2014 IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)

Implementation of Single Neuron Using Various Activation Functions with FPGA

S.Jeyanthi[1], M. Subadra[2]
[1]ME-Applied Electronics, [2]Professor & Head
[1,2,] Department of Electronics & Communication Engg, Einstein College of Engineering, Tirunelveli-12
[1]sjeyanthi12@yahoo.in,[2]subadra_m@yahoo.com

***Abstract*-A main characteristic element of any artificial neural network (ANN) is an activation function since their results are used as a starting point for any complex ANN application. This paper is about design and implementation of single neuron with three different types of activation function sigmoid, linear and threshold activation function. Activation function package is written using VHDL and single precision floating point representation method is used to design the single neuron. The simulation result is taken using Xilinx ISE14.5i and the performance of different activation functions are analyzed and compared in terms of device utilization and CPU time.**

Keywords- Artificial neuron, activation function, FPGA VHDL.

I. INTRODUCTION

Most attractive properties of ANNs is the capability to adapt their behavior by changing characteristics of the modeled system [1]. ANN's decision making ability and performance depend on several key factors such as the structure of network, type of activation functions in the neurons, period of learning, and number of stimulus used during the learning process [2]. Therefore designing of special activation functions is necessary to simplify the network structure and also to accelerate convergence time and learning process [3]. ANN real time applications are possible when low cost, high speed neural network computation is made realizable using hardware. This hardware realization preserves the inherent parallelism of ANN [4].

This work concentrates on designing of digital artificial neuron using different activation functions. Activation functions have the following characteristics [5].

1. On-chip/off-chip
2. Analog/digital
3. Computation method (threshold/look-
4. up table)
5. Clock and data transfer rates
6. Cascadability

The selection of topology depends upon the above mentioned characteristics. The hardware topology, and the types of activation function determines the quality of the network, speed of conversion and the efficiency of the synaptic weights updates [6]. As a result, a careful selection of the activation function has a huge impact on the network performance [7]. Thus this paper aims at the implementation of various activation functions and also to analyze the performance in terms of device utilized and CPU time.

In this paper neuron model with three types of activation function namely sigmoid, linear and threshold is designed using VHDL. Section II reviews the literature for implementation of different activation functions. Section III describes the neuron model with different activation function and Section IV discusses the results obtained in performance analysis of activation function.

II. LITERATURE REVIEW

Sibi et al (2013) performed an analysis of the different sigmoid functions like sigmoid symmetric, sigmoid symmetric stepwise, sigmoid stepwise activation functions for classification problem using Resilient propagation algorithm. They carried out the Training activity for mushroom dataset with an expected error of 0.0999. They calculated the total number of epochs and the error at last epoch. They achieved an 0.0003930641 error for sigmoidal activation function [3].

Manish Panicker et al (2012) proposed an effective FPGA implementation method for log sigmoidal and bipolar sigmoidal function. The hardware is designed for 32 bit fixed point arithmetic. They used the piecewise linear approximation to approximate sigmoid activation function and second order approximation, to approximate bipolar sigmoid activation function. They modeled the hardware using Verilog HDL and simulated using MODELSIM 6.5. Xilinx 10.1 ISE was used to implement the designs in Spartan 3 FPGA. The maximum mean square error value was found to be 0.00187 for bipolar sigmoid activation [18].

Thamer Jamel et al (2012) proposed an approximation method for sigmoid activation function using simple optimization circuit, which is well suited for FPGA implementation. Their proposed method gave the similar

ISBN No. 978-1-4799-3914-5/14/$31.00 ©2014 IEEE

limiting function as sigmoid one, with simple digital circuit. They compared the Output of neuron using MATLAB and FPGA and they have obtained approximately equal outputs for both methods [8].

Sahin et al (2011) described an implementation of different exponential activation functions and compare their performances in terms of in terms of clock period and latency. The activation functions were designed and coded in VHDL and mapped to Virtex-6 using Xilinx's ISE Web Pack 12.1 EDA tool. They have used Xilinx's CORDIC design to calculate e^x between $-\pi/4$ and $\pi/4$. They have achieved the result of LogSig units with higher clock periods than RadBas and TanSig because of its cascaded adder design [2].

