ENERGY EFFICIENT, SCALABLE NEUROMORPHIC COMPUTING WITH GROWTH TRANSFORM NEURAL NETWORK

Chakrabartty, Shantanu, Gangopadhyay, Ahana

Markiewicz, Gregory

T-019112

Technology Description

Engineers in Prof. Shantanu Chakrabartty's laboratory have developed Growth Transform Neural Network (GTNN), a flexible system for designing scalable neuromorphic processors for use in deep learning systems and support vector machines. GTNN frames the neuromorphic system as a network-level energy optimization problem, thereby providing a framework to develop scalable and energy efficient analog machine learning algorithms.

At the fundamental level, a single action potential generated by a biological neuron is not optimized for energy and consumes significantly more power than an equivalent floating-point operation in a Graphical Processing Unit (GPU) or a Tensor Processing Unit (TPU). Yet a population of coupled neurons in the brain, using Giga coarse neural operations (or spikes) can learn and implement diverse functions compared to an application-specific deep-learning platform that typically uses Peta 8-bit/16-bit floating-point operations or more. GTNN addresses this neuron-to-network energy-efficiency gap by exploiting dynamics at the level of a single neuron in a population. The formulation mimics learning networks using a fully-coupled, analog neuromorphic architecture designed as a unified dynamical system that encodes information using short-term and long-term network dynamics. This system is ultra-energy-efficient, while optimizing a learning or task objective in real time. Thus, GTNN could offer an elegant approach for designing chip architectures comprising millions or billions of neurons for real-time learning.
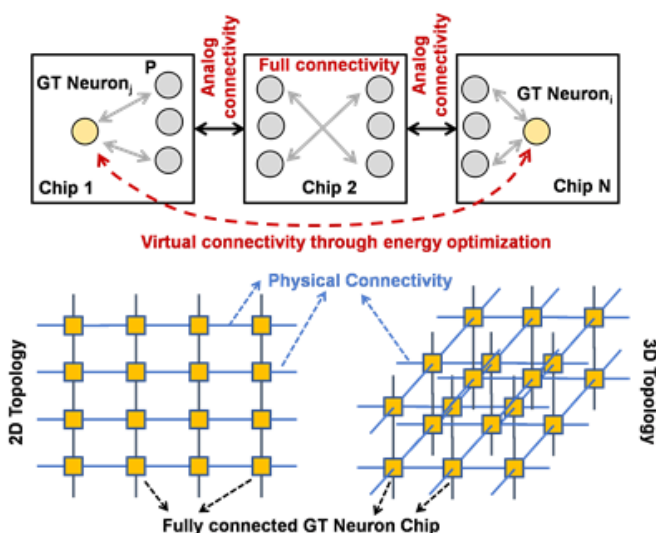
The GTNN formulation enables a neuron located on a chip

to communicate with a neuron on another chip using analog signals without being physically connected to each other. As a result, scalable two-dimensional and three-dimensional analog topologies are possible, which then self-optimize for energy-consumption while performing learning tasks in real-time.

Stage of Research

The inventors have developed the GTNN model and corresponding software toolkit. They demonstrated the model for different types of single neuron and population dynamics, including a spiking associative memory network that uses fewer spikes than conventional architectures while maintaining high recall accuracy at high memory loads.

Publications

- Gangopadhyay A, Mehta D, & Chakrabartty S. (2020). A spiking neuron and population model based on the growth transform dynamical system. *Frontiers in Neuroscience*, 14:425.
- Chatterjee O & Chakrabartty S. (2020). Resonant machine learning based on complex growth transform dynamical systems. *IEEE:TNNLS*, 32(3): 1289-1303.

Applications

- **Neuromorphic computing –** hardware and software for machine learning applications that demand high energy efficiency

Key Advantages

- **Energy efficient:**
  - exploits population dynamics to improve overall system energy-efficiency while optimizing learning/objectives in real time
  - uses fewer spikes than traditional architecture for associative memory
- **Scalable:**
  - potentially enables scaling to billions of neurons because network dynamics do not require explicit spike routing
  - encompasses all the neurons in the network (both visible and hidden), with the potential for easier and more effective training of hidden neurons in deep networks
- **Flexible, independent optimization:**
  - incorporates different neural dynamics that have been observed in electrophysiological recordings
  - provides independent control and optimization of different components of neurodynamics

Patent Application – US 20200401876

Related Web Links - AIM Laboratory