Supplementary Information

Parallel Optical Computing Capable of 100-Wavelength Multiplexing

Xiao Yu[1,†], Ziqi Wei[2,3,†], Fangyuan Sha[1,†], Xinyu Wang[1], Yanqi Chu[1], Zhen Wang[1], Xilin Han[1,*], Hongwei Wang[4], Shulan Yi[1], Yuhu Cheng[1,3], Guangwei Hu[4,*] and Peng Xie[1,3,*]

[1]Wangzhijiang Innovation Center for Laser, Aerospace Laser Technology and System Department, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai 201800, China
[2]School of Physics, Peking University, Beijing 100871, China
[3]Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
[4]School of Electrical and Electronic Engineering, Nanyang Technological University, 308232, Singapore
[†]These authors contributed equally to this work.

*Correspondence: hanxilin@siom.ac.cn, guangwei.hu@ntu.edu.sg, pengx@siom.ac.cn

S1: Single soliton formation

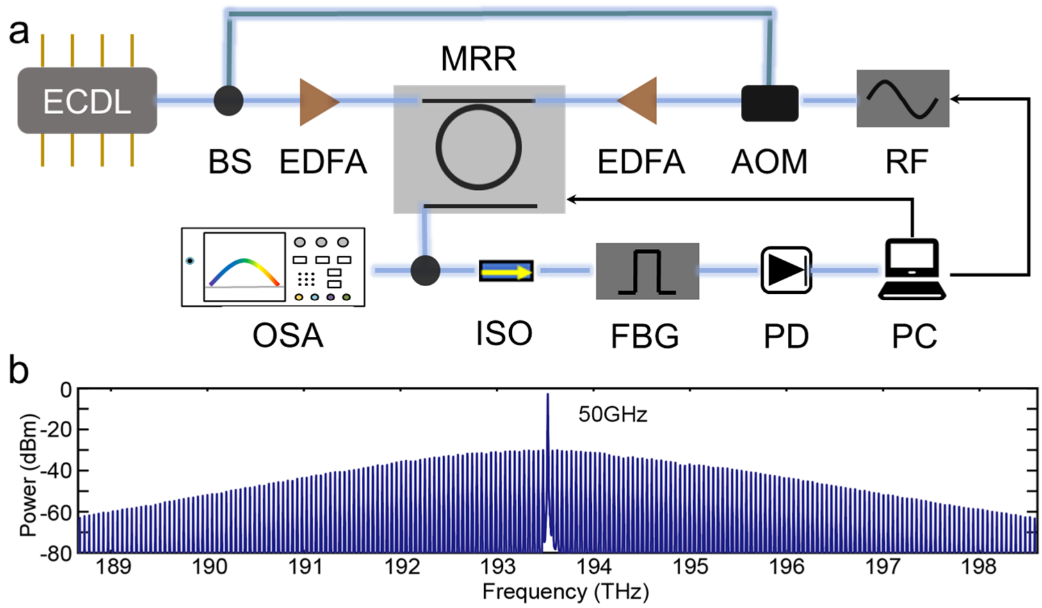

Fig.S1 The generation of a single soliton microcomb. (a) Schematic of experimental setup; (b) Optical spectrum of single soliton microcomb. ECDL, External cavity diode laser; BS: beam splitter; AOM: Acousto-Optic Modulator; EDFA, Erbium-doped fiber amplifier; OSA, Optical spectrum analyzer; MRR, micro-ring resonator; PD, Photodetector; ISO, Isolator; FBG: fiber bandpass grating; PC: personal computer.

The experimental setup for single soliton microcomb generation is illustrated in Fig. S1(a). To stabilize the intracavity optical field within the thermal equilibrium regime, an auxiliary laser is frequency-shifted by 150 MHz through radio frequency (RF) drive of an Acousto-Optic Modulator (AOM). The microcavity, thermally stabilized by an external temperature controller, exhibits a coupling loss of 3 dB per facet. The pump laser operates at 1550.12 nm wavelength, generating a soliton in the TM_{00} mode with a quality factor (Q) of 1.6 million. Both pump and auxiliary lasers are amplified to 1W. To achieve robust program-controlled soliton microcomb generation, we implement a three-step algorithm:

① **Pump suppression:** The output port incorporates a fiber Bragg grating bandpass filter to eliminate residual pump light. The initial modulated frequency on the AOM is set to 150 MHz. Both pump and auxiliary lasers are amplified to 1 W output power.

② **Thermal tuning and RF tuning for soliton formation:** Precise adjustment of the microcavity's operating temperature facilitates soliton formation. Increasing the modulated frequency and operation temperature of the microcavity ensures the single soliton formation.

③ **Stabilize single soliton microcomb:** When the output power reaches the single-soliton existence region, the program activates automatic stabilization via power feedback. If unsuccessful, the system iteratively re-executes Step ①.

