Optical Fiber Communication Conference 2021 © OSA 2021

Neuromorphic Photonics for Optical Communication Systems

Chaoran Huang[1] , Thomas Ferreira de Lima[1], Shinsuke Fujisawa[2], Alexander N. Tait[1], Hsuan-Tung Peng[1], Bhavin J. Shastri[3], Ting Wang[2] and Paul. R. Prucnal[1]

[1]Princeton University, Princeton, NJ 08542, USA
[2]NEC Laboratories America, Inc.,4 Independence Way, Princeton, NJ 08540, USA
[3]Queens University, Kingston, ON K7L 3N6, Canada
email addresses: chaoranh@princeton.edu

Abstract: Neuromorphic photonics creates processors 1000× faster than electronics while consuming less energy. We will discuss the role of neuromorphic photonics in optical communications, review existing approaches, and outline the required technologies to evolve this field. © 2021 The Author(s)

1. Introduction

The world is witnessing an explosion of Internet traffic. The wide deployment of digital signal processing (DSP) technologies in optical communication systems are one of the most significant advances that have sustained the data traffic growth over the last decade. The key to advancing DSP relies on the consistent improvement in CMOS technologies. However, as embodied in Moores law, hardware scaling of application-specific integrated circuit (ASIC) based DSP chip is practically unsustainable, and it can hardly accommodate the exponentially increasing data traffic in the future. In parallel, many efforts are focused on developing new algorithms to increase transmission capacity. Machine learning algorithms, especially using neural network models, have been shown to outperform traditional DSP algorithms in performing many functions in optical communication systems [1]. To date, the benefits of machine learning algorithms are largely validated using conventional computers and conducted offline. Real-time implementation of machine learning algorithms on ASIC chips remains a grand challenge due to the high throughput, low energy, and low latency requirement in optical communication systems.

The emerging field of neuromorphic photonics promises to solve these challenges by creating radical new hardware platforms that emulate the underlying neural network model with photonic devices and circuits. Photonics has unmatched feats for interconnects and communications which can negate the bandwidth and interconnectivity trade-offs that electronic hardware fundamentally suffers. As a result, algorithms running on neuromorphic photonic hardware (i.e., photonic neural networks (PNNs)) could break performance limitations in electronics, and gain advantages in speed, latency, and power consumption in solving intellectual tasks that are unreachable by conventional digital electronic platforms. This could be a major benefit for optical communication systems. This paper will provide a rationale for PNNs as a compelling alternative to process optical communication signals. We will review PNN approaches and their applications in long-haul and short-reach optical communication systems.

2. Photonic neural networks (PNNs)

The general idea of neuromorphic photonics is to create photonic systems to mimic neural network models, i.e., neurons interconnected by massive physical synapses with co-integrated non-volatile memories. As a result, neural network computations can be carried out by the governing physical dynamics of the photonic systems. Neural network models highlight the need for high-degree physical interconnections (i.e., neuron-to-neuron and neurons-to-memory communications). PNNs can outperform electronic systems fundamentally because photonics has unmatched feats for interconnects. First, photonic waveguides can increase interconnectivity by carrying many parallel signals simultaneously through wavelength-division multiplexing (WDM), resulting in an aggregate bandwidth of up to 10 Tb/s per single waveguide. Second, photonics performs linear operations, i.e., matrix multiplication, at the speed of light, providing much lower latency than electronic MAC operations. Third, photonic matrix multiplication consumes almost 'zero' dynamic energy. Static power cost, caused by laser wall-plug efficiency, waveguide losses, energy consumed in maintaining the weights, can be optimized with high-efficient photonic devices. Moreover, matrix multiplication with photonics is purely analog and does not require any logic operations, meaning that the static power cost is shared by a large number of MAC operations. The associated energy costs using photonics are currently at the order of fJ per MAC with silicon photonics and aJs per MAC with sub-λ nanophotonics. On the other hand, energy cost in google TPU is hundreds of fJ per MAC [2]. PNNs are a growing field of interest and enable new domains of information processing applications. The interested readers are directed to recent report reviewing a full coverage of this area [3].

