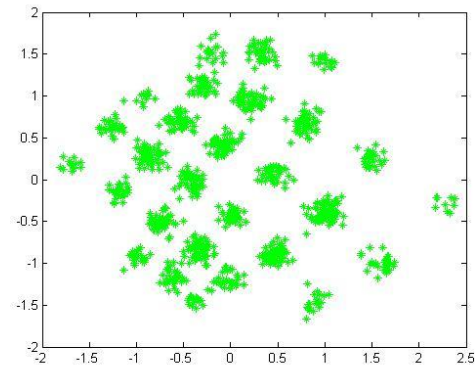


Fig. 6. The distribution of input data $y(n)$ of Equalizer network structure. Thus, only one proposed unit at hidden layer is capable to classify testing symbols clearly. The output of equalizer in figure 7 and 8 shows that only two data in HON2 equalizer and approximately 10 data in HON1 equalizer are misclassified among 10,000 testing data. In experiments we have also found that on increasing number of units at hidden layer, both equalizers are able to classify symbols with SNR less than 23dB.

C. Ionosphere data

This radar data was collected by a system in Goose Bay, Labrador and available in Johns Hopkins University

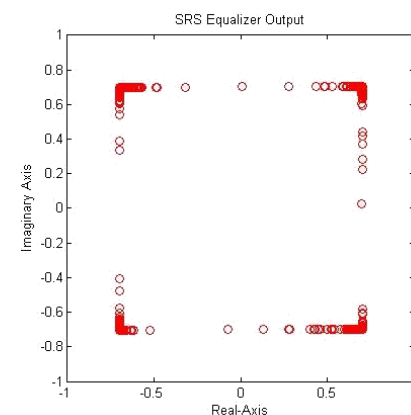


Fig. 7. Eye diagram of the outputs of the HON1 Equalizer



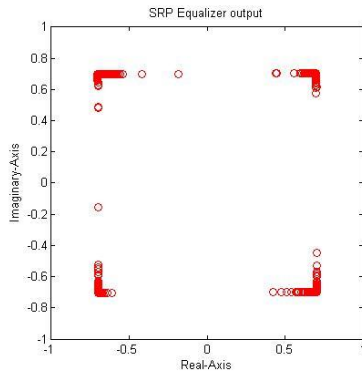


Fig. 8. Eye diagram of the outputs of the HON2 Equalizer

Ionosphere database [11]. This classification problem has been solved by three network. The training and testing performance of different network is shown in table VI. “Good” class is interpreted as $1+j*0$ and “bad” class is interpreted as $-1+j*0$.

D. Two Dimension Transformations

In our experiments we have tested the performance of both proposed new neuron models for composition of two or three transformations over a set of standard geometric figures like circle, square, ellipse and parabola and obtained fine results. We have taken 2-1-2 network where hidden layer consists any of only one proposed unit while output neurons are conventional summing neuron. First input of input layer takes a set of points lying on locus of figure and second input is reference point of input figure, similarly the first neuron of output layer gives the locus of transformed figure and second output is reference point of transformed figure. In all cases network is trained by a set of points lying on the line passing through a reference point (may be origin) and tested over standard geometric figures.

Examples presents learning and generalization of mapping over basic curves of geometry. Transformations of various curves are shown in figures, Blue color figure represent the input test figure, desired output is shown by Green color and actual results are shown in red color.

Translation and Scaling: First we investigate the behavior of the network which learned similarity transformation and translation in combination.

$$Z = \Psi1(z) = \alpha z + B \tag{12}$$

The mapping $\Psi1$ in equation (12) dilates or contracts the vector z by a factor α and displaces by means of the vector B C which gives the distance and direction of displacement on complex plane.

Rotation and Translation: Here we investigate the behavior of the network which learned the composition of rotation and Translation.

$$Z = \Psi2(z) = A_z + B \tag{13}$$

Where $A = e^{i\tau}$ in equation (13) rotate the vector z by τ in counterclockwise direction and displaces by means of the vector B C which gives the distance and direction of displacement on complex plane.

