

Scalable reservoir computing on coherent linear photonic processor

Mitsumasa Nakajima ¹✉, Kenji Tanaka ¹ & Toshikazu Hashimoto¹

Photonic neuromorphic computing is of particular interest due to its significant potential for ultrahigh computing speed and energy efficiency. The advantage of photonic computing hardware lies in its ultrawide bandwidth and parallel processing utilizing inherent parallelism. Here, we demonstrate a scalable on-chip photonic implementation of a simplified recurrent neural network, called a reservoir computer, using an integrated coherent linear photonic processor. In contrast to previous approaches, both the input and recurrent weights are encoded in the spatiotemporal domain by photonic linear processing, which enables scalable and ultrafast computing beyond the input electrical bandwidth. As the device can process multiple wavelength inputs over the telecom C-band simultaneously, we can use ultrawide optical bandwidth (~5 terahertz) as a computational resource. Experiments for the standard benchmarks showed good performance for chaotic time-series forecasting and image classification. The device is considered to be able to perform 21.12 tera multiplication-accumulation operations per second ($\text{MAC} \cdot \text{s}^{-1}$) for each wavelength and can reach petascale computation speed on a single photonic chip by using wavelength division multiplexing. Our results are challenging for conventional Turing-von Neumann machines, and they confirm the great potential of photonic neuromorphic processing towards peta-scale neuromorphic super-computing on a photonic chip.

¹NTT Device Technology Laboratories, 3-1 Morinosato-Wakamiya, Atsugi, Kanagawa 243-0198, Japan. ✉email: mitsumasa.nakajima.wc@hco.ntt.co.jp

Nowadays, machine learning techniques are advancing at a tremendous speed¹, and their applications for artificial intelligence (AI) systems are penetrating society. This is motivating the development of special-purpose AI hardware such as application-specific integrated circuits (ASICs) and field-programmable gate arrays (FPGAs)^{2,3}, which provide much faster and more energy-efficient computational resources. Recently, photonic implementations of artificial neural networks (ANNs) are attracting interest because they have great potential to reduce operational power, increase speed, and reduce latency beyond what is possible in electronic computing^{4–21}. Optical circuits can perform a large-scale multiply-accumulate (MAC) operation—a dominant factor in ANN computation—with ultrahigh processing speed thanks to their ultrawide bandwidth (terahertz region) and inherent parallelism in space, time, phase, and wavelength domains. As this operation is executed simply by light propagation and interference, the principal energy consumption is very small. Although the basic operation of this alternative computation was originally proposed in the early 1970s^{22–24}, the research was suspended due to the rapid development of electronic large-scale-integration (LSI) technology and an AI winter. The revival of AI through deep learning technology has led to a rethinking of neuromorphic photonic systems. Recent progress in photonic integrated circuits²⁵, which were developed for telecom applications, enables us to implement photonic ANN compactly⁴. This will boost the research of photonic ANN for alternative computational devices.

Coherent linear photonic processors^{4,26,27} are key engines for implementing such a computation system on compact optical chips. These processors are composed of an array of Mach–Zehnder interferometers (MZIs), which enables various types of optical topologies. As pioneered by Shen et al.,⁴ it is possible to map the mathematical description of a neural network onto a photonic chip with external nonlinear devices. Up to now, photonic circuits have been reported for various ANN models, such as fully connected multilayer perceptrons^{4–6}, spiking neural networks⁷, convolutional neural networks⁸, and recurrent neural networks, including reservoir computing (RC)^{9–21}.

Among them, RC^{28–30} is gaining attention for its affinity with optical elements and its excellent performance. The standard deep neural network (DNN) requires fine tuning of each weight through the use of the error back propagation algorithm³¹, which in turn requires highly accurate and uniform large-scale integration of tunable optical elements, which is a very challenging issue for fabrication. In addition, the training time of a photonic ANN is generally much longer than that with electrical devices due to the slow response of the phase shifters in the MZIs (on the order of milliseconds for thermo-optic devices). Therefore, standard photonic ANNs are effective only at the inference stage. On the other hand, in the RC framework, only output weights are trained, which is carried out by using a linear regression scheme. The other weights are fixed randomly. Thus, there is no need for any fine tuning of the optical system under training. As the training time is determined by forward propagation in RC, it can be accelerated by using photonics. In spite of the simple training, DNNs based on photonics have shown excellent performance comparable to that of standard DNNs on a series of benchmark tasks, such as speech recognition²⁰, economic forecasting³², action detection in movie data³³, and telecom signal compensation^{18,19,28}. Apart from optics, many physical implementations have been reported such as spintronic devices³⁴.

