Supplementary Materials for 42 TOPS/mm^2 Photonic Convolutional Processor Empowered by a Soliton Comb and a Micro-Disk Resonator Array

Shanshan Cheng[1,2†], Yiran Guan[1†], Chenye Qin[2], Zheng Dai[1], Kunpeng Jia[2*], Wei Liang[3], Shi-ning Zhu[2], Zhenda Xie[2*], Jianping Yao[1*]

1 Microwave Photonics Research Laboratory, School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada
2 National Laboratory of Solid State Microstructures, School of Electronic Science and Engineering, College of Engineering and Applied Sciences, School of Physics, Key Laboratory of Intelligent Optical Sensing and Manipulation, Ministry of Education, and Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, China
3 Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou 215123, China

**jiakunpeng@nju.edu.cn*
**xiezhenda@nju.edu.cn*
**jpyao@uottawa.ca*
[†]*These authors contribute equally*

1. Frequency analysis of the data stream based on different flattening strategies

The original figure is first transformed into one-dimensional data based on the kernel size and the stride in both horizontal and vertical directions. The corresponding electronic waveform is then generated by the arbitrary waveform generator (AWG) at a data rate of 64 GBaud. According to the Nyquist sampling principle, the bandwidth of the generated signal is approximately 32 GHz. Such a high bandwidth poses a significant challenge to electronic components, including cables, oscilloscopes, photodetectors (PD), and electro-optic (EO) modulators. To address this challenge, the experiment investigates various data flattening strategies that optimize the frequency distribution of the generated signal and convolution result.

In the experiment, we employ two data flattening methods, as illustrated in Fig. S1. The original image, as shown in Fig. S1a, is first segmented into multiple submatrices—each corresponding to the kernel size—by sliding along either the horizontal or vertical direction. Within each submatrix, the two-dimensional data is converted into a one-dimensional sequence using a predefined traversal order shown in Fig. S1b and Fig. S1c. These methods determine how the pixel or data values within the kernel are arranged in the resulting 1D vector. Zero padding is not applied at the boundaries of the input image. The data array will eventually be loaded onto multiple frequency channels. When these frequency-multiplexed signals propagate through a dispersion module, they undergo wavelength-dependent delays, thereby achieving temporal separation of the data streams as shown in Fig. S1d and Fig. S1e. Fig. S1d and Fig. S1e show the data streams obtained from the Flatten-1 and Flatten-2 methods, respectively, with both horizontal and vertical strides set to 1.

We check the frequency distribution of the convolution result with different flattening strategies. Fig. S2a1 and Fig. S2b1 show two commonly used kernels extracting the vertical and horizontal edges of the image,

respectively. The convolution results under these two kernels exhibit different frequency distributions under different data flattening methods. As shown in Fig. S2a2 and Fig. S2a3, the energy of the convolution result between the kernel in Fig. S2a1 and the data stream obtained using the Flatten-1 method is primarily concentrated in the low-frequency region, in contrast to the result from the Flatten-2 method. Conversely, the other kernel shows the opposite behavior, with the Flatten-2 method yielding a convolution result that concentrates more energy in the low-frequency components. These two strategies enable the desired feature map to be obtained while alleviating the reliance on high-bandwidth electronic components.

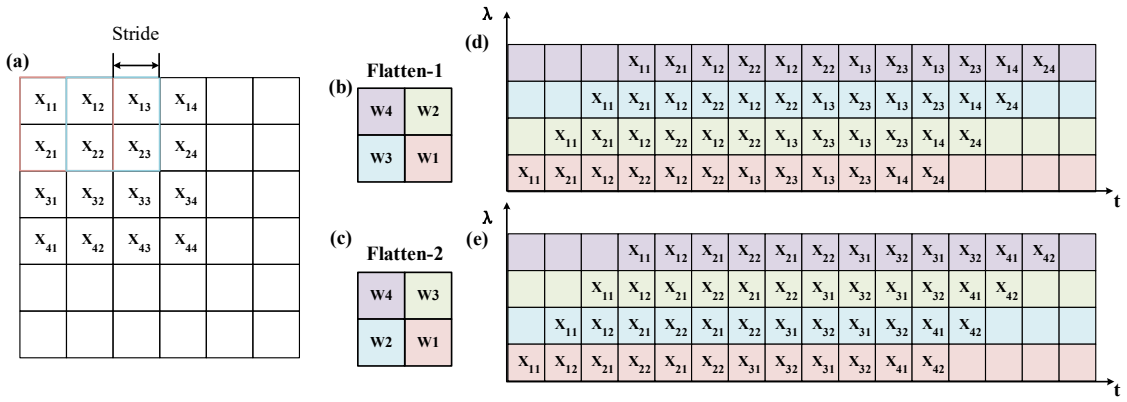

Fig. S1. Diagram of the one-dimensional data generation process using two different flattening methods.

