

Conceptual Design Study of Neural Network Dataflow Machine

Hideki KAWAGUCHI and Chenxu WANG

*Dep. of Information and Electronic Eng., Muroran Institute of Technology,
Muroran, 050-8585, Japan*

Abstract

To aim to achieve portable, low cost and low power consumption high-performance computing (HPC) in practical use of neural network (NN), this paper presents a conceptual design study of dedicated computer for the NN processing. Then, dataflow architecture is adopted to develop the dedicated computer to be extremely high-performance by highly parallel processing. The topology of the NN and signal flow in the forward process are suitable for hardware acceleration, and it seems to readily construct the hardware circuit of the NN. In particular, dataflow architecture hardware circuits for both of the forward process and backward propagation process are considered in this work.

1 Introduction

The neural network (NN) technology is now widely used not only in advanced AI technologies but also in daily life products. It is known that the inference process can be carried out quickly after the training process is completed, on the other hand, heavy computation is needed for the training process using huge amount of training data. Therefore high-performance computing (HPC) for the training process of the NN is strongly required. Then, portable, low cost and low power consumption HPC for the NN processing is essential from viewpoints of mobile applications such as independent control robots or drones. We here discuss a method of dedicated computer [1], [2] for the portable HPC of the NN processing employing reconfigurable computing technology. Then dataflow architecture [3] is adopted for the NN machine to achieve extremely high-performance computation by highly parallel processing.

2 Hardware Configuration of NN Dedicated Computer

An example of data signal flow of typical NN scheme is depicted in Fig.1. It is assumed that the NN consists of input (4 nodes), output (3 nodes) and two hidden layers (4 nodes). In the forward process, values of signal $y_i^{(n)}$ at the i -th-node of the n -th layer are calculated as follows,

$$y_i^{(n)} = \text{ReLU}(s), \quad s = \sum_{j=0}^N w_{ji} y_j^{(n-1)} \quad (1)$$

where Rectified Linear Unit (ReLU) function is adopted as an activation function. On the other hand, in the backward propagation process, the connection weights $w_{ji}^{(n)}$ are updated as follows,

$$w_{ji}^{(n)} = w_{ji}^{(n)} - \varepsilon y_j^{(n-1)} z_i^{(n)} \quad (2)$$

The connection weights $w_{ji}^{(n)}$ should be stored in data registers, because $w_{ji}^{(n)}$ are referred during forward and backward processes, and are modified during the backward propagation process.

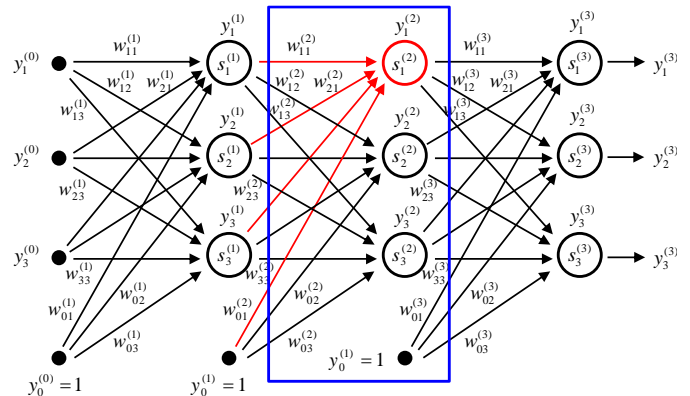


Figure 1: Typical neural network signal flow

A configuration of logic circuits related to the registers for the connection weightings $w_{ji}^{(n)}$ is depicted in Fig.2(a). Outputs of this module are calculated as $w_{ji}^{(n)} y_j^{(n-1)}$ in the forward process, and $w_{ji}^{(n)} z_i^{(n)}$ in the backward propagation process. In addition, $w_{ji}^{(n)}$ is updated according to (2) in the backward process. The one node calculation for $y_j^{(n)}$ in the n -th layer marked by red-color in Fig.1 can be executed by summing up outputs of the module of Fig.2(a) and ReLU function as shown in Fig.2(b). To combine the circuits of Fig.2(b) for all nodes and to connect with same inputs, calculations of one layer forward process can be executed in a single clock cycle by the hardware circuits of Fig.2(c). Finally, hardware circuits for the whole NN forward process can be constructed by cascade connection of the circuits of Fig.2(c) for three-layers (Fig.3). The backward propagation process is also executed in the same circuit to control the data signal flow appropriately. All of circuits are designed by using a hardware description language, VHDL, and validity of the designed circuits are confirmed by VHDL logic circuit simulations.

