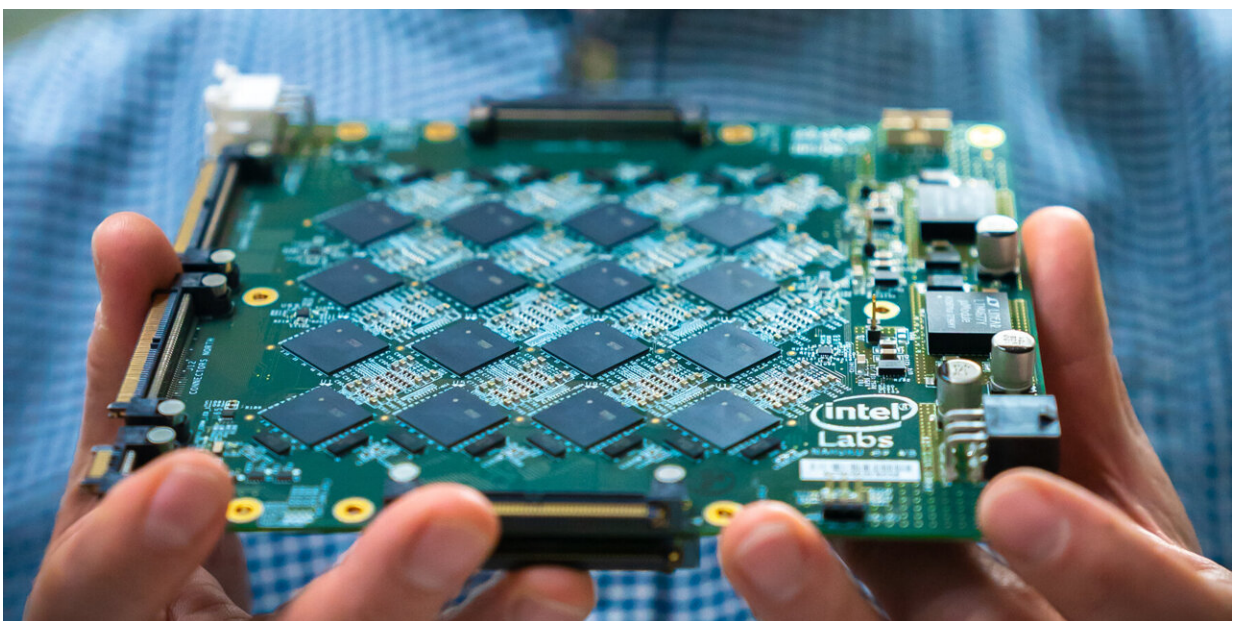# A neuromorphic computing architecture that can run some deep neural networks more efficiently

June 14 2022, by Ingrid Fadelli

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

As artificial intelligence and deep learning techniques become increasingly advanced, engineers will need to create hardware that can run their computations both reliably and efficiently. Neuromorphic computing hardware, which is inspired by the structure and biology of the human brain, could be particularly promising for supporting the

operation of sophisticated deep neural networks (DNNs).

Researchers at Graz University of Technology and Intel have recently demonstrated the huge potential of neuromorphic computing hardware for running DNNs in an experimental setting. Their paper, published in *Nature Machine Intelligence* and funded by the Human Brain Project (HBP), shows that neuromorphic computing hardware could run large DNNs 4 to 16 times more efficiently than conventional (i.e., non-brain inspired) computing hardware.

"We have shown that a large class of DNNs, those that process temporally extended inputs such as for example sentences, can be implemented substantially more energy-efficiently if one solves the same problems on neuromorphic hardware with brain-inspired neurons and neural network architectures," Wolfgang Maass, one of the researchers who carried out the study, told TechXplore. "Furthermore, the DNNs that we considered are critical for higher level cognitive function, such as finding relations between sentences in a story and answering questions about its content."

