Deep Learning Processors Based on Si Photonic Circuits

TAKENAKA Mitsuru

We have conducted research on deep learning processors using programmable Si photonic integrated circuits. By integrating compound semiconductors and phase-change materials into Si photonic circuits, we have successfully achieved precise control and measurement of optical phase within the circuits, as well as highly sensitive optical intensity detection. This enables precise control of large-scale programmable photonic circuits with high speed and low power consumption. Additionally, by using the newly proposed photonic circuits, we have demonstrated the potential to accelerate learning through on-chip error back-propagation, in addition to inference, greatly contributing to the early realization of deep learning processors based on Si photonic circuits. Here, we present our recent research on deep learning processors based on Si photonic circuits.

Introduction

Over the past 50 years, the miniaturization of transistors has progressed to the point of reaching 3 nm, and research and development for further miniaturization is ongoing worldwide[1]. However, as semiconductor miniaturization becomes even more challenging, it is believed that improving computing performance through optoelectronic fusion will be necessary for the beyond 2 nm generation. In particular, the development of artificial intelligence (AI), represented by ChatGPT, has been remarkable, and new computing technologies that go beyond conventional principles are truly sought after. Against this background, optical operations using large-scale Si photonic circuits has attracted much attention[2-3]. By using programmable Si photonic circuits[4], arithmetic operations such as multiply-accumulate (MAC) operation can be performed using light, and it is expected that the performance of deep learning will be significantly improved. Active research is being conducted worldwide for this purpose. However, various challenges remain to achieve large-scale programmable photonic circuits, as shown in Figure 1. In order to program photonic circuits, precise control of the phase of optical signals in the circuit is required. However, the currently mainstream optical phase shifter that uses heater heating consumes a large amount of power and is not suitable for large-scale integration[5]. In addition, in order to precisely control parameters such as optical phase, it is necessary to integrate numerous optical power monitors for measuring circuit operations, but there are challenges in sensitivity and optical insertion loss. High-speed and low-power operation is required for photodetectors that read out optical operations as electrical signals, but sensitivity and parasitic capacitance are also challenges.

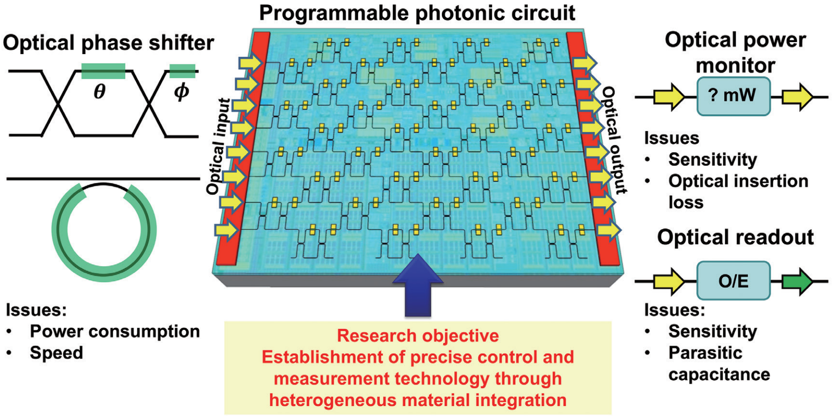


Figure 1 Tehcnology issues in programmable photonic circuits.

To overcome these technological challenges, we have conducted research to establish precise control and measurement techniques for photonic circuits by integrating compound semiconductors and phase-change materials heterogeneously into Si photonic circuits. Additionally, we have developed unique programmable photonic circuits capable of performing both learning and inference functions in the optical domain. In this paper, we present our recent research, which advances the early realization of deep learning processors through optoelectronic fusion.

Optical phase shifters

An optical phase shifter is a device that can manipulate the phase of an optical signal as it passes through a photonic circuit. Since the programming of a photonic circuit relies on optical phase tuning, it is essential to integrate numerous optical phase shifters into the photonic circuit. To control optical phase, a thermo-optic phase shifter, as depicted in Figure 2(a), is widely employed in Si photonic circuits due to its simple fabrication procedure and low optical insertion loss[6]. When heating a Si waveguide by current injection, the refractive index of Si increases due to the thermo-optic effect, resulting in an optical phase shift. However, it requires approximately 10 – 20 mW of heating power to achieve a π phase shift, which hinders large-scale integration. Moreover, thermal crosstalk between neighboring thermo-optic phase shifters makes precise control of optical phase difficult. As a substitute for power-hungry thermo-optic phase shifters, various device candidates are being intensely investigated around the world[7-10].

