Extended Abstracts of the 2018 International Conference on Solid State Devices and Materials, Tokyo, 2018, pp491-492

Artificial neural networks based on Mach-Zehnder interferometers in a Si PIC

Yichen Shen[1,2]

[1] Massachusetts Institute of Technology, Department of Phyiscs
[2] Lightelligence Inc.
[1] 77 Massachusetts Avenue
Cambridge 02139, USA
Phone: +1-410-336-1337 E-mail: ycshen@mit.edu
[2] 268 Summer Street
Cambridge 02139, USA
Phone: +1-410-336-1337 E-mail: yichen@lightelligence.ai

Abstract
Artificial Neural Networks (ANN) have dramatically improved performance for many machine learning tasks. We demonstrate new architectures based on Photonic Integrated Circuits for carrying out ANN and other statistical processing tasks. I will discuss how it can enable orders of magnitude enhancement in computational speed, latency and power efficiency over state of the arts electronics.

Computers that can learn, combine, and analyze vast amounts of information quickly, efficiently, and without the need for explicit instructions are a powerful tool for handling large datasets. Indeed, "deep learning" algorithms have received an explosion of interest in both academia and industry for their utility in image recognition, language translation, decision making problems, and more [1]. Traditional central processing units (CPUs) are far suboptimal for implementing these algorithms [2]; and a growing effort in academia and industry has been put towards the development of new hardware architectures tailored towards applications in artificial neural networks and deep learning [3]. Graphical Processing Unit (GPUs), Application Specific Integrated Circuits (ASICs) and field-programmable gate arrays (FPGAs), have enabled both energy efficiency and speed enhancements for learning tasks. In parallel, hybrid optical-electronic systems that implement spike processing and reservoir computing have been shown. However, the computational speed and power efficiency achieved with these hardware architectures are still limited by electronic clock rates and Ohmic losses.

Optical neural networks offer a promising alternative approach to microelectronic and hybrid optical-electronic implementations. In fact, Artificial Neural Networks are a promising optical-assisted computing paradigm because (1) *they rely heavily on fixed matrix multiplications*: linear transformations (and certain non-linear transformations) can be performed at the speed of light and detected at rates exceeding 100 GHz [4] in photonic networks, and in some cases, with minimal power consumption. (2) *they have weak requirements on resolution*: in general, statistical computing are more robust to error and can be operated at much lower resolution: optimum for analog computing and (3) *once a neural network is trained, the architecture can be passive*, the computation on the optical signals will be performed without additional energy input.

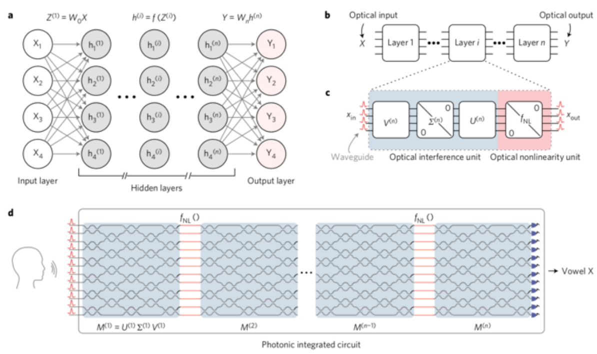

Figure 1 | General architecture of the ONN. a, General artificial neural network architecture composed of an input layer, a number of hidden layers and an output layer. b, Decomposition of the general neural network into individual layers. c, Optical interference and nonlinearity units that compose each layer of the artificial neural network. d, Proposal for an all-optical, fully integrated neural network. [6]

Artificial neural network architecture contains an input layer, at least one hidden layers, and an output layer. In each layer, information propagate through neural network via linear combination (e.g. matrix multiplication) followed by nonlinear activation function applied to the result from linear combination. In training an artificial neural network model, data are fed into the input layer, and output is calculated through the forward propagation step. Then the parameters are optimized through the back propagation procedure. The key component optical neural network (ONN) computing is the optical interference unit, depicted in Fig. 1. Which is used to perform arbitrary *unitary* matrix multiplication on the input optical signal. The unitary matrix can be obtained using a network of Mach-Zehnder interferometers. Mathematically, it can be rigorously proved that any arbitrary unitary matrix can be represented by the network of Mach-Zehnder interferometers [5].

In this talk, I will show that some main frame statistical computing algorithms can be effectively carried out by the Si PIC, I will also discuss the engineering challenges and opportunities for photonics community.

Acknowledgements

I would like to express sincere thanks to all the collaborators for this project.

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature 521, 436–444 (2015).

[2] S. K. Esser et al, “Convolutional networks for fast, energy-efficient neuromorphic computing,” Proceedings of the National Academy of Sciences 113, 11,441–11,446 (2016).

[3] C. Mead, “Neuromorphic electronic systems,” Proceedings of the IEEE 78, 1629–1636 (1990).

[4] L. Vivien et al, “Zero-bias 40gbit/s germanium waveguide photodetector on silicon,” Opt. Express 20, 1096–1101 (2012).

[5] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, “Experimental realization of any discrete unitary operator,” Phys. Rev. Lett. 73, 58–61 (1994).

[6] Y. Shen et al, “Deep Learning with Coherent Nanophotonic Circuits,” Nat. Phot. 11, 441 (2017).