iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35601

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI June 2021- Available at www.ijraset.com

Analysis of Neuromorphic Computing Systems and its Applications in Machine Learning

Pranava Bhat[1], Sumukha K R[2], Nagaraj G Cholli[3]
Department of Information Science and Engineering, R.V. College of Engineering, Bengaluru, India

Abstract: ***The domain of engineering has always taken inspiration from the biological world. Understanding the functionalities of the human brain is one of the key areas of interest over time and has caused many advancements in the field of computing systems. The computational capability per unit power per unit volume of the human brain exceeds the current best supercomputers. Mimicking the physics of computations used by the nervous system and the brain can bring a paradigm shift to the computing systems. The concept of bridging computing and neural systems can be termed as neuromorphic computing and it is bringing revolutionary changes in the computing hardware.***

Neuromorphic computing systems have seen swift progress in the past decades. Many organizations have introduced a variety of designs, implementation methodologies and prototype chips. This paper discusses the parameters that are considered in the advanced neuromorphic computing systems and the tradeoffs between them. There have been attempts made to make computer models of neurons. Advancements in the hardware implementation are fuelling the applications in the field of machine learning. This paper presents the applications of these modern computing systems in Machine Learning.

Keywords: Neuromorphic Computing, Neurons, Synapses, Spiking Neural Networks, SpiNNaker, Machine Learning

I. INTRODUCTION

In current von-Neumann architecture, the memory is separated from CPU and the data transfer during computation between them is the major cause of bottleneck, hindering performance. To overcome this, physical computers have to mimic the human brain and nervous system. This style of computing is called Neuromorphic Computing, where the physics of computation used in the computer chips is the same as that of the human nervous system. This is different from Neural Networks (software mimic) as Neuromorphic Computing is aimed at hardware implementation.

Computers are built on a fundamental architecture for their functioning. Currently, the majority of the computers follow a system where the storage(memory) is kept separate from the processing unit. This system is called the von-Neumann architecture.

This means that there is a need for transfer of data between memory and CPU and vice versa during various cycles of operation, which puts a limit on maximum achievable performance for a given set of hardware in this system and can cause an overload on the transfer buses, leading to occurance of a bottleneck situation.

This is in contrast to how a human nervous functions which is efficient and consumes a negligible amount of energy. Observing this and the slowing down of Moore's Law [1], scientists and engineers have thought and gone on to design computation systems which mimic the human brain and nervous system. They are trying to simulate and implement the biological aspect of the human nervous system, which is able to solve computationally demanding and intensive problems. It is an upcoming interdisciplinary area of research which focuses more on hardware implementation of the introduced concept.

The main area to focus for implementation in the neuromorphic system are the 'synapses', which are gaps between nerves in the human nervous system which are key in passing of messages in the human body as electric currents. Similarly, various compounds (compounds of silicon) are being tested which can electronically simulate these synapses and neurons, by observing the flow of electric current in these synapses (semiconductor junctions) and their capacity to retain current. This ability to generate a gradient of signals from one neuron to another, and retain memory in the form of ions dispossessed in the junction simultaneously is crucial in working of neuromorphic systems. When a large number of such neurons and synapses are working simultaneously, the neuromorphic system on a chip would be highly energy efficient.

Neuromorphic computing is an interdisciplinary area of research, aiming to emulate the computational principles of biological neural systems and to utilize such principles to solve complex and computationally intensive problems. The ultimate goal of this discipline is to design and construct silicon models of neurons and synapses in the human nervous system. The basic physics governing the current flow across a biological junction (synapses - close connections between two nerve cells) and a semiconductor junction in an ion channel are the same. This observation has led to the origin of the neuromorphic computing domain [2].

The aim in Neuromorphic Computing is to try to mimic the nerves in our body by using precise electric currents, which flow across a Synapse. The receiving computer neuron gets activated in a certain or performs some kind of computation based on the number and kind of ion it received.

This ability to transmit a gradient of understanding from neuron to neuron and to have them all working together simultaneously means that neuromorphic chips could eventually be more energy efficient.

II. METRICS OF COMPARISON FOR NEUROMORPHIC COMPUTING SYSTEMS

With the advancement in Artificial Intelligence and development of complex algorithms, suitable hardware that can perform heavy computations in an efficient way is needed. Areas like Quantum computing, carbon nanotubes are being explored. Neuromorphic hardware systems are one of the promising solutions in this regard. It aims to provide cognitive abilities to a computing machine. This system mimics the biological nervous systems. Below are some of the key parameters of these neuromorphic systems and these metrics are used in comparative analysis of modern hardware architectures.

Hardware metrics such as process size, count of neurons and synapses, clock speed, package dimensions, weight, memory are easy to obtain through product documentations. These parameters are specific to a neuromorphic hardware. However they do not provide sufficient information to make comparisons between hardwares with respect to their performance in a task [3].