As an energy-efficient optical phase shifter, we have investigated a metal-oxide-semiconductor (MOS) capacitor-based device using heterogenous integration of III-V semiconductor on Si, as depicted in Figure 2(b)[11]. Using direct wafer bonding technology, a thin n-type InGaAsP is bonded onto a p-type Si rib waveguide with an Al_2O_3 bonding interface[12]. We employed an Al_2O_3 layer deposited through atomic layer deposition (ALD) to passivate the InGaAsP surface, which is essential for the MOS-based optical phase shifter. The ALD Al_2O_3 layer also facilitates a hydrophilic surface, enabling wafer bonding without the need for a plasma surface activation process. When a gate voltage (Vg) is applied between the InGaAsP membrane and Si waveguide, electrons and holes accumulate at the InGaAsP and Si MOS interfaces, respectively. The electron-induced refractive index change in InGaAsP is more than ten times greater than in Si Moreover, the hole-induced absorption in InGaAsP, which is significantly greater than in Si, can be completely removed since hole accumulates only on the Si side. As a result, efficient, low-loss optical phase modulation is achieved using the InGaAsP/Si hybrid MOS capacitor. The energy consumption required to sustain a π phase shift, predominantly influenced by gate leakage current through a MOS capacitor, is less than 1 nW[13]. This value is more than one million times smaller than that of a thermo-optic phase shifter. Therefore, we anticipate a significant reduction in the power consumption of a programmable PIC through the utilization of III-V/Si hybrid MOS optical phase shifters.

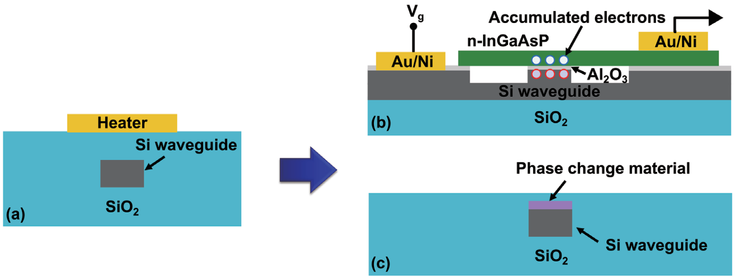


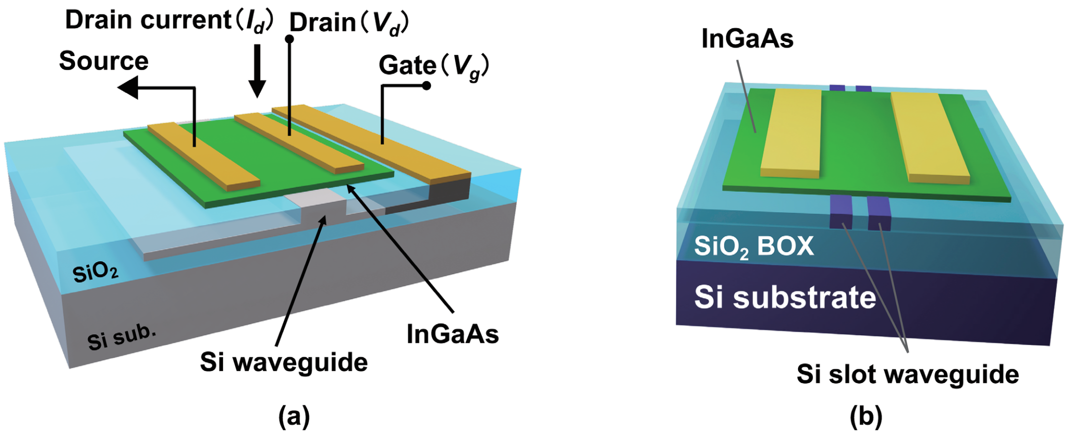
Figure 2 (a) Optical phase shifter based on thermo-optic effect. (b) Optical phase shifter based on heterogeneous integration of III-V compound semiconductor and Si. (c) Optical phase shifter using phase change material.

Non-volatile optical phase shift is also crucial to achieve in-memory computing, which can help overcome the von Neumann bottleneck. In addition to ferroelectric materials, phase change materials are promising candidates for achieving non-volatile optical phase shift[14]. A significant contrast in optical constants exists between the amorphous and crystalline states of a phase-change material. By depositing a phase change material on a Si waveguide as shown in Figure 2(c), this large contrast in optical constants can be used for optical modulation. $Ge_2Sb_2Te_5$ (GST), widely employed in rewritable optical disks, was initially introduced into a Si waveguide for non-volatile optical intensity modulation at near-infrared wavelengths. However, optical phase modulation is not achievable due to the significant absorption in GST. To address this issue, we have proposed to use GST in mid-infrared (MIR) wavelengths[15]. We experimentally demonstrated the band-edge absorption of GST can be reduced using a longer operating wavelength and achieved 2.6 dB loss for a p phase shifter at a 2.32 µm wavelength. To further reduce optical loss, we have developed $Ge_2Sb_2Te_3S_2$ (GSTS), where the bandgap energy of GSTS is increased through sulfur doping into GST[16]. Using GSTS, we achieved 0.29 dB of π phase shift at a 2.34 µm wavelength, which is one of the lowest values among phase change material-based optical phase shifters. We applied the GSTS optical phase shifter to a ring resonator and successfully tuned the resonance wavelength of the resonator without significant degradation in its Q-factor. Owing to the non-volatility in the phase change material-based optical phase shifter, the standby power consumption can be zero after tuning optical phase, enabling the large-scale integration of optical phase shifters into a programmable photonic circuit.

Optical power monitor and photodetector

The monitoring of optical power in a Si waveguide is crucial for a programmable photonic circuit. Since initial phase errors are unavoidable due to fabrication variations in Si waveguides, the initial state of a programmable photonic circuit must be calibrated in some way. Monitoring optical outputs by tuning optical phase shifters on a chip is the most common method for indirectly estimating the circuit state. However, this method is applicable only to specific circuits. A more direct method of measuring the circuit state is to embed numerous optical power monitors directly into a photonic circuit. A Ge photodetector can be easily integrated with a Si photonic circuit; however, tapping waveguides are additionally required for power monitoring, which increases optical loss and complicates the circuit further. Reading a small change in the conductance of a Si waveguide due to the increase in free carriers resulting from optical absorption is another potential method for power monitoring. This method enables low insertion loss; however, its low sensitivity necessitates a phase-sensitive detection circuit, which complicates the overall circuit design.

To address the challenges in power monitoring, we have proposed a waveguide-coupled phototransistor, as shown in Figure 3(a)[17]. A thin InGaAs membrane is bonded onto a Si waveguide to serve as an absorber. When the light signal propagates through the Si waveguide, the evanescent light is absorbed in the InGaAs layer. The absorbed light generates photocarriers in the InGaAs layer, which are detected as a photocurrent between the source and drain terminals. Unlike conventional photodetectors, we propose utilizing the Si waveguide as a gate



(a)

(b)

Figure 3 (a) Waveguide-coupled phototransistor for optical power monitoring.
(b) Waveguide-coupled photodetector with InGaAs membrane on Si slot waveguide.

electrode by applying a gate voltage. Here, we employ an Al_2O_3 bonding interface between the InGaAs and Si layers to achieve superior surface passivation, enabling electron accumulation at the InGaAs MOS interface. By applying a positive gate voltage to the InGaAs layer through the Si gate electrode, the InGaAs layer functions as a transistor channel, controlling the drain current flowing between the source and drain terminals. Through transistor operation, the photocurrent is amplified, leading to a significant increase in responsivity. Previously reported waveguide phototransistors have utilized metal gate electrodes, which are positioned far from the channel to minimize optical loss caused by the electrodes. Consequently, gate controllability is compromised. By employing a Si waveguide as a gate electrode, we can position the Si gate electrode underneath the InGaAs channel, enabling effective gating without a significant increase in optical loss. Leveraging these advantages in our proposal, we have achieved approximately 10^6 A/W responsivity, which represents the highest value among waveguide-coupled phototransistors. We found that the switching speed ranges from 1 µs to 100 µs, which is sufficient for power monitor applications. Thanks to its high responsivity, we can reduce the device length to minimize optical insertion loss. Hence, we can achieve a high-responsivity, transparent optical power monitor without the need for additional complicated electrical circuits.

