TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 15816 - 15819

FPGA Implementation of Radial Basis Function and Recurrent Neural Network for Speech Recognition

Rajasekar B[1], Peram Ashok Reddy[2], Panjugula Tejeswara Reddy[2], Balasankar Karavadi[1], Mathan N [3]
[1]Associate Professor, Department of ECE, Sathyabama Institute of Science and Technology, Chennai
[2]B.E., ECE, Sathyabama Institute of Science and Technology, Chennai
[3]Assistant Professor, Department of ECE, Sathyabama Institute of Science and Technology, Chennai
Email: rajasekar.ece@sathyabama.ac.in

Article Info
Volume 82
Page Number: 15816 - 15819
Publication Issue:
January-February 2020

Article History
Article Received*: 18 May 2019*
Revised*: 14 July 2019*
Accepted*: 22 December 2019*
Publication*: 28 February 2020*

Abstract

Neural networks can catch needlessly deterministic nonlinear and for machine learning non-parametric models it has merits such as fast, highly accurate computation when compared to other probable computing. In this paper, feasibility of Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) and Radial Basis Function Neural Network (RBFNN) for speech recognition is studied. The stochastic incline descent (SGD) method is applied to update the parameters of RBFNN and the temporal classification method is used for training the LSTM-RNN, whereas FSM is used to reluctant the hardware resource usage. The automatic proposal mechanization is done by Simulink and Xilinx is used to verify the Verilog code and it is implemented in FPGA-SPARTAN-6..

Keywords: *ANN, RBFNN, LSTM-RNN, SGD, FSM*

I. INTRODUCTION

Speech recognition is a major area of research. Speech signals are non-linear in nature and it requires non-stationary analysis. Conventionally, non-stationary signals are decomposed using Fourier Transform, spectral transform or any other unitary transform and the features are aggregated. These aggregated features are used for decision making. However, the performance is strongly dependent on the features. Hence in recent years, the paradigm has shifted to Deep Learning Neural Networks. For one Dimensional analysis, Long Short-Term Memory (LSTM) is used for speech recognition. Recurrent Neural Networks are used for non-stationary signal analysis. A combination of these two networks result in efficient classification of speech signals. In this paper, LSTM-RNN is implemented for speech recognition and LSTM-RNN is implemented using FPGA-SPARTAN-6. Considerable research is carried out in this area.

Jose Albin et al (2015) proposed an automated speech recognition system using feedforward neural networks. In this work, the features are extracted and are then used for training and testing the classifier. Sensitivity of the proposed system is high for trained dataset but not for test dataset. Vinothkumar et al (2015) used non-stationary signal analysis for extracting features from biometric signals which can be further used for classification. Sak et al (2014) used recurrent projection to improve the performance of LSTM for acoustic signal modeling. yıldırım, Özal (2018) used the sub band co-efficient as inputs to LSTM for ECG classification. Yun-Long Kong ID (2018) used LSTM for detecting disturbances in satellite images. Salma Alhagry et al (2017) used LSTM for emotion detection from EEG signals. From the literature, it is evident that LSTM can be used for analysis, prediction and classification on non-stationary signals. In this paper, LSTM is used for speech recognition. LSTM is implemented in FPGA-SPARTAN-6.

This paper is organized as follows: An overview of Recurrent Neural Networks (RNN) is provided in Section II. RTL view and schematic are dealt in Section III. Section IV deals with FPGA implementation. Section V concludes the works and provides the future directions.

Published by: The Mattingley Publishing Co., Inc.

TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 15816 - 15819

II. RECURRENT NEURAL NETWORK (RNN)

In order to perform speech recognition, the features are extracted and are then fed to neural networks. Of the various architectures, Recurrent Neural Networks (RNN) are used in this work.