Figure S1(b) shows the optical spectrum of the generated single-soliton microcomb, spanning the entire S+C+L band. This broadband spectral coverage enables 100-channel wavelength-multiplexing optical computing.

S2. WDM Data Loading

To demonstrate broadband parallel optical computing, each subsystem requires special modifications. Frequency lines of the soliton microcomb source are filtered out by the dense multiplexing element (DWDM). The WDM part is a commercial module commonly used in DWDM optical communication. Each microcomb line is regarded as an individual channel. The microcomb line is divided into N channels, where N is the number of input ports. Each channel is modulated, and all frequency components of the same channel are gathered by one multiplexer. The multiplexed channels enter the MZI network port-by-port. By increasing the scale of modulators and selected microcomb lines, the subsystem is feasible to demonstrate 100 wavelengths multiplexing and matrix encoding. In this work, the proof-of-concept setup was designed to balance flexibility while maintaining functionality. The setup is further optimized for full integration, enabling minimized system size and operational noise in future implementations.

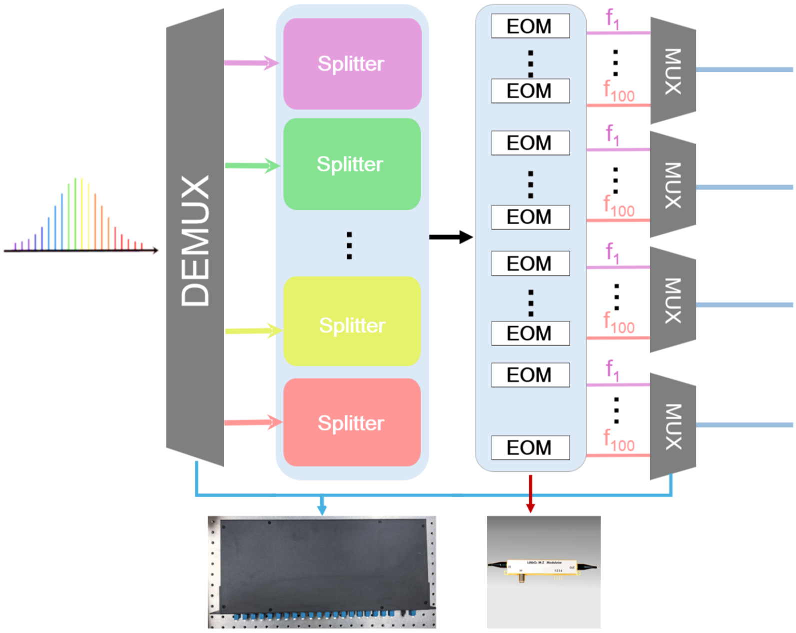

Fig. S2 Schematic of input vectors encoding subsystem

S3. Optical Computing Chip Design

The optical computing chip is fabricated on a Silicon-on-Insulator (SOI) platform using a 90 nm CMOS-compatible process to ensure high consistency across components. The splitters in the MZI employ broadband multimode interferometers (MMIs), optimized for uniform splitting ratios across the operational wavelength range (1530–1570 nm), thereby fulfilling the requirements for parallel optical computing. To increase coupling efficiency, a tapered waveguide is designed to match the mode with the single mode fiber array. The coupling loss of the chip is about 3 dB/facet. The optical computing chip exhibits a baseline transmission loss of 10 dB inherent to its 16×16 mesh architecture. The loss can be improved by using low-loss material platforms and optimized photonic components.

The phase shifters in the MZIs are implemented via resistive heaters patterned atop the waveguide layer. To mitigate thermal crosstalk, the driving power for each phase shifter is maintained at the milliwatt level, and additional thermal isolation structures are integrated around the MZIs to further suppress cross-channel thermal interference.

The half-wave voltage (V_{π}) of a single MZI is designed to be 1.5 V, enabling precise phase control with low voltage driving electronics.

S4. MZI mesh driving system

The MZI mesh is driven by a homemade high precision 64 channels matrix voltage source which provides 0.001 V precision from 0 V to 10 V. The temperature of the system is controlled by a TEC. The MZI mesh is initialized through a series of scans and the scanned transmission trace is plotted in Figure S3. The trace is fitted with an $\phi \sim \cos(V^2)$ function. The control parameters of the MZI mesh are then corrected by a fine tune algorithm to optimize its control precision.

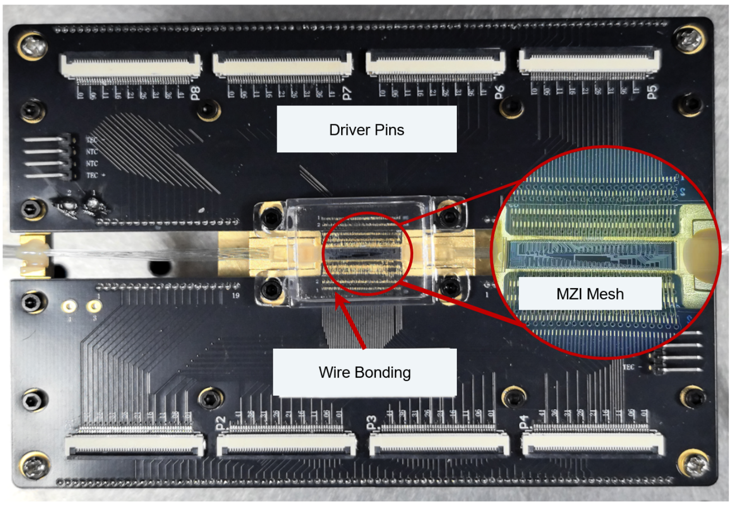

Fig. S3 16×16 Optical computing board

S5. Transmission Spectrum

Fig.S4 shows the input and output spectra based on our MZI mesh. The results demonstrate the bandwidth of MZI mesh beyond 40 nm, which can support over 100-wavelength multiplexing in parallel optical computing with ~0.4 nm channel spacing.

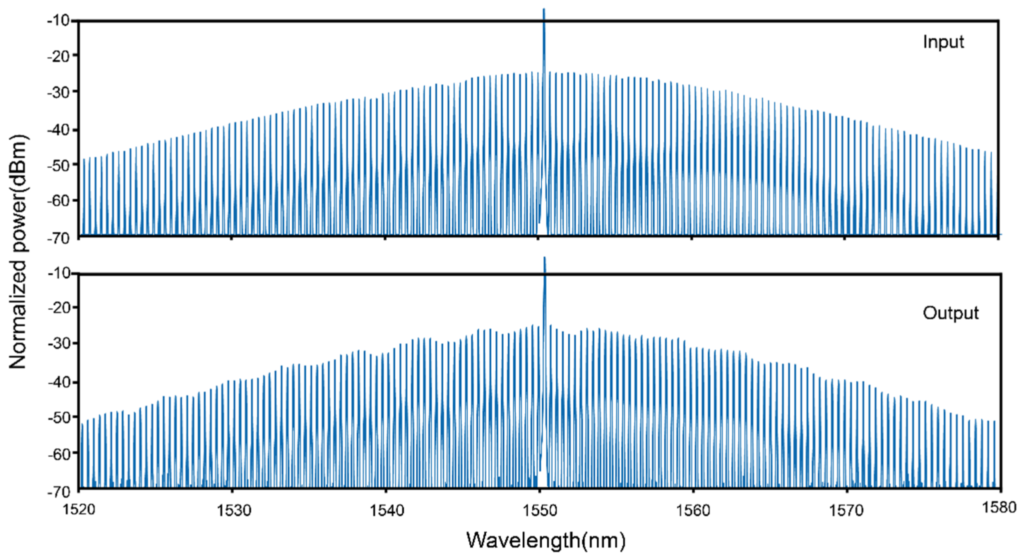

Fig. S4 Input (upper) and output (lower) spectrum of optical computing signal

S6. Fidelity and Consistency Measurement

The fidelity measurement for the optical computing system is conducted using a precise port-to-port scanning method. This approach involves examining the system's response by sequentially applying signal vectors to each input port and recording the corresponding output. The port-to-port scanning technique effectively illustrates the actual matrix configured on the optical computing system, providing a clear visual representation of the system's operational accuracy. The process is detailed as follows:

①**Input Signal Vectors:** The input to the system consists of unit vectors, which are applied sequentially: [1 0 0 0 0]^^{T}, [0 1 0 0 0]^^{T},⋯,[0 0 0 0 1]^{T}. This sequential application allows for a comprehensive line-by-line examination of the matrix configured on the optical computing system.

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} & u_{15} \\ u_{21} & u_{22} & u_{23} & u_{24} & u_{25} \\ u_{31} & u_{32} & u_{33} & u_{34} & u_{35} \\ u_{41} & u_{42} & u_{43} & u_{44} & u_{45} \\ u_{51} & u_{52} & u_{53} & u_{54} & u_{55} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \\ u_{41} \\ u_{51} \end{bmatrix} \tag{1}$$

②**Matrix Configurations:** Two types of matrices are configured for testing the system's fidelity: a diagonal matrix and a random matrix S_1, S_2. The diagonal matrix is initially set to an identity matrix using the MZI mesh configuration, ensuring that each input port corresponds directly to a specific output port.