Proposition 5.2: Let vector of length $r, R; (-1 \leq r, R \leq 1)$ and argument $\vartheta, \varphi; (0 \leq \vartheta, \varphi \leq 2\pi)$. The CVNN has learned

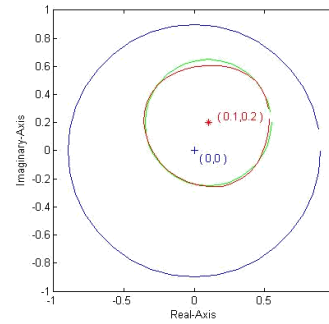


Fig. 9. Translation and Scaling with HON2 neuron

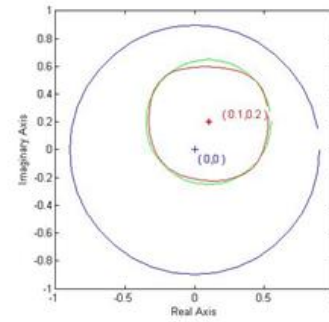


Fig. 10. Translation and Scaling with HON1 neuron

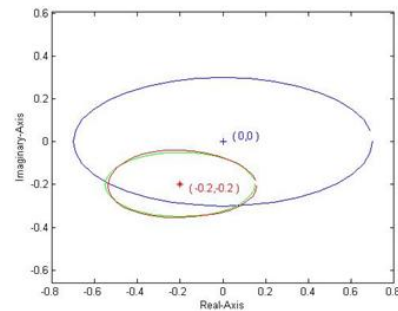


Fig. 11. Translation and Scaling with HON2 neuron

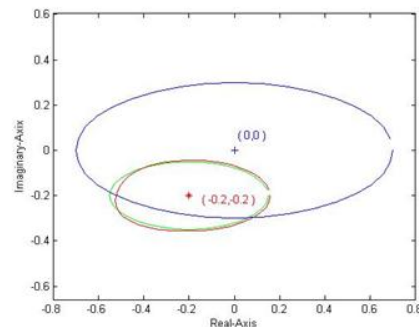


Fig. 12. Translation and Scaling with HON1 neuron

the mapping $\Psi2 : C \rightarrow C$, over a set of points, from $r e^{i\vartheta}$ to $(r e^{i(\vartheta+\tau)} + \beta e^{i\varphi})$. The input test points are set of points $R e^{i\varphi}$ on locus of geometric curve then network generate locus of curve with points $(R e^{i(\varphi+\tau)} + \beta e^{i\sigma} + [T heError])$. Where $[T heError]$ denotes the difference between actual and expected output test points in transformation. Note that network has learned angle of rotation τ and displacement distance β in direction of σ .



Example The learning patterns z are rotated counterclockwise over $\pi/2$ radians and translated by $B = (0 + i*0.2)$. So such 19 training inputs lay on the line $y = x$, ($-0.636 \leq x \leq 0.636$) and training output lie on the line $y = -x + 0.2$, ($-0.636 \leq x \leq 0.636$). As shown in figure 13 and 14, test input points lay on the parabola $y^2 = 4px$ with vertex at origin and focus at distance $p = 0.2$, desired output should map on the parabola $x^2 = 4p(y - 0.2)$, which is rotated by $\pi/2$ and vertex displaced at $(0,2)$.

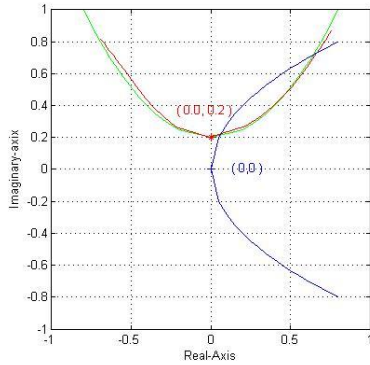


Fig. 13. Translation and Rotation with HON2 neuron

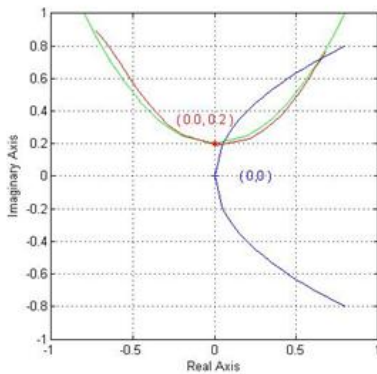


Fig. 14. Translation and Rotation with HON1 neuron

Scaling and Rotation: Here we investigate the behavior of the network which learned the composition of rotation and scaling.

$$Z = \Psi_3(z) = A_z \tag{14}$$

Where $A = \alpha e^{i\tau}$ in equation (14) rotate the vector z by τ in counterclockwise direction and dilates or contracts by a factor α on complex plane.