Isa Z.Saad et al (2010) investigated the different activation functions in MLP networks in terms of the accuracy performances. They performed the analysis of MLP classification by comparing the 6 different activation functions for same data set of breast cancer thyroid diseases. Results of classification showed that the MLP network with hyperbolic tangent activation is suitable for MLP network to classify data of two classes and neuronal function is applicable in MLP network for classifying data of three classes [4].

Banuelos saucedo et al (2003) developed a FPGA (Field Programmable Gate Array) based digital implementation of a McCulloch-Pitts type of neuron model with three types of non-linear activation function: step, ramp-saturation, and sigmoid. They implemented these three models using XC4005ECP84 FPGA. The results were obtained for the operation speed and the percentage of used resources. They concluded that step function and the ramp-saturation function require a very few configurable logic blocks (CLB) while compared with the sigmoid function [1].

From literature review sigmoid activation function with approximation method [8], threshold and linear activation functions are taken for analyzing the performance of implementation of neuron using VHDL and XILINX ISE 14.5i. These activation functions are chosen since it have the advantages of simple mathematical model, low resource utilization and speed while compare to other activation functions like tangential and Gaussian activation function.

III. NEURON MODEL

The basic processing elements of neural networks are called artificial neurons. While comparing the artificial neuron to the biologically neuron, the effects of the synapses are represented by connection weights that modulate the effect of the associated input signals, and the nonlinear characteristic exhibited by neurons is represented by an activation function (transfer function) [9].

Activation functions are the mathematical formula, which is used to determine the output of a processing node (neuron). If the weighted input sum value is greater than the threshold value, then the neuron produces output. Figure.1 shows Fundamental structure of neuron.

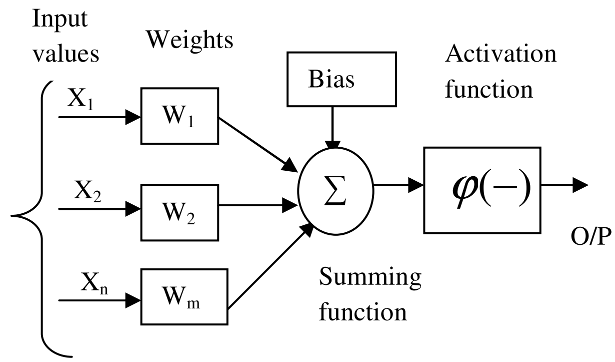

Figure.1 Fundamental structure of neuron

Each neuron contains a set of links, describing the neuron inputs(X_1,X_2,.. X_n) with weights W_1,W_2, …, W_m, an adder function for computing the weighted sum of the inputs

$$u=\sum_{j=1}^{m} w_j x_j \tag{1}$$

and an activation function φ for limiting the amplitude of neuron output and bias 'b'

$$y=\varphi(u+b) \tag{2}$$

Several different functions may be used as an activation function. Activation functions for the hidden units are needed to introduce non-linearity into the networks [12]. For the output units, activation functions should be chosen to be suited to the distribution of the target value. Different activation functions are acquired for different networks so that it results in better performances [13]. There are two types of activation functions in neural network, linear and nonlinear activation function.

A. *Threshold activation function*

This function is mainly used in single layer networks. It limits the output of neuron to one of the two values. It is also known as Heaviside function [10]. The following equation express gives the threshold AF.

Binary threshold

$$y=\begin{cases}1; & x \geq 0\\ 0; & x<0\end{cases}$$

(3)

Bipolar threshold

$$y=\begin{cases}1; & x \geq 0\\ -1; & x<0\end{cases}$$

B. *Linear activation function*

It is also known as identity function. Different types of linear function expressed in table.1. It is a flow-through mapping of the perceptron's potential to its output. These functions have differentiable and unbounded characteristics. Figure.2 represents the linear activation function graph.

TABLE.1 DIFFERENT TYPE OF LINEAR ACTIVATION FUNCTION.