However, most conventional demonstrations have been limited to implementations of a middle recurrent layer called a “reservoir”. The scalability for chip integration is highly limited due to their simple architectures (less than 23 neurons in the reservoir^{14–16}). In addition, the bandwidth is limited by the electrical components

used for the nonlinear activation functions and input–output (I/O) frontend, such as response-lasers (gigahertz order)^{12,15} and spatial light modulators (hertz order)¹⁷. As the slowest one determines the operating time, the potential bandwidth of light (>THz) is limited to the electrical bandwidth. As a result, the computation speed, the distinguishing feature of photonic RC, is still considered to be below or at the same level as that of commercial electrical hardware such as central processing units (CPUs) or graphics processing units (GPUs). Note that there has not been a clear discussion in the literature. Our estimations are discussed in “Discussion” section.

In this paper, we propose a scalable photonic implementation of an RC system based on coherent linear photonic processors driven by an optical pulse source. In contrast to previous approaches, both the input and reservoir weights are optically encoded in the spatiotemporal domain, which enables scalable integration on a compact chip. The ultrafast optical pulse source up-converts the input signals to the higher frequency region beyond the electrical bandwidth. In the linear photonic processor, we can operate the photonic signals beyond the electrical bandwidth limitation. Although the nonlinear activation functions are only implemented in the input and output, the complex-valued evolution in coherent systems ensures rich dynamics comparable to that of incoherent nonlinear systems as described in the ref.¹³. We also demonstrate parallel processing based on wavelength division multiplexing (WDM). This enables the use of ultrawide optical bandwidth (>THz) as a computational resource, which boosts the computation efficiency beyond the computation bottleneck remaining from the I/O bandwidth. Experiments with the standard benchmarks showed good performance for chaotic time-series forecasting and image classifications with ultrafast processing speed of 17.1 ns per image. The device can achieve ten TMAC/s for each wavelength and can reach petascale computation speed on a single photonic chip by using WDM. Our results are challenging for conventional Turing–von Neumann machines, and they significantly advance photonic neuromorphic processing towards peta-scale neuromorphic super-computing on a photonic chip.

Results

Basic operation principle. Figure 1a shows the proposed device architecture for photonic RC. In this system, the radio-frequency (RF) input signal $u(t)$ with modulation time interval T is coded to the amplitude of the electromagnetic field of optical pulse, $\delta(t)$, by an optical Mach–Zehnder modulator (MZM). Then, the ultrafast optical pulse up-converts the RF-signal to the optical frequency, and this electro-optic conversion results in a sinusoidal nonlinearity. The modulated optical amplitude $u'(t)$ is described as

$$u'(t) = \delta(t) \cdot |\sin[\pi/2\{\gamma u(t) + V_o/V_\pi\}]|, \quad (1)$$

where V_o is offset voltage, and γ and V_π are the characteristic voltages of the MZM. Then, the converted signals $u'(t)$ are input to first stage of the linear optical circuits for the spatiotemporal input-masking. In this part, they are split into N -branches and then transmitted through different delay lines with a delay differential θ . The θ is set to satisfy the relationship $\theta = T/N$, where N is the virtual node count in a single optical cavity^{11,12,35}. Then, they are weighted by optical cross connecting units. This means that the input signals spread along time and space division with complex-valued weights. In general, the masked response $s_l(t)$ with continuous time representation is described as

$$s_l(t) = \int_{-\infty}^t h_l(\tau) u'(t - \tau) d\tau, \quad (2)$$

where $h_l(\tau)$ is an impulse response to the l th output port for the input mask circuit. For the comparison with the digital mask

