WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Neuromorphic Computing Hardware: A Review

Shivu Chauhan

Mechanical Engineering Department, Thapar Institute of Engineering and Technology, Punjab, India
Corresponding Author Email : schauhan1_be18@thapar.edu

ABSTRACT

CNNs have now established themselves as the de-facto standard for distinguishing between visual features in an image while capturing its spatial and temporal characteristics. Computational capabilities were a limiting factor in model performance in computer vision tasks a decade ago; however, in the foreseeable future and the current circumstances, the challenge is to develop environmentally friendly, power-efficient algorithms. Developing nature-inspired computer algorithms is one such technique. This is the main reason why, although having lower performance than CNNs, spiking neural networks (SNNs) have attracted a lot of interest. SNNs are closer to the present neural networks in that they mirror the mammalian visual cortex. Unlike multi-layer perceptrons, they only fire when the specified membrane potential is breached. The human brain, which has been optimizing itself for a million years, consumes the same amount of power as a typical laptop. Researchers have previously used spike-based computation to develop neuromorphic energy efficient microchips. Using neuromorphic computer processors instead of traditional GPUs might be more environmentally friendly and efficient. Furthermore, developing on-device AI is a current issue. SNNs might be just an ideal solution for this problem. From a classification standpoint, this paper presents an in-depth attention of deep learning-based simulation studies of SNNs.

Keywords: Neumann Computing Architecture, Hardware

Hardware

The conventional von Neumann computing architecture that separates the processing and the memory unit cannot handle the complexity of deep learning models and their various paradigms. Neuromorphic computing hardware consists of neurons and synapses to storing data and computation with a neural network to communicate with them efficiently. CMOS-based chips developed by IBM and Intel (see Appendix) for SNN computations are limited to large-scale modular computing units with increasing neurons and synapses. This study reiterates that significant advances are required to increase the number of onboard neurons and synapses to mimic the complex operations of the human brain like cognitive processing, sophisticated motor control, learning, and abstraction. Moreover, CMOS-based systems are limited by their room-scale size. Thus, we need to look into Memristor-based alternatives due to their synaptic-like behaviour.

The Memristor, i.e., 'Resistor + memory,' is a building block of synaptic devices that requires high integration density, low latency, low power, and non-volatile memory capable of mimicking the brain's learning and forgetting functions. These capabilities are imperative for the implementation of synaptic learning used in neuromorphic computing. We find Memristors as promising candidates for synaptic learning applications. Some of the Hybrid CMOS-Memristor devices are investigated below:

1. **Phase Change RAM (P-Ram):** It is a nonvolatile ram that utilizes the difference in resistivity levels between the amorphous phase (high resistivity) and the crystalline phase (low resistivity) between its two electrodes. On the application of voltage or current to the electrodes, the phase of the material is changed. To 'Set' is the phase change from amorphous to crystalline phase through successive voltage or current pulses. While 'Reset' is the phase change from 'Crystalline' to 'Amorphous' achieved by applying a large current pulse for a short time. PRAM has many desirable characteristics like high speed, low energy consumption suitable for implementing artificial synapses in machine learning algorithms. However, its inherent resistance drift phenomenon in the amorphous phase destroys its stability and causes power consumption problems that are not desirable in the long run. PRam needs to resolve its reliability issues.

2. **Resistive RAM (ReRam):** It is a non-volatile RAM in which the resistance changes according to the applied voltage [1]. It uses an insulator that produces a metal defect which forms a conductivity path hence a 'Set' operation. This conductivity path or filament is removed by another voltage hence a 'RESET' operation. ReRam offers significant advantages by reducing the physical space to express multiple resistances as resistance is developed by creating oxygen vacancies through metal ions. However, the conductivity path developed in ReRAMs introduces variations in a magnitude spread over the path and causes reliability issues in implementing long-term memory effects.

3. **Atomic Switch Network (ASN):** It is a network-based synaptic device that exhibits short-term and long-term memory. This device is typically composed of a network of interfacial atomic switches, which are self-assembled and randomly connected. Ag2S metal-insulator-metal (MIM) interface is widely used to form the switches [1]. On the application of input voltage, the conductance in the device is changed and is sustained for some time called (STP) short term plasticity, mimicking human neurons, however, the duration of conductance is random and most of the time uncontrollable [1] hence ASN are potential candidates for synaptic devices if they overcome these limitations.
   New synaptic memory devices exhibit issues that make older design principles obsolete.The resistance variation of ReRAM and PRAM introduces errors over time hence a trade-off between energy consumption and induced error. We propose that a high error rate may be suitable in some hardware systems as long as the system's core functionality is met, in our case, lesser energy consumption to protect the environment.

We believe hybrid systems of memristor layers embedded on CMOS substrate is the solution to neuromorphic hardware engineering[2]. They will enable computing and learning processes by combining memristors on spiking processors to fire neurons in silicon chips after attaining a specific threshold value[2]. PCB design techniques [3] can assemble chips with memristors, allowing to substantially scale up the number of neurons and synapses in a neural system. However, there would be limitations relating to the density of memristor layers and their onboard programmability.

Conclusion and Future Research Directions

In this paper we discussed the current gaps between available RAM (Hybrid CMOS-Memristor Architecture) options and the ideal hardware required for Neuromorphic engineering. We also proposed hybrid Memristor-CMOS based PCB designs which can substantially increase the number of neurons and synapses hence mimicking the brain more efficiently.

References

[1] Park, J. Neuromorphic Computing Using Emerging Synaptic Devices: A Retrospective Summary and an Outlook. *Electronics* 2020, *9*, 1414. https://doi.org/10.3390/electronics9091414

[2] Camuñas-Mesa, L.A.; Linares-Barranco, B.; Serrano-Gotarredona, T. Neuromorphic Spiking Neural Networks and Their Memristor-CMOS Hardware Implementations. *Materials* 2019, *12*, 2745. https://doi.org/10.3390/ma12172745

[3] Zamarreno-Ramos C, Linares-BarrancoA, Serrano-Gotarredona T, Linares-Barranco B. Multicasting mesh AER: a scalable assembly approach for reconfigurable neuromorphic structured AER systems. Application to ConvNets. IEEE Trans Biomed Circuits Syst. 2013 Feb;7(1):82-102. doi: 10.1109/TBCAS.2012.2195725. PMID: 23853282.

Appendix

A. CMOS based SNN Specific Neuromorphic chips

1. **IBM's TrueNorth chip** contains 16M neurons(onboard) and 4B synapses (onboard) [2] but does not offer on-chip learning like the human brain. TrueNorth consumes around 25pJ of energy per connection compared to 10fJ of the human brain [2].
2. **Intel's Loihi chip** contains 131k on-chip neurons and 126M on-chip synapses. It offers on-chip learning and consumes around 81pJ of energy per connection.
3. ETHZ-INI developed the **ROLLS chip** which included 256 neurons and 128 k on-line learning synapses. It has been updated to the Dynamic Neuromorphic Asynchronous Processor (DYNAPs) with 1 K neurons and 64 k on-line learning synapses [2].
4. **Stanford's NeuroGrid** uses subthreshold analogue neural circuits however it has been updated with the BrainDrop chip prototype which is a single-core chip planned to be part of 1M neurons Brain Storm system.

B. Comparative Analysis of various CMOS based SNN Chips

Platform	***Brain***	***TrueNorth***	***Loihi***	***Rolls***	***NeuroGrid***
Technology	Natural	Digital	Digital	Mixed-Signal sub threshold	Analog sub threshold
Transistors	-	5.4 B	2.07B	12.2M	23M
neurons(on chip)	-	1M	131K	256	65K
synapses(on chip)	-	256M	126M	128K	100M
neurons(board)	10^{11}	16M	-	-	1M
synapses(board)	10^{15}	4B	-	-	4B
Energy consumed	10fJ	25pJ	81pJ	>77fJ	100pJ
On-chip learning	Yes	No	Yes	Yes	No