2. Feature maps of other handwritten digit numbers

In addition to the digit "9" detailed in the main manuscript, we also demonstrate convolutional results for various other handwritten digits. The experimental results are presented in Fig. S3, and the corresponding simulated results are presented in Fig. S4. A clear consistency can be observed between the experimental results and the target structure across different input images.

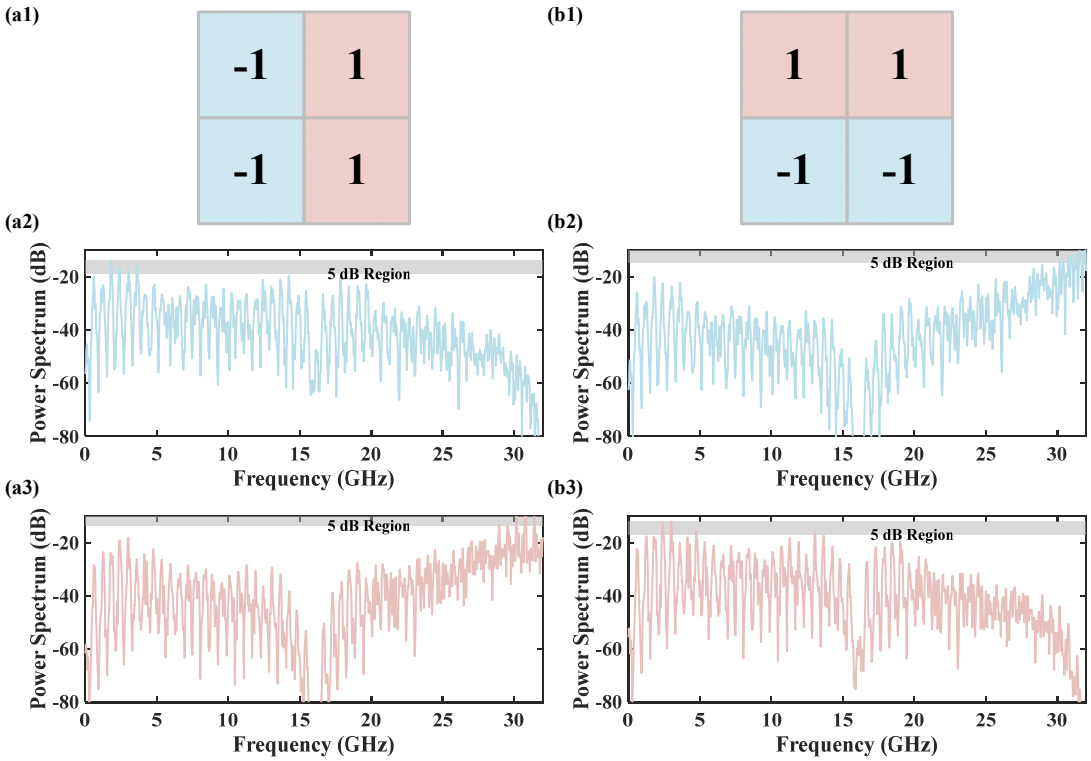

Fig. S2. Convolution kernels a1 [-1, 1; -1, 1] and b1 [1, 1; -1, -1]. (a2-a3) Frequency spectra of the convolution result between the convolution kernel [Fig. S2a1] and data stream generated using the Flatten-1 method and the Flatten-2 method, respectively. (b2-b3) Spectra of the convolution result between the convolution kernel [Fig. S2b1] and data stream generated using the Flatten-1 method and Flatten-2 method, respectively.