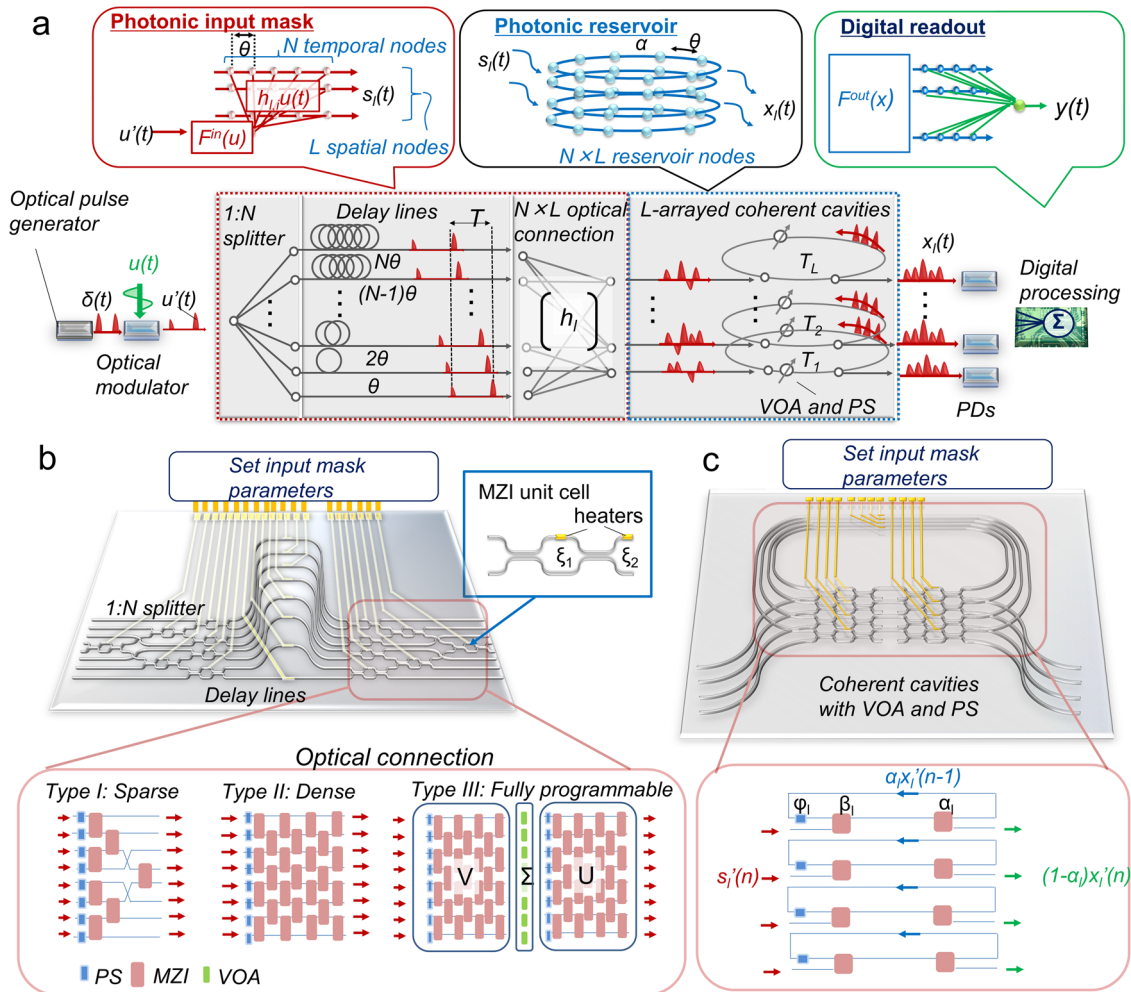


Fig. 1 Scalable reservoir computer (RC) on coherent photonic processor. **a** Schematics of proposed reservoir computing architecture. The electric input signal $u(t)$ with modulation time interval T is coded on optical pulse, $\delta(t)$ by optical modulator. This electro-optic conversion results in a sinusoidal nonlinearity (F^{in}). In the optical circuit, both the input and reservoir weights are encoded in the spatiotemporal domain. L -array of optical cavities with N temporal nodes acts as RC system with $N \times L$ virtual nodes. The output signals from optical circuit are detected by photodetectors (PDs), which generate nonlinear conversion (F^{out}). The detected signals $|x_l|^2$ are weighted and summed by digital processing unit to obtain final output $y(t)$. **b, c** Schematic of waveguide layout for **b** input mask circuit, and **c** reservoir circuit. Three types of possible installations for optical connection in the input mask circuit are also shown in **c**. The input weights h_l and reservoir parameters $\alpha_l, \beta_l, \varphi_l$ can be reconfigured by tuning the phase shifter (PS), Mach-Zehnder interferometer (MZI) and variable optical attenuator (VOA).

operation^{11,12,35}, we consider a discretized time $t(n,i)$ corresponding to each interval of duration θ and T : $t(n,i) = nT + i\theta$, where $n \in \mathbb{Z}$ and $i \in [1; N]$. By considering the ideal optical impulse with repeating time of T , the discretized expression of (1) is described as