NAME	FUNCTION
Linear	$y = x$
Positive linear	$y = \begin{cases} x; & x \geq 0 \\ 0; & x < 0 \end{cases}$
Saturating linear	$y = \begin{cases} 1; & x \geq 0 \\ 0; & x < 0 \\ x; & 0 \leq x \leq 1 \end{cases}$
Symmetric Saturating linear	$y = \begin{cases} 1; & x \geq 0 \\ 1; & x < 0 \\ x; & 0 \leq x \leq 1 \end{cases}$

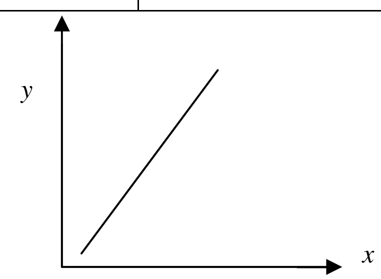

Figure.2 linear activation function

C. Sigmoid activation function

Sigmoid activation functions are the differentiable functions that map the perceptron's potential to a range of values, such as 0 to 1. This is the most commonly used soft-limiting activation function which satisfies this differentiable requirement of in back propagation neural networks. The following figure.3 shows the sigmoidal activation function graph.

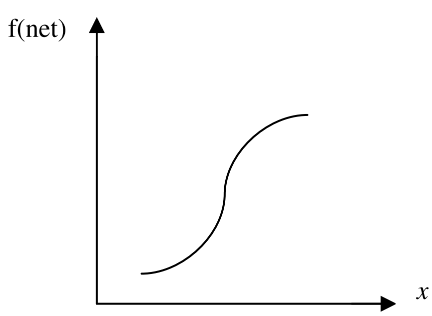

Figure.3 sigmoidal activation function

Sigmoid activation function can be represented as follows which using the logistic function equation

$$f(net) = 1/(1+e^{-net}) \qquad (5)$$

This function contains exponential term which is difficult to calculate and direct implementation is also difficult therefore approximation functions are used to implement the sigmoid function on hardware. In this work following approximation function is used [8].

$$f(net) = 1/2\left[\frac{net}{|net|} + 1\right] \qquad (6)$$

The following data flow graph (figure.4) is used to configure the hardware. This circuit takes as its input value net which represents the result of sum of product and return the output f(net). This approximation function is chosen based upon network performance and precision of accuracy [3]. The design of equation requires the divider circuit and the absolute circuit.

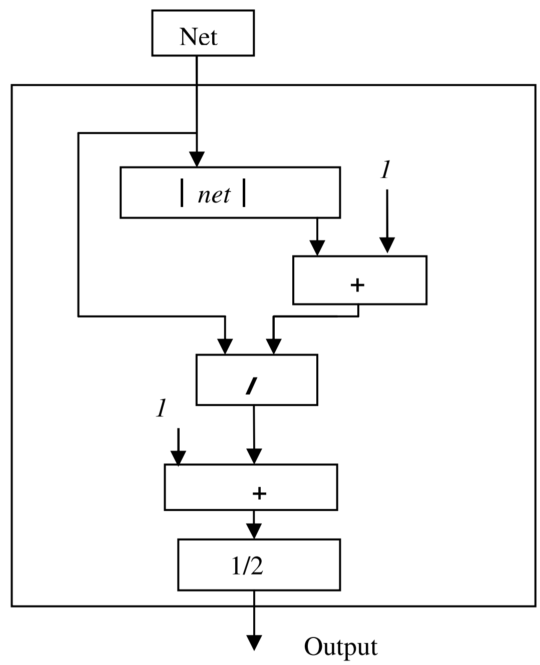

Figure.4 sigmoidal activation function with approximation function implementation.

IV. RESULTS AND DISCUSSIONS

Digital artificial neuron is designed as in section III with three type activation function using VHDL and XILINX ISE 14.5i. 32 bit single precision floating point data representation is used for accuracy. Simulation result of all activation functions are taken using XILINX 14.5i and corresponding results during performance analysis result are compared in terms of hardware resource utilization and speed.

Figure 5 & 6 shows a behavioral simulation result of bipolar and binary threshold AF.