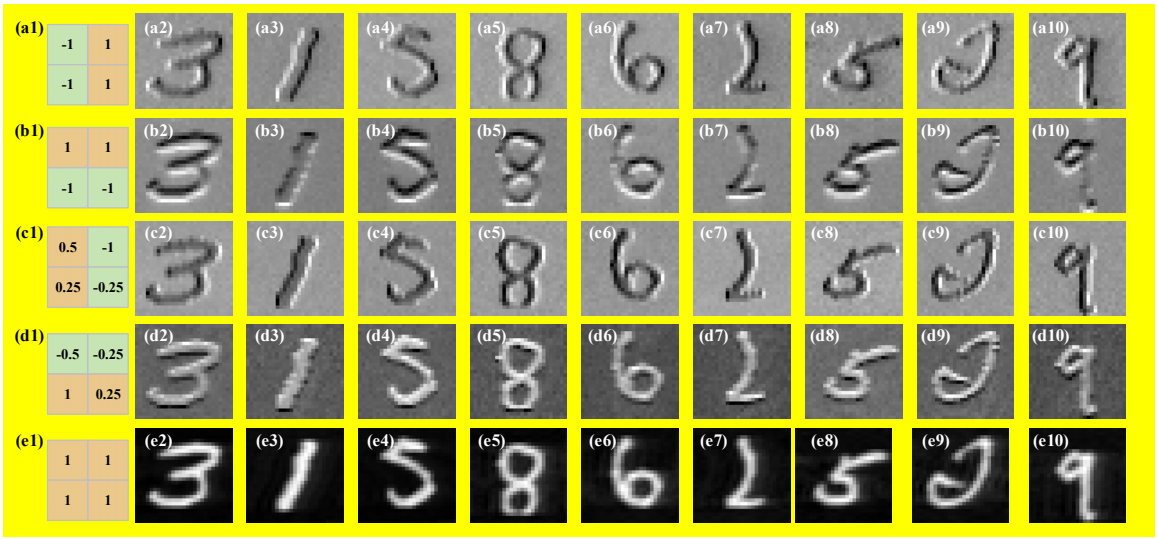

Fig. S3. (a1-e1) Different convolution kernels with [-1, 1; -1, 1], [1, 1; -1, -1], [0.5, -1; 0.25, -0.25], [-0.5, -0.25; 1, 0.25] and [1, 1; 1, 1]. (a2–a10), (b2–b10), (c2–c10), (d2–d10), and (e2–e10) Experimental feature maps obtained using the convolution kernels shown in (a1)–(e1), respectively.

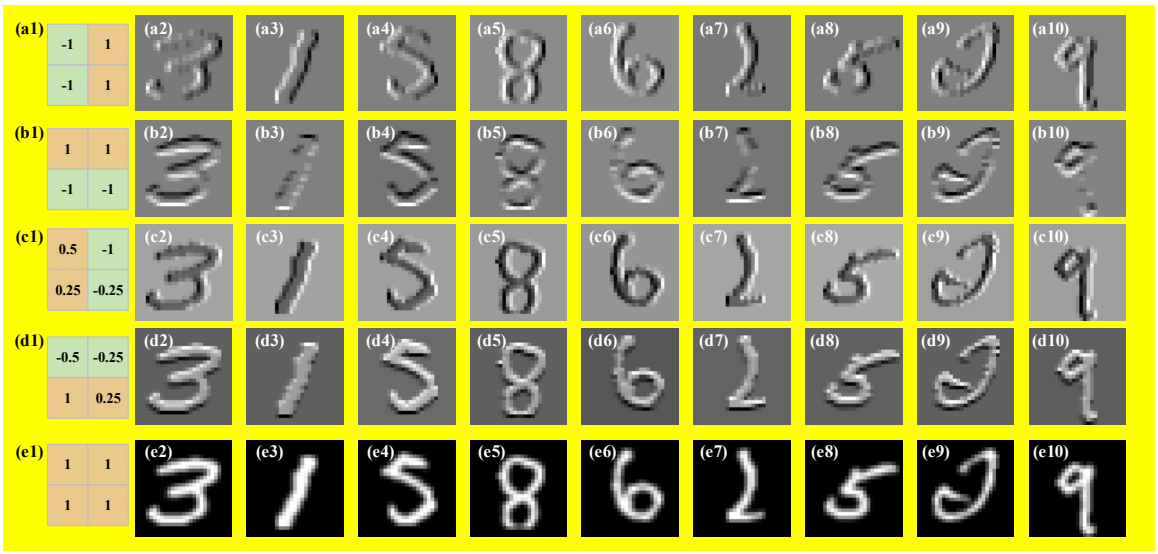

Fig. S4. (a1-e1) Different convolution kernels with [-1, 1; -1, 1], [1, 1; -1, -1], [0.5, -1; 0.25, -0.25], [-0.5, -0.25; 1, 0.25] and [1, 1; 1, 1]. (a2–a10), (b2–b10), (c2–c10), (d2–d10), and (e2–e10) Target feature maps obtained using the convolution kernels shown in a1-e1, respectively.

3. Energy consumption of the Photonic Convolutional Processor

The energy consumption mainly contains four parts: the microcomb generation module, EDFA, EOM, and on chip mesh structure. The detailed analysis of the energy consumption is shown in Table S1. The packaged microcomb module operates with a current source, a temperature controller, and a DC bias. The DC voltage is applied to a silicon sheet to tune the feedback phase. Since it doesn't form a closed electrical circuit, the energy consumption of the DC bias can be neglected. The light source of this photonic convolutional processor, with 205.5 mW power consumption, exhibits a low energy consumption characteristic.