$$S_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, S_2 = \begin{bmatrix} 0.215 & 0.404 & 0.004 & 0.231 & 0.146 \\ 0.257 & 0.510 & 0.005 & 0.161 & 0.068 \\ 0.495 & 0.131 & 0.243 & 0.027 & 0.104 \\ 0.120 & 0.031 & 0.395 & 0.194 & 0.260 \\ 0.146 & 0.045 & 0.367 & 0.308 & 0.134 \end{bmatrix}. \tag{2}$$

③**Fidelity Calculation:** The fidelity of the system is quantified using the Frobenius norm to compare the measured matrix *A* against the programmed matrix *S*. The fidelity *F* is calculated using the formula:

$$\boldsymbol{F} = 1 - \frac{\|\boldsymbol{D}\|_F}{\|\boldsymbol{S}\|_F} \tag{3}$$

Here, $\|\boldsymbol{D}\|_F$ is the Frobenius norm of the difference matrix, $\|\boldsymbol{S}\|_F$ is the Frobenius norm of the programmed matrix. $\boldsymbol{F}$ is the fidelity of the optical computing system.

④**Results:** For the identity matrix S_1 configuration, the calculated average fidelity is 0.861. For the random matrix S_2 configuration, as shown in Fig. 4b of the manuscript, the average fidelity is 0.907.

These results confirm the system's ability to accurately replicate the intended matrix configurations, indicating a high level of programmability and potential for replacing linear layers in neural networks.

S7. Parallel Computing Results

The slight slope of power difference with respect to wavelength observed in Fig.3c is attributed to the dispersion of waveguide. The dispersion can be expressed as $\Delta = 2\pi \cdot \delta\lambda \cdot L / \lambda^2$, where $\delta\lambda$ is the wavelength range of parallel optical computing, L is the difference in optical path and lambda is the central wavelength of the microcomb source. When the dispersion error is relatively small, it can be corrected by using the first-order compensation from the Taylor expansion of the error function. In practice, the MZI mesh is calibrated at 1550 nm. The dispersion correction is conducted by a post-processing method on the output array to avoid introducing complex algorithms consuming extra computility, as the system will be used to provide computility. The method sets the input of two channels close to 1530 nm and 1570 nm at the same level. Due to the dispersion error, the output power at both channels is different. In each output port, we extract the power difference from 1530 nm to 1570 nm as the baseline and the baseline takes zero at 1550 nm. A wavelength-dependent compensation term is then added to the output ports according to their wavelength. The output power is corrected by subtracting the baseline. The port-to-port scanning result after correction is shown in Fig.S5-S7. One can find that the vector-matrix multiplication remains consistent over a broad spectral range.

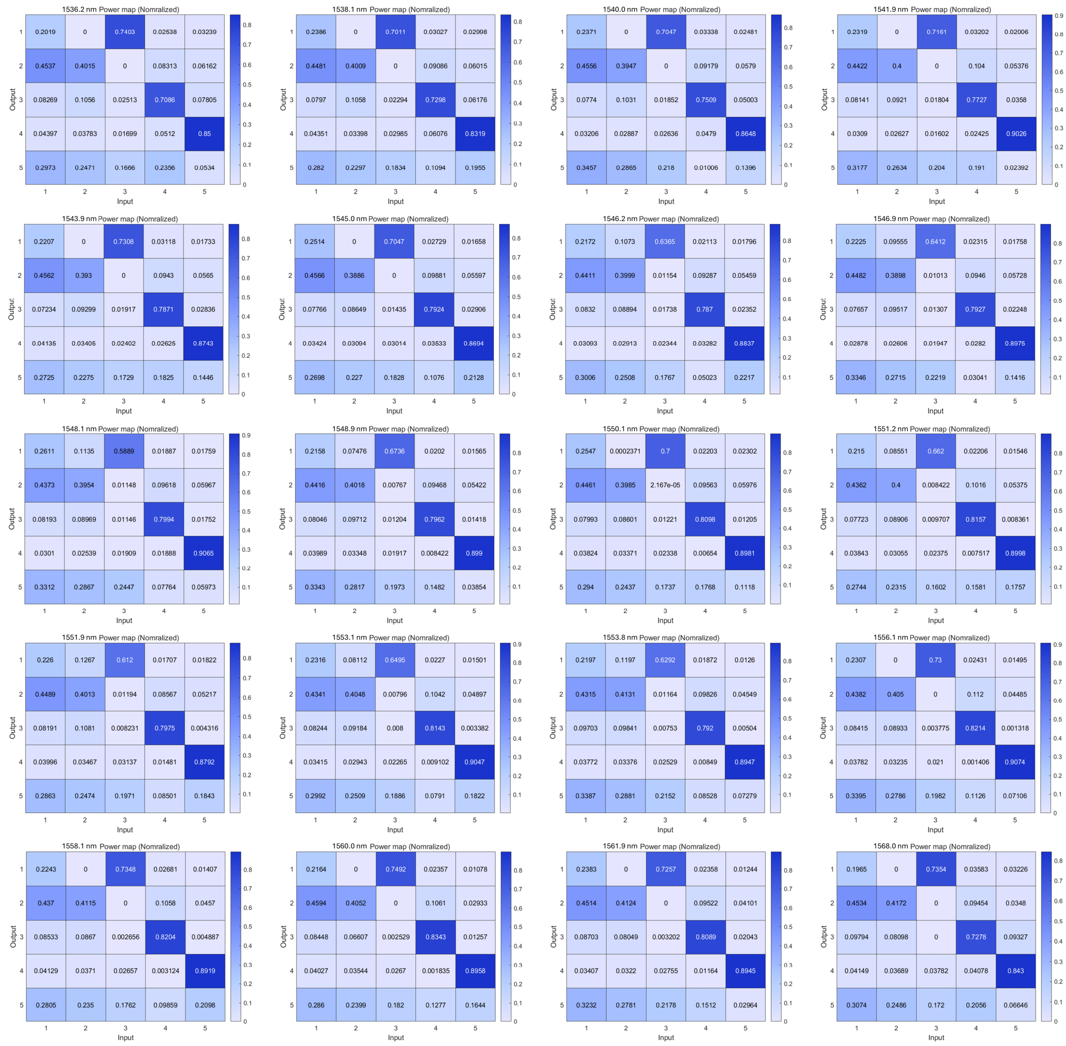

Fig.S5 Port-to-port power scanning of a random matrix I from 1536 nm to 1568 nm.

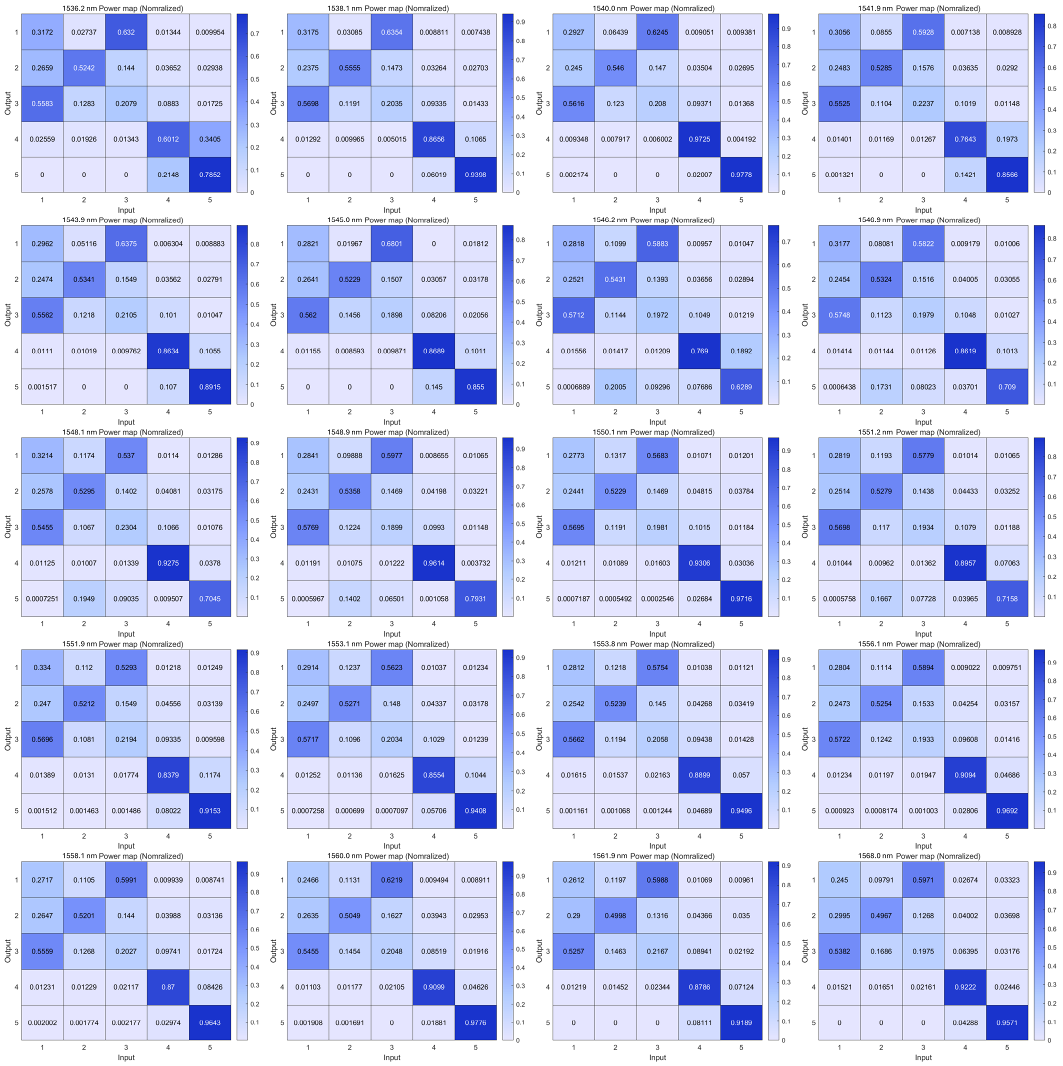

Fig.S6 Port-to-port power scanning of a random matrix II from 1536 nm to 1568 nm.

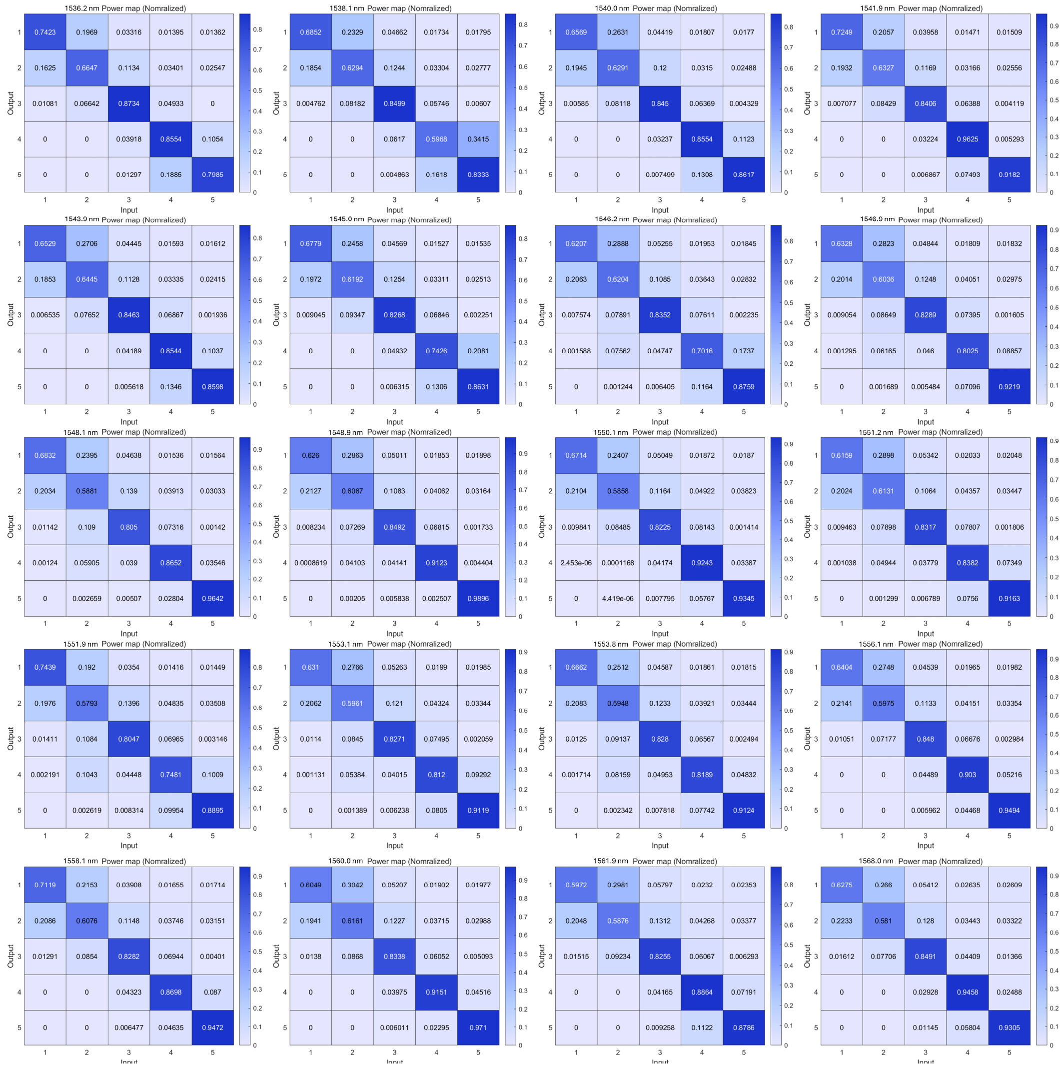

Fig.S7 Port-to-port power scanning of a random matrix III from 1536 nm to 1568 nm.

S8. MZI component bit precision estimation

The optimized bit precision of the MZI component is estimated by its extinction ratio. An MZI in the mesh is scanned over voltage and the variation of output power is recorded in Fig. S8. The extinction ratio can be estimated as the ratio between the minimal and maximal optical power, which is 0.0069. The extinction ratio is smaller than 2^{-7}. Therefore, the optimized bit precision of the MZI component can reach 7 bits.

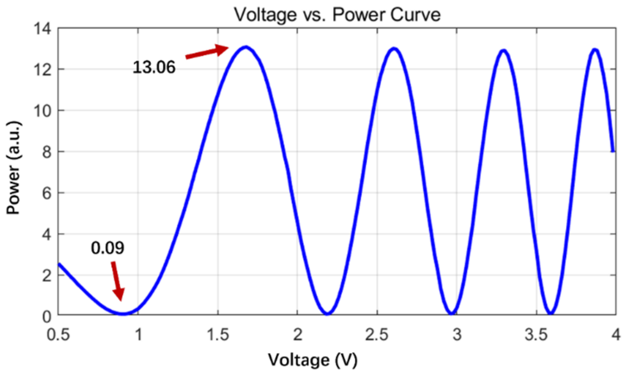

Fig S8 Voltage scan of single MZI component

S9 Estimation of power efficiency

We provide an estimation of the power consumption of the system. To make the power efficiency calculation comparable to literature records, the power efficiency of the optical computing system is calculated by the protocol mentioned in papers [1,2]. For cascaded MZI meshes, the computility is calculated by

$$\mathbf{C} = \mathbf{f_{mod}} \times \mathbf{N} \times \mathbf{N} \times \mathbf{WDM\ Channels} \times \mathbf{2} \tag{4}$$

where f_{mod} is the modulation frequency, N is the size of the matrix. In this work, the microcomb source can support 50GHz modulation bandwidth and 100 WDM channels, and the optical computing core can support a 16×16 matrix, which corresponds to the computility of 2560 TOPS.

The energy consumption and computility-power efficiency are calculated by the following equation [2]:

$$\mathbf{P_{total}} = \mathbf{N}\left(\frac{\mathbf{1}}{\mathbf{N}}\frac{\mathbf{h\nu}}{\mathbf{\eta}}\mathbf{max}\left(\mathbf{2}^{\mathbf{2N_b+1}}, \frac{\mathbf{C_d V_r}}{\mathbf{e}}\right)\mathbf{f_{mod}} + \mathbf{P_{rec}} + \mathbf{E_{mod}N_b f_{mod}}\right) \times \mathbf{WDM\ Channels}, \tag{5}$$

$$\mathbf{Power\ efficiency} = \frac{\mathbf{Computation\ throughput\ (TOPS)}}{\mathbf{Power\ consumption\ (W)}} \tag{6}$$

where $N = 16$, $P_{rec} = 150$ mW, $E_{mod}N_b f_{mod} = 350$ mW@50GHz, WDM channels is 100, $\frac{1}{N}\frac{h\nu}{\eta}max\left(2^{2N_b+1}, \frac{C_d V_r}{e}\right)f_{mod}$ is negligible. The result $P_{total} = 800$ W. Thus, computility-power efficiency can reach 3.2 TOPS/W. Here are the comparisons of some typical works reporting efficiencies in Table 1.

Table 1 Comparison of typical characteristics of recent optical computing work

Year	Journal Information	Technique	MAC per cycle	Available Channels	Max Achievable Computility	Max Power Efficiency
2024	Nature	WM [3]	27	1	16.2 TOPS	⩾1 TOPS/W
2022	Nature	MRR [4]	30	1	/	/
2023	Nat. Commun.	MRR [1]	/	/	⩾2.5 TOPS	⩾2.38 TOPS/W
2022	Nat. Commun.	MRR-WDM [5]	12	4	/	/
2021	Nature	Xbar-WDM [2]	64	8	⩾2 TMAC/s	⩾0.4 TOPS/W
2025	Nat. Commun.	MZI-TDM [6]	32	16	⩾1.25 TOPS	⩾1.21 TOPS/W
2025	Nature	MZI [7]	4096	1	8.19 TOPS@1GHz	⩾2.38 TOPS/W
2025	Nature	MZI [8]	16384	1	65.6 TOPS@2GHz	⩾1.75 TOPS/W
2024	Nat. Commun.	MRR [9]	/	/	211 TOPS@60GHz	/
2025	**This Work**	**MZI-WDM**	**25600**	**100**	**2560 TOPS@50GHz**	**⩾3.2 TOPS/W**

* MZI: Mach-Zehnder interferometer; MRR: Microring resonator; MDM: Mode division multiplexing; WDM: Wavelength division multiplexing; Xbar: Crossbar; TDM: Time division multiplexing; WM: Weighted matrix; CVOCA: Complex-valued optical convolution accelerator.

Overall, optical computing from research to application, there are three bottlenecks consisting of computing chip-scale, modulation frequency and computing parallelism. The breakthrough of computing parallelism can furtherly improve computility by 2-3 orders of magnitude, when computing chip-scale and modulation frequency reach the physical limitation.

S10 Robustness of conventional convolutional neural network

The dispersion compensation ensures the matrix consistency with ~90% consistency which ensures all the wavelengths can perform the task with relatively consistent performances. To verify the applicability of parallel computing to artificial intelligence algorithms such as CNN, we examine the robustness of a digital CNN running on the computer and the result is depicted in Fig.S9. The result suggests that a 10% difference in multi-layer CNN would not induce prominent variations in imaging classifications. The classification accuracy drops by 2.5%. Our preliminary proof-of-concept experiment in image classification is also verified via the cascade broadband MZI network in Ref. 20. In application aspects, our computing architecture has better compatibility with all the MZI meshes and demonstrates the prominent advantages in parallel information processing. This characteristic enriches the applications in the telecommunication system and data center.

The simulation method is as follows: we perform a CNN using PyTorch on a personal computer with two layers of convolution kernel (8 and 16 nodes) and 2 layers of linear kernel (400 and 64 nodes) on the MNIST dataset. The convolution kernels are perturbed randomly with different error levels. We run a Monte-Carlo test at each error level and the accuracy drop with respect to noise level is depicted in Fig. R16. When the noise level is smaller than 0.3, the average accuracy drops lower than 5%. Therefore, the current 90% consistency is sufficient for CNN applications.

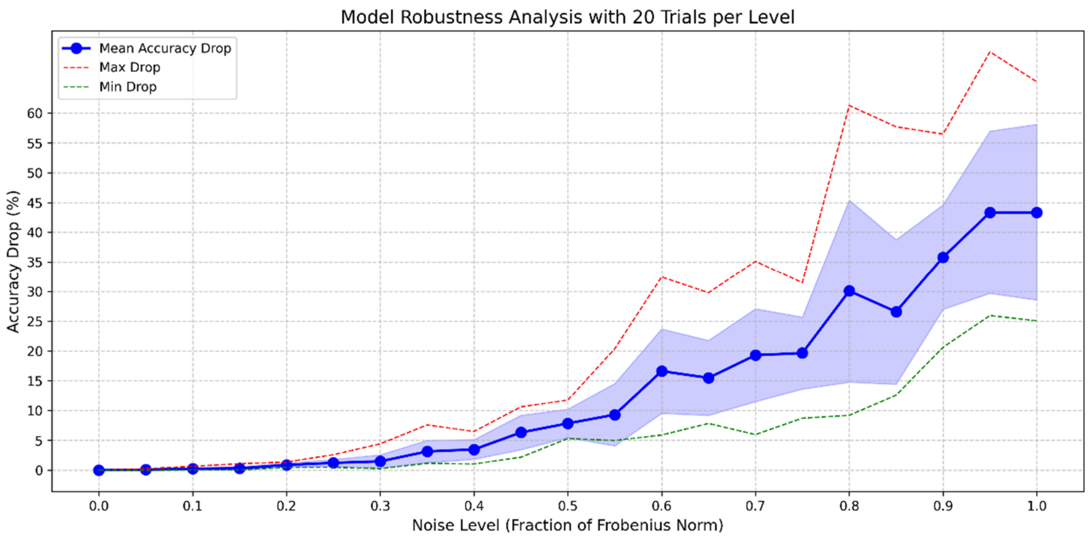

Fig.S9 Recognition accuracy of 4-layer CNN with perturbed convolution kernel.

Reference:

[1] B. Bai et al., Microcomb-based integrated photonic processing unit, Nat. Commun. **14**, 66 (2023).

[2] J. Feldmann et al., Parallel convolutional processing using an integrated photonic tensor core, Nature **589**, 52 (2021).

[3] B. Dong et al., Partial coherence enhances parallelized photonic computing, Nature **632**, 55 (2024).

[4] F. Ashtiani, A. J. Geers, and F. Aflatouni, An on-chip photonic deep neural network for image classification, Nature **606**, 501 (2022).

[5] S. Xu, J. Wang, S. Yi, and W. Zou, High-order tensor flow processing using integrated photonic circuits, Nat. Commun. **13**, 7970 (2022).

[6] S. Hong et al., Versatile parallel signal processing with a scalable silicon photonic chip, Nat.

Commun. **16**, 288 (2025).
[7] S. Hua et al., An integrated large-scale photonic accelerator with ultralow latency, Nature **640**, 361 (2025).
[8] S. R. Ahmed et al., Universal photonic artificial intelligence acceleration, Nature **640**, 368 (2025).
[9] D. Wang, Y. Nie, G. Hu, H. K. Tsang, and C. Huang, Ultrafast silicon photonic reservoir computing engine delivering over 200 TOPS, Nat. Commun. **15**, 10841 (2024).