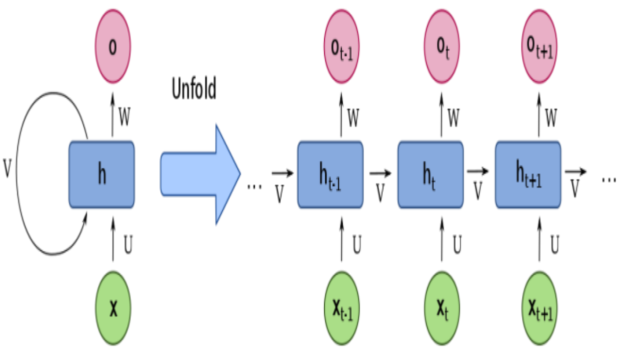

Fig 1: RNN Diagram

A Recurrent Neural Network (RNN) is, essentially, a regular ANN where some neurons (especially in the hidden layer) have feedback connections to themselves, i.e. their outputs are fed as inputs. The relevance of this different structure is the possibility to retain sequence information about the data. Before, each incoming data point only contributed to the training of the network, but the information about the correlation between themselves and the data points that preceded them did not influence the training step. The temporal relationship is disregarded and each data point considered conditionally independent of any other. This is not necessarily true, and in fact, there are many cases where the correlation between data points is high for those closely spaced in time, such as in video signals, audio signals, or other kinds of temporal sequences of data. Therefore, the feedback connection of the neuron to himself acts as a kind of memory element that takes into account in the present decision, the history of decisions previously taken, and hence the previous data.

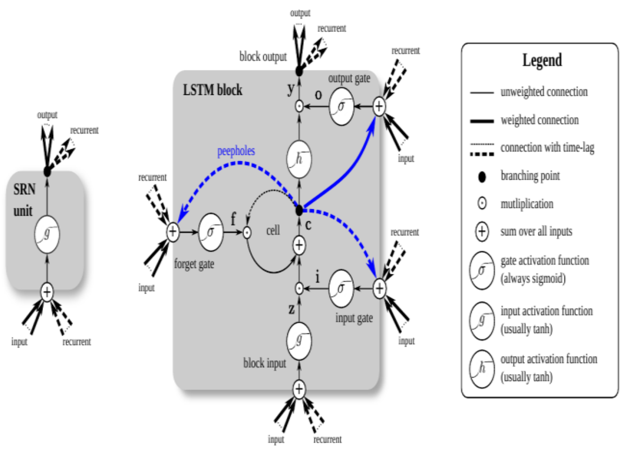

Fig 2: A complete LSTM neuron

Complexity of RNN lies in the gradient descent algorithm that does not global minimum. Hence in order to train the data, LSTM cell as shown in Fig. 3 is used. Forget unit, memory unit, input modulation unit, output gate, memory cell and hidden state are the various components of a single LSTM cell. Forget unit and input modulation unit provide the required information for further processing while training the data. Having trained the network, the next task is to test and implement the network. In this case, the input signals are fed directly to the network and speech recognition is performed.

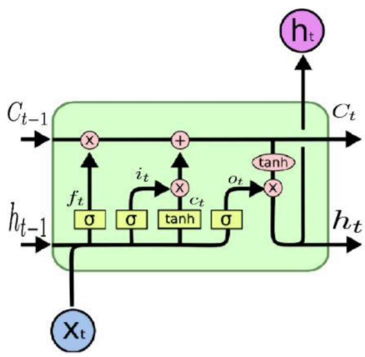

Fig 3: Single LSTM structure (courtesy: Yun-Long Kong et al (2018))

In order to perform FPGA implementation, it is necessary to formulate the mathematical equations of a single LSTM cell and hence the layers of LSTM. These equations are then implemented in Xilinx.

Published by: The Mattingley Publishing Co., Inc.

TEST Engineering & Management
January - February 2020
ISSN: 0193 - 4120 Page No. 15816 - 15819

III. RTL VIEW AND RTL SCHEMATIC

The RTL view of LSTM architecture is shown in Fig. 4 and RTL schematic for LSTM is shown in Fig.5 respectively. The proposed algorithm consists of the following steps: speech signal inputs, parameter identification and signal classification. Initially behavioral model program is written and is executed. Later test bench is provided. Performance is studied in terms of bonded IOBs, EUT-FF pairs and LUTS (Table 1). An effective design must involve in minimum utilization of the available hardware.