Typical energy consumption of the low-power EDFA is about 2 W, while the optical power required in this experiment is 15.5 dBm, which can be readily provided by such an EDFA. Furthermore, the output power requirement from the EDFA can be greatly reduced by replacing the grating coupler with an edge coupler on the convolutional chip.

Each convolutional unit requires approximately 16 mW. In the experiment, two 4 × 4 convolutional kernels were employed, corresponding to 32 microdisks in total, leading to an overall tuning power of 512 mW. The overall power consumption of the optical components in the photonic convolutional processor is approximately 2.72 W, corresponding to an electrical energy efficiency of ~0.66 pJ per operation.

As the arbitrary waveform generator and the employed oscilloscope can be replaced by an integrated high-speed DAC and ADC, their energy consumption can be small. For example, a state-of-the-art high-speed DAC[1] operating at 100-112 Gb/s has an energy efficiency of approximately 1.36 pJ/bit. Based on this value, the DAC power consumption at a data rate of 64 Gb/s is estimated to be 1.36 pJ/bit × 64 Gb/s ≈ 87 mW.

A state-of-the-art 8-bit time-interleaved ADC operating at a sampling rate of 24-72 GS/s has an energy efficiency of 2.0-3.3 pJ per conversion[2]. Based on this value, the total power consumption is estimated to be 235 mW at 72 GS/s.

Table S1 Energy consumption of the components in the photonic convolutional processor.

Components	Voltage (V)	Current (A)	Power (W)
Current source for microcomb	1.5	0.125	0.1875
Temperature controller for microcomb	0.180	0.101	0.018
DC bias for microcomb	1	N/A	N/A
EDFA	N/A	0.768	2
DC bias for modulator	5.59	N/A	N/A
Micro disk unit	1.17	0.014	0.016
mesh structure	N/A	N/A	0.016×32 0.512
Total power consumption			2.72

4. High Q fiber FP cavity

The fabrication process of the fiber Fabry–Pérot cavity consists of three main steps:

Step 1: A 0.5 cm-long ceramic ferrule serves as the mechanical housing for the fiber FP cavity.

Step 2: Carefully cleave a commercial few-mode optical fiber and insert it into the ceramic ferrule with precise alignment.

Step 3: Both fiber end facets are mechanically polished to achieve subwavelength surface roughness and subsequently coated with dielectric Bragg mirrors providing reflectivity of above 99.99% over the 1500-1600 nm wavelength range.

Fig. S5. The fabrication process of the Fiber FP cavity.

The transmission spectrum of the fiber FP cavity is measured using a tunable semiconductor laser (CTL 1550, Toptica). The measured resonance linewidth is 912.4 kHz, corresponding to an exceptionally high quality factor of $Q = 2.12\times10^8$.

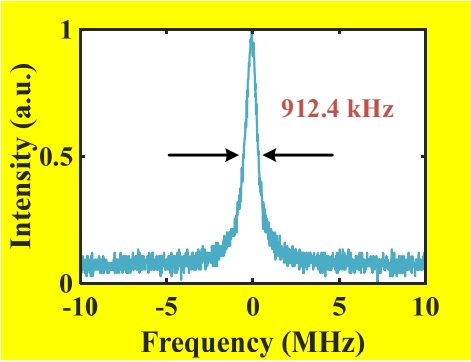

Fig. S6. The transmission spectrum of the Fiber FP cavity.

5. Linewidth characterization of the pump laser

To evaluate the frequency and phase fluctuations of the pump laser, its linewidth is measured using the correlated delayed self-heterodyne method[3]. The experimental setup is shown in Fig. S7. After amplification, the laser is split into two arms. In one arm, the optical frequency is shifted by 40 MHz using an acousto-optic modulator, while in the other arm, the laser passes through a 1 km single-mode fiber to eliminate coherence between the two paths. Two polarization controllers are employed to optimize the interference intensity. Two arms are then recombined by a 1×2 fiber coupler, and the resulting beat signal is detected by a photodetector (PD) and analyzed with a phase noise analyzer (PNA). The measured phase noise power spectral density (PSD) $S_{\Delta\varphi}(f)$ of pump laser after self-injection locking is shown in Fig. S8.

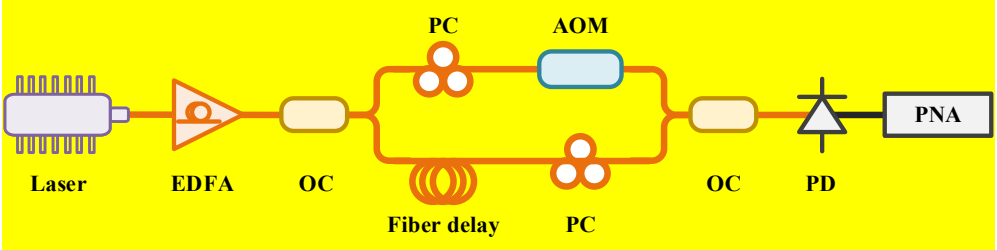

Fig. S7. Experimental setup for laser linewidth characterization. EDFA, erbium-doped fiber amplifier. OC, optical coupler. PC, polarization controller. AOM, Acousto-optic modulator. PD, photodetector. PNA, phase noise analyzer.

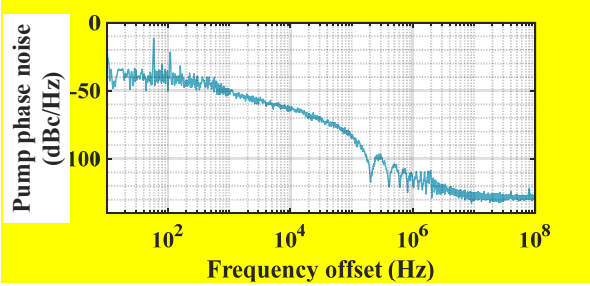

Fig. S8. Measured phase noise of the beat signal between the two optical paths using the delayed self-heterodyne scheme.

According to the measured power spectral density of the beat-note phase noise $S_{\Delta\varphi}(f)$, the power spectral density of the laser phase noise can be calculated by

$$S_{\phi}(f) = S_{\Delta\varphi}(f)\frac{1}{4\sin^2(\pi f\tau)},$$

where τ is the delay time and f is the frequency offset. The power spectral density of the laser frequency noise $S_v(f)$(with units of Hz^2/Hz) is directly related to that of the phase noise through

$$S_v(f) = f^2 S_{\phi}(f),$$

The fundamental linewidth of the pump laser can be calculated from the power spectral density of the frequency noise at the white noise floor. The corresponding mathematical relationship is given by

$$\Delta v = \pi S_w,$$

where Δv is the fundamental linewidth, and S_w is the power spectral density value of the laser frequency noise at the white noise floor.

6. Potential Scalability of the Photonic Computing Core

In the current experiment, 8 comb lines with a frequency spacing of 60 GHz are employed to perform four 2×2 convolution operations, enabling a computational density of 42 TOPS/mm². Each comb line can be reused for two different kernels by activating two different microdisk resonators. As shown in Fig. 3, microdisk resonators highlighted in orange and green are independent of those highlighted in blue and purple, as the two groups do not share the same optical path. Under the current chip structure, increasing the number of comb lines from 8 to 16 and activating all microdisk resonators would double the computing density to 84 TOPS/mm², as the computing throughput is doubled while the footprint remains fixed.

The optical delay line in the experiment employs a 2 km single mode fiber, whose low-loss transmission window (>100 nm) fully covers the operating bandwidth of the comb source. The electric I/O devices, including a commercial Lithium Niobate modulator and a balanced photodetector, also support over 100 nm bandwidth, which provides sufficient margin beyond the bandwidth required for future scaling.

The measured optical transmission response of the mesh structure, including two grating couplers, three crossbars, and four microdisk resonators on chip, is shown in Fig. S9. The bandwidth limitation is mainly

determined by the free spectral range (FSR) of the microdisk resonator, which is approximately 18 nm. It provides sufficient margin beyond the 7.2 nm bandwidth required to achieve a computational density of 84 TOPS/mm². For future scaling to larger microdisk meshes (e.g., 8×8 or 16×16), if the output ports from different microdisk regions are mutually uncorrelated, wavelength-division multiplexing (WDM) can be adopted to reuse the same set of wavelength channels across different regions. This approach enables increased parallelism and higher computing throughput within a limited wavelength range.

When scaling to larger microdisk mesh structures, additional loss is mainly contributed by the crossbar waveguide, which is approximately 0.08 dB per crossbar waveguide. Increased loss reduces the detected optical power and may therefore degrade the overall signal-to-noise ratio (SNR), impacting the robustness and scalability of the system. This can be mitigated by optimizing the waveguide-microdisk coupling toward critical coupling and by replacing grating couplers with edge couplers with index-matching, which would reduce the coupling loss from ~6.5 dB to ~1.5 dB per facet. From Fig. S9, the insertion loss varies by 1.66 dB from 1543 nm to 1561 nm. This wavelength-dependent non-uniformity can be mitigated by resonance tuning of the microdisk resonators.

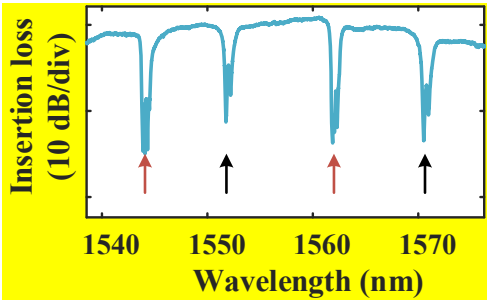

Fig. S9. Measured optical transmission response of the microdisk resonator. The red arrows indicate the fundamental modes, and the black arrows indicate the higher-order modes.

7. Device Structure

The relative vertical positions of a microdisk resonator, the heater, and the metal layer are indicated in Fig. S10.

Fig. S11 illustrates the layout of the microdisk resonator array in which the key geometrical parameters used in the design are highlighted. The spacing between adjacent microdisks is 12.14 μm. The disk pitch is 49.71 μm in the horizontal direction and 61.85 μm in the vertical direction.

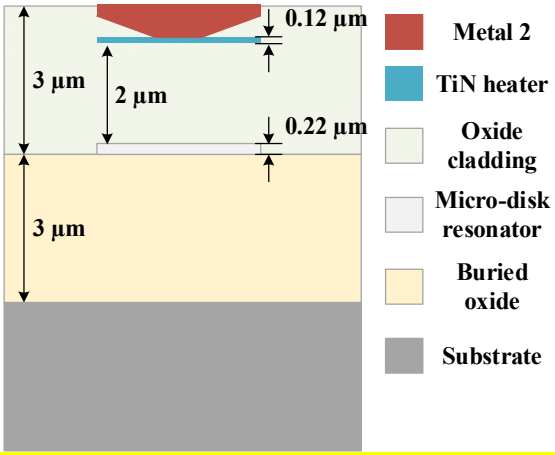

Fig. S10. Cross-sectional schematic of a thermo-optically tunable microdisk resonator.

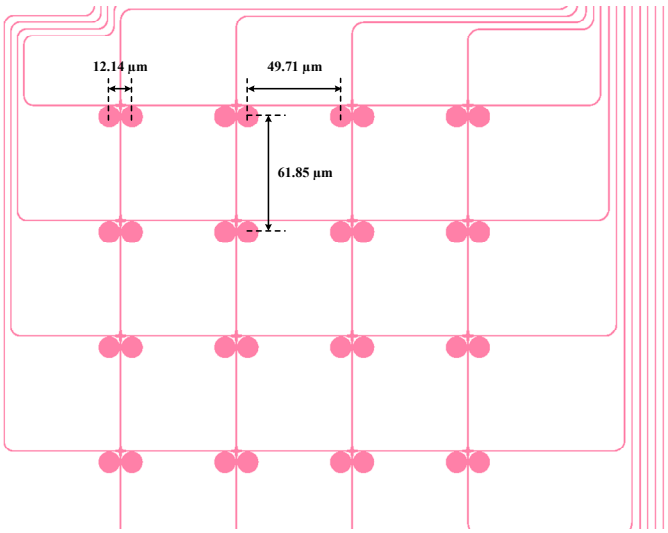

Fig. S11. Layout of optical waveguides, microdisk array on the chip. The electrical components and metal layers are hidden.

8. Configuration and Precision Evaluation of Kernel Weights

In the experiment, the kernel weight is configured by applying a driving voltage to the microheater integrated on the microdisk resonator. To enable accurate and stable kernel-weight control, a lookup table is first established to calibrate the relationship between the applied voltage and the resulting kernel weight. In addition, a closed-loop feedback calibration link is implemented to compensate for drift and stabilize the kernel weights during operation.

A tunable voltage is applied to the microdisk resonator, inducing a voltage-dependent shift of the resonance wavelength. Initially, the resonance peak is located on the blue side of the fixed laser wavelength. As the applied voltage increases, the resonance peak continuously shifts from the blue side to the red side

and eventually moves completely away from the laser wavelength. The output power is monitored using the photodetector (PD), and the output voltage of the PD is recorded. And the kernel weight is obtained from the measured PD voltage through normalization. By correlating the applied voltage with the kernel weight, a lookup table is established. Fig. S12 illustrates the process of establishing the lookup table.

Fig. S13 describes the closed-loop architecture employed to compensate for the variations of the kernel weights. A light wave from the input source is coupled into the microdisk resonators, and the output optical spectrum is monitored by an optical spectral analyzer (OSA). A power extraction module is then used to retrieve the power of each wavelength channel from the measured spectrum. These extracted powers are compared with the corresponding reference values to generate error signals. The error signals are fed into a proportional-integral-derivative (PID) controller, which produces control signals to drive the voltage source. By tuning the voltage applied to the microdisk resonator, the system is able to dynamically compensate for the spectral response and enable a stable multi-channel configuration of the kernel weights. In addition, a pre-established lookup table is used to provide an initial voltage setting for coarse calibration, enabling rapid determination of the operating point before the feedback loop performs fine adjustment.

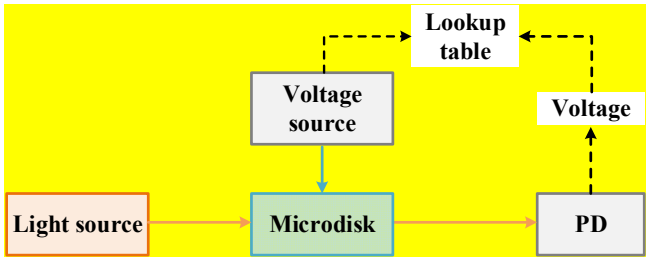

Fig. S12. Schematic of the lookup-table establishment process. PD, photodetector.

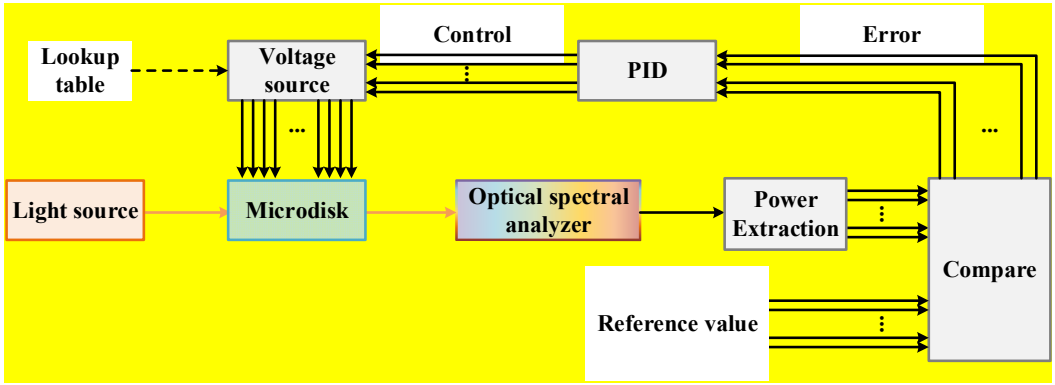

Fig. S13. Closed-loop feedback architecture for compensating for variations in the kernel weights.

Fig. S14 shows that the resonant frequency of the microdisk resonator shifts slightly toward lower frequencies as the applied voltage increases. Based on the measured resonance shift, the dependence of the kernel weight on the applied voltage is derived, as shown in Fig. 2b4. By numerically differentiating the measured weight-voltage curve, the maximum slope is extracted to be approximately 11.7 V^{-1}. The tuning

accuracy of the voltage source (National Instruments) is 0.1 mV. Therefore, the corresponding tuning accuracy of the kernel weight is estimated to be $\Delta w \approx 11.7 \times 0.1\ \text{mV} \approx 0.0012$.

The weight precision N_b is calculated by

$$N_b = log_2\left(\frac{\mu_{max} - \mu_{min}}{2\sigma}\right),$$

where μ_{max} and μ_{min} are the maximum and minimum values of the weight, respectively, and σ represents the maximum deviation of the measured weight from the target value. As the optical frequency fluctuation of the comb lines is below the hertz level, the comb lines can be regarded as stable optical sources. Consequently, in the experiments, the deviation of the kernel weight mainly originates from resonance shifts induced by voltage fluctuations. The voltage fluctuation of the voltage source is less than 0.1mV, and the maximum value of the deviation of the weight is 0.0012. Using the conservative worst-case criterion, the effective weight precision is 9.7 bits. The resonance shift of the microdisk resonator was monitored over 1 hour, as shown in Fig. S15. No significant drift is observed, indicating good thermal stability of the system.

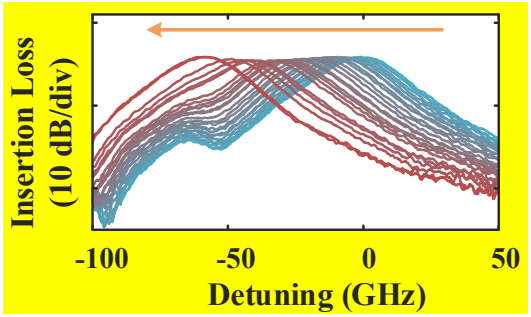

Fig. S14. The optical spectrum measured at the drop port of a microdisk resonator. The resonance frequency shifts as a function of the applied voltage from 60 mV to 680 mV. The orange arrow represents the increased applied voltage.

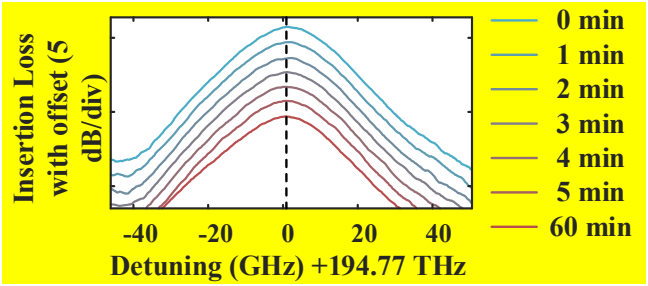

Fig. S15. Insertion loss spectra measured at different times. For clarity, the traces are vertically offset by 1dB.

9. SNR Analysis for Convolution Fidelity

To quantitatively evaluate the signal-to-noise ratio (SNR) of each wavelength channel, we use a convolution kernel [0, 0; 0, 1], where only one frequency channel is enabled and detected at the photodetector. The corresponding time-domain waveform, as shown in Fig. S16 is recorded by an oscilloscope, exhibiting an SNR of 22.2dB. The SNR is calculated by

$$\mathrm{SNR} = 20log_{10}\left(\frac{V_{signal,rms}}{V_{noise,rms}}\right),$$

where $V_{signal,rms}$ is the Root Mean Square of the signal voltage, and $V_{noise,rms}$ is the Root Mean Square of the noise voltage.

Based on the measured temporal waveform, we add an additive white Gaussian noise (AWGN) at different noise levels to generate noisy waveforms with different SNR levels, which are then used for image reconstruction to evaluate the impact of SNR on the convolution fidelity. The reconstructed images with different SNRs are shown in Fig. S17. This reflects how the convolution output progressively degrades as the SNR decreases.

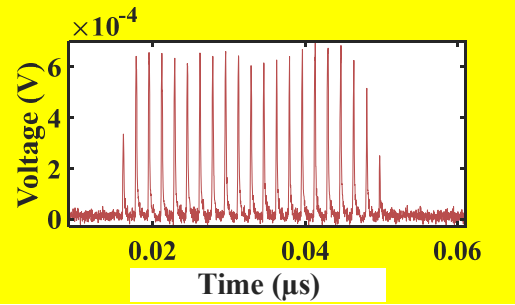

Fig. S16. Measured temporal convolution waveform obtained using a convolution kernel [0, 0; 0, 1].

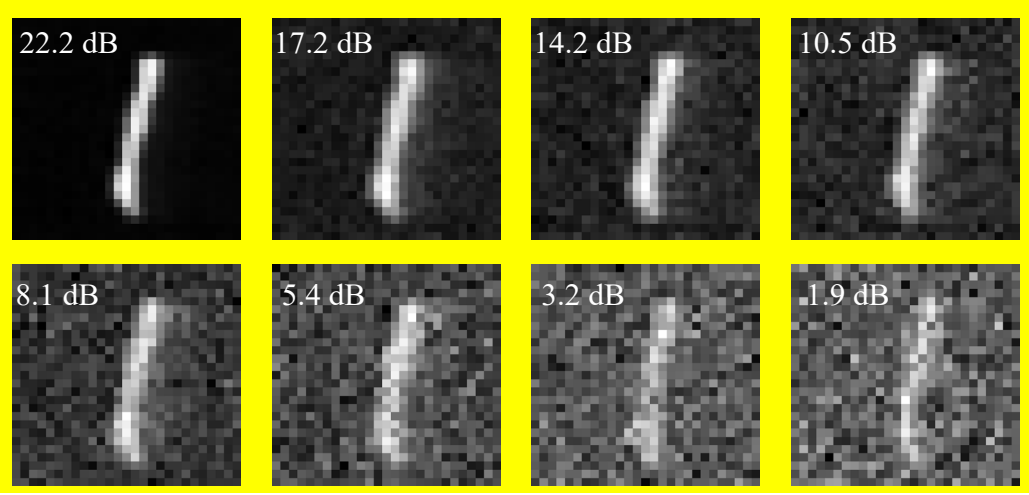

Fig. S17. Reconstructed figures under different SNR levels by adding an additive white Gaussian noise at different noise levels to the measured waveform [Fig. S16]. A clear degradation in reconstruction quality is observed as the SNR decreases.

References

1 Kossel, M. A. *et al.* in *2021 IEEE International Solid-State Circuits Conference (ISSCC).* 130-132 (IEEE).

2 Kull, L. *et al.* A 24–72-GS/s 8-b Time-Interleaved SAR ADC With 2.0–3.3-pJ/Conversion and >30 dB SNDR at Nyquist in 14-nm CMOS FinFET. *IEEE J. Solid-State Circuits* **53**, 3508-3516 (2018).

3 Camatel, S. & Ferrero, V. Narrow linewidth CW laser phase noise characterization methods for coherent transmission system applications. *J. Lightwave Technol.* **26**, 3048-3055 (2008).