On-chip Training Silicon Photonic Circuits for Novel Classification Computing

Guangwei Cong[1], Noritsugu Yamamoto[1], Takashi Inoue[1], Yuriko Maegami[1], Morifumi Ohno[1], Shota Kita[2], Shu Namiki[1], and Koji Yamada[1]

[1]*National Institute of Advanced Industrial Science and Technology (AIST), 16-1, Onogawa, Tsukuba, Ibaraki 305-8569, Japan.*

[2]*NTT Basic Research labs., 3-1, Morinosato Wakamiya, Atsugi-shi, Kanagawa 243-0198, Japan*

gw-cong@aist.go.jp ORCID: 0000-0003-3519-635X

***Abstract*—In this invited talk, we review our work on on-chip training silicon photonic circuits for novel classification computing based on support vector machine like projection-based principle and its recent progress.**

Keywords—integrated photonic circuits, optical computing, silicon photonics, classification

I. Introduction

Exploiting optical/photonic systems to perform computing tasks has been gaining wide research attention over the past several years [1], which has been witnessed as the second boom of optical computing since the final two decades of the last century. The research motivation lies in two facts: (1) with CMOS technology approaching to the physical limits, it is becoming more and more difficult to obtain significant performance enhancement for digital processors by downsizing CMOS devices; and (2) at the same time, AI models based on deep learning are growing at an exponential rate, which highly requires specifically designed accelerators for much more efficient processing of AI workloads to reduce energy consumption and achieve faster performance. Thus, it is expected to utilize optical/photonic technologies to realize such AI accelerators [2]. Since pure optical computer is still not practical, optical/photonic technologies usually contribute to improve computing systems in two ways: high-speed optical interconnects and photonics-assisted computing. Running super-big AI models must involve in huge numbers of computing operations and memory access operations in distributed processors and memories, which induce significant delays and power consumption. Optical links between these computing resources can enable high speed and low power data movement. The photonics-assisted computing is usually implemented in electronics-photonics hybrid systems where the part of photonics can complete specific computing tasks such as matrix multiplication by optical propagation, which obviously has speed and power advantages over traditional digital processors based on transistors. So far, various novel digital and analog optical/photonic computing systems have been demonstrated, including neural network [3,4], reservoir computing [5], kernel convolution [6], arithmetic logic [7], Ising machine [8], etc. However, for the target using optical/photonic systems to accelerate neural network computing, some challenges remain unsolved: the scale is still small, all-optical dense layer connection via integrated nonlinear activation functions, on-chip training for large scale, in-line high throughput computing due to heavy processing by intermediate or afterward electronics, etc. Recently, we proposed and demonstrated a photonic circuit for classification computing, which implements the support vector machine like projection-based principle by utilizing the electro-optical nonlinearity (cosine phase-amplitude relation) of Mach-Zehnder interferometers (MZI) [9]. The electronics parts were used for the data input and output (IO) only and the computing was done only by optical propagation inside the interferometer circuits. This photonic circuits can perform nonlinear classification without using nonlinear activation functions. We also demonstrated direct on-chip training for this device without using pre-trained network models. In this talk, we will introduce our work about this photonic classifier device and its recent progress.

II. Photonic Chip and Measurement Setup

The fabricated chip and its packaged module are shown in Fig. 1. Its topology was explained in detail in [9]. It was fabricated on a 220-nm photonic silicon-on-insulator wafer with a 3-μm buried oxide layer at the AIST-SCR 12-inch CMOS line utilizing ArF liquid immersion photolithography [10]. It consists of thermo-optic phase shifters and MZIs. The heater is TiN and the electrode is AlCu alloy. The ground pads are indicated in Fig. 1. This chip was packaged with fiber arrays for optical IOs, and wire bonding was done for electrical IOs. For measurement, the laser was input to the chip after being tuned to TE polarization. The optical powers from eight output ports were measured by a multi-channel photodetector array. A computer read these powers and controlled multi-channel direct-current electrical sources to tune all phase shifters on chip. The computer ran the training algorithms to learn the phase status for all phase shifters, i.e., weights in form of voltage, for a specific computing task. Once the phase status is learned, it is kept as it is for continuous inference application for the new data input. The data is input from electrical domain to the phase shifter of MZIs, as indicated by the arrows in Fig. 1 for Boolean or Iris classification computing. Once the data is applied, the computing will be completed immediately when the light propagates out of the chip to the photodetectors, which results in low power and low latency. The classification results are marked out by the maximum-power (brightest) port position.

