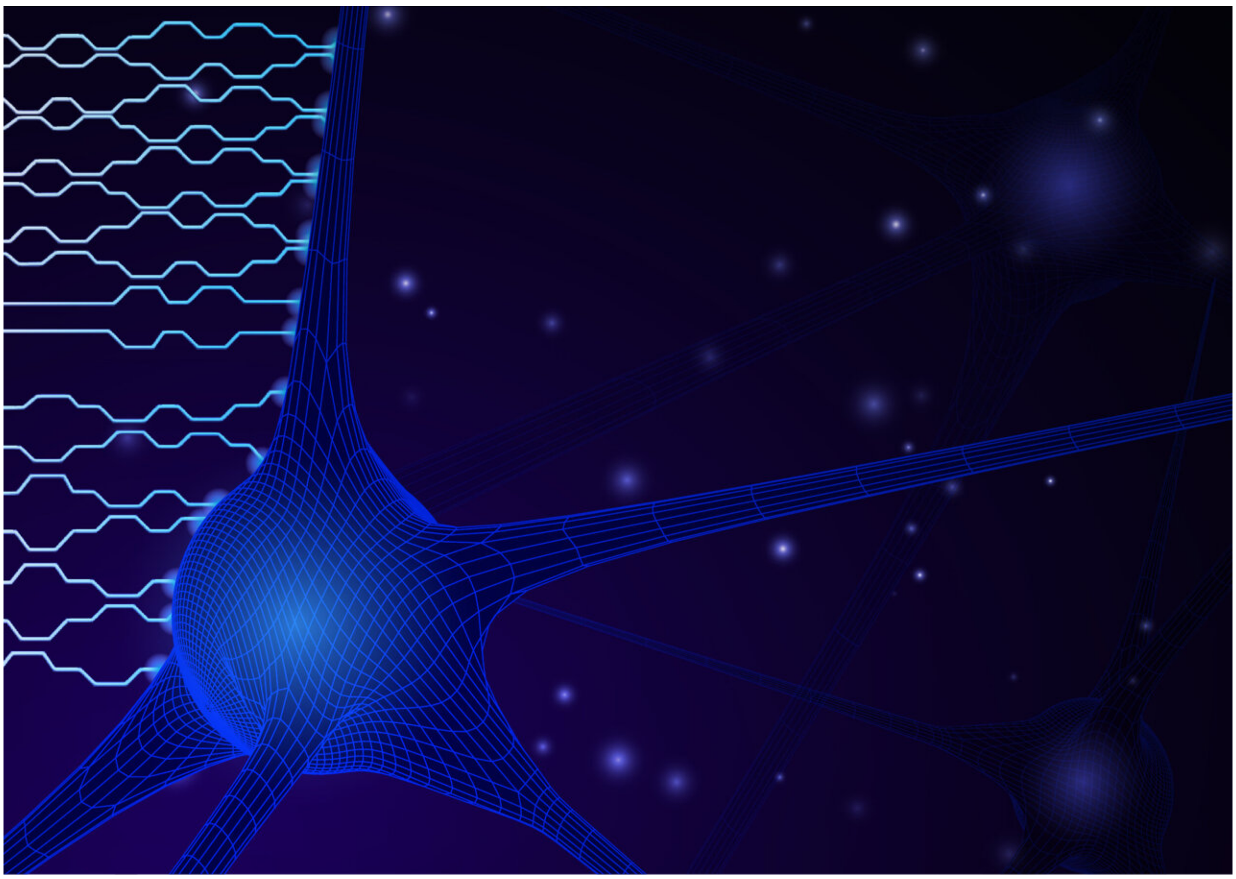
Neural networks on photonic chips: Harnessing light for ultra-fast and low-power artificial intelligence

May 2 2023



Neural networks on photonic chips. Credit: Politecnico di Milano

Neural networks are distributed computing structures inspired by the structure of a biological brain and aim to achieve cognitive performance comparable to that of humans but in a much shorter time.

These technologies now form the basis of machine learning and artificial intelligence systems that can perceive the environment and adapt their own behavior by analyzing the effects of previous actions and working autonomously. They are used in many areas of application, such as speech and image recognition and synthesis, autonomous driving and augmented reality systems, bioinformatics, genetic and molecular sequencing, and high-performance computing technologies.

Compared to conventional computing approaches, in order to perform complex functions, neural networks need to be initially "trained" with a large amount of known information that the network then uses to adapt by learning from experience. Training is an extremely energy-intensive process and as computing power increases, the neural networks' consumption grows very rapidly, doubling every six months or so.

Photonic circuits are a very promising technology for neural networks because they make it possible to build energy-efficient computing units. For years, the Politecnico di Milano has been working on developing programmable photonic processors integrated on silicon microchips only a few mm^2 in size for use in the field of data transmission and processing, and now these devices are being used to build photonic neural networks.

"An artificial neuron, like a biological neuron, must perform very simple mathematical operations, such as addition and multiplication, but in a neural network consisting of many densely interconnected neurons, the energy cost of these operations grows exponentially and quickly becomes prohibitive. Our chip incorporates a photonic accelerator that allows calculations to be carried out very quickly and efficiently, using a programmable grid of silicon interferometers. The calculation time is equal to the transit time of light in a chip a few millimeters in size, so we are talking about less than a billionth of a second (0.1 nanoseconds)," says Francesco Morichetti, Head of the Photonic Devices Lab of the Politecnico di Milano.

"The advantages of photonic neural networks have long been known, but one of the missing pieces to fully exploit their potential was network training.. It is like having a powerful calculator, but not knowing how to use it. In this study, we succeeded in implementing training strategies for photonic neurons similar to those used for conventional neural networks. The photonic 'brain' learns quickly and accurately and can achieve precision comparable to that of a conventional neural network, but faster and with considerable energy savings. These are all building blocks for artificial intelligence and quantum applications," adds Andrea Melloni, Director of Polifab the Politecnico di Milano micro and nanotechnology center.

In addition to applications in the field of neural networks, this device can be used as a computing unit for multiple applications where high computational efficiency is required, e.g., for graphics accelerators, mathematical coprocessors, data mining, cryptography and quantum computers.

The work is published in the journal *Science*.

More information: Sunil Pai et al, Experimentally realized in situ backpropagation for deep learning in photonic neural networks, *Science* (2023). DOI: 10.1126/science.ade8450

Provided by Polytechnic University of Milan