Workloads like multivariate classification systems can be executed on these hardware for performance analysis. The Large Scale Visual Recognition Challenge (ILSVRC) benchmark is a popular data set for workloads in the ANN community. Executing specific workloads on hardware helps to analyse such as power, latency, throughput and accuracy.

A. Neurons and Synapses

The human brain has billions of neurons and trillions of synapses. Neurons produce neural signal impulse and synapses allow a neuron to deliver a signal to another neuron. In neuromorphic hardwares, a core that processes the data and the structure that connects these cores in parallel for the transmission of information can be considered as neurons and synapses respectively. The number of neurons and synapses in each core of a neuromorphic processor is a point of consideration while developing the hardware architectures that are aiming to mimic the human brain [4].

B. Power

Energy consumptions of neuromorphic systems are projected to be extremely low, comparable to the human nervous system. Parameters like number of used cores, number of active synapses and spike rate affect the power usage of an architecture. Different hardware configurations and diagnostic utility components such as nvidia-smi can be employed during classifier test and training to get power measurements.

C. Latency

The delay between occurrence of a stimuli and the nervous cell responding to it is called spike latency. Information transfer in neuromorphic systems in the form of spikes involves a delay. This is referred to as latency. Some AI tasks which have hard deadlines must get rapid response from the system. Low latency architectures can be deployed at edge devices that can process continuous streams of data [5].

D. Throughput

Performance of neuromorphic systems is commonly measured in terms of throughput (Giga Operations Per Second - GOPS). Approaches such as in-memory computing (IMC) that reduces the data movement between memory and processor significantly improves the throughput.

E. Accuracy

The accuracy of results obtained from neuromorphic computing systems are of utmost importance during comparative analysis between them. The resolution of synaptic weights in the architecture has an impact on the system accuracy. The test data and the methodology must remain same for all the systems under analysis [6]. These metrics must be well defined for a neuromorphic computing system. The Key Performance Indicators (KPIs) for an architecture depends on the problem it is aiming to solve. In certain cases, there is a tradeoff between the parameters. Effort for producing high throughput often results in high energy consumption and low accuracy. Also care must be taken to maintain the same test environment for different systems.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI June 2021- Available at www.ijraset.com

III. COMPARATIVE ANALYSIS OF EXISTING NEUROMORPHIC COMPUTING SYSTEMS

Current neuromorphic hardware, still in primitive developmental stages, is constructed as a grid of neurons and the interconnect junctions serve as the synapses. The following section provides a description and comparative analysis of the existing neuromorphic computing systems.

A. True North

TrueNorth On-chip neuromorphic computing system has developed by SyNAPSE, having the following features:
No. of Neurons: ~ 1million
No. of Synapses: ~256 million
Design: 4096 neurosynaptic cores, with each core containing 256 neurons and 256 × 256 synapses using SRAM-based crossbar array. It is the largest single chip built within the neuromorphic community.
Power Consumed: ~72mW. This is significantly lower compared to other systems [7].

B. SpiNNaker

Developed in 2018 by University of Manchester based on custom designed application-specific integrated circuit (ASIC), the SpiNNaker chip is more flexible and is usable for generic applications. Some features of the SpiNNaker chip are:
No. of Neurons: ~ 850,000
Design: 47 ARM core chips, based on 130nm architecture.
Power Consumed: 75W
A custom designed ASIC called SpiNNaker chip which contains 18 ARM processor nodes is the core hardware element. 47 of these chips make a complete SpiNNaker board. The ultimate product is an assembly of 1200 such boards [8].

C. Neurogrid

Neurogrid is also a custom ASIC which uses Very large-scale integration (VLSI) based on analog circuits to mimic neurons and synapses. The neurons are designed to implement a quadratic integrate-and-fire (QIF) model with the help of synapses. It was built by Stanford University in 2014. It has the following features:
No. of Neurons: ~ 1million
Design: Each chip is a 256x256 mesh of neurons, based on 180nm silicon technology.
Power Consumed: 3W
A neurocore chip which uses analog VLSI and digital asynchronous VLSI to implement neurons and spike-based communications respectively is the core hardware element of Neurogrid. The routing architecture uses a tree topology and each neuron is implemented using a quadratic integrate-and-fire (QIF) model [9].

D. Intel Loihi

The name of the neuromorphic chip designed by Intel in 2017 is Loihi. This chip has been designed to be more provisional towards implementation of Spiking Neural Networks (SNNs), and has the capability to retain memory, and thus is able to learn and train itself over a period of usage, through qualitative inputs. Hence it is more suitable for indeterminate environments where computational requirements demand autonomous operation and constant learning. It has the following features:
No. of Neurons: ~ 130,000
Design: 128 neurosynaptic cores, assembled on a 14nm chip. These chips can be programmed and triggered for better efficiency, based on requirement.
Power Consumed: ~72mW [10].