To further reduce the optical insertion loss of the presented phototransistor, we proposed employing an InP membrane instead of an InGaAs membrane[18]. Since the bandgap energy of InP is greater than the photon energy of interest, InP is essentially transparent for light at a 1.55 µm wavelength. Consequently, photodetection using InP is not feasible. However, we have observed substantial photocurrent in the InP/Si waveguide-coupled phototransistor. We expect that a small optical absorption in InP related to the crystal defects in InP is amplified through the phototransistor operation. As a result, the responsivity of greater than 1 A/W is experimentally demonstrated even using InP. Since the strong band-edge absorption is not used, we achieved an optical insertion loss of less than 0.25 dB. Since we do not rely on strong band-edge absorption, we have achieved an optical insertion loss of less than 0.25 dB. Additionally, we have monolithically integrated the waveguide-coupled phototransistor and MOS-based optical phase shifter by bonding the same InP membrane onto the Si waveguides. The output from a Mach-Zehnder interferometer controlled with the MOS-based optical phase shifter is monitored by the InP/Si phototransistor, exhibiting the feasibility of power monitoring a programmable photonic circuit.

High-speed and low-power photodetectors are also crucial for reading the results of optical operations performed by a programmable photonic circuit. If low-power MOS or phase change material-based optical phase shifters are utilized, the power required for the optical operation itself becomes negligibly small. Consequently, the power consumption of a programmable photonic circuit for deep learning is primarily dominated by optical input and output interfaces. Therefore, the low power consumption of the optical receiver system is crucial for reducing the operating power of deep learning. A conventional optical receiver system necessitates a transimpedance amplifier to convert photocurrent into a voltage signal, resulting in significant electricity consumption. To resolve this issue, the receiver-less system was proposed, in which a transimpedance amplifier is replaced by a load resistor[19]. However, to obtain sufficient voltage across the load resistor, its resistance must be greater than 1 kΩ. In this scenario, the operation speed is greatly limited by the RC constant of the receiver system. To achieve an operation speed greater than 10 GHz, the capacitance of a photodetector must be reduced to approximately 1 fF. However, achieving such a small capacitance is challenging without compromising responsivity, as there is a trade-off relationship between responsivity and capacitance in a photodetector. To overcome this trade-off relationship, we have proposed a III-V/Si hybrid photodetector, wherein a thin InGaAs absorber is bonded onto a Si slot waveguide, as illustrated in Figure 3(b)[20]. A Si slot waveguide enables the confinement of propagating light into a narrow gap between two Si rails. Consequently, optical absorption in the InGaAs is enhanced, allowing us to reduce the device length to lower the capacitance without sacrificing responsivity. We experimentally demonstrated a low capacitance of 1.9 fF and a high responsivity of 1.0 A/W at a wavelength of 1.55 µm wavelength simultaneously. The capacitance is expected to be further reduced to less than 1 fF by eliminating the parasitic capacitance of the electrodes. Hence, the presented photodetector contributes to the realization of a receiver-less system for a low-power readout of a programmable photonic circuit.

Photonic crossbar array

A programmable photonic circuit based on nested Mach-Zehnder interferometers arranged in a triangle or rectangular mesh is the most common circuit topology for deep learning applications because of its universality in optical operations. However, the size of a single Mach-Zehnder interferometer is large, which makes it unsuitable for large-scale integration. To overcome this issue, we have proposed photonic crossbar array based on ring resonators, as depicted in Figure 4[21]. Add-drop ring resonators, which can selectively drop a specific wavelength signal into a drop waveguide, are arranged in a lattice fashion. As indicated in Figure 3(a), each ring resonator is tuned so that drop wavelengths do not overlap either in row and column. For the inference operation, a multiple-wavelength optical signal serving as an input vector is injected from the left-hand side of the crossbar. Subsequently, each wavelength signal is dropped at its respective ring resonator. Through optical phase shifters integrated with ring resonators, the amount of the dropped signal can be tuned, akin to the weight of a matrix for multiplication. Consequently, the multiplication between each input vector element and matrix weight can be executed. Finally, the dropped multiple-wavelength signal are simultaneously converted into electrical signal, which corresponds to the add operation. As a result, we can perform multiply-accumulation operation in optoelectronic fashion with the photonic crossbar array. One of the unique features of the proposed crossbar array is that it can be also used for accelerating learning[22]. As shown in Figure 4(b), when a multiple-wavelength optical signal serving as an error vector is injected from the top side of the crossbar, multiply-accumulation operation between the error vector and transposed matrix used for inference can be obtained from the right-hand side of the circuit. This multiply-accumulate operation is utilized in backpropagation for the learning of neural networks. Hence, the photonic crossbar array enables on-chip backpropagation in the optical domain. To show the feasibility of our proposal, we fabricated the photonic crossbar array using Si photonics platform as shown in Figure 4(c). We have successfully demonstrated inference and learning operations using this photonic crossbar array chip. In this demonstration, we utilized thermo-optic phase shifters to tune the ring resonators on the crossbar array. To reduce the power consumption of the phase shifters and mitigate thermal crosstalk between the ring resonators, we are now aiming for the integration of MOS-based or phase change material-based optical phase shifters with the crossbar array.