3. PNNs for optical communications

PNNs are well suited for optical communications because the optical signals are processed directly in the optical domain. This innovation avoids prohibitive energy consumption overhead and speed reduction in ADCs, especially in data center applications. Ref [4] shows 40 folds of power reduction when ADCs are not needed if inputs are optical signals. In parallel, many PNN approaches are inspired by optical communication systems, making PNNs naturally suitable for processing optical communication signals. For example, we proposed synaptic weights and neuron networking architecture based on the concept of WDM to enable fan-in and weighted addition [5] (Fig. 1(a)). This architecture can provide a seamless interface between PNNs and WDM systems. It can be applied as a front-end processor to address inter-wavelength or inter-mode crosstalk problems that DSP usually lacks the bandwidth or computing power to process (e.g., fiber nonlinearity compensation in WDM systems). Moreover, PNNs combine high-quality waveguides and photonic devices that have been initially developed for telecommunications. We first proposed a scalable silicon PNN, composing microring resonator (MRR) banks for synaptic weighting and O/E/O neurons to produce machine learning activation functions [6]. The MRR weight bank can be re-purposed as standard WDM filters, and the O/E/O neurons use typical silicon photodetector and modulator. Thankfully, the optimization of associated devices in PNNs can utilize the fruits of the entire silicon photonic ecosystem that is paramountly driven by telecommunications and data center applications. Therefore, PNNs, by default, can support fiber optic communication rates and enables real-time processing.

In order to truly demonstrate photonics can excel over DSP, careful considerations are required to identify different application scenarios (i.e., long-haul, short-reach) and system requirements (i.e., performances, energy). Continuous research is needed to improve photonic hardware and to develop hardware-compatible algorithms. Here, we discuss several approaches to train and apply PNNs for optical communications.

Neuromorphic approach: we developed a PNN based on the so-called ”neuromorphic” (i.e., neuron isomorphic) approach, aiming to map physical models of optoelectronic systems to abstract models of neural networks [5]. By doing so, PNNs can leverage existing machine learning algorithms (i.e., back-propagation) to train the PNN and then map training results from simulations to photonic hardware. The concept is known as heterogeneous computing and is shown in Fig. 1(c). We applied this approach to implement a feed-forward neural network model that can learn the nonlinear perturbation and compensate the nonlinear distortions in a 10800 km submarine fiber transmission link [7]. We trained the network with neural network programming tools (i.e., TensorFlow) and directly reconfigured the silicon photonic chip. A proof-of-concept experiment demonstrates that the silicon PNN can produce a similar Q factor improvement compared to the neural network simulated on a computer. This work validates that PNNs can be used as a heterogeneous device enabling fast inference tasks in optical communications.

We also proposed a photonic architecture enabling all-to-all continuous-time recurrent neural networks (RNNs) [5], as shown in Fig. 1. RNNs can resemble optical fiber transmission systems: the linear neuron-to-neuron connections with internal feedback is analog to linear multiple-input multiple-output (MIMO) fiber channel with dispersive memory. With neuron nonlinearity, RNNs can be ideally used to approximate all types of linear and nonlinear effects in a fiber transmission system and compensate for different transmission impairments. RNNs, consisting of many feedback connections, are considered to be computationally expensive for digital hardware and require at least milliseconds to conduct a single inference. Contrarily, in photonic RNN, the feedback operations are simply done by busing the signals on photonic waveguides, allowing photonic hardware capable of converging to the solution within microseconds. This architecture also adopts the neuromorphic approach and thus allows to train PNNs externally using standard machine learning algorithms.

Reservoir computing (RC): RC is a class of RNNs that consist of a reservoir of randomly connected neurons followed by a readout layer. Contrary to the neuromorphic approach, inside the reservoir, the weight connections are fixed. Only the weights in the readout layer are trained with supervised learning. Reservoirs do not need a known isomorphism between the physical hardware and a neural network model. This feature makes the circuit simple to construct. Neuromorphic hardware, on the other hand, has the burden of establishing an isomorphic mapping to neural network models. In return, neuromorphic systems can leverage existing machine learning algorithms to guarantee particular behaviors. L. Appeltant et al. proposed an optical RC using only a single nonlinear node with delayed feedback connections as shown in Fig. 1(d), making such photonic systems very cost-efficient to implement [8]. Despite their simple circuits, photonic RCs have demonstrated many functions, including dispersion and nonlinear compensations for both IM/DD and coherent optical systems. The interested reader is directed towards the recent review papers in this area [9].

Adaptive online learning: intra-data-center links constitute the largest volume among all market segments in data center interconnects. Multi-mode fibers are used to scale the information capacity with orthogonal optical modes. However, modal crosstalk limits the achievable data rate and transmission distance. Within the distance