IV. APPLICATIONS IN MACHINE LEARNING

The deep neural networks are usually fully connected, i.e each unit in a network layer receives input from each unit in the previous layer and sends signals to each unit in the subsequent layer. These deep learning based neural networks have achieved groundbreaking results in many real world problems but their functioning is not similar to the human neural system as their name suggests. The biologically inspired mathematical model of spiking neurons gave rise to Spiking Neural Networks (SNNs) which emulate natural neural networks that exist in biological brains. SNNs implemented on neuromorphic hardwares offer many benefits such as low power consumption, high throughput and faster information processing.

This is achieved by communicating and processing information in a parallel fashion. SNNs are ideal for processing spatio-temporal data and perform event-driven information processing [11][12][13].

Existing neuromorphic computing systems require specialized knowledge for operating them and they are currently under use only in advanced research groups. In the medium term, power efficient and small scale neuromorphic systems can be used to implement cognitive functionalities in smartphones. Neuromorphic systems may enable the integration of cognitive functions into a wide range of intelligent consumer products.

The neuromorphic technology is still young. Initiatives such as the Human Brain Project may provide new insights on the understanding of information processing inside the human brain. Some leading hardware manufacturing organizations have aimed to explore AI on neuromorphic systems in commercial and defence sectors. Areas such as self-driving cars, computational physics simulations and problems such as identifying an individual's face out of random images are being explored by neuromorphic systems. In today's smartphones, data processing and computation for deep learning applications such as speech recognition occurs in the servers on cloud. Running complex and huge machine learning models on small devices by making all the computations locally can be made possible by having neuromorphic computing systems on devices.

V. CONCLUSIONS

Non-traditional computing fields are explored actively in recent times and Neuromorphic computing is an exciting hardware model in this domain. It has many benefits over typical von Neumann systems in terms of power efficiency, throughput and accuracy. The way of computing in such systems is similar to the information processing of the human brain. The paper discusses the metrics of importance in neuromorphic systems. The intrinsic and extrinsic metrics can be measured for comparative analyses between different architectures.

Neuromorphic Computing can become a game-changing model in the field of Machine Learning, with exponential powers of computation as compared to levels achieved by current computing architectures.The programming aspect of the new type of hardware in neuromorphic computing remains a challenge. Nevertheless, there is huge potential to be extracted from this form of computation, with a huge scope of related areas of research, which are to be explored.There are many advanced systems in this domain and they offer a lot of potential applications in the field of Machine Learning. Neuromorphic computing systems are an emerging technology and the challenge is to bring them onboard for everyday engineering challenges. If the neuromorphics move from research labs to commercial applications, it can be revolutionary.

VI. ACKNOWLEDGMENT

We thank R.V. College of Engineering, Bengaluru and its faculty for the support and guidance in writing this paper.

REFERENCES

[1] Reynolds, John J., "Reservoir Computing in an Evolutionary Neuromorphic Framework. " PhD diss., University of Tennessee, 2019. https://trace.tennessee.edu/utk_graddiss/5785

[2] Moradi, S., & Manohar, R. (2018). The impact of on-chip communication on memory technologies for neuromorphic systems. Journal of Physics D: Applied Physics, 52(1), 014003. doi:10.1088/1361-6463/aae641

[3] Vineyard, Craig, Severa, William, Koc, Cetin, Green, Sam, and Green, Sam. Benchmarking Event-Driven Neuromorphic Architectures.. United States: N. p., 2019. Web. doi:10.1145/3354265.3354278.

[4] Calimera, Andrea et al. "The Human Brain Project and neuromorphic computing." Functional neurology vol. 28,3 (2013): 191-6. doi:10.11138/FNeur/2013.28.3.191

[5] Marti, Daniel, et al. "Energy-efficient neuromorphic classifiers." Neural computation 28.10 (2016): 2011-2044.

[6] Benchmarking neuromorphic vision: Lessons learnt from computer vision Tan C., Lallee S., Orchard G. (2015) Frontiers in Neuroscience, 9 (OCT) , art. no. 374

[7] Merolla, Paul A et al. "Artificial brains. A million spiking-neuron integrated circuit with a scalable communication network and interface." Science (New York, N.Y.) vol. 345,6197 (2014): 668-73. doi:10.1126/science.1254642

[8] Furber S, Galluppi F, Temple S and Plana L 2014 The spiNNaker project Proc. IEEE 102 652–65