In their tests, Maass and his colleagues evaluated the energy-efficiency of a large neural network running on a neuromorphic computing chip created by Intel. This DNN was specifically designed to process large letter or digit sequences, such as sentences.
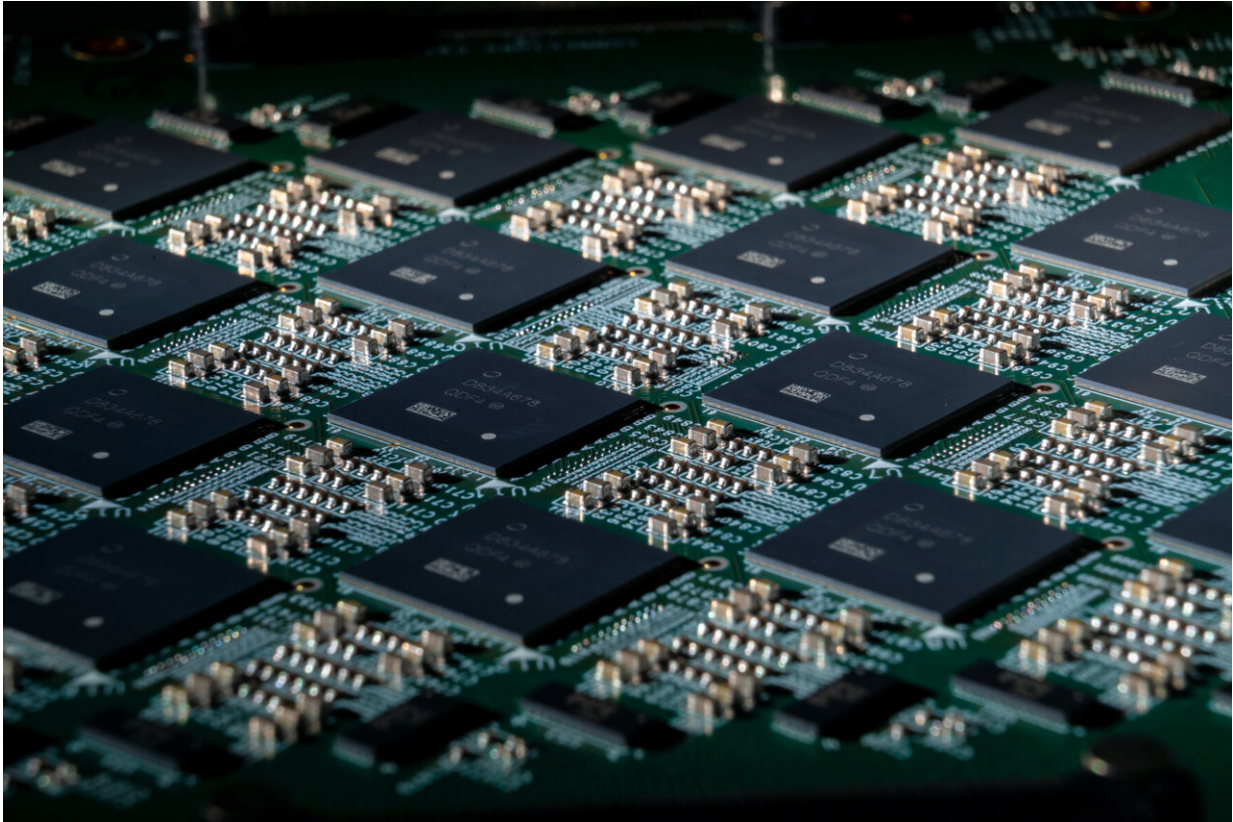
The close-up shows an Intel Nahuku board, each of which contains eight to 32 Intel Loihi neuromorphic research chips. Credit: Tim Herman/Intel Corporation

The researchers measured the energy consumption of the Intel neuromorphic chip and a standard computer chip while running this same DNN and then compared their performances. Interestingly, the researchers found that adapting the neuron models contained in computer hardware so that they resembled neurons in the brain enabled new functional properties of the DNN, improving its energy-efficiency.

"Enhanced energy efficiency of neuromorphic hardware has often been conjectured, but it was hard to demonstrate for demanding AI tasks," Maass explained. "The reason is that if one replaces the artificial neuron

models that are used by DNNs in AI, which are activated 10s of thousands of times and more per second, with more brain-like 'lazy' and therefore more energy-efficient spiking neurons that resemble those in the brain, one usually had to make the spiking neurons hyperactive, much more than neurons in the brain (where an average neuron emits only a few times per second a signal). These hyperactive neurons, however, consumed too much energy."

Many neurons in the brain require an extended resting period after being active for a while. Previous studies aimed at replicating biological neural dynamics in hardware often reached disappointing results due to the hyperactivity of the artificial neurons, which consumed too much energy when running particularly large and complex DNNs.
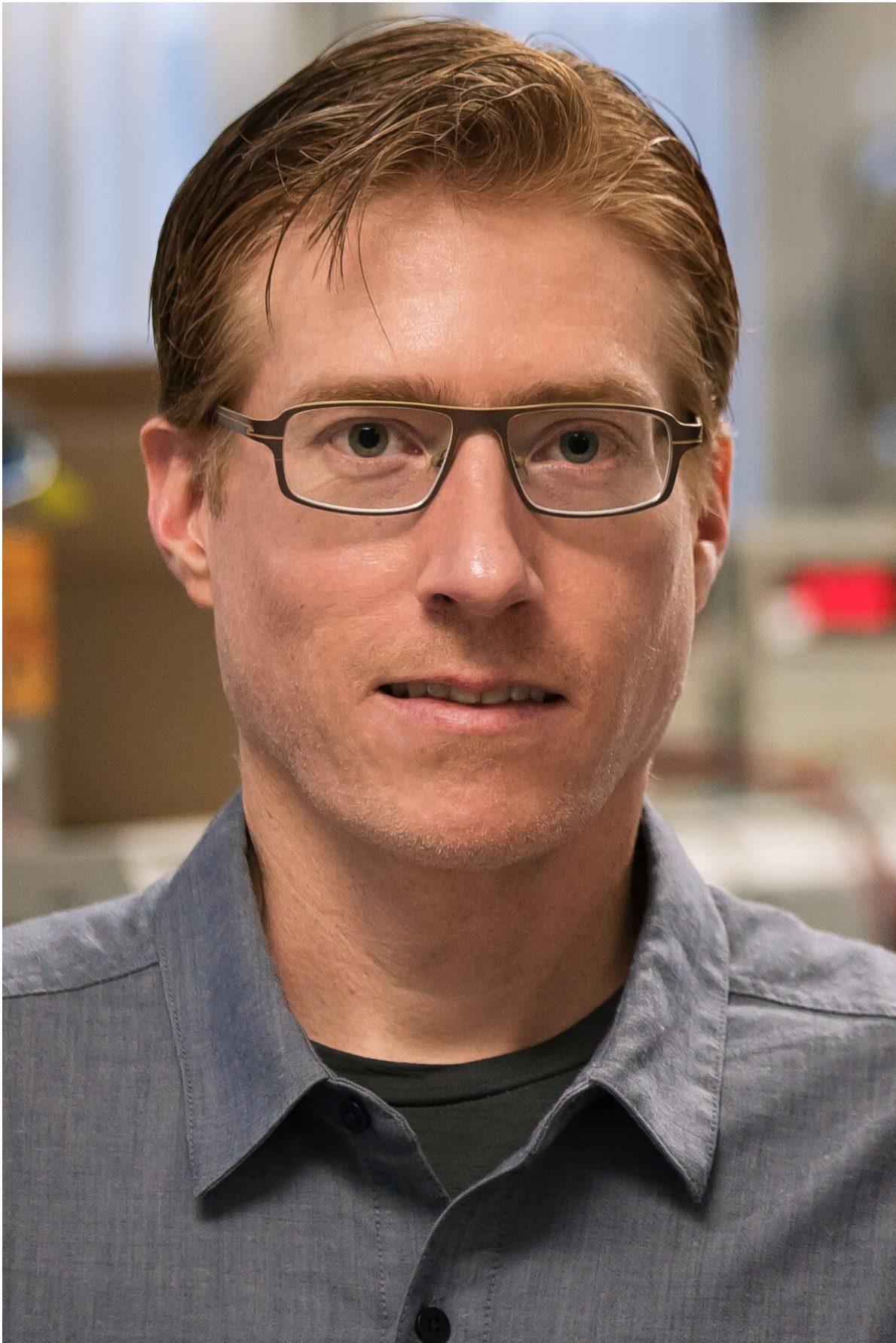
Mike Davies is Director of Intel's Neuromorphic Computing Lab, which focuses on research efforts to enable next generation devices and autonomous systems guided by the principles of biological neural computation including sensing, robotics, healthcare, and large-scale AI applications. Credit: Tim Herman/Intel Corporation

In their experiments, Maass and his colleagues showed that the tendency of many biological neurons to rest after spiking could be replicated in neuromorphic hardware and used as a "computational trick" to solve time series processing tasks more efficiently. In these tasks, new information needs to be combined with information gathered in the recent past (e.g., sentences from a story that the network processed beforehand).

"We showed that the network just needs to check which neurons are currently most tired, i.e., reluctant to fire, since these are the ones that were active in the recent past," Maass said. "Using this strategy, a clever network can reconstruct based on what information was recently processed. Thus, 'laziness' can have advantages in computing."

The researchers demonstrated that when running the same DNN, Intel's neuromorphic computing chip consumed 4 to 16 times less energy than a conventional chip. In addition, they outlined the possibility of leveraging the artificial [neurons](#)' lack of activity after they spike, to significantly improve the hardware's performance on time series processing tasks.
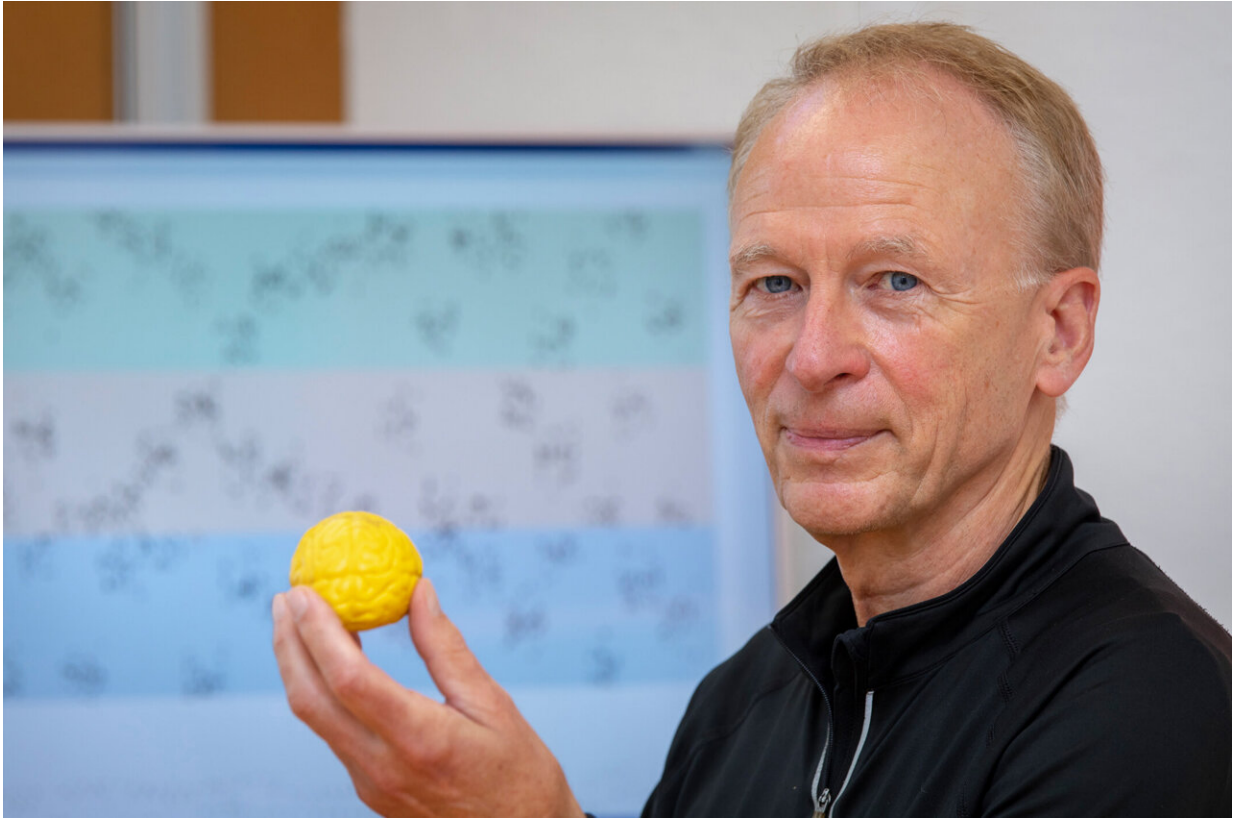
TU Graz scientist Wolfgang Maass deals with the question of how the information processing in the human brain works and how this could be used for computer applications. Credit: Lunghammer—TU Graz

In the future, the Intel chip and the approach proposed by Maass and his colleagues could help to improve the efficiency of neuromorphic computing hardware in running large and sophisticated DNNs. In their future work, the team would also like to devise more bio-inspired strategies to enhance the performance of neuromorphic chips, as current hardware only captures a tiny fraction of the complex dynamics and functions of the human brain.

"For example, human brains can learn from seeing a scene or hearing a sentence just once, whereas DNNs in AI require excessive training on

zillions of examples," Maass added. "One trick that the brain uses for quick learning is to use different learning methods in different parts of the brain, whereas DNNs typically use just one. In my next studies, I would like to enable neuromorphic hardware to develop a 'personal' memory based on its past 'experiences,' just like a human would, and use this individual experience to make better decisions."

More information: Arjun Rao et al, A Long Short-Term Memory for AI Applications in Spike-based Neuromorphic Hardware, *Nature Machine Intelligence* (2022). DOI: 10.1038/s42256-022-00480-w

© 2022 Science X Network