Proposition 5.3: Let vector of length $r, R; (-1 \leq r, R \leq 1)$ and argument $\vartheta, \phi; (0 \leq \vartheta, \phi \leq 2\pi)$. The CVNN has learned the mapping $\Psi_3 : \mathbb{C} \rightarrow \mathbb{C}$, over a set of points, from $r e^{i\vartheta}$ to $\alpha r e^{i(\vartheta+\tau)}$. The input test points are set of points $R e^{i\phi}$ on locus of geometric curves then network generate locus of curve with points $(\alpha R e^{i(\phi+\tau)} + [\text{The Error}])$. Where [The Error] denotes the difference between actual and expected output test points in transformation. Note that network has learned angle of rotation τ and scaling factor α .

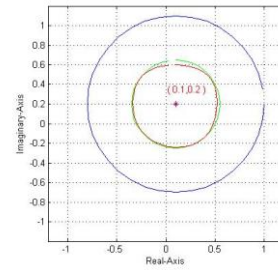


Fig. 15. Scaling and Rotation with HON2 neuron

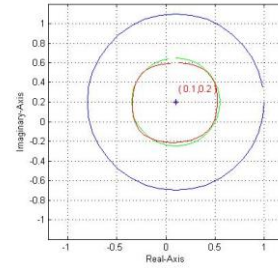


Fig. 16. Scaling and Rotation with HON1 neuron

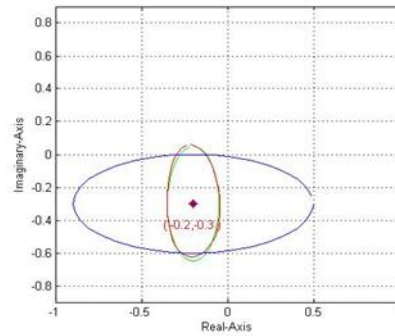


Fig. 17. Scaling and Rotation with HON2 neuron

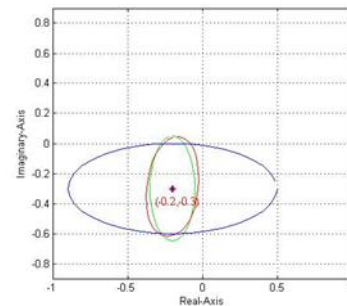


Fig. 18. Scaling and Rotation with HON1 neuron

Scaling, Rotation and Translation: Here we investigate the behavior of the network which learned the composition of all three Transformations.

$$Z = \Psi(z) = A_z + B \tag{15}$$

Figure 21 and 22 shows this transformation using network with HON2 and HON1 units respectively.

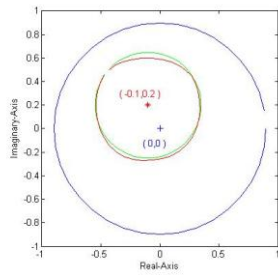


Fig. 19. Scaling,Rotation and Translation with HON2 neuron

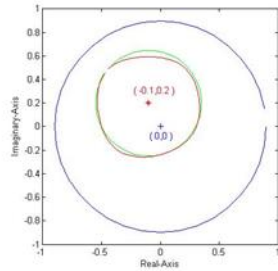


Fig. 20. Scaling,Rotation and Translation with HON1 neuron

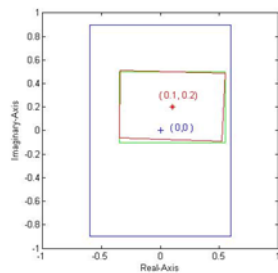


Fig. 21. Scaling,Rotation and Translation with HON2 neuron

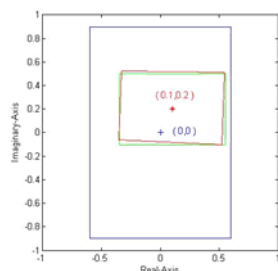


Fig. 22. Scaling, Rotation and Translation with HON1 neuron

VI. CONCLUSION

The performance and efficiency of both complex valued single neuron models along with their use in network has been tested on wide variety of benchmark and real life problems. The models have been compared with each other and with other work not only on the basis of their connectivity but also on the basis of their functional ability. Apart from these models it was also seen that new generalized definition of 'multiplicative PE' in complex domain have boosted up the computational power of neuron and surely it will promote for future research. All simulation results have shown the superiority of non-linear aggregation over linear one. Both

equalizers performed well specially for severely non-linear distorted and rapidly varying channels.