$$s_l(n,i) = h_{l,i}u(n). \tag{3}$$

$$h_{l,i} = m_{l,i}\exp(j\Psi_{l,i}),$$

where $m_{l,i} \in [0; 1]$ and $\Psi_{l,i} \in [0; 2\pi]$ are the amplitude and phase delay from i th delay line to l th spatial. Thus, the input mask circuit acts as all-optical complex-valued spatiotemporal input weight generator. The matrix shape $h_{l,i}$ and its programmability depend on the type of optical cross connects. More detailed explanations for the input mask circuit are described in the Supplementary Note 1 and Supplementary Fig. 1. The masked optical signals $s_l(t)$ are input to the second stage of the optical processor. In this part, the integrated L -array of coherent cavities acts as spatially parallel delay-based optical reservoir with complex-valued evolution. As discussed in the refs.^{36,37}, a parallel reservoir computer increase

the virtual node count in the reservoir, which enhance the performance. The evolution equation of the complex-valued amplitude inside the l th reservoir cavity is given by

$$x_l(t) = \alpha_l \exp(j\varphi_l)x_l(t - T_1) + \beta_l s_l(t), \tag{4}$$

where T_1 is the roundtrip time, α_l is the feedback gain that is reconfigured through a programmable optical attenuator inside the cavity, β_l is the transmission coefficient of the input fiber coupler, j is an imaginary unit, and φ_l is the phase detuning of the cavity. The subscript l denotes that the parameters are for l th cavity. The T_1 is typically set to $T_1 = TN/(N-q)$, where $0 \leq q < N$. Then, the continuous time evolution in (2) can be approximated by the following discrete time evolution equations:

$$\begin{cases} x_l(n,i) = \alpha_l \exp(j\Delta\varphi_l)x_l(n-1,i-q) + \beta_l s_l(n,i) & (q < i \leq N) \\ x_l(n,i) = \alpha_l \exp(j\Delta\varphi_l)x_l(n-2,N+i-q) + \beta_l s_l(n,i) & (0 \leq i \leq q) \end{cases} \tag{5}$$

Thus, the parallel optical cavities act as parallel reservoir cavities. Although the nonlinear activation functions are only

implemented in the input and output, the complex-valued evolution in coherent systems ensures rich dynamics comparable to that in incoherent nonlinear systems, as described in¹³. The signals from the reservoir cavity are directly detected by photodiodes (PDs) installed after each cavity. Their dynamics are sampled by using an analog-to-digital converter or oscilloscope with a sampling interval of θ . Their discretized dynamic responses are considered as the squared norm of complex-valued virtual node response $|x_i(n, i)|^2$, where $n' \in [1; N' (=N\theta/\theta)]$ and N' are the measurable nodes. The outputs $y(n')$ are obtained from weighted summation of $|x_i(n', i)|^2$, which is described as:

$$y(n') = \sum_l \sum_i^N \omega_{l,i} |x_i(n', i)|^2, \quad (6)$$

where $\omega_{l,i}$ are trainable read-out weights, which are determined by minimizing the mean square error using the Tikhonov regularization or standard gradient descent method. We can obtain the $y(n')$ value with time interval T , which is the same as the input one.

The equilibrium network architecture is also shown in Fig. 1a. As the proposed architecture can integrate the spatiotemporal RC system on a compact chip, it is more scalable than the conventional on-chip integration^{13–15}. In contrast to previous digital pre-processing^{11,12,35}, we can set the short virtual node interval θ beyond the RF input sampling rates by setting a short delay difference and using an ultrafast pulse source. For instance, we can achieve the ultra-wideband optical computing on a single photonic chip by employing the femtosecond ($>$ THz) or attosecond pulse lasers ($>$ PHz). However, in this condition, the measurement of the reservoir state and post processing are difficult to perform due to the output sampling rate θ . Thus, it remains as the final bottleneck. Although the output bandwidth still remains bottleneck to access the ultrawide bandwidth of light, we can overcome it by using WDM and thus use the ultrawide optical bandwidth ($>$ THz) as a computational resource to boost the computation efficiency much more. The details are described in subsection “Parallel data processing using WDM”.

Integrated coherent linear photonic processor. As the proposed system requires coherent interference for the MAC operation, we need to accurately maintain the phases and delays in the optics. In addition, the system needs functional optics, including couplers, phase shifters, variable optical attenuators (VOAs), and MZIs. Thus, it is very difficult to build this optics with bulk fiber optics. Here, we employ our photonic platform technology, called the planar lightwave circuit (PLC), to integrate such a complex system on a compact chip. The PLC is a silica-based waveguide technology that has excellent features for composing functional optical devices: low transmittance loss (~ 0.02 dB/cm), mass fabrication using standard wafer processes, excellent stability (>10 -year operation³⁸), many available lineup functional components, including interferometers and delay lines³⁹, and a wide operating bandwidth (visible⁴⁰ to mid-infrared wavelengths^{38,39}). Thanks to these features, PLC devices have already been installed as optical systems for optical fiber links and have also been utilized as advanced circuits for fundamental science such as a quantum photonics⁴¹ and an optical lattice clock⁴². Here, we apply this technology to the photonic RC for the first time.