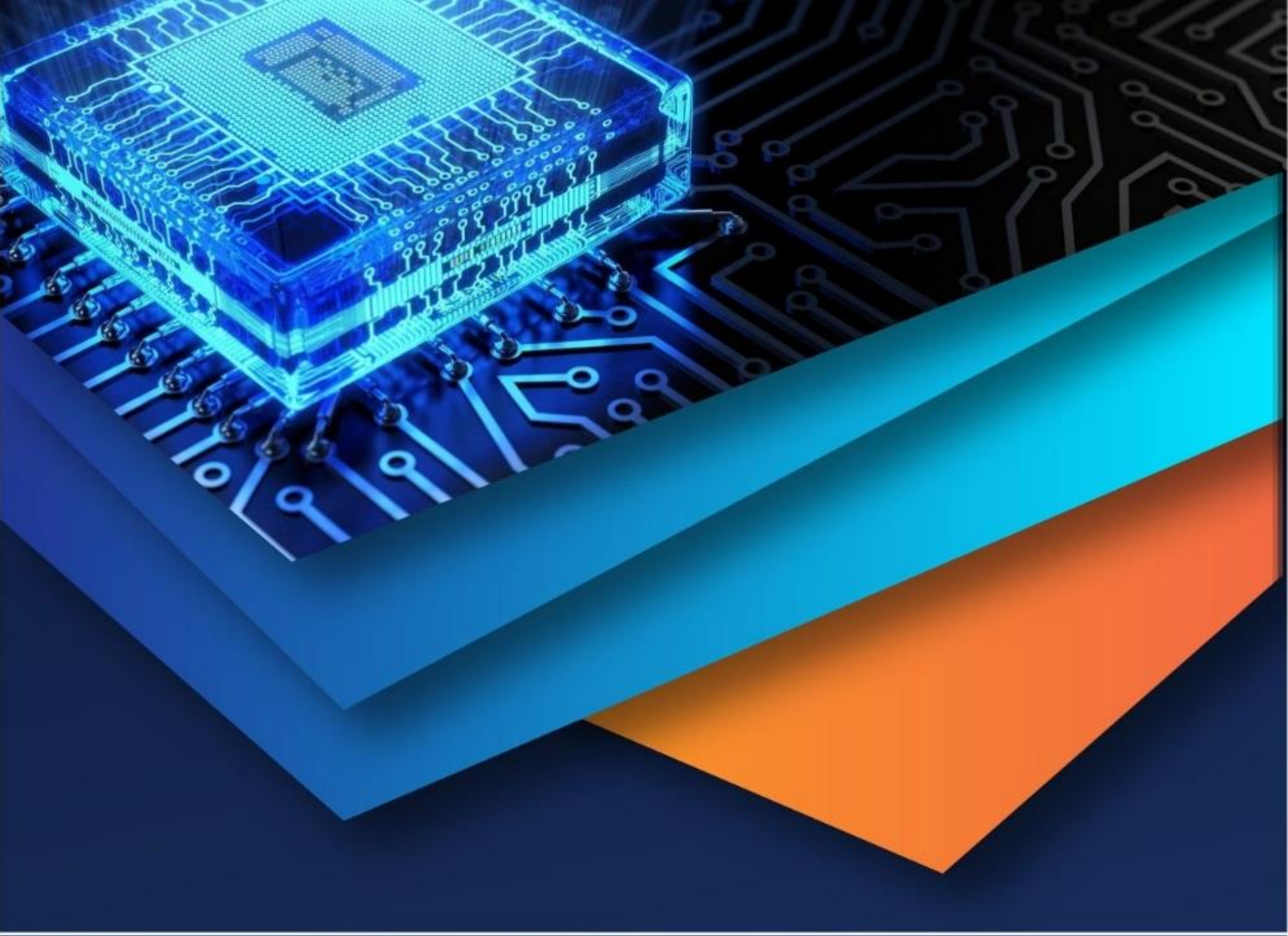
[9] Benjamin B V, Gao P, McQuinn E, Choudhary S, Chandrasekaran A R, Bussat J, Alvarez-Icaza R, Arthur J, Merolla P and Boahen K 2014 Neurogrid: a mixed-analog digital multichip system for large-scale neural simulations Proc. IEEE 102 699–716

[10] M. Davies et al., "Loihi: A Neuromorphic Manycore Processor with On-Chip Learning," in IEEE Micro, vol. 38, no. 1, pp. 82-99, January/February 2018, doi: 10.1109/MM.2018.112130359.

[11] Neuromorphic Computing, Beyond Today's AI, https://www.intel.in/content/www/in/en/research/neuromorphic-computing.html

[12] Ghosh-Dastidar, S. and Adeli, H. (2009). Spiking neural networks. International journal of neural systems, 19(04):295–308

[13] Maass, W. (1997). Networks of spiking neurons: the third generation of neural network models. Neural networks, 10(9):1659–1671



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY