ACS Photonics

pubs.acs.org/journal/apchd5

Perspective

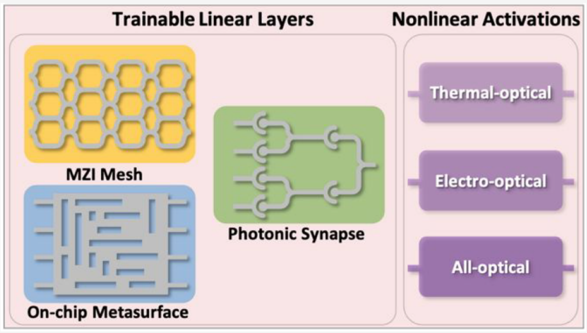
Integrated Photonic Neural Networks: Opportunities and Challenges

Kun Liao, Tianxiang Dai, Qiuchen Yan, Xiaoyong Hu,* and Qihuang Gong

Cite This: ACS Photonics 2023, 10, 2001−2010

Read Online

ABSTRACT: Photonic neural networks benefit from the use of photons to perform intelligent inference computing with ultrafast and ultralow energy consumption in ultra-high-throughput, providing the efficient photonic hardware for the new generation of intelligent computing, and the effective way to support large-scale integration for on-chip all-optical computing chips. With the rapid development of photonic neural networks, demands for efficient computation power have increased dramatically. However, the weak and impractical optical nonlinear activations, the lack of suitable configurations for integrated photonic hardware, and proper optical storage mediums pose challenges to this field. In this Perspective, we propose our current point of view and a suggestive roadmap in the field of integrated photonic platform for optical neural networks. Throughout the discussion, we highlight recent progresses meeting with major challenges. We also identify some next challenges still ahead to realize integrated photonic neural networks capable of matching the current computational power of graphic cards.



KEYWORDS: *integrated photonics, optical neural network, metasurface, spiking neural network, optical nonlinear activation, on-chip training*

■ INTRODUCTION

Inspired by biological brains, as one of the artificial intelligence algorithms, neural networks have the excellent inference performance in the execution of inference tasks due to the superior learning ability.[1,2] So far, neural networks have shown the outstanding performance and vigorous application prospect in various fields such as speech recognition,[3,4] image classification,[5,6] computer vision,[7,8] and natural language processing.[9,10] Nearly two decades past the realization of the limits for Moore's law,[11−14] processors based on electronic hardware have finally hit the bottleneck of unsustainable performance growth. Yet, demands for high-performance computing have dramatically increased due to the proliferation of artificial intelligence.[15−17] Such demands have resulted in significantly increased energy consumption in this field,[18−21] which is becoming the primary limitation for deployment and further development. In fact, one of the largest neural networks required 3.14E23 floating-point operations per second (FLOPS) of computation[22] to train. These computations usually take the form of a matrix multiplication,[23] which is greatly limited in speed and energy efficiency by the nature of the electronic devices due to the limited bandwidth and electronic signals easy to interfere with each other,[24,25] forcing a binary form of low level operations with careful consideration of signal integrity. In contrast, photonic processors that compute with photons instead of electrons display some extraordinary properties, such as an ultrawide communication bandwidth, ultrahigh processing frequency, and ultralow energy consumption.[26−28] Apart from these single-threaded advantages, additional dimensions division multiplexing of light field such as wavelength[29,30] and spatial mode[31−33] enable multithread processing with almost no extra computing overhead, leading to a significant acceleration against traditional electronic computers.

The conventional photonic computing framework under the von-Neumann architecture aims to form individual photonic devices with a preoptimized performance to a photonic integrated circuit and, further, to construct the photonic computing chip, which owns regular physical structures and clear physical principles, but a single function and a large integrated crosstalk.[34−37] In addition, the trade-off between ultrafast response time and ultralow energy consumption has brought great challenges to the large-scale integration of

Special Issue: Photonics in China

Received: September 28, 2022
Revised: January 24, 2023
Accepted: January 26, 2023
Published: February 7, 2023

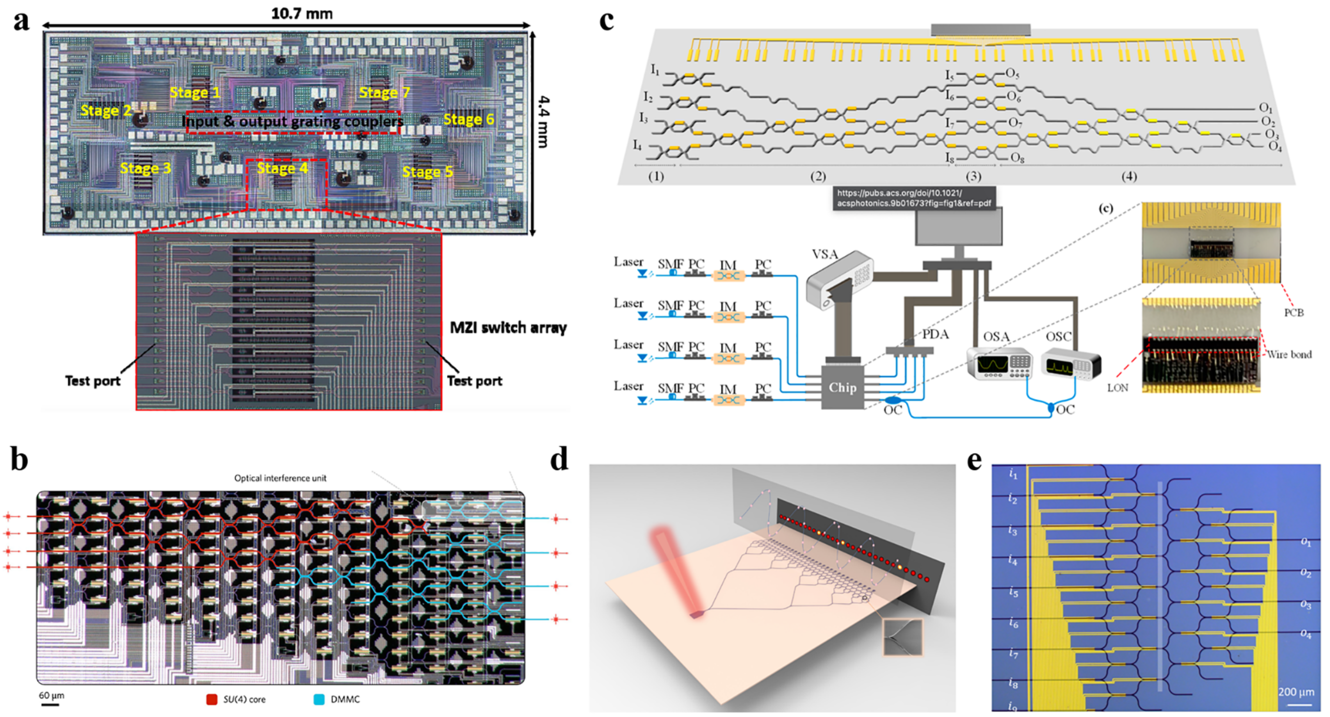


Figure 1. On-chip waveguide-based IPNNs. (a) Optical image of the fabricated 16 × 16 optical switch after an electrical and optical package. Reprinted with permission from ref 53. Copyright 2016 Optical Society of America. https://creativecommons.org/licenses/by/4.0/. (b) Optical micrograph illustration of the experimentally demonstrated optical interference unit, realizing both matrix multiplication (red highlighted) and attenuation (blue highlighted) fully optically. Reprinted with permission from ref 54. Copyright 2017 Springer Nature. (c) Reconfigurable optical signal processing circuits capable dynamically configuring functions and schematic of the experimental setup. Reprinted with permission from ref 55. Copyright 2020 American Chemical Society. (d) Schematic diagram of the all-optical transcendental equation solver based on the convolutional neural network. Reprinted with permission from ref 56. Copyright 2021 Chinese Academy of Sciences. https://creativecommons.org/licenses/by/4.0/. (e) Optical micrograph of the locally connected IPNNs with nine input ports (i_1−i_9) and four output ports (o_1−o_4). Reprinted with permission from ref 57. Copyright 2022 De Gruyter. https://creativecommons.org/licenses/by/4.0/.

conventional photonic chips following the von Neumann architecture. Thus, conventional photonic computing is more suitable for processing computing tasks with specific functions.[38,39] Compared with the conventional photonic computing scheme, the emerging photonic computing framework, the neural network, which came into being under the non-von Neumann framework has the characteristics of processing, in memory, multiple functions and the ability to perform intelligent inference tasks.[30,40,41] Many schemes for optical neural networks (ONNs) have been explored in the past decades. Before the rising trend of integrated waveguides made possible by advances in complementary metal oxide semiconductor (CMOS) fabrication,[42] ONNs usually take the form of free-space diffractions.[43−46] The core component for reconfigurable free space diffractive ONNs is spatial light modulator (SLM),[47] which operates as the programmable input light controller for the neural network. For free-space ONNs, crosstalk noise is normally negligible, thus enabling analogue computation,[48,49] which gives birth to the earliest steps of optical computing, including optical matrix multipliers and optical diffractive neural networks.[50−52] However, such implementations are not viable for deployment due to the requirement for SLM and free space.

Integrated photonic neural networks (IPNNs) are much more compact for high integration, yet face their own problems as well, including a weak optical nonlinearity, the lack of suitable configurations for photonic hardware, and proper optical storage mediums. In this Perspective, we highlight the recent advancements of IPNNs in China from aspects of different proposed configurations for photonic hardware, implementations of on-chip optical nonlinear activations, and training methods of IPNNs, and we also discuss some next challenges still ahead to realize high performance IPNNs with strong computing power.

CONFIGURATIONS FOR IPNNS

IPNNs display the possibility of miniaturizing and integrating with current computational devices. The on-chip integrated photonic platform provides promising configurations available for computation with great scalability, high computational density, and easy fabrication. Although simple structures such as fixed couplers are often explored on a larger scale, they are usually of lower density and are hard to expand to adapt to conventional requirements. Thus, more compact, novel structures suitable for photonic hardware are being actively explored. Here, we focus on several representative configurations developed in recent years.

On-Chip Waveguide-Based IPNNs. On-chip waveguides are the core structures of integrated photonics. The simplest way to utilize waveguides for computation is by using fixed couplers as Mach−Zehnder interferometers (MZIs) and phase shifters with a strong electro-optic effect/thermo-optic effect/all-optical nonlinear effect. The combination of two arms of the coupler with different phase modulations converts phase modulation into intensity modulation, enabling optical computation equivalent to complex addition and multi-

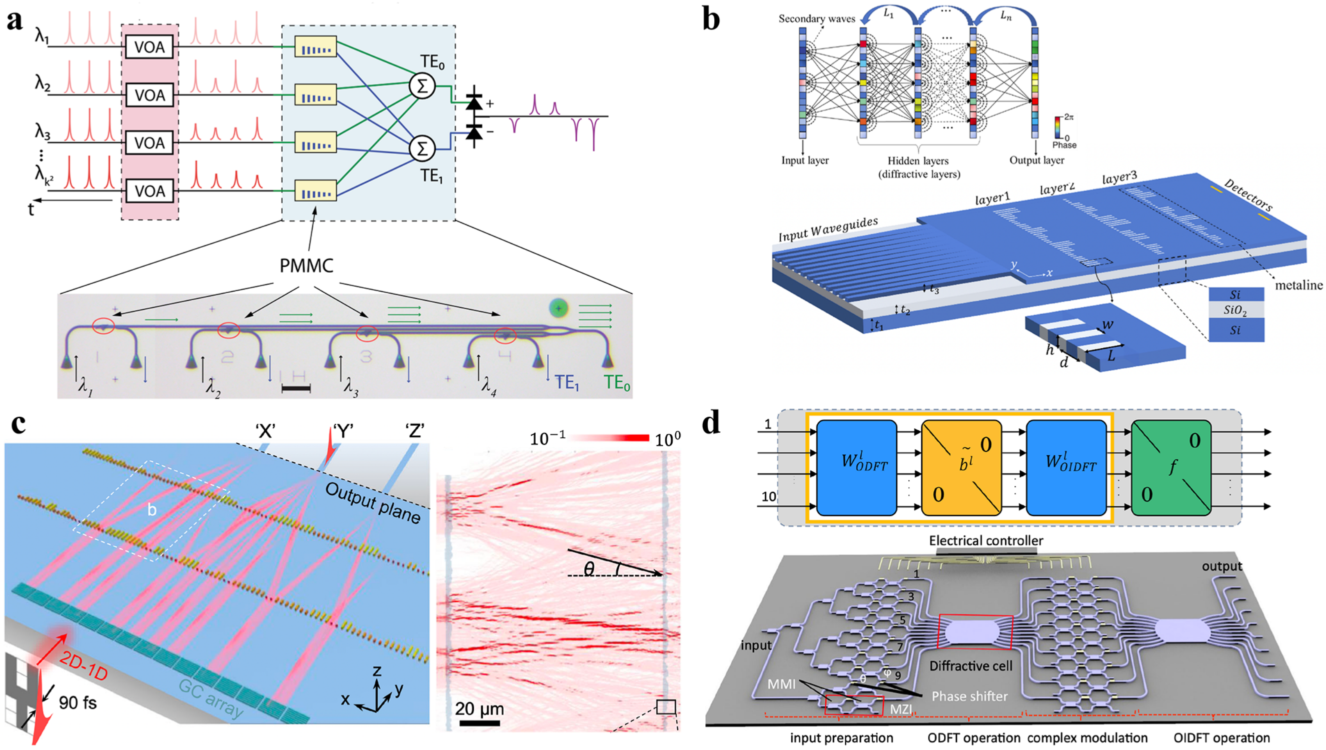


Figure 2. On-chip dielectric metasurface-based IPNNs. (a) Schematic of optical convolution for image processing. A patch of pixels of an image is encoded as optical pulses; the output in TE$_0$ and TE$_1$ are summed incoherently and measured with photodetectors (top). Optical microscope image of the photonic core consisting of four PMMCs with four input channels (bottom). Reprinted with permission from ref 76. Copyright 2021 Springer Nature. https://creativecommons.org/licenses/by/4.0/. (b) Schematic of diffractive ONN: each point on a given layer acting as a secondary source of a wave (top). Each of the three slots on a given layer act as a neuron with a complex-valued transmission coefficient (bottom). Reprinted with permission from ref 77. Copyright 2021 Optical Society of America. https://creativecommons.org/licenses/by/4.0/. (c) Schematics of the 1D multilayer diffractive metasurface system. As an example, a pattern "Y" with less than 90 fs pulse duration is coupled to the integrated metasystem through the input grating coupler array (left). Subwavelength structure manifested diffraction between metasurface layers (right). Reprinted with permission from ref 78. Copyright 2022 Springer Nature. https://creativecommons.org/licenses/by/4.0/. (d) Schematic of the IDNN operates on complex-valued inputs using coherent light (top). Schematic of the experimental device with four functional parts: input signal preparation; implementing ODFT operation; modulating amplitude/phase in the Fourier domain; implementing OIDFT operation (bottom). Reprinted with permission from ref 79. Copyright 2022 Springer Nature. https://creativecommons.org/licenses/by/4.0/.

plication. Different from neural network in electronic computers, where the weight values and intermediate results are not bounded, the optical signals propagating through waveguides must follow the rules of energy conservation, putting additional limits on the function of the IPNNs.

In 2016, Zhou et al. proposed an optical switch with a large number of channels,[53] as shown in Figure 1a. which functions with the equivalence of a neural network in binary values. Yet this device may not fully take advantage of the low noise of optical waveguides, as it supports more precision than binary signals. In 2017, Soljačić et al. proposed a new architecture for a fully optical feed-forward neural network,[54] as shown in Figure 1b, which contains 56 programmable MZIs using analog light signals for computation. The authors demonstrated its utility for vowel recognition with a computational ability equivalent to a 4 × 4 matrix multiplication. However, available space on the device for holding the controlling electrodes limits further scaling of the device channels. Besides, the weight of the network was trained via computer simulation, then configured by manual operation, which is time-consuming and becomes a serious challenge when the network is expanded to a larger scale. Note that they also discussed the feasibility of on-chip optimization with the gradient of weights based on the forward propagation. Based on a similar structure, Dong et al. proposed a silicon photonic signal processor in 2020, which also functions as the linear layer for a neural network, with capabilities of dynamically configuring the functions for reconfigurable processors, as shown in Figure 1c. The proposed structure contains 48 phase shifters and 20 MZIs, demonstrated with three functions, including multi-channel optical switching, optical multiple-input-multiple-output descrambler, and a tunable optical filter.[55] Such networks demonstrate the possibility of IPNNs to meet the requirements of simple computing tasks, but the most widely deployed artificial neural networks are in computer vision tasks, where a substantially larger number of channels are required. In 2021, Hu et al. proposed an advanced photonic neural network with 27 channels and a structural resemblance of the convolutional neural network,[56] as demonstrated in Figure 1d. This scheme can extend the functions to realize transcendental equation solvers, multifunctional logic gate operators, and half-adders, with better scalability capabilities. However, although this configuration contains three layers, it is not an orthogonal representation of the neural network, as there is a lack of nonlinear activation function. Details about integrated optical activation functions will be further discussed in the later section. But it is worth noting that in 2022, Hu et al. demonstrated a five-layer integrated locally connected

https://doi.org/10.1021/acsphotonics.2c01516
ACS Photonics 2023, 10, 2001−2010

photonic neural network with an all-optical saturated absorption nonlinear activation layer,[57] as shown in Figure 1e. This network contains an activation layer after the second linear layer, allowing it to deal with complex tasks such as solving for the eigenvalues of a matrix.

As for now, however, even the best optimized MZI structures are hard to directly scale up for conventional usage. If we count one MZI to match a single float point operation, a typical neural network would take at least 1 billion to be functional,[58,59] which is not applicable with any type of MZI design due to the amount of phase shifters used to control the MZIs. In addition, conventional neural networks usually contain dozens of layers.[60−63] as for the scheme of all-optical activation, strong optical nonlinearity results in strong energy loss, leading to almost no output with the same depth in an IPNN with current structural patterns, unless some external optical pumping is applied to the system.

On-Chip Dielectric Metasurface-Based IPNNs. Large-scale, highly integrated, low-power hardware is becoming increasingly important for realizing IPNNs with advanced optical computing capabilities. Traditional MZIs have limited scalability and consume excessive power. Here, the optical metasurfaces offer new possibilities for realizing light−matter interactions by compressing the light fields at subwavelength scales.[64−67] First, it is worth noting that functional optical computing components based on metasurfaces have been reported, such as computational imaging[68] and compact logic operator.[69] Then, it is also of great research significance to use metasurfaces to replace the components of traditional chips for computing.[70,71] Moreover, machine learning may enrich the metasurface devices by exploiting machine-learning techniques for improving optical system design and hardware.[72,73] Optical on-chip dielectric metasurfaces, as artificially designed electromagnetic interfaces,[74,75] can manipulate the degree of freedom of on-chip optical signal transmission, and they have the potential to realize on-chip integrated photonic computing with a compact footprint, broadband, and low loss. In recent years, a series of research work on PNNs based on integrated metasurface platforms has been gradually reported, with the feature structures interpreted as neurons with dense coupling, with the ability to stack multiple computations in limited footprints.

In 2021, Li et al. demonstrated a multimode photonic computing core,[76] which is a metasurface programmable waveguide mode converter based on phase-change material $Ge_2Sb_2Te_5$ (GST) to achieve phase gradient control. This phase-change metasurface mode converter (PMMC) exploits the large refractive index change of the GST during the phase transition to control the switching of the two spatial modes of the waveguide, TE_0 and TE_1 modes. A 2 × 2 PMMC array (Figure 2a) was built and used as a programmable kernel to implement a multimodal optical convolutional neural network. It can perform image processing and recognition tasks with high precision. This multimode photonic core, with a wide operating bandwidth and a compact device footprint, holds promise for large-scale PNNs with ultrahigh computational throughput. In 2021, Cheng et al. proposed a PNN composed of 1D multilayer diffractive metasurfaces,[77] as shown in Figure 2b. They considered the difference in the effective refractive index of the same slit at different positions for light input at different angles and the mutual interference between adjacent grooves of different lengths for the input light at the same angle. Thus, three silicon slots were used as a single neuron, thereby minimizing the impact of the above two factors and making the phase delay induced by the silicon slot more accurately approximate the value of the pretrained neuron. This scheme has a simple structure design, and the manufacturing process is compatible with the CMOS process, which is convenient for large-scale and low-cost manufacturing. Similarly, in 2022, Gu et al. experimentally fabricated and demonstrated this 1D multilayer diffractive metasurface system,[78] which was fabricated on SOI substrates by one-step lithography and etching (Figure 2c). The demonstrated spatial pattern classifier in a two-layer metasystem achieves accuracies of 92% and 89% under narrowband continuous wave excitation and broadband femtosecond pulse excitation, respectively. The proposed diffractive element system provides an alternative machine learning architecture for photonic integrated circuits. In 2022, Liu et al. demonstrated a silicon-based scalable integrated diffractive neural network consisting of two ultracompact diffractive units and N MZIs for parallel Fourier transform, convolution operations, and optical calculations for specific applications,[79] as shown in Figure 2d. The adopted on-chip compact diffractive unit enables the footprint and power consumption of the proposed architecture to be reduced from quadratic scaling to linear scaling of the input data dimension required for MZI-based PNN architectures. This solution contributes to the realization of energy-efficient, ultracompact, and large-scale integrated optical computing chips.

Given that on-chip dielectric metasurfaces as a unique way to manipulate the degrees of freedom of on-chip light fields demonstrate efficient, compact optical connections for IPNNs and are compatible with on-chip CMOS fabrication processes, they provide a feasible way for the efficient implementation of neural networks on integrated photonic platforms.

Integrated Photonic Spiking Neural Networks. Spiking neural networks (SNN),[80,81] which compute with light pulses instead of steady beams, tell a different story from the network configurations mentioned above. Such networks do not require a large array of unit devices, as signals can propagate through the structure multiple times and are a closer resemblance to how neurons generate electrical pulses based on membrane voltage.

The spiking neuron has several typical characteristics: (i) The spiking neuron only emits an output pulse when the power of the perturbation signal exceeds a certain threshold. (ii) The spiking neuron can be excited with several sufficiently closely spaced subthreshold pulses by integrating their combined perturbation. (iii) After an excitation, the neuron needs to relax to its steady-state before it can be triggered again. (iv) The output spike of the previous neuron should be strong enough to excite the next neuron, which is the foundation of forming a multilayer SNN. (v) The arrival of a stimulus can stop the spiking activity, which plays an important role in the spike-timing-dependent plasticity (STDP) learning rules.[82] Photonic synapses should meet these requirements to be able to construct a photonic spiking neural network.

In 2021, Wang et al. proposed a mechano-photonic artificial synapse implemented with a graphene/MoS_2 heterostructure and an integrated triboelectric nanogenerator.[83] This artificial synapse is capable of generating spiking patterns by controlling the charge transfer/exchange between the graphene and the MoS_2 layer. Moreover, other features, such as postsynaptic photocurrents, persistent photoconductivity, and photosensitivity, are controlled by the triboelectric potential. In 2021,

https://doi.org/10.1021/acsphotonics.2c01516
ACS Photonics 2023, 10, 2001−2010

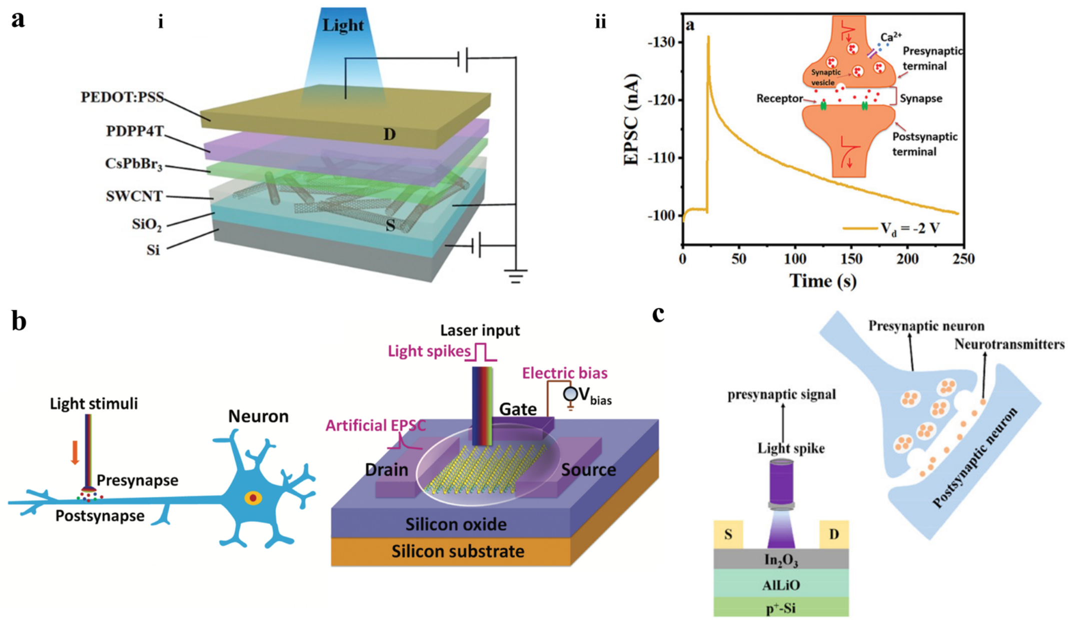


Figure 3. Integrated photonic spiking neural networks. (a) (i) 3D schematic of the vertical organic field-effect transistors device, where PDPP4T and $CsPbBr_3$ QDs are sandwiched between the source-drain electrodes. (ii) EPSC (excitatory postsynaptic current) triggered by a presynaptic light spike. The inset shows a schematic diagram of the EPSC generation process in biological synapses. Reprinted with permission from ref 84. Copyright 2021 John Wiley and Sons. (b) Schematic image of a neuron subject to light stimuli (left) and a schematic of the photonic MoS_2 synapse (right). Reprinted with permission from ref 85. Copyright 2019 Royal Society of Chemistry. (c) Schematic diagram of the device and synapse, where the UV light spike is considered as the presynaptic signal. Reprinted with permission from ref 86. Copyright 2020 American Chemical Society.

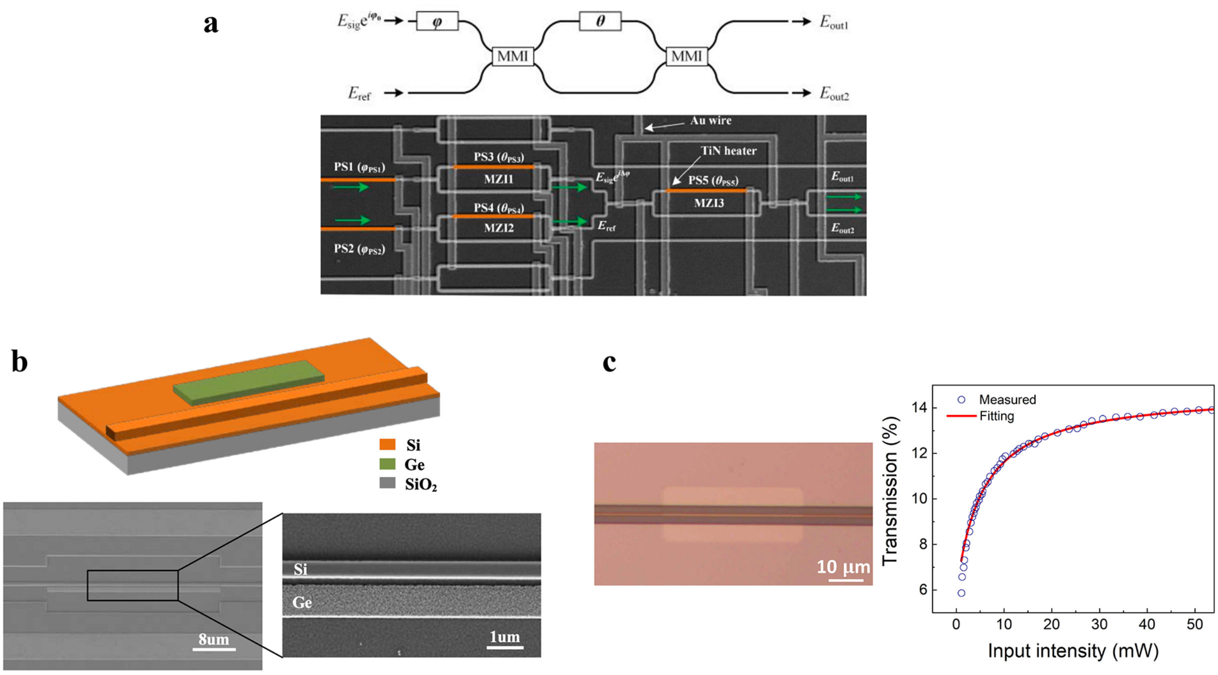


Figure 4. Implementations of on-chip optical nonlinear activations. (a) Optical nonlinear circuit based on a programmable MZI (top). Micrograph of the on-chip circuit (bottom). Reprinted with permission from ref 94. Copyright 2020 The Authors. (b) The schematic diagram of the optical activation function device (top). SEM image of fabricated device (bottom). Reprinted with permission from ref 96. Copyright 2022 IEEE. (c) The optical nonlinear activation layer of single-layer graphene covered on the silicon waveguide with the length of 40 μm: the optical micrograph of the optical nonlinear activation cell (left). The transmission curve with fitting parameters of saturated absorption effect of α_S = 0.088, α_{NS} = 0.852, and I_s = 5.446. Reprinted with permission from ref 57. Copyright 2022 De Gruyter. https://creativecommons.org/licenses/by/4.0/.

Huang et al. proposed a type of optical synapse based on vertical organic field-effect transistors,[84] as shown in Figure 3a. This scheme shows superior low-voltage capabilities, allowing for an obvious response with 1.3 fJ per spike. Another kind of

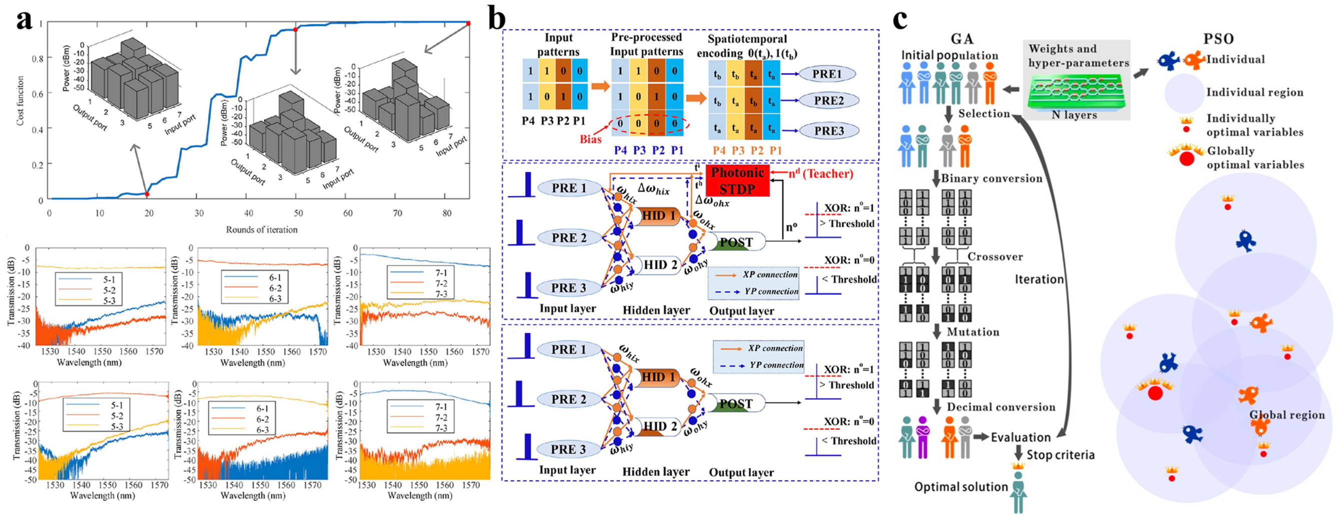


Figure 5. Training methods of IPNNs. (a) Cost function dependent on the rounds of iteration in training process (top). Transmission spectra for the signal and noise of the optical links in the shown routing states (bottom). Reprinted with permission from ref 55. Copyright 2020 American Chemical Society. (b) The architecture of the proposed multilayer photonic SNN for solving the XOR task. Reprinted with permission from ref 97. Copyright 2019 Optical Society of America. https://creativecommons.org/licenses/by/4.0/. (c) The flowcharts of the learning algorithms for the ONNs based on GA and PSO algorithms. Reprinted with permission from ref 98. Copyright 2021 IEEE.

optical synapsis is integrated with electrical ones, where in 2019, Wan et al. proposed a photoelectronic hybrid-integrated synaptic device based on a 2D MoS_2 phototransistor gated by an electric double-layer biopolymer electrolyte (Figure 3b).[85] This synapse is functional with both the synaptic potentiation and depression effects, enabling a dynamic transition from a high-pass to a low-pass photonic filter, showing flexibility for computation. Although not entirely required, a better function for the artificial optical synapse would require a shift between short-term and long-term memories, which in 2020, Zhang et al. proposed an all-inorganic light-stimulated In_2O_3/AlLiO thin-film transistor,[86] as shown in Figure 3c. This artificial synapse achieved a simulation of the associative learning behaviors under light stimulation.

However, with promising abilities, none of these spiking synapses shown above are incorporated in an actual neural network outside of simulation, and experimental testing of the designed network configuration is yet to be demonstrated. In summary, the proposed spiking synapses are in a much more primary state and would require extensive research to catch up with other approaches.

IMPLEMENTATIONS OF ON-CHIP OPTICAL NONLINEAR ACTIVATIONS

The implementations of IPNNs mentioned above are purely linear in nature, which limits the computational power of such devices to be a single matrix multiplication from inputs to outputs. On electronic computers, a nonlinear activation function is usually applied after each linear layer,[87,88] similar to neurons.[89−91] For deeper structures, nonlinear activation is vital for higher performance.[92,93] In recent years, some nonlinear elements have been proposed for nonlinear activation layers of IPNNs, including thermo-optic modulation schemes, electro-optic modulation schemes, and all-optical nonlinear modulation schemes.

In 2020, Tian et al. proposed and experimentally demonstrated a thermo-optic modulated reconfigurable MZI-based on-chip optical nonlinear unit for ONN,[94] as shown in Figure 4a, which does not require optical-electrical conversion or high input power. This nonlinear unit can be reconfigured to perform multiple types of nonlinear activation functions, making it one of the candidate scheme for future ONNs. In 2022, Chu et al. proposed a mechanism for programmable low-threshold optically controlled nonlinear activation function based on integrated on-chip Ge−Si photodetectors and silicon photoswitches.[95] They experimentally change of the slope of the nonlinear activation function by adjusting the photodetector bias voltage, validating the generation of linear unit (ReLU) and sigmoid functions at a low threshold power of 0.2 mW. They claim that these devices are suitable for low-power ONNs. In 2022, Dong's group proposed an all-optical realization of a nonlinear activation function based on the hybrid integration of germanium (Ge) and silicon by using the plasmonic dispersion effect and carrier absorption effect of Ge at 1550 nm (Figure 4b).[96] This scheme has a large operating bandwidth with the response frequency of 70 MHz, a low loss of 4.28 dB, and a low threshold power of 5.1 mW. In 2022, Hu's group proposed and experimentally verified the performance of a single-layer graphene with optical nonlinear saturable absorption effect covered on a silicon waveguide as a nonlinear activation layer in an on-chip integrated locally connected photonic neural network (Figure 4c).[57] Compared with the configuration without the proposed nonlinear layer, there is about 30% improvement in accuracy in the task of solving the eigenvalues of second-order real symmetric matrices. This work lays the foundation for a new generation of intelligent on-chip integrated all-optical computing.

In general, the successive advent of various types of activation functions with reconfigurable, low power consumption, high bandwidth, and low loss provide a solid foundation for the further development of on-chip IPNNs with powerful computing power.

TRAINING METHODS OF IPNNS

For IPNNs with steady optical inputs and outputs, the network is usually trained with gradient descending methods. For

https://doi.org/10.1021/acsphotonics.2c01516
ACS Photonics 2023, 10, 2001−2010

neural networks in computers, given the relationship between network weights and outputs, it is easy to obtain the gradient of weights used for minimizing the loss value. Using such gradient values, by applying various gradient descent methods, the network weights are adjusted step by step toward higher performance.[88] There are two approaches in obtaining the gradient value. The first approach is to use numerical methods to simulate the IPNN, and compute the gradient via the simulation. With this method, the simulated network is trained on a computer, with the optimized network weights applied on the physical network after training. The second approach is to obtain the gradient by experimentally measuring the difference in network output after modifying the network weight by a small value, using this to conduct on-chip training.[54]

In 2020, Dong et al. proposed and experimentally verified a self-configuring and reconfigurable silicon photonic signal processor. All functions are programmed through self-configuration without any internal structure information by using the deep learning improved numerical gradient descent algorithm,[55] as shown in Figure 5a. For IPNNs that work in a spiking pattern, an entirely different approach for training is used, namely, the spike timing dependent plasticity method, which optimizes the spiking neurons based on the past record of how often their activities coincide. In 2019, Xiang et al. applied this method in training simulated spiking ONNs built by using a vertical-cavity, surface-emitting laser with an embedded saturable absorber,[97] as shown in Figure 5b. This algorithm is tested on the classical XOR task, with only two inputs and one output. Another method that is available for training IPNNs is neuroevolution, derived from the evolution process that imitates the biological brain, which generates a group of networks, picks out the better networks, and uses them to generate the next group. In 2021, Xu et al. proposed neuroevolution methods for training IPNNs, as shown in Figure 5c. They showed the calculated results support the fact that these training algorithms are competitive with gradient-based learning algorithms.[98]

■ CONCLUSION AND OUTLOOK

In this Perspective, we have highlighted the recent advancements of IPNNs from several aspects which have been continuously concerned. Despite the promising potentials in high speed and high energy efficiency computation, IPNNs still face challenges concerning the scalability to meet requirements for real-world tasks. (1) As the main configuration, on-chip MZIs are difficult to scale up to computer vision and other tasks that require a large number of inputs. (2) On-chip metasurface-based IPNNs come with much more compact optical connections and efficient computation, while it is hard to optimize the specified high-dimensional transfer matrix and hard to achieve reconfiguration of weight parameters for all degrees of freedom. (3) Spiking networks with neuromorphic designs demonstrate excellent results by using temporal encoding of input meeting with these challenges. However, they are in a much more primary state and still need sufficient experimental demonstration of the full network configuration. (4) As for the on-chip optical nonlinear activation functions, various types of activation functions have been successively reported. In order to achieve an effective increase in computing power, it is necessary to continue to develop optical activations with reconfigurability, low power consumption, high bandwidth, and low loss. Especially, as for the all-optical activations utilizing the third-order nonlinear effect of materials, the dissipation of energy indicates that a full optical activation will run out of energy after a rather shallow depth. The solution to this lies in either some way to refill the energy or to develop optical nonlinearity with an independence to optical intensity. (5) In addition, the training of IPNNs still relies on simulation, but a stable and efficient scheme of on-chip training is yet to be realized. The efficient combination of novel configurations specified for the integrated photonic platform with light-field-driven nonvolatile materials will hopefully address this challenge.

■ AUTHOR INFORMATION

Corresponding Author

Xiaoyong Hu − *State Key Laboratory for Mesoscopic Physics and Department of Physics, Collaborative Innovation Center of Quantum Matter, Beijing Academy of Quantum Information Sciences, Nano-optoelectronics Frontier Center of Ministry of Education, Peking University, Beijing 100871, China; Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China; Peking University Yangtze Delta Institute of Optoelectronics, Nantong, Jiangsu 226010, China*; orcid.org/0000-0002-1545-1491; Email: xiaoyonghu@pku.edu.cn

Authors

Kun Liao − *State Key Laboratory for Mesoscopic Physics and Department of Physics, Collaborative Innovation Center of Quantum Matter, Beijing Academy of Quantum Information Sciences, Nano-optoelectronics Frontier Center of Ministry of Education, Peking University, Beijing 100871, China*; orcid.org/0000-0003-2872-6359

Tianxiang Dai − *State Key Laboratory for Mesoscopic Physics and Department of Physics, Collaborative Innovation Center of Quantum Matter, Beijing Academy of Quantum Information Sciences, Nano-optoelectronics Frontier Center of Ministry of Education, Peking University, Beijing 100871, China*; orcid.org/0000-0002-9403-7511

Qiuchen Yan − *State Key Laboratory for Mesoscopic Physics and Department of Physics, Collaborative Innovation Center of Quantum Matter, Beijing Academy of Quantum Information Sciences, Nano-optoelectronics Frontier Center of Ministry of Education, Peking University, Beijing 100871, China*

Qihuang Gong − *State Key Laboratory for Mesoscopic Physics and Department of Physics, Collaborative Innovation Center of Quantum Matter, Beijing Academy of Quantum Information Sciences, Nano-optoelectronics Frontier Center of Ministry of Education, Peking University, Beijing 100871, China; Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China; Peking University Yangtze Delta Institute of Optoelectronics, Nantong, Jiangsu 226010, China*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsphotonics.2c01516

Funding

This work was supported by the National Key Research and Development Program of China (2018YFB2200403) and the National Natural Science Foundation of China (NSFC) (91950204, 92150302).

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Jain, A. K.; Mao, J.; Mohiuddin, K. M. Artificial neural networks: A tutorial. *Computer* **1996**, *29* (3), 31−44.
(2) Basheer, I. A.; Hajmeer, M. Artificial neural networks: Fundamentals, computing, design, and application. *J. Microbiol. Methods* **2000**, *43* (1), 3−31.
(3) Lippmann, R. P. Review of neural networks for speech recognition. *Neural Comput.* **1989**, *1* (1), 1−38.
(4) Nassif, A. B.; Shahin, I.; Attili, I.; Azzeh, M.; Shaalan, K. Speech recognition using deep neural networks: A systematic review. *IEEE Access* **2019**, *7*, 19143−19165.
(5) Egmont-Petersen, M.; de Ridder, D.; Handels, H. Image processing with neural networks-a review. *Pattern Recogn.* **2002**, *35* (10), 2279−2301.
(6) Rawat, W.; Wang, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.* **2017**, *29* (9), 2352−2449.
(7) Zhou, Y.-T.; Chellappa, R. *Artificial Neural Networks for Computer Vision*; Springer Science & Business Media, 2012; Vol. 5.
(8) Srinivas, S.; Sarvadevabhatla, R. K.; Mopuri, K. R.; Prabhu, N.; Kruthiventi, S. S.; Babu, R. V. A taxonomy of deep convolutional neural nets for computer vision. *Front. Robot. AI* **2016**, *2*, 36.
(9) Goldberg, Y. A primer on neural network models for natural language processing. *J. Artif. Intell. Res.* **2016**, *57*, 345−420.
(10) Li, H. Deep learning for natural language processing: Advantages and challenges. *Natl. Sci. Rev.* **2018**, *5* (1), 24−26.
(11) Lundstrom, M. Moore's law forever? *Science* **2003**, *299* (5604), 210−211.
(12) Thompson, S. E.; Parthasarathy, S. Moore's law: The future of si microelectronics. *Mater. Today* **2006**, *9* (6), 20−25.
(13) Cavin, R. K., III; Lugli, P.; Zhirnov, V. V. Science and engineering beyond moore's law. *Proc. IEEE* **2012**, *100*, 1720−1749.
(14) Waldrop, M. M. The semiconductor industry will soon abandon its pursuit of moore's law. Now things could get a lot more interesting. *Nature* **2016**, *530* (7589), 144−147.
(15) Ghahramani, Z. Probabilistic machine learning and artificial intelligence. *Nature* **2015**, *521* (7553), 452−459.
(16) Barredo Arrieta, A.; Diaz-Rodriguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; Chatila, R.; Herrera, F. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inform. Fusion* **2020**, *58*, 82−115.
(17) Glover, F. Future paths for integer programming and links to artificial-intelligence. *Comput. Oper. Res.* **1986**, *13* (5), 533−549.
(18) Kaytez, F.; Taplamacioglu, M. C.; Cam, E.; Hardalac, F. Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *Int. J. Elec. Power* **2015**, *67*, 431−438.
(19) Rahman, A.; Srikumar, V.; Smith, A. D. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Appl. Energy* **2018**, *212*, 372−385.
(20) Kim, T.-Y.; Cho, S.-B. Predicting residential energy consumption using CNN-LSTM neural networks. *Energy* **2019**, *182*, 72−81.
(21) Ahmad, M. W.; Mourshed, M.; Rezgui, Y. Trees vs neurons: Comparison between random forest and ann for high-resolution prediction of building energy consumption. *Energy Buildings* **2017**, *147*, 77−89.
(22) Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J. D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A. Language models are few-shot learners. *NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems*; ACM, 2020; Article No. 159, pp 1877−1901.
(23) Sarle, W. S. Neural networks and statistical models. *Proceedings of the 19th Annual SAS Users Groups International Conference*, New York, Analytical System Institute, April 1994, SAS, 1994; pp 1538−1550.
(24) Schwabe, R.; Zelinger, S.; Key, T.; Phipps, K. Electronic lighting interference. *IEEE Ind. Appl. Mag.* **1998**, *4* (4), 43−48.
(25) Salahuddin, S.; Ni, K.; Datta, S. The era of hyper-scaling in electronics. *Nat. Electron.* **2018**, *1* (8), 442−450.
(26) Von Bieren, K. Lens design for optical fourier transform systems. *Appl. Opt.* **1971**, *10* (12), 2739−2742.
(27) Xu, X.; Tan, M.; Corcoran, B.; Wu, J.; Boes, A.; Nguyen, T. G.; Chu, S. T.; Little, B. E.; Hicks, D. G.; Morandotti, R.; et al. 11 TOPS photonic convolutional accelerator for optical neural networks. *Nature* **2021**, *589* (7840), 44−51.
(28) Willner, A. E.; Khaleghi, S.; Chitgarha, M. R.; Yilmaz, O. F. All-optical signal processing. *J. Lightwave. Technol.* **2014**, *32* (4), 660−680.
(29) Shu, H.; Chang, L.; Tao, Y.; Shen, B.; Xie, W.; Jin, M.; Netherton, A.; Tao, Z.; Zhang, X.; Chen, R.; Bai, B.; Qin, J.; Yu, S.; Wang, X.; Bowers, J. E. Microcomb-driven silicon photonic systems. *Nature* **2022**, *605* (7910), 457−463.
(30) Feldmann, J.; Youngblood, N.; Karpov, M.; Gehring, H.; Li, X.; Stappers, M.; Le Gallo, M.; Fu, X.; Lukashchuk, A.; Raja, A. S.; Liu, J.; Wright, C. D.; Sebastian, A.; Kippenberg, T. J.; Pernice, W. H. P.; Bhaskaran, H. Parallel convolutional processing using an integrated photonic tensor core. *Nature* **2021**, *589* (7840), 52−58.
(31) Wang, J.; Yang, J.-Y.; Fazal, I. M.; Ahmed, N.; Yan, Y.; Huang, H.; Ren, Y.; Yue, Y.; Dolinar, S.; Tur, M.; Willner, A. E. Terabit free-space data transmission employing orbital angular momentum multiplexing. *Nat. Photonics* **2012**, *6* (7), 488−496.
(32) van Uden, R. G. H.; Correa, R. A.; Lopez, E. A.; Huijskens, F. M.; Xia, C.; Li, G.; Schülzgen, A.; de Waardt, H.; Koonen, A. M. J.; Okonkwo, C. M. Ultra-high-density spatial division multiplexing with a few-mode multicore fibre. *Nat. Photonics* **2014**, *8* (11), 865−870.
(33) Ouyang, X.; Xu, Y.; Xian, M.; Feng, Z.; Zhu, L.; Cao, Y.; Lan, S.; Guan, B.-O.; Qiu, C.-W.; Gu, M.; Li, X. Synthetic helical dichroism for six-dimensional optical orbital angular momentum multiplexing. *Nat. Photonics* **2021**, *15* (12), 901−907.
(34) Sun, C.; Wade, M. T.; Lee, Y.; Orcutt, J. S.; Alloatti, L.; Georgas, M. S.; Waterman, A. S.; Shainline, J. M.; Avizienis, R. R.; Lin, S.; Moss, B. R.; Kumar, R.; Pavanello, F.; Atabaki, A. H.; Cook, H. M.; Ou, A. J.; Leu, J. C.; Chen, Y.-H.; Asanović, K.; Ram, R. J.; Popović, M. A.; Stojanović, V. M. Single-chip microprocessor that communicates directly using light. *Nature* **2015**, *528* (7583), 534−538.
(35) Atabaki, A. H.; Moazeni, S.; Pavanello, F.; Gevorgyan, H.; Notaros, J.; Alloatti, L.; Wade, M. T.; Sun, C.; Kruger, S. A.; Meng, H.; Al Qubaisi, K.; Wang, I.; Zhang, B.; Khilo, A.; Baiocco, C. V.; Popović, M. A.; Stojanović, V. M.; Ram, R. J. Integrating photonics with silicon nanoelectronics for the next generation of systems on a chip. *Nature* **2018**, *556* (7701), 349−354.
(36) Marpaung, D.; Yao, J.; Capmany, J. Integrated microwave photonics. *Nat. Photonics* **2019**, *13* (2), 80−90.
(37) Wan, N. H.; Lu, T.-J.; Chen, K. C.; Walsh, M. P.; Trusheim, M. E.; De Santis, L.; Bersin, E. A.; Harris, I. B.; Mouradian, S. L.; Christen, I. R.; Bielejec, E. S.; Englund, D. Large-scale integration of artificial atoms in hybrid photonic circuits. *Nature* **2020**, *583* (7815), 226−231.
(38) Liu, W.; Li, M.; Guzzon, R. S.; Norberg, E. J.; Parker, J. S.; Lu, M.; Coldren, L. A.; Yao, J. A fully reconfigurable photonic integrated signal processor. *Nat. Photonics* **2016**, *10* (3), 190−195.
(39) Ying, Z.; Feng, C.; Zhao, Z.; Dhar, S.; Dalir, H.; Gu, J.; Cheng, Y.; Soref, R.; Pan, D. Z.; Chen, R. T. Electronic-photonic arithmetic logic unit for high-speed computing. *Nat. Commun.* **2020**, *11* (1), 2154.
(40) Feldmann, J.; Youngblood, N.; Wright, C. D.; Bhaskaran, H.; Pernice, W. H. P. All-optical spiking neurosynaptic networks with self-learning capabilities. *Nature* **2019**, *569* (7755), 208−214.
(41) Ashtiani, F.; Geers, A. J.; Aflatouni, F. An on-chip photonic deep neural network for image classification. *Nature* **2022**, *606* (7914), 501−506.
(42) Doylend, J. K.; Knights, A. P. The evolution of silicon photonics as an enabling technology for optical interconnection. *Laser Photonics Rev.* **2012**, *6* (4), 504−525.

(43) Lin, X.; Rivenson, Y.; Yardimci, N. T.; Veli, M.; Luo, Y.; Jarrahi, M.; Ozcan, A. All-optical machine learning using diffractive deep neural networks. *Science* **2018**, *361* (6406), 1004−1008.

(44) Zhou, T.; Lin, X.; Wu, J.; Chen, Y.; Xie, H.; Li, Y.; Fan, J.; Wu, H.; Fang, L.; Dai, Q. Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit. *Nat. Photonics* **2021**, *15* (5), 367−373.

(45) Qian, C.; Lin, X.; Lin, X.; Xu, J.; Sun, Y.; Li, E.; Zhang, B.; Chen, H. Performing optical logic operations by a diffractive neural network. *Light: Sci. Appl.* **2020**, *9* (1), 59.

(46) Liu, C.; Ma, Q.; Luo, Z. J.; Hong, Q. R.; Xiao, Q.; Zhang, H. C.; Miao, L.; Yu, W. M.; Cheng, Q.; Li, L.; Cui, T. J. A programmable diffractive deep neural network based on a digital-coding metasurface array. *Nat. Electron.* **2022**, *5* (2), 113−122.

(47) Armitage, D.; Anderson, W.; Karr, T. High-speed spatial light modulator. *IEEE J. Quantum Electron.* **1985**, *21* (8), 1241−1248.

(48) Tsai, F.-C. F.; O'Brien, C. J.; Petrović, N. S.; Rakić, A. D. Analysis of optical channel cross talk for free-space optical interconnects in the presence of higher-order transverse modes. *Appl. Opt.* **2005**, *44* (30), 6380−6387.

(49) Hu, W.; Li, X.; Yang, J.; Kong, D. Crosstalk analysis of aligned and misaligned free-space optical interconnect systems. *J. Opt. Soc. Am. A* **2010**, *27* (2), 200−205.

(50) Goodman, J. W.; Dias, A.; Woody, L. Fully parallel, high-speed incoherent optical method for performing discrete fourier transforms. *Opt. Lett.* **1978**, *2* (1), 1−3.

(51) Mosca, E. P.; Griffin, R. D.; Pursel, F. P.; Lee, J. N. Acoustooptical matrix-vector product processor: Implementation issues. *Appl. Opt.* **1989**, *28* (18), 3843−3851.

(52) Sun, C.-C.; Chang, M.-W.; Hsu, K. Y. Matrix−matrix multiplication by using anisotropic self-diffraction in batio3. *Appl. Opt.* **1994**, *33* (20), 4501−4507.

(53) Lu, L.; Zhao, S.; Zhou, L.; Li, D.; Li, Z.; Wang, M.; Li, X.; Chen, J. 16 × 16 non-blocking silicon optical switch based on electro-optic mach-zehnder interferometers. *Opt. Express* **2016**, *24* (9), 9295−9307.

(54) Shen, Y.; Harris, N. C.; Skirlo, S.; Prabhu, M.; Baehr-Jones, T.; Hochberg, M.; Sun, X.; Zhao, S.; Larochelle, H.; Englund, D.; Soljačić, M. Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **2017**, *11* (7), 441−446.

(55) Zhou, H.; Zhao, Y.; Wang, X.; Gao, D.; Dong, J.; Zhang, X. Self-configuring and reconfigurable silicon photonic signal processor. *ACS Photonics* **2020**, *7* (3), 792−799.

(56) Liao, K.; Chen, Y.; Yu, Z.; Hu, X.; Wang, X.; Lu, C.; Lin, H.; Du, Q.; Hu, J.; Gong, Q. All-optical computing based on convolutional neural networks. *Opto-Electron. Adv.* **2021**, *4* (11), 200060.

(57) Liao, K.; Li, C.; Dai, T.; Zhong, C.; Lin, H.; Hu, X.; Gong, Q. Matrix eigenvalue solver based on reconfigurable photonic neural network. *Nanophotonics* **2022**, *11* (17), 4089−4099.

(58) Iandola, F. N.; Han, S.; Moskewicz, M. W.; Ashraf, K.; Dally, W. J.; Keutzer, K. Squeezenet: Alexnet-level accuracy with 50× fewer parameters and <0.5 MB model size. *arXiv:1602.07360 [cs.CV]* **2016**, na.

(59) Alom, M. Z.; Taha, T. M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, M. S.; Van Esesn, B. C.; Awwal, A. A. S.; Asari, V. K. The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv:1803.01164 [cs.CV]* **2018**, na.

(60) Targ, S.; Almeida, D.; Lyman, K. Resnet in resnet: Generalizing residual architectures. *arXiv:1603.08029 [cs.LG]* **2016**, na.

(61) Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. *arXiv:1512.00567 [cs.CV]* **2016**, na.

(62) Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning. *AAAI* **2017**, *31*, na.

(63) Wu, Z.; Shen, C.; Van Den Hengel, A. Wider or deeper: Revisiting the resnet model for visual recognition. *Pattern. Recognit.* **2019**, *90*, 119−133.

(64) Cheben, P.; Halir, R.; Schmid, J. H.; Atwater, H. A.; Smith, D. R. Subwavelength integrated photonics. *Nature* **2018**, *560* (7720), 565−572.

(65) Sun, S.; He, Q.; Xiao, S.; Xu, Q.; Li, X.; Zhou, L. Gradient-index meta-surfaces as a bridge linking propagating waves and surface waves. *Nat. Mater.* **2012**, *11* (5), 426−431.