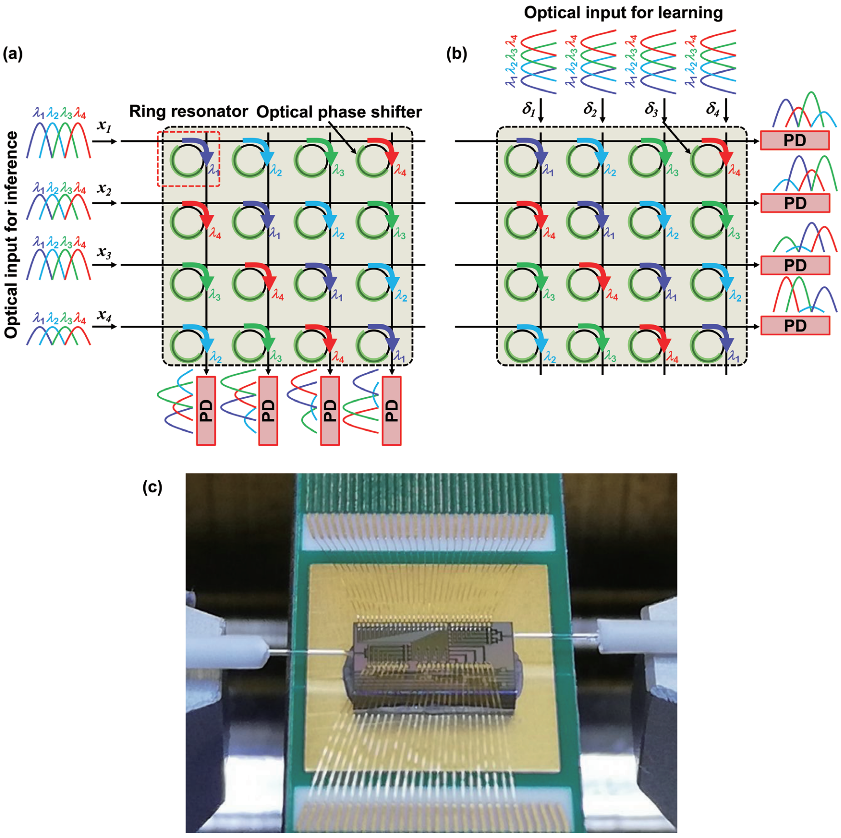


Figure 4 Photonic crossbar array based on ring resonators.
(a)Inference operation. (b) Learning operation with on-chip backpropagation. (c)Photograph of photonic crossbar array chip fabricated by Si photonics platform.

Conclusion

In this paper, we have explored the possibilities of a programmable photonic circuit based on Si photonics in advancing computing technologies, particularly in the domain of deep learning processors. As Moore's law begins to slow down, the fusion of electronics and photonics becomes increasingly crucial for overcoming the limitations posed by conventional computing architectures. Our research, which is based on innovative materials integration and new circuit architectures such as a photonic crossbar array, will contribute to the realization of large-scale programmable photonic circuits. With ongoing research and development efforts, we anticipate the creation of deep learning processors that harness the power of light to redefine the boundaries of computational efficiency and intelligence.

References

[1] W. Cao *et al.*, "The future transistors," *Nature*, vol. 620, no. 7974, pp. 501–515, Aug. 2023.

[2] Y. Shen *et al.*, "Deep learning with coherent nanophotonic circuits," *Nat. Photonics*, vol. 11, no. 7, pp. 441–446, Jun. 2017.

[3] B. J. Shastri *et al.*, "Photonics for artificial intelligence and neuromorphic computing," *Nat. Photonics*, vol. 15, no. 2, pp. 102–114, Jan. 2021.

[4] W. Bogaerts *et al.*, "Programmable photonic circuits," *Nature*, vol. 586, no. 7828, pp. 207–216, Oct. 2020.

[5] M. Takenaka *et al.*, "III–V/Si Hybrid MOS Optical Phase Shifter for Si Photonic Integrated Circuits," *J. Lightwave Technol.*, vol. 37, no. 5, pp. 1474–1483, Mar. 2019.

[6] S. Liu *et al.*, "Thermo-optic phase shifters based on silicon-on-insulator platform: state-of-the-art and a review," *Frontiers of Optoelectronics*, vol. 15, no. 1, p. 9, Apr. 2022.

