Biological Neural Networks as the Forefront of AI Processing

Written by Edward Lin

Abstract

Graphic processing units (GPUs) are a major component of artificial intelligence (AI) processing power. As AI becomes more sophisticated and more processing power is needed to run these intellectual models, a growing concern of energy demand and extensive AI training becomes an increasing concern. To create more sophisticated machine learning algorithms, scientists in the field of neuromorphic computing studied the brain for its ability to efficiently process and store information. Finding a way to incorporate the brain directly into computing may create novel algorithms to meet the increasing demands of information processing.

Introduction

Graphic processing units (GPUs) are chips within a computer that process data simultaneously, performing parallel computations. Making approximately 36,500 calculations per second, GPUs consist of three major components that use basic arithmetic operations for neural network processing. The GPU's tensor core is the forefront of critical AI operations and AI learning, which uses geometric transformations and large matrix computations for AI neural network optimization. As science further progresses, we obtain more data about the world, which requires more processing power. This exponential growth in data requires increasingly more powerful processing capabilities to interpret this data. However, the current processing power of modern-day computational systems are inefficient and expensive, as they have a high energy and resource demand for running the GPUs and regulating their temperature as shown in Figure 1. Since the late 1980s, scientists and engineers have been studying the structural organization within the brain to help them optimize information processing, giving rise to the field of neuromorphic computing. The processing speed of the brain is similar to that of a supercomputer, and outperforms a supercomputer in terms of energy efficiency, spatial optimization, memory, and storage. Rather than mimicking and studying how brain structures optimize information processing, it may be effective to explore the direct integration of these systems, also known as "reverse

	Frontier supercomputer (June 2020)	Human brain
Speed	1.102 exaFLOPS	~1 exaFLOPS (estimate)
Power requirements	21 MW	10–20 W
Dimensions	680 m^2 (7,300 sq ft)	1.3–1.4 kg (2.9–3.1 lb)
Cost	$600 million	Not applicable
Cabling	145 km (90 miles)	850,000 km (528,000 miles) of axons and dendrites
Memory	75 TB/s read; 35 TB/s write; 15 billion IOPS flash storage system, along with the 700 PB Orion site-wide Lustre file system	2.5 PB (petabyte)
Storage	58 billion transistors	125 trillion synapses, which can store 4.7 bits of information each

The Hewlett Packard Enterprise Frontier, or OLCF-5, is the world's first exascale supercomputer, hosted at the Oak Ridge Leadership Computing Facility (OLCF) in Tennessee. It is compared here with the human brain. For sources see (6–11).

Figure 1. Comparison Between Supercomputer and Human Brain

neuromorphic computing". Implementing the biological neuronal networks within the brain directly into a computer's processing units can maximize efficiency, leading to greater information processing and more sophisticated machine learning algorithms (Kagan et. al, 2023). As the brain receives new information, it reorganizes itself and forms new connections. The incorporation of self-restructuring capabilities to AI opens more dynamic approaches to AI training. In addition, the incorporation of neuroplasticity functions in GPUs may allow neural networks to adjust how they process inputs without rigorous retraining. Such approaches give engineers the freedom to develop more powerful algorithms with deeper, more advanced neural networks. The implementation of neuroplasticity would enable AI models to process complex, continuous flows of data more effectively, allowing the energy-efficient and adaptable data processing of the brain to be manifested in GPUs.

> **The implementation of neuroplasticity would enable AI models to process complex, continuous flows of data more effectively, allowing the energy-efficient and adaptable data processing of the brain to be manifested in GPUs.**

Biological Neural Networks (BNNs)

Biological Neural Networks BNNs, or biological neural networks, are clusters of neurons connected to each other through axons and dendrites (in a sense, they can be described as miniature brains with the most basic complexity). Through these axonal and dendritic connections, BNNs are able to demonstrate characteristics that mirror those of artificial intelligence: computational performance and network plasticity (the ability of neurons to arrange themselves based on the stimulation received). The vast quantity of connections within a BNN enables the biological structure to undergo parallel processing across multiple neuronal signaling pathways and allows stimuli to be distributed vastly. When integrated within computing systems, BNNs have the ability to not only process the information and exhibit a response, but also rearrange themselves according to the stimulus, displaying plasticity and small amounts of memory (Dranias et. al, 2014). Memory consolidation in BNNs can be categorized into two types of memory processes, fading memory and hidden memory. Fading memory relies on the firing activity in response to a stimulus, lasting only for a brief moment, while hidden memory depends on synaptic plasticity to strengthen the connections between neurons. Hidden memory allows the BNN to retain information for prolonged periods. However, the retention of this information is disrupted or restructured when the BNN receives a high-intensity electrical input (stimuli).