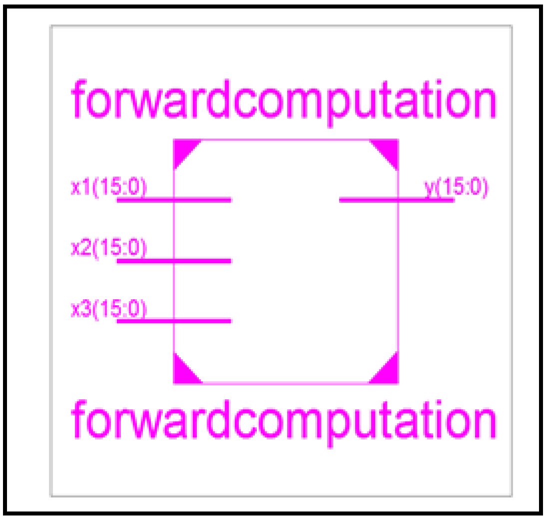

Fig 4: RTL view LSTM structure

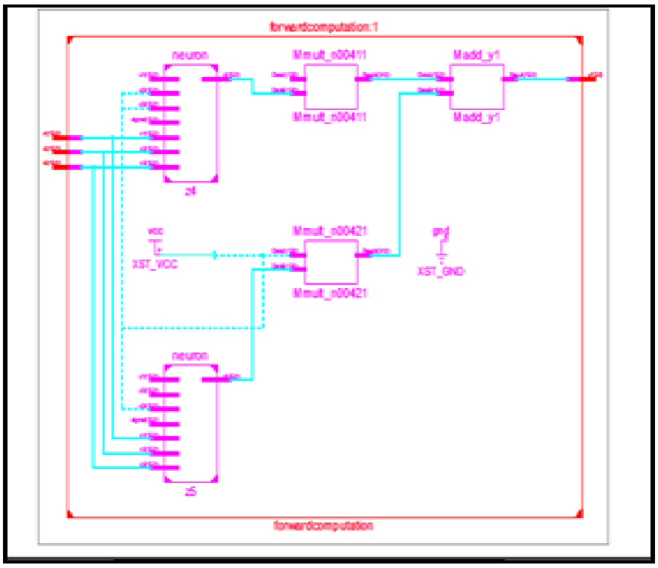

Fig 5: RTL schematic LSTM structure

Table 1: Design Utilization Summary of LSTM (AREA ANALYSIS)

Logic Utilization (Number of)	Used	Available	Utilization Percentage
Bonded IOBs	96	102	94%
Fully used LUT-FF pairs	0	87	0%
Slice LUTs	87	5720	1%
No. of DSP48A1s	11	16	68%

From the last column of Table 1, it is found that the proposed architecture uses 94% of bonded IOBs and 68% of DSPs.

IV FPGA IMPLEMENTATION

FPGA can be used for the implementation of simulated algorithms (Kavi Camilin et al (2016), Mathan et al (2018), Rahman et al (2016)). Having tested the efficiency of the proposed LSTM algorithm, the next task is to implement in FPGA. Of the various FPGA boards, FPGA-SPARTAN-6 (Fig. 6) is used as it has more logic units to support speech signal analysis and speech recognition.

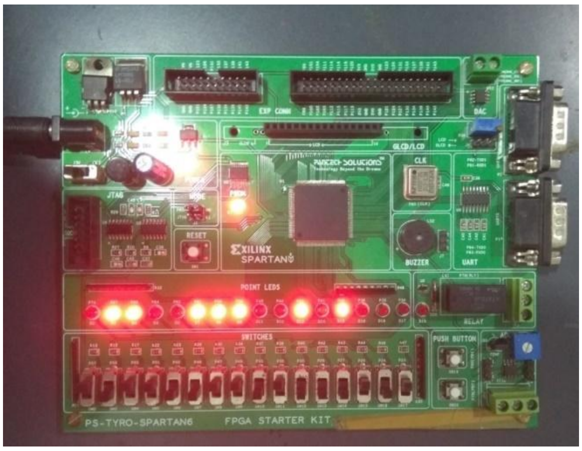

Fig 6: FPGA Implementation of LSTM structure

In order to understand the efficency of the proposed technique, accuracy, area and time consumed are calculated and are compared with the conventional Radial Basis Feed Forward (RBFN). The readings are lisetd in Table 2. From the Table 2, the proposed architecture outperforms

Published by: The Mattingley Publishing Co., Inc.