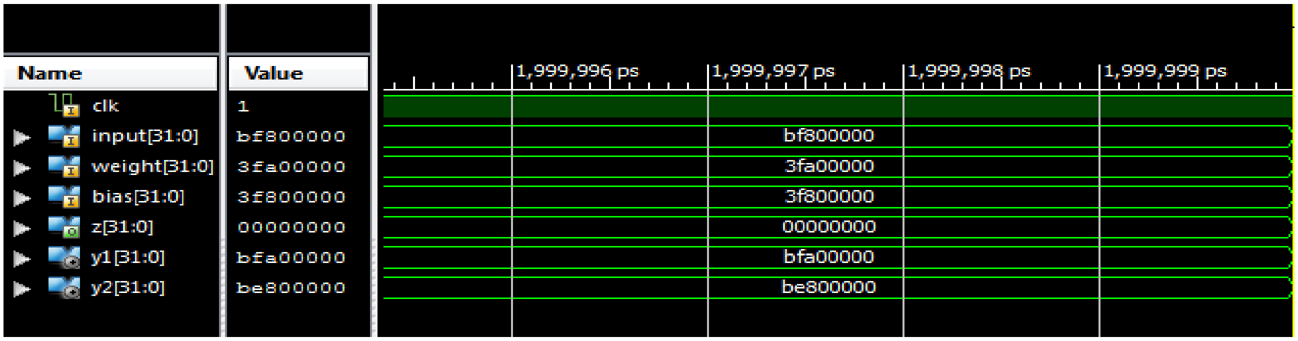

Figure 5.Simulation result bipolar threshold

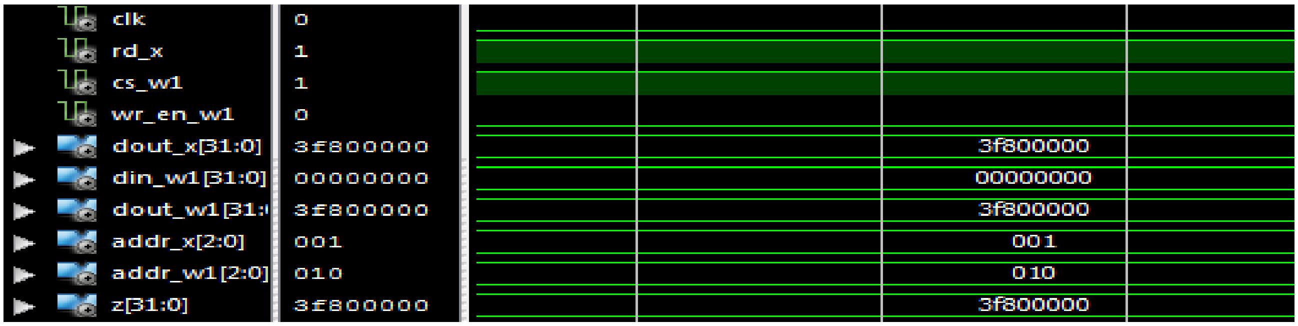

Figure 6.Simulation result binary threshold

Table.2 shows an AF output for different input, weight values. Where the bias value is 3f800000 (1) are represented by hexa decimal notation of single precision floating numbers (IEEE 754) and equivalent integer values. Output column shows that an activation of neuron based upon the different activation function mentioned in section III.

TABLE.2. OUTPUT OF SINGLE NEURON WITH DIFFERENT ACTIVATION FUNCTION

ACTIVATION FUNCTIONS	INPUT		WEIGHT		OUTPUT	
	Floating point	**Equivalent Integer**	**Floating point**	**Equivalent Integer**	**Floating point**	**Equivalent Integer**
Binary threshold. Y=1; x≥0 Y=0; x<0	3f200000	0.625	3fa00000	1.25	3f800000	1
	Bf800000	-1	3fa00000	1.25	00000000	0
Bipolar threshold. Y=1; x≥0 Y=-1; x<0	3f200000	0.625	3fa00000	1.25	3f800000	1
	Bf800000	-1	3fa00000	1.25	Bf800000	-1
Linear. Y=x	3f200000	0.625	3fa00000	1.25	40020000	2.25
Positive linear. Y= x; x≥0 Y= 0; x<0	3f200000	0.625	3fa00000	1.25	3fe40000	1.78125
	Bf800000	-1	3fa00000	1.25	Be800000	-0.25
Sigmoid. y = $1/(1+e^{-x})$	3f800000	1	3f800000	1	3f400000	0.75

linear and sigmoid AF simulation result is shown in figure 7 and 8.