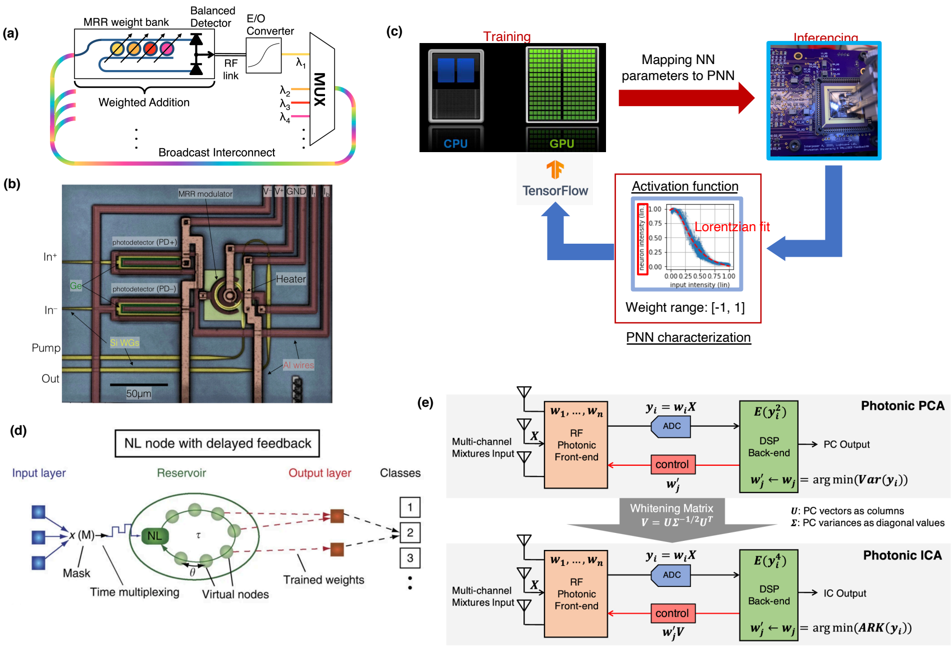

Fig. 1. (a) WDM based photonic neural network [5]; (b) OEO silicon photonic neuron [3]; (c) concept of heterogeneous computing with PNN [7]; (d) optical reservoir computing [8]; (e) photonic blind source separation [11]

of intra-data-center links (<300 m), crosstalk mostly results from a random unitary mode coupling [10]. Thus the crosstalk can be compensated by applying an unitary matrix that reverts the channel response. DSP-based MIMO algorithm is powerful to estimate the channel response; however, DSP is typically avoided in data center links due to its high latency, power consumption, and upgrading cost. Modal crosstalk can be solved using photonic matrix multipliers, together with a photonic implementable algorithm to estimate the channel response, for example, Ref [10]. The modal crosstalk compensation problem can be formulated as separating an unknown mixture of unknown but independent signals. We demonstrated a photonic system to de-mixing signals in the optical domain with blind source separation (BSS), an unsupervised learning algorithm [11]. Photonic BSS only needs to observe and analyze statistical properties. Key merits stemming from this statistical approach include being able to adaptively learning the fiber dynamics that randomly drift and automatically update its outputs, because the statistical property is time-invariant. Besides, photonic BSS only needs to sample the signals deeply below the Nyquist frequency, meaning that BSS can be implemented with slow electronics (sub-MHz), which significantly reduces the power consumption that concerns conventional DSP.

4. Discussion and conclusions

Applying PNNs to optical communication systems can potentially address many challenges of signal processing in terms of bandwidth, latency, and power consumption. PNNs operate in the analog domain, which has a direct effect on noise and noise propagation. Therefore, we need to add more consideration to maintain the signal-to-noise ratio (SNR) for optical communication systems. At the system level, for example, O/E/O neurons are preferred because they can restore the signal's SNR by converting the information to another external and clean optical pump. At the device level, future studies should focus on creating low-loss and high-efficient devices to minimize the insertion loss of PNNs. As photonics integration technology matures, we expect PNNs to one day usher a paradigm change in signal processing for optical communications.

References

1. F. N. Khan *et al.*, J. Light. Technol. **37**, 493–516 (2019).
2. M. A. Nahmias *et al.*, IEEE J. Sel. Top. Quantum Electron. **26**, 1–18 (2019).
3. B. J. Shastri *et al.*, Nat. Photonics **15**, 102–114 (2021).
4. M. Miscuglio and V. J. Sorger, Appl. Phys. Rev. **7**, 031404 (2020).
5. A. N. Tait *et al.*, J. Light. Technol. **32**, 4029–4041 (2014).
6. A. N. Tait *et al.*, Phys. Rev. Appl. **11**, 064043 (2019).
7. C. Huang *et al.*, in *OFC Postdeadline Papers 2020,* (Optical Society of America), p. Th4C.6.
8. L. Appeltant *et al.*, Nat. communications **2**, 1–6 (2011).
9. D. Brunner *et al.*, J. applied physics **124**, 152004 (2018).
10. A. Annoni *et al.*, Light. Sci. & Appl. **6**, e17110–e17110 (2017).
11. P. Y. Ma *et al.*, Opt. Lett. **45**, 6494–6497 (2020).