3 Conclusion

This paper has proposed a conceptual design of dedicated computer for the neural network to aim to achieve portable HPC for practical use of AI technologies. The VHDL design of individual module circuits and circuit simulation are carried out. The next stage is entire circuit simulation and implementation of the designed circuit in the FPGA, which will be presented in near future.

References

- [1] J.R.Marek, M.A.Mehalic and A.J.Terzuoli, Conf. Proc. 8th Annu. Rev. Progress in Applied Computational Electromagnetics, (1992), pp.546-553.
- [2] H.Kawaguchi, K.Takahara, D.Yamauchi, IEEE Trans. on Magnetics, Vol.38, Issue:2, (2002), pp.689-692
- [3] S.Matsuoka, K.Ohmi, H.Kawaguchi, IEICE Trans. Electron., Vol.E86-C, No.11 Nov. (2003), pp.2199-2206.

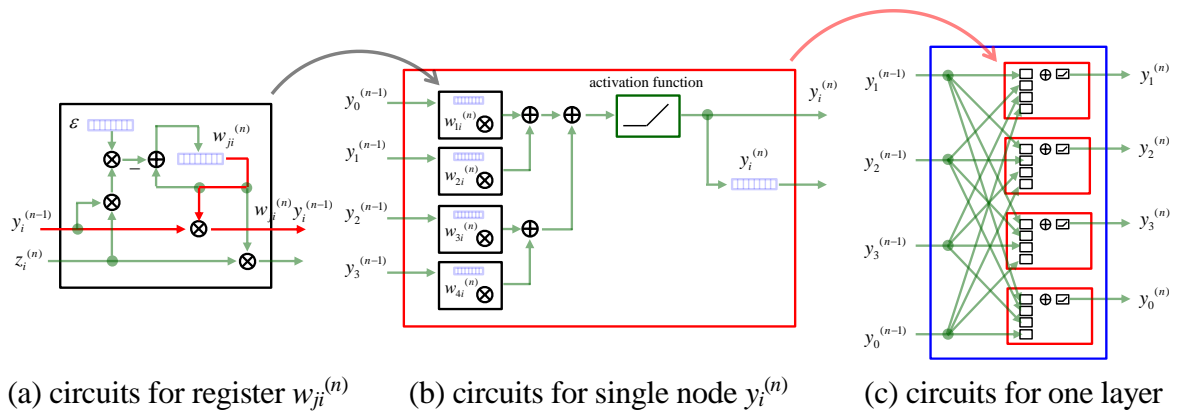


Figure 2: Hardware modules of NN forward process

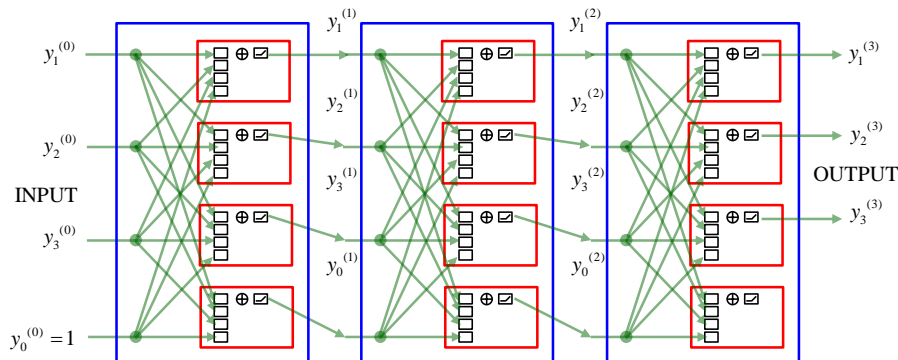


Figure 3: Hardware configuration of NN forward process