[7] M. Dong *et al.*, "High-speed programmable photonic circuits in a cryogenically compatible, visible–near-infrared 200 mm CMOS architecture," *Nat. Photonics*, vol. 16, no. 1, pp. 59–65, Dec. 2021.

[8] J. Geler-Kremer *et al.*, "A ferroelectric multilevel non-volatile photonic phase shifter," *Nat. Photonics*, vol. 16, no. 7, pp. 491–497, May 2022.

[9] D. U. Kim *et al.*, "Programmable photonic arrays based on microelectromechanical elements with femtowatt-level standby power consumption," *Nat. Photonics*, pp. 1–8, Nov. 2023.

[10] R. Tang *et al.*, "Non volatile hybrid optical phase shifter driven by a ferroelectric transistor," *Laser Photon. Rev.*, vol. 17, no. 11, Nov. 2023.

[11] J.-H. Han, F. Boeuf, J. Fujikata, S. Takahashi, S. Takagi, and M. Takenaka, "Efficient low-loss InGaAsP/Si hybrid MOS optical modulator," *Nat. Photonics*, vol. 11, no. 8, pp. 486–490, Jul. 2017.

[12] J.-H. Han, M. Takenaka, and S. Takagi, "Study on void reduction in direct wafer bonding using Al_2O_3/HfO_2 bonding interface for high-performance Si high-k MOS optical modulators," *Jpn. J. Appl. Phys.*, vol. 55, no. 4S, 04EC06, Mar. 2016.

[13] Q. Li, J.-H. Han, C. P. Ho, S. Takagi, and M. Takenaka, "Ultra-power-efficient 2 × 2 Si Mach-Zehnder interferometer optical switch based on III-V/Si hybrid MOS phase shifter," *Opt. Express*, vol. 26, no. 26, p. 35003, Dec. 2018.

[14] C. Ríos *et al.*, "Integrated all-photonic non-volatile multi-level memory," *Nat. Photonics*, vol. 9, no. September, pp. 725–732, 2015.

[15] Y. Miyatake *et al.*, "Non-volatile Compact Optical Phase Shifter based on $Ge_2Sb_2Te_5$ operating at 2.3 µm," Opt. Mater. Express, vol. 12, no. 12, pp. 4582–4593, Oct. 2022.

[16] Y. Miyatake *et al.*, "Proposal of low-loss non-volatile mid-infrared optical phase shifter based on $Ge_2Sb_2Te_3S_2$," *IEEE Trans. Electron Devices*, vol. 70, no. 4, pp. 2106–2112, Apr. 2023.

[17] T. Ochiai *et al.*, "Ultrahigh-responsivity waveguide-coupled optical power monitor for Si photonic circuits operating at near-infrared wavelengths," *Nat. Commun.*, vol. 13, no. 1, p. 7443, Dec. 2022.

[18] T. Akazawa, K. Sumita, S. Monfray, F. Boeuf, K. Toprasertpong, S. Takagi, M. Takenaka, "Transparent in-line optical power monitor integrated with MOS optical phase shifter using InP/Si hybrid integration," *European Conference on Optical Communication (ECOC2023)*, We.D.4.5, Glasgow, UK, 1–5 October 2023.

[19] C. Debaes et al., "Receiver-less optical clock injection for clock distribution networks," *IEEE J. Sel. Top. Quantum Electron.*, vol. 9, no. 2, pp. 400–409, Mar. 2003.

[20] T. Akazawa *et al.*, "Low-Capacitance Ultrathin InGaAs Membrane Photodetector on Si Slot Waveguide Toward Receiverless System," *IEEE Trans. Electron Devices*, vol. 69, no. 12, pp. 7184–7189, Dec. 2022.

[21] S. Ohno, K. Toprasertpong, S. Takagi, and M. Takenaka, "Si microring resonator crossbar arrays for deep learning accelerator," *Jpn. J. Appl. Phys.*, vol. 59, no. SG, GGE04, Feb. 2020.

[22] S. Ohno, R. Tang, K. Toprasertpong, S. Takagi, and M. Takenaka, "Si Microring Resonator Crossbar Array for On-Chip Inference and Training of the Optical Neural Network," *ACS Photonics*, vol. 9, no. 8, pp. 2614–2622, Aug. 2022.

Dr. TAKENAKA Mitsuru

Professor,
School of Engineering,
Department of Electrical Engineering
and Information Systems,
The University of Tokyo