SUMMARY OF CONTRIBUTIONS OF THE DISSERTATION

PhD candidate : Le Minh Fellows code: 2027005
Major : Electronics Engineering Major code: 2020
Dissertation title : Research on optimizing and implementing the neuromorphic computing system applying memristors
Supervisor one : Assoc. Prof. Truong Ngoc Son, PhD
Supervisor two :

Summary of theoretical and academic contribution of the dissertation:

Neuromorphic computing systems aim to implement artificial neural networks and deep learning models directly on integrated circuits, simulating the information-processing behaviors of the biological brain, instead of relying on traditional software architectures. Neuromorphic computing systems have been successfully implemented on VLSI platforms based on CMOS technology. However, CMOS technology is currently approaching its size limit. Researching new circuit elements and new technologies for neuromorphic computing systems has attracted lots of interest. Memristors have the characteristics of changing the resistance when enough current passes through and the ability to remember the resistance when the power is off that are similar to those of the synapse in biological brain modeling. Therefore, memristors are widely applied to designing neuromorphic computing systems, replacing traditional CMOS technology. In addition to the ability to program and store, memristors also enable the execution of key calculations in artificial neural networks (multiplication, addition) within the memristors at the circuit level (in-memory computing), allowing the acceleration of processing in artificial neural networks and replacing the design of computing-circuits on VLSI. The implementation of weight sets and in-memory computations often utilizes two complementary memristor crossbar arrays. This is a common architecture and has been proved to be efficient in implementing neuromorphic computing systems, enabling the storage and computation of signed weights or the execution of similarity-scoring functions based on the XNOR operation. The twin crossbar array architecture was then proposed to save power consumption when data are binary images with the low density of “1” bit. In order to decrease the number of memristors to reduce factors such as power consumption and process-variation of memristors, the single crossbar array architecture was proposed based on the minimization of the XNOR function and was considered as a potential architecture for neuromorphic computing system designs.

Memristor array architectures have been demonstrated to be efficient in designing neuromorphic computing systems. However, these architectures were evaluated under different conditions and constraints. To choose an optimal architecture, an overall analysis and evaluation under the same number of conditions is required. The first contribution of the dissertation is to provide the general analysis and evaluation results of the three memristor crossbar array architectures applied in implementing neuromorphic computing systems, to determine the potential architecture that satisfies many design aspects.

Based on the analysis and evaluations, the dissertation proves that the single memristor crossbar architecture demonstrates more advantages than the other two architectures. However, it has limitations for binary data with low density. The dissertation proposes a solution to optimize the single crossbar array architecture by analyzing the mathematical model of the XNOR function. The designs and simulations at circuit level indicate that the optimized memristor crossbar architecture addresses the impact of data density on the performance of the memristor array, while retaining outstanding advantages such as the saving of power and number of memristors within the array, the tolerance of noise and memristance variation. This is the second contribution of the dissertation that is significant in applying memristor crossbar array architecture in implementing neuromorphic computing systems.

The optimized memristor array architecture is then applied to designing an XNOR neural network (XNOR-Net) with multi-layers. The XNOR neural network is a model of binary neural network that is efficiently deployed on resource-limited devices. In the design, the dissertation applies the optimized memristor crossbar array architecture to hidden layers. On the output layer, by analyzing, the dissertation proves that it is efficient to apply the simplified single memristor crossbar array architecture to minimize the peripheral circuit, save power and area without affecting the recognition error of the circuit. This is also a contribution of the dissertation to applying the memristor array architecture for designing deep-learning neural networks on integrated-circuit platforms.

HCMC,   /   /

PhD candidate

(Sign and name)

Le Minh