Demonstrating significant energy savings using neuromorphic hardware

24 May 2022, by Christoph Pelzl

One of Intel's Nahuku boards, each of which contains eight to 32 Intel Loihi neuromorphic chips. Credit: Tim Herman/Intel Corporation

For the first time TU Graz's Institute of Theoretical Computer Science and Intel Labs demonstrated experimentally that a large neural network can process sequences such as sentences while consuming four to sixteen times less energy while running on neuromorphic hardware than non-neuromorphic hardware. The new research based on Intel Labs' Loihi neuromorphic research chip that draws on insights from neuroscience to create chips that function similar to those in the biological brain.

The research was supported by The Human Brain Project (HBP), one of the largest research projects in the world with more than 500 scientists and engineers across Europe studying the human brain. The results of the research are published in published in Nature Machine Intelligence.

Human brain as a role model

Smart machines and intelligent computers that can autonomously recognize and infer objects and relationships between different objects are the subjects of worldwide artificial intelligence (AI) research. Energy consumption is a major obstacle on the path to a broader application of such AI methods. It is hoped that neuromorphic technology will provide a push in the right direction. Neuromorphic technology is modeled after the human brain, which is highly efficient in using energy. To process information, its hundred billion neurons consume only about 20 watts, not much more energy than an average energy-saving light bulb.

In the research, the group focused on algorithms that work with temporal processes. For example, the system had to answer questions about a previously-told story and grasp the relationships between objects or people from the context. The hardware tested consisted of 32 Loihi chips.

Loihi research chip: Up to sixteen times more energy-efficient than non-neuromorphic hardware

"Our system is four to sixteen times more energy-efficient than other AI models on conventional hardware," says Philipp Plank, a doctoral student at TU Graz's Institute of Theoretical Computer Science. Plank expects further efficiency gains as these models are migrated to the next generation of Loihi hardware, which significantly improves the performance of chip-to-chip communication.

"Intel's Loihi research chips promise to bring gains in AI, especially by lowering their high energy cost," said Mike Davies, director of Intel's Neuromorphic Computing Lab. "Our work with TU Graz provides more evidence that neuromorphic technology can improve the energy efficiency of today's deep learning workloads by re-thinking their implementation from the perspective of biology."
Mimicking human short-term memory

In their neuromorphic network, the group reproduced a presumed memory mechanism of the brain, as Wolfgang Maass, Philipp Plank's doctoral supervisor at the Institute of Theoretical Computer Science, explains that "experimental studies have shown that the human brain can store information for a short period of time even without neural activity, namely in so-called 'internal variables' of neurons. Simulations suggest that a fatigue mechanism of a subset of neurons is essential for this short-term memory."

"Recurrent neural structures are expected to provide the greatest gains for applications running on neuromorphic hardware in the future," said Davies. "Neuromorphic hardware like Loihi is uniquely suited to facilitate the fast, sparse and unpredictable patterns of network activity that we observe in the brain and need for the most energy efficient AI applications."

Symbiosis of recurrent and feed-forward network

The researchers link two types of deep learning networks for this purpose. Feedback neural networks are responsible for "short-term memory."