Multielectrode Arrays

MEAs, or multielectrode arrays, help researchers integrate neurons within computational devices. Structurally, they are a flat surface with multiple microelectrodes embedded in an array placed underneath a BNN that is in a petri dish. Each electrode independently records extracellular activity. The recorded signals are then digitized and processed by a computer to interpret neural activity. The computer then generates an electrical stimulus that is sent as a response to the neurons to evoke another response, generating a real-time feedback loop.

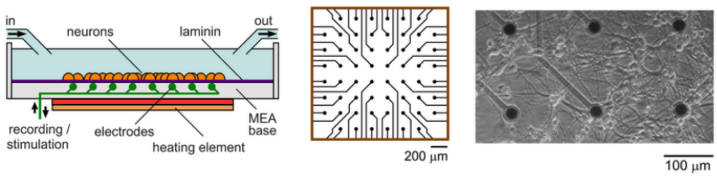


Figure 2. Multielectrode array structure (left to right) schematic, 60 electrode array, cultured neurons

Ping-Pong Experiment

Through an experiment revolving around ping-pong, researchers connected a BNN to an MEA that provided input of the game's environment, allowing the BNN to control a ping-pong paddle in real-time. The neurons received structured feedback through successful (positive reinforcement) and missed hits (negative reinforcement), allowing the BNN to self-correct from an optimal to a more accurate dynamic state. Through this, the BNN could retain task-specific information for brief intervals, displaying a positive correlation between branching ratio (neuroplasticity) and task performance. This enhanced BNN performance for that one specific task. When researchers no longer provided feedback to the BNN, its "hit-to-miss ratio" decreased (Habibollahi and Kagan, 2023). Once connected to a computing system through a multielectrode array, BNNs rearranged themselves until they reached an optimal dynamic state to efficiently process the structured stimuli. Despite reaching an optimal state, BNNs displayed low accuracy but a high level of computing, thus requiring feedback. Feedback was only in regards to one specific task, allowing BNNs to process information efficiently and with high accuracy for one task, making BNNs optimal for computing one task over and over again. However, if required to switch tasks, BNNs needed to rearrange themselves and be given the correct feedback to perform a task swiftly (Habibollahi and Kagan, 2023).

Neuronal Capabilities of Matrix Processing

The ping-pong experiment's success with parallel processing and real-time adaptation points to the potential of larger neural systems. In the brain, neural processing in the cerebral cortex occurs at intervals of a few milliseconds. Despite being slower than a computer, the cortex is able to compensate for this difference through its ability to process vast amounts of information in parallel (Ballard, 1986). Resembling GPUs, this parallel processing in the brain is how the brain receives sensory information and executes motor skills (Sigman and Dehaene, 2008). In addition, hierarchical (information flowing through successive layers of processing) and modular organization (each cortex is divided into specialized regions for particular tasks) of the brain is essential for matrix computations. Present artificial intelligence neural networks follow a similar information flow, but creating a processing system consisting of specialized BNNs can structure information for maximal efficiency and scalability. This hypothetical computational system, much like the brain, will have areas of the computer's processing unit distinguished by function. Each area will have their own specialization of information processing. As information is received, the information will be decomposed and sent to the right processing subregions, taking into advantage a BNN's fading memory. Information processing can then be split into smaller, manageable bits of information that each subregion of BNNs can process, allowing for scalability and efficiency.

Ethics

Despite using only BNNs, which in comparison to a human brain is minuscule in proportion, it is possible to create a sentient life force when such an organization is scaled to an extent. Eventually, if such biotechnological methods are implemented, one will create structures that contain a neuron count that mirrors certain large mammalian organisms. In the ping-pong game, researchers were able to notice a certain degree of self awareness within the BNN, despite being comparatively small to that of a simple organism's (Kagan et. al, 2022). However, if neurons were integrated into technology, will such a system eventually have a form of consciousness? Due to exponential growth of technology, the size of BNN integrated systems will also have to increase to address this demand. But once a certain number of neurons are integrated and interacting, it can eventually have its own life force. It is essential to bring up such topics when discussing the potential future direction advancement of biocomputational devices to prevent the unethical exploitation of sentient beings as inanimate objects. Furthermore, ethical considerations regarding the creation of such conscious systems must be explored, drawing a line as to when a system has developed consciousness.