Sigmoid simulation result gives more precise output between 0 to 1 values, compare than other activation functions.

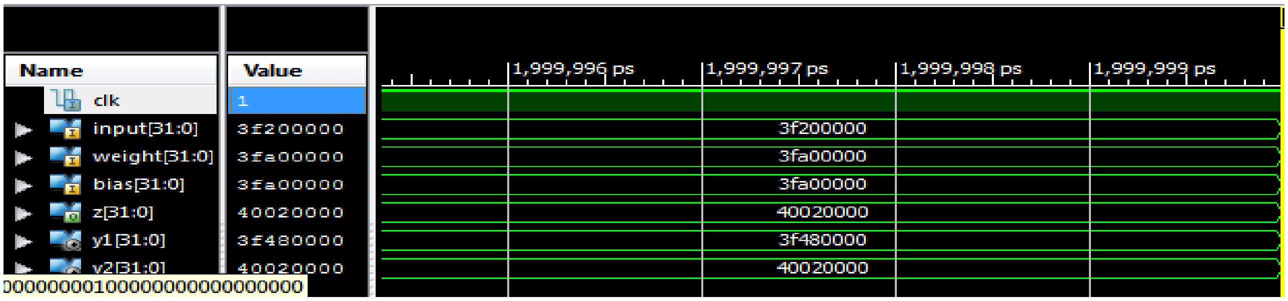

Figure 7.Simulation result of Linear function

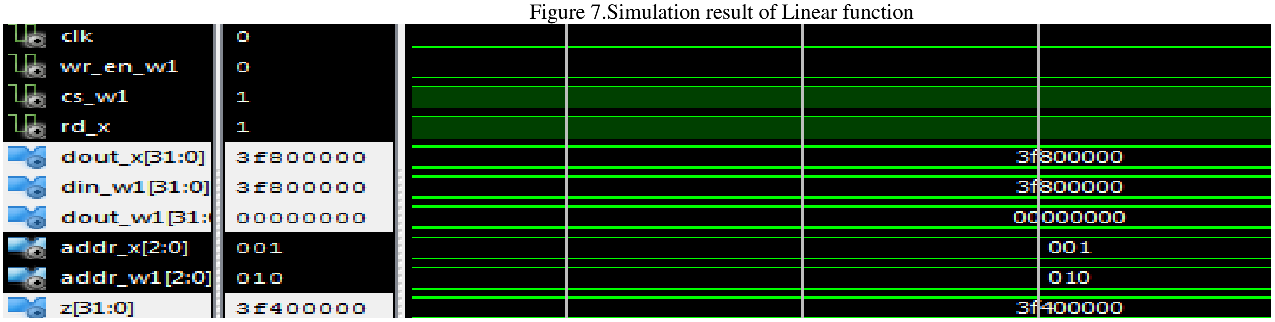

Figure 8.Simulation result of sigmoid function

TABLE.3. HARDWARE UTILIZATION

Function	Threshold	Linear	Sigmoid
Number of slices	7%	16%	84%
Number of slice FFs	14%	15%	2%
Number of 4 i/p LUTs	12%	67%	81%
Total CPU time	4409.54s	235.27s	6409.54s
Memory usage	262596 KB	238148KB	462596KB
Frequency	156.079MHz	67.547MHz	5.441MHz

Hardware utilization of different activation function is given in table.3 From table it can be seen that sigmoidal function requires more hardware resources than the threshold and linear function.

V. CONCLUSION

This work has the conclusion of successful implementation of single neuron with different activation function using VHDL and XILINX ISE 16.5i. Advantages of different activation function are discussed with performance analysis and sigmoidal activation functions with approximation method provide the best approximation with lowest error while compromising the hardware resource.

VI. FUTURE WORK

It is planned to realize, multilayer perceptron ANN using different activation function with on-line back propagation algorithm to solve standard XOR problem using FPGA with floating point precision.