January - February 2020
ISSN: 0193 - 4120 Page No. 15816 - 15819

the conventional network in terms of accuracy, area and computational time.

Table 2: Comparison of output analysis of RBFN-NN & LSTM-NN

SPECIFICATION	ACCURACY (RECOGNIZE OF SIGNALS)	AREA ANALYSIS	TIMING REPORT
RBFN-NN	75 (OUTPUTS)	353 (LUT'S)	36.965 (ns)
LSTM-NN	4974 (OUTPUTS)	87 (LUT'S)	46.654 (ns)

V CONCLUSION

This paper implicates a digital hardware implementation of an RBFNN with an SGD-based learning mechanism and brings a new LSTM method of recurrent neural network which is an application of speech recognition which exhibits the results of fast computational characteristic and excellent accuracy on EDA simulation environment. Both the designs are implemented in the Xilinx ISE tool and verified the outputs, thus showing that LSTM is efficient in terms of RBF neural network. From the result analysis, we found that the accuracy of signal reorganization is improved in LSTM than RBFN, and also the area is reduced in LSTM, and timing is more in LSTM due to the timing required to observe the signal from the input signal. These neural networks are used in speech reorganization, face reorganization and also image processing techniques.

REFERENCES

[1] Albin, A. Jose, Nandhitha, N. M.; Roslin, S. Emalda, ART Network Based Text Independent Speaker Recognition System for Dynamically Growing Speech Database, Advances in Intelligent Systems and Computing, Volume 327 Page 473-480, 2015

[2] Vinothkumar.C and Paul Joseph Parakkal.B, “Multimodal biometric authentication process for high secured border control”, Research Journal of Pharmacy and Technology, 8(9), 2015, PP 1264-1268.

[3] Sak, H., Senior, A., & Beaufays, F. (2014). Long Short-Term Memory Recurrent Neural Network architectures for large scale acoustic modeling. In *Proc. interspeech*.

[4] yıldırım, Özal. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. Computers in Biology and Medicine. 96. 10.1016/j.compbiomed.2018.03.016.

[5] Yun-Long Kong ID, Qingqing Huang, Chengyi Wang, Jingbo Chen, Jiansheng Chen and Dongxu He, Long Short-Term Memory Neural Networks for Online Disturbance Detection in Satellite Image Time Series, Remote Sens. 2018, 10, 452; doi:10.3390/rs10030452

[6] Salma Alhagry, Aly Aly Fahmy, Reda A. El-Khoribi, Emotion Recognition based on EEG using LSTM Recurrent Neural Network, International Journal of Advanced Computer Science and Applications, Vol. 8, No. 10, 2017, pp. 355-358.

[7] Kavi Camilin, M., Susan Sangeetha, K., Rajasekar, B.FPGA implementation of ASK, BPSK and QPSK modulator using hardware co-simulation, Research Journal of Pharmaceutical, Biological and Chemical Sciences, 2016.

[8] Mathan N, Vadivel M and, Jayashri S, "Performance metrics on ultra-low-power polyphase decimation filter using carbon nanotube Field-effect transistor technology ", International Journal of Computer-Aided Engineering and Technology, 2018 Vol.10 No.3, pp.209 – 217.

[9] Rahman, A., Kuriakose, E., Rajasekar, B., Low power CMOS design technique for power switches gating. Research Journal of Pharmaceutical, Biological, and Chemical Sciences, 2016.

Published by: The Mattingley Publishing Co., Inc.