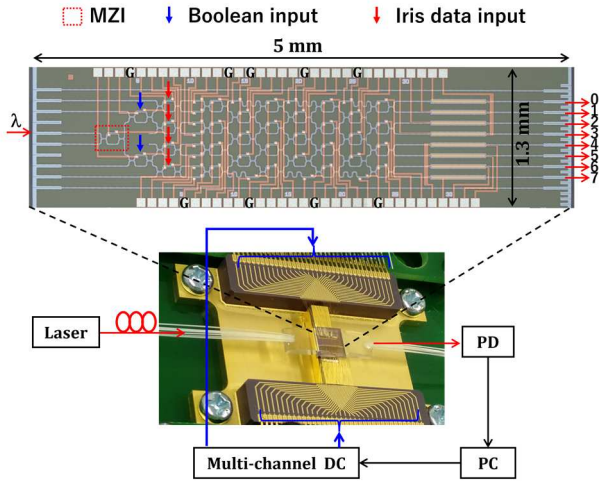

Fig. 1. Photonic chip, packaged module, and measurement setup.

III. On-chip Training and Classification Computing

We adopted two algorithms, the forward propagation algorithm and bacteria foraging algorithm, for on-chip training in our demonstration, which were run by the computer in Fig. 1. Fig. 2 shows the flow chart of the training program that was implemented by Python using PyTorch, an open-source machine learning framework [11]. The device is encapsulated as a python class module (PhotonicClassifier) in which both algorithms are defined as the member functions. The training can be easily done by declaring the model and calling the algorithm function.

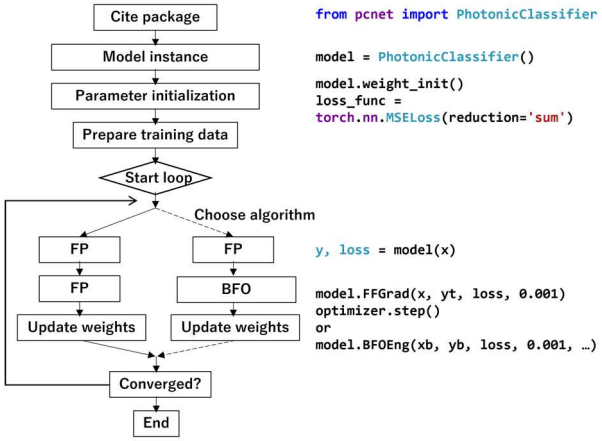

Fig. 2. Flow chart of the on-chip training, including forward propagation algorithm and bacteria foraging algorithm.

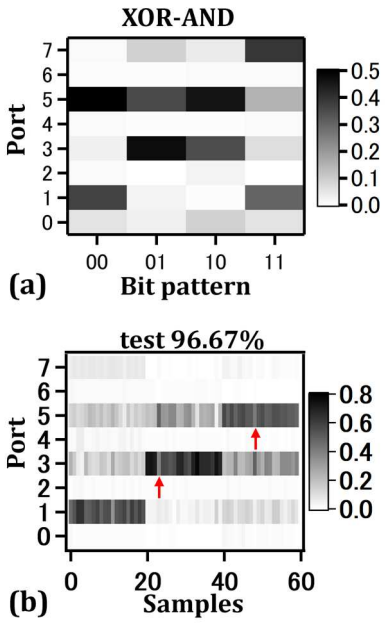

Fig. 3. (a) Simultaneous Boolean computing of XOR and AND logics. (b) Verification results of Iris classification.