Conclusion

Although the idea of integrating neurons into the processing system of computers seems efficient and beneficial to the further advancement of AI, it is still fairly new and many caveats have yet to be considered. For example, how would BNNs be integrated at a scale into computers/super computers that will provide enough processing power to match that of the computer? How much energy and resources would it take to maintain the functionality of a BNN once integrated into the computer's processing? Such topics are hard to answer and current research on the topic is in its infancy. In addition, with the global emergence of AI made public in only the past few years. But the benefits of such computational systems are apparent and mitigate the flaws of current information processing systems in computers.The ability of the brain to process information in a hierarchical and modular manner, as well as BNN's plasticity and hidden memory, can allow information processing systems to be optimally efficient. Without the rigid structure of AI neural networks, such a system allows for fluidity of adaptability, while taking on an energy-efficient approach.

References

1. Smirnova, L., Caffo, B. S., Gracias, D. H., Huang, Q., Pantoja, I. E. M., Tang, B., Zack, D. J., Berlinicke, C. A., Boyd, J. L., Harris, T. D., Johnson, E. C., Kagan, B. J., Kahn, J., Muotri, A. R., Paulhamus, B. L., Schwamborn, J. C., Plotkin, J., Szalay, A. S., Vogelstein, J. T., . . . Hartung, T. (2023). Organoid intelligence (OI): the new frontier in biocomputing and intelligence-in-a-dish. Frontiers in Science, 1. https://doi.org/10.3389/fsci.2023.1017235

2. Roberts, T. P., Kern, F. B., Fernando, C., Szathmáry, E., Husbands, P., Philippides, A. O., & Staras, K. (2019). Encoding Temporal Regularities and Information Copying in Hippocampal Circuits. Scientific reports, 9(1), 19036. https://doi.org/10.1038/s41598-019-55395-1

3. Kagan, B. J., Gyngell, C., Lysaght, T., Cole, V. M., Sawai, T., & Savulescu, J. (2023). The technology, opportunities, and challenges of Synthetic Biological Intelligence. Biotechnology Advances, 68, 108233. https://doi.org/10.1016/j.biotechadv.2023.108233

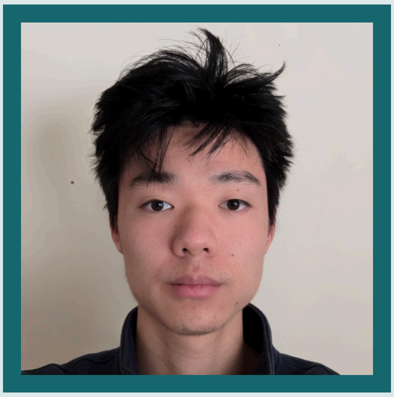
4. Dranias, M. R., Westover, M. B., Cash, S., & VanDongen, A. M. (2015). Stimulus information stored in lasting active and hidden network states is destroyed by network bursts. Frontiers in integrative neuroscience, 9, 14. https://doi.org/10.3389/fnint.2015.00014

5. Sigman, M., & Dehaene, S. (2008). Brain mechanisms of serial and parallel processing during dual-task performance. The Journal of neuroscience : the official journal of the Society for Neuroscience, 28(30), 7585–7598. https://doi.org/10.1523/JNEUROSCI.0948-08.2008

6. Nelson, M. E., & Bower, J. M. (1990). Brain maps and parallel computers. Trends in Neurosciences, 13(10), 403–408. https://doi.org/10.1016/0166-2236(90)90119-u

7. Ballard, D. H. (1986). Cortical connections and parallel processing: Structure and function. Behavioral and Brain Sciences, 9(1), 67–90. https://doi.org/10.1017/s0140525x00021555

8. Kagan, B. J., Kitchen, A. C., Tran, N. T., Habibollahi, F., Khajehnejad, M., Parker, B. J., Bhat, A., Rollo, B., Razi, A., & Friston, K. J. (2022). In vitro neurons learn and exhibit sentience when embodied in a simulated game-world. Neuron, 110(23), 3952–3969.e8. https://doi.org/10.1016/j.neuron.2022.09.001



About the Author

Edward is a Sophomore at the University of Illinois majoring in Neural Engineering. Through his studies, he aspires to implement biological mechanisms/systems into computers and explore AI-neural network connections. Some of his interests include playing volleyball, filming, and going on road trips. After graduation, he hopes to attend graduate school.