ACKNOWLEDGEMENTS

This work was supported by the Department of Science and Technology (DST), Government of India under SERB Fast track Scheme for Young Scientists. The authors would like to acknowledge DST, India , management and Principal of Einstein College of Engineering, Tamil Nadu, India for providing the facility and necessary support to design and implement this work.

REFERENCES

[1] M. A. Banuelos saucedo, J. Castillo Hernandez and Quintana Thierry,"Implementation of a neuron model using FPGAs", Journal of Applied Research and Technology, pp.248-255, 2003.

[2] Sahin, I. Koyuncu, "Design and Implementation of Neural Networks Neurons with RadBas, LogSig and TanSig Activation Functions on FPGA", Electronics And Electrical Engineering, ELEKTRONIKA,2012.

[3] P.Sibi, s.allwyn and p.siddarth, "Analysis of different activation functions using back propagation neural networks", journal of theoretical and applied information technology, pp.1264-1268, 2013.

[4] Isa Z.Saad, S.Omar, M.K.Osman and K.A.Ahmad, "Suitable MLP Network Activation Functions for Breast Cancer and Thyroid Disease Detection", International Conference on Computational Intelligence, Modelling and Simulation, pp.39 - 44,2010.

[5] Jagath C. Rajapakse, Amos R. Omondi, "FPGA implementation of Neural Networks", Spriger.ISBN-10 0-387-28487-7 (e-book), 2006.

[6] Bekir and Vehbi Olgac, "Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks", International Journal of Artificial Intelligence And Expert Systems (IJAE),pp.111-122.

[7] Alin Tisan, Stefen Oniga, Daniel MIC and Attila Buchman,"Digital Implementation Of The Sigmoid Function For FPGA Circuits", ACTA Technica NAPOCENSIS Electronics and Telecommunication, 2009.

[8] Thamer and khamma, "Implementation of a sigmoid activation function for neural network using FPGA. 13th scientific conference of almamoon university college, Iraq, 2011.

[9] Djalal Eddine Khodja, Aissa Kheldoun and Larbi Refoufi, "Sigmoid Function Approximation For ANN Implementation In FPGA Devices", Proceedings of the 9th WSEAS International conference on cir- cuits,systems, electronics, pp.112- 116,2010.

[10] Bhaskar, VHDL Primer. PTR Prentice Hall, 2003.

[11] Sathish Kumar, "Neural Networks: A Classroom Approach", Tata McGraw-Hill Publishing Company Limited, New Delhi, 2004.

[12] K.Basterret, and J.M.Tarela, "Approximation of sigmoid function and the derivative for artificial neurons", Advances in neural networks and applications, WSES Press, Athens, pp.397– 401, 2001. [15] Manish Panicker, C.Babu, (2012) "Efficient FPGA Implementation of Sigmoid and Bipolar Sigmoid Activation Functions for Multilayer Percep-trons", IOSR Journal of Engineering, pp. 1352-1356, 2012,

[13] Janardan Misra and IndranilSaha, "Artificial neural networks in hardware: A survey of two decades of progress ", Elsevier, Neuro-computing 74 (2010) pp.239–255.

[14] Ashkan Hosseinzadeh Namin, Karl Leboeuf, Roberto Muscedere, Huapeng Wu, and Majid Ahmadi," Efficient Hardware Implementation of the Hyperbolic Tangent Sigmoid Function", IEEE, 2009.

[15] Emad A. M. Andrews Shenouda," A Quantitative Comparison of Different MLP Activation Functions in Classification", Springer, pp.849 – 857, 2006.

[16] Valeriu Beiu, Jan Peperstraete, Joos Vandewalle and Rudy Lau-wereins, "Close Approximations of Sigmoid Functions by Sum of Steps for VLSI Implementation of Neural Networks", The Scientific Annals, Section Informatics, 1994.

[17] [17] Ashkan Hosseinzadeh Namin, Karl Leboeuf, Huapeng Wu, and Majid Ahmadi, "Artificial Neural Networks Activation Function HDL Coder", IESEE, 2009.

[18] [18] Manish Panicker, and C.Babu," Efficient FPGA Implementation of Sigmoid and Bipolar Sigmoid Activation Functions for Multilayer Perceptrons", IOSR Journal of Engineering, 2012.