(66) Xie, Y.; Wang, W.; Chen, H.; Konneker, A.; Popa, B.-I.; Cummer, S. A. Wavefront modulation and subwavelength diffractive acoustics with an acoustic metasurface. *Nat. Commun.* **2014**, *5* (1), 1−5.

(67) Luo, X. Subwavelength optical engineering with metasurface waves. *Adv. Opt. Mater.* **2018**, *6* (7), 1701201.

(68) Luo, Y.; Zhao, Y.; Li, J.; Çetintaş, E.; Rivenson, Y.; Jarrahi, M.; Ozcan, A. Computational imaging without a computer: Seeing through random diffusers at the speed of light. *eLight* **2022**, *2* (1), 4.

(69) Zhao, Z.; Wang, Y.; Ding, X.; Li, H.; Fu, J.; Zhang, K.; Burokur, S. N.; Wu, Q. Compact logic operator utilizing a single-layer metasurface. *Photonics Res.* **2022**, *10* (2), 316−322.

(70) Chen, Z.; Segev, M. Highlighting photonics: Looking into the next decade. *eLight* **2021**, *1* (1), 2.

(71) Luo, X.; Hu, Y.; Ou, X.; Li, X.; Lai, J.; Liu, N.; Cheng, X.; Pan, A.; Duan, H. Metasurface-enabled on-chip multiplexed diffractive neural networks in the visible. *Light: Sci. Appl.* **2022**, *11* (1), 158.

(72) Zhu, R.; Qiu, T.; Wang, J.; Sui, S.; Hao, C.; Liu, T.; Li, Y.; Feng, M.; Zhang, A.; Qiu, C.-W.; Qu, S. Phase-to-pattern inverse design paradigm for fast realization of functional metasurfaces via transfer learning. *Nat. Commun.* **2021**, *12* (1), 2974.

(73) Li, L.; Zhao, H.; Liu, C.; Li, L.; Cui, T. J. Intelligent metasurfaces: Control, communication and computing. *eLight* **2022**, *2* (1), 7.

(74) Chen, H.-T.; Taylor, A. J.; Yu, N. A review of metasurfaces: Physics and applications. *Rep. Prog. Phys.* **2016**, *79* (7), No. 076401.

(75) Lin, D.; Fan, P.; Hasman, E.; Brongersma, M. L. Dielectric gradient metasurface optical elements. *Science* **2014**, *345* (6194), 298−302.

(76) Wu, C.; Yu, H.; Lee, S.; Peng, R.; Takeuchi, I.; Li, M. Programmable phase-change metasurfaces on waveguides for multimode photonic convolutional neural network. *Nat. Commun.* **2021**, *12* (1), 96.

(77) Fu, T.; Zang, Y.; Huang, H.; Du, Z.; Hu, C.; Chen, M.; Yang, S.; Chen, H. On-chip photonic diffractive optical neural network based on a spatial domain electromagnetic propagation model. *Opt. Express* **2021**, *29* (20), 31924−31940.

(78) Wang, Z.; Chang, L.; Wang, F.; Li, T.; Gu, T. Integrated photonic metasystem for image classifications at telecommunication wavelength. *Nat. Commun.* **2022**, *13* (1), 2131.

(79) Zhu, H. H.; Zou, J.; Zhang, H.; Shi, Y. Z.; Luo, S. B.; Wang, N.; Cai, H.; Wan, L. X.; Wang, B.; Jiang, X. D.; Thompson, J.; Luo, X. S.; Zhou, X. H.; Xiao, L. M.; Huang, W.; Patrick, L.; Gu, M.; Kwek, L. C.; Liu, A. Q. Space-efficient optical computing with an integrated chip diffractive neural network. *Nat. Commun.* **2022**, *13* (1), 1044.

(80) Tavanaei, A.; Ghodrati, M.; Kheradpisheh, S. R.; Masquelier, T.; Maida, A. Deep learning in spiking neural networks. *Neural networks* **2019**, *111*, 47−63.

(81) Ghosh-Dastidar, S.; Adeli, H. Spiking neural networks. *Int. J. Neural. Syst.* **2009**, *19* (04), 295−308.

(82) Bi, G.-q.; Poo, M.-m. Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **1998**, *18* (24), 10464−10472.

(83) Yu, J.; Yang, X.; Gao, G.; Xiong, Y.; Wang, Y.; Han, J.; Chen, Y.; Zhang, H.; Sun, Q.; Wang, Z. L. Bioinspired mechano-photonic artificial synapse based on graphene/MoS_2 heterostructure. *Sci. Adv.* **2021**, *7* (12), eabd9117.

(84) Chen, T.; Wang, X.; Hao, D.; Dai, S.; Ou, Q.; Zhang, J.; Huang, J. Photonic synapses with ultra-low energy consumption based on vertical organic field-effect transistors. *Adv. Opt. Mater.* **2021**, *9* (8), 2002030.

(85) Jiang, J.; Hu, W.; Xie, D.; Yang, J.; He, J.; Gao, Y.; Wan, Q. 2d electric-double-layer phototransistor for photoelectronic and spatiotemporal hybrid neuromorphic integration. *Nanoscale* **2019**, *11* (3), 1360−1369.
(86) Jiang, D.; Li, J.; Fu, W.; Chen, Q.; Yang, Y.; Zhou, Y.; Zhang, J. Light-stimulated artificial synapse with memory and learning functions by utilizing an aqueous solution-processed In2O3/AlLiO thin-film transistor. *ACS Appl. Electron. Mater.* **2020**, *2* (9), 2772−2779.
(87) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. *arXiv:1512.03385 [cs.CV]* **2016**, na.
(88) LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86* (11), 2278−2324.
(89) Forti, M.; Nistri, P. Global convergence of neural networks with discontinuous neuron activations. *IEEE Trans. Circuits Syst.* **2003**, *50* (11), 1421−1435.
(90) Cessac, B.; Samuelides, M. From neuron to neural networks dynamics. *Eur. Phys. J-Spec. Top.* **2007**, *142* (1), 7−88.
(91) Sheela, K. G.; Deepa, S. N. Review on methods to fix number of hidden neurons in neural networks. *Math. Probl. Eng.* **2013**, *2013*, 1.
(92) Sharma, S.; Sharma, S.; Athaiya, A. Activation functions in neural networks. *Int. J. Eng. Appl. Sci. Technol.* **2020**, *4* (12), 310−316.
(93) Agostinelli, F.; Hoffman, M.; Sadowski, P.; Baldi, P. Learning activation functions to improve deep neural networks. *arXiv:1412.6830 [cs.NE]* **2014**, na.
(94) Li, Q.; Liu, S.; Zhao, Y.; Wang, W.; Tian, Y.; Feng, J.; Guo, J. Optical nonlinear activation functions based on mzi-structure for optical neural networks. *2020 Asia Communications and Photonics Conference (ACP) and International Conference on Information Photonics and Optical Communications (IPOC)*, Beijing, China, Oct 24−27, 2020, ACP/IPOC, 2020; pp 1−3.
(95) Huang, Y.; Wang, W.; Qiao, L.; Hu, X.; Chu, T. Programmable low-threshold optical nonlinear activation functions for photonic neural networks. *Opt. Lett.* **2022**, *47* (7), 1810−1813.
(96) Li, H.; Wu, B.; Tong, W.; Dong, J.; Zhang, X. All-optical nonlinear activation function based on germanium silicon hybrid asymmetric coupler. *IEEE J. Sel. Top. Quantum Electron.* **2023**, *29*, 1−6.
(97) Zhang, T.; Wang, J.; Dan, Y.; Lanqiu, Y.; Dai, J.; Han, X.; Sun, X.; Xu, K. Efficient training and design of photonic neural network through neuroevolution. *Opt. Express* **2019**, *27* (26), 37150−37163.
(98) Xiang, S.; Ren, Z.; Zhang, Y.; Song, Z.; Guo, X.; Han, G.; Hao, Y. Training a multi-layer photonic spiking neural network with modified supervised learning algorithm based on photonic stdp. *IEEE J. Sel. Top. Quantum Electron.* **2021**, *27* (2), 1−9.

https://doi.org/10.1021/acsphotonics.2c01516
ACS Photonics 2023, 10, 2001−2010