This article explored the generalized mapping over standard geometric structures. The Single Units presented here are strong enough to solve many hard problems like Detections of Symmetry, Ionosphere data etc and only one unit in hidden layer of a three layered network is able to solve Channel Equalization problem and perform composition of transformations. Although in comparison of two models HON2 model perform better than HON1 but present study shows that both of these neuron models improve learning capabilities in terms of convergence speed and performance with same or less number of neurons in the network as compared to standard CVNN. They also requires same or less number of connections and parameters for same or better performance than in conventional NN and hence significantly reduces overall burden of network. This novel work will definitely provide timely and in-depth analysis on the sustained development distributed management technologies with a clear interdisciplinary approach.

REFERENCES

1. D Mandic and V S L Goh, "Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models", Wiley 2009.
2. C. Koch, "Biophysics of Computation: Information Processing in Single Neurons", Oxford University Press, 1999.
3. H. C. Tuckwell, "Introduction to Theoretical Neurobiology", Cambridge University Press, 1988.
4. B K Tripathi and P K Kalra, "On Efficient Learning Machine with Root Power Mean Neuron in Complex Domain", IEEE Transaction on Neural Network, vol. 22, no. 05, pp. 727-738, May 2011.
5. Vivek Srivastava, B K Tripathi and V K Pathak, "Hybrid computation model for intelligent system design by synergism of modified EFC with neural network", International Journal of Information Technology and Decision Making, World Scientific, vol 14, no 1, pp 17-41, January 2015.
6. B W Mel, "Information Processing in dendritic tree", Neural Computation, vol.6, pp. 1013-1085, 1995.
7. C. Koch and T. Poggio, "Multiplying with synapses and neurons", Single Neuron Computation, Academic Press: Boston, Massachusetts, pp.315-315, 1992.
8. B A Arcas, A L Fairhall and W. Bialek, "What can a single neuron compute ?", Advances in Neural Information Processing, T. Leen, T. Dietterich and V. Tresp (Eds.), pp. 75-81, 2001.
9. M. Fajjul Amin and K Murase, "Single-layered complex-valued neural network for real-valued classification problems", Neurocomputing, 2008.
10. C C Lee, P C Chung, J R Tsai and C I Chang, "Robust Radial Basis Function Neural Network", IEEE transaction on Systems, Man and Cybernetics-Part B: Cybernetics, vol.29, No.6, Dec. 1999.
11. C. L Blake and C J Merz, "UCI Repository of machine learning database, <http://www.ics.uci.edu/mealmn/MLRepository.html>", University of California, Department of Information and computer science, 1998.
12. F. Piazza and N. Benvenuto, "On the complex Backpropagation Algorithm", IEEE Transaction on Signal Processing, Vol.40, No.4, pp.967-969, April 1992.
13. G. Tesauro and B. Janssens, "Scaling Relationships in Backpropagation Learning", Complex Systems 2, Pages 39-84 1988.
14. Cha S kassam, "Channel Equalisation using Complex Valued Radial Basis function networks", IEEE. Journal Select. Areas Commun, vol13, pp.122-131, Jan, 1995.
15. Michael Schmitt, "On the Complexity of Computing and Learning with multiplicative Neural networks", Neural Computation, 14, 241-301 (2001).

16. R. Durbin and D. E. rumelhart, "Product unit: A computationally powerful and biologically plausible extension to back-propagation net-works", *Neural Computation*, vol.1, pp. 133-142, 1989.
17. Tripathi B K, Chandra B and Kalra P K, "Complex Generalized-Mean Neuron Model and its Applications", *Applied Soft Computing* (Elsevier), vol. 11, no. 01, pp. 768 -777, January 2011.
18. T. Nitta, "Orthogonality of Decision Boundries in Complex-valued Neural Networks", *Neural Computation*, Vol.16No.1,pp.73-97,2004.
19. Xiaoming Chen and Songsong Li, "A Modified Error Function for the Complex-Value Back propagation Neural Network", *Neural Information Processing*, vol.8,No.1, 2005.
20. T. Nitta, "An analysis of thefundamental structure of complex-valued neurons", *Neural Processing Letters* 12 : 239-246,2000