We performed on-chip training for two computing tasks: Boolean logics and Iris dataset classification [9]. For the former, the input is the bit pattern, and the output is the logic value. For example, XOR is a linearly inseparable problem that is usually used to test the nonlinear classification models [12]. For the latter, it is to judge the Iris flower species from four flower's structural parameters. For the Boolean logic, an example we demonstrated was the simultaneous computing of XOR and AND. The two bits were input to two MZIs indicated in Fig. 1. We assigned the port 1 and 3 to stand for the logic values 0 and 1 for XOR, respectively. Similarly, the port 5 and 7 were assigned to stand for the logic values 0 and 1 for AND. We trained the photonic chip to output the maximum optical power to the corresponding port in response to the input bit pattern. After the phase status of all on-chip phase shifters was learned, the bits could be continuously input for verification with fixing the voltages applied to all phase shifters, and the optical propagation completed computing. Since optical propagation is much faster than the data input speed (e.g., using modulators), the high throughput can be guaranteed. Fig. 3(a) shows the verification results of XOR and AND. For example, when we input the bit pattern of "11", the optical powers are high at the ports 1 and 7, indicating the results of "0" for XOR and "1" for AND. Besides the Boolean computing, we also demonstrated the Iris dataset classification. Similarly, we assigned the ports 1, 3, and 5 to present the three kinds of Iris flowers, Setosa, Versicolor, and Virginica, respectively. The four parameters were input to the four MZIs indicated in Fig. 1 after normalization. On-chip training was performed using 90 samples by both algorithms mentioned above. The training performance of both algorithms were compared in [9]. After the training was completed, we verified another 60 samples. These samples were input in an order from the species one to three, each having 20 samples. The verified result is shown in Fig. 3(b). The optical powers are obviously separated into three groups, indicating that the maximum power occurs at the ports 1, 3, and 5, corresponding to the input sample species. The classification accuracy is about 96.7%.

Acknowledgment

This work was supported by Japan Science and Technology Agency (JST), CREST Grant Number JPMJCR15N4 and JPMJCR21C3, Japan. The authors thank all staffs at the AIST-SCR station for device fabrication.

References

[1] C. Li, et al, "The challenges of moder computing and new opportunities for optics," PhotoniX, vol. 2, no. 20, Sept. 2021.

[2] K. Kitayama, et al., "Novel frontier of photonics for data processing-photonic accelerator, " APL Photonics, vol. 4, pp. 090901, Aug. 2019.

[3] Y. Shen, et al., "Deep learning with coherent nanophotonic circuit," Nat. Photon., vol. 11, pp. 441–446, Jun. 2017.

[4] H. Zhang, et al., "An optical neural chip for implementing complex-valued neural network," Nat. Commun., vol. 12, no. 457, Jan. 2021.

[5] K. Vandoorne, et al., "Experimental demonstration of reservoir computing on a silicon photonics chip," Nat. Commun., vol. 5, no. 3541, Mar. 2014.

[6] X. Y. Xu, et al., "11 TOPS photonic convolutional accelerator for optical neural networks," Nature, vol. 589, pp. 44–51, Jan. 2021.

[7] Z. Ying, et al., "Electronic-photonic arithmetic logic unit for high-speed computing," Nat. Commun., vol. 11, no. 2154, May 2020.

[8] M. Prabhu, et al., "Accelerating recurrent Ising machines in photonic integrated circuits," Optica, vol. 7, no. 5, pp. 551-558, May 2020.

[9] G. Cong, et al., "On-chip bacterial foraging training in silicon photonic circuits for projection-enabled nonlinear classification," Nat. Commun., vol. 13, no. 3261, June 2022.

[10] K. Yamada, et al., "A 300-mm-wafer silicon photonics technology for ultra-low-energy optical network systems," in Asia Communications and Photonics Conference, Guangzhou, China, 2017, S4H.3.

[11] https://pytorch.org/.

[12] T. W. Hughs, et al., "Training of photonic neural networks through *in situ* backpropagation and gradient measurement," Optica, vol. 5, no. 7, pp. 864–871, Jul. 2018.