As a first demonstration, we integrated the above described linear optical system with $N = 32$ and $L = 16$ into the PLC. Thus, there are 512 virtual neurons in the circuit. Figure 1b show a schematic of the PLC layout for the input mask. For simplicity, the case for the RC with $N = 8$ is illustrated. The delay difference θ in the input mask circuit is set to ~ 8.3 ps ($1/120$ GHz) by adjusting the waveguide length. As this delay value corresponds

to the optical path length of ~ 2.49 mm, it can be accurately implemented to the PLC by using a standard lithography technique (submicrometer accuracy). For the splitting and weighting of the signal, the MZIs are integrated in the mask circuits. Each MZI has two heaters to adjust its phase by using the thermos-optic effect. The MZIs apply the following unitary conversion:

$$U = \begin{pmatrix} e^{i\xi_2} \sin(\xi_1/2) & e^{i\xi_2} \cos(\xi_1/2) \\ \cos(\xi_1/2) & -\sin(\xi_1/2) \end{pmatrix}, \quad (7)$$

where ξ_1 and ξ_2 are the phase shift of MZIs shown in the inset of Fig. 1b. By cascading the MZIs, we can realize various $N \times N$ optical topologies. Here, we illustrate three types of possible installations at the bottom of Fig. 1b. The simplest case is to use a mirror of a $1:N$ variable splitter as shown in Type I in Fig. 1b⁴³. This filter supports the arbitrary dense connection to the center output port with only an $(N - 1)$ array of MZIs. However, the connection to the other spatial port becomes sparse. This means that the h,l for $l = L/2$ is dense and fully programmable, but the other ports are sparse and not programmable. The dense connection can be realized by using an $N(N - 1)/2$ array of MZIs as shown in Type II in Fig. 1b, which is called a universal unitary operator^{44,45}. Although this architecture can realize dense and fully programmable connections to the output ports, only unitary conversion is supported. The arbitrary matrix connection is realized by combining $2N(N - 1)$ MZIs as shown in Type III in Fig. 1b^{4,46,47}. This network uses a physical instantiation of the singular value decomposition, which is a factorization of any matrix (M) as $M = U\Sigma V^\dagger$, where U is an $N \times N$ unitary matrix; Σ is an $N \times N$ diagonal, rectangular matrix of nonnegative real numbers; and V is an $N \times N$ unitary matrix. Here, two universal unitary circuits (U, V^\dagger) are connected by a column of single MZIs that are used as variable attenuators implementing Σ . Although the densely connected input mask is preferred to achieve better RC performance, the dense unitary conversion requires 496 MZIs for $N = 32$, which is challenging for the first trial due to the complex wiring and operation. Thus, in this study we employed Type I, which only requires 31 MZIs.

Figure 1c show a schematic of reservoir coherent cavities. For simplicity, the case for RC with $L = 4$ is illustrated in this figure. A VOA, phase shifter, and variable coupler is installed in each cavity. A previous coherent cavity-based RC¹³ was implemented in an optical fiber ring with a length of over 200 m. In contrast, we can integrate the system on a compact optical chip thanks to the small θ value. The previous fiber cavity was unstable in terms of temperature and vibration, and thus required the huge thermal isolation and an external feedback operation to stabilize the device. On the other hand, our cavity can be compactly implemented on an optical chip by using silica-based optical waveguide technology. Moreover, it is highly stable and does not require any feedback operation thanks to the telecom-grade PLC technology. The cavity length is set to $(6.0 + 0.4 \times l)$ cm, where l is the cavity number. The corresponding roundtrip time is $(\sim 290 + 20 \times l)$ ps. In the reservoir PLC, the coupling ratio to the passive cavity is set to $10/90$, which corresponds to β in Eq. (11) is $(0.1)^{1/2} = 0.316$. The $\alpha_1 \in [0, \alpha_{1,\max}]$ and $\varphi \in [0, 2\pi]$ for each optical layer is also reconfigurable by tuning the MZIs on the RC chips. The achievable $\alpha_{1,\max}$ value depends on the cavity number due to the loss increase from the intersection points, which is almost completely determined by $\alpha_{1,\max} = 0.87 - 0.021 \times l$. This dependence could be reduced by optimizing the waveguide design. Figure 2 shows the appearance of our optical RC circuit fabricated by the PLC technique. The input mask and reservoir circuit were fabricated on the separate chips for the ease of characterization. These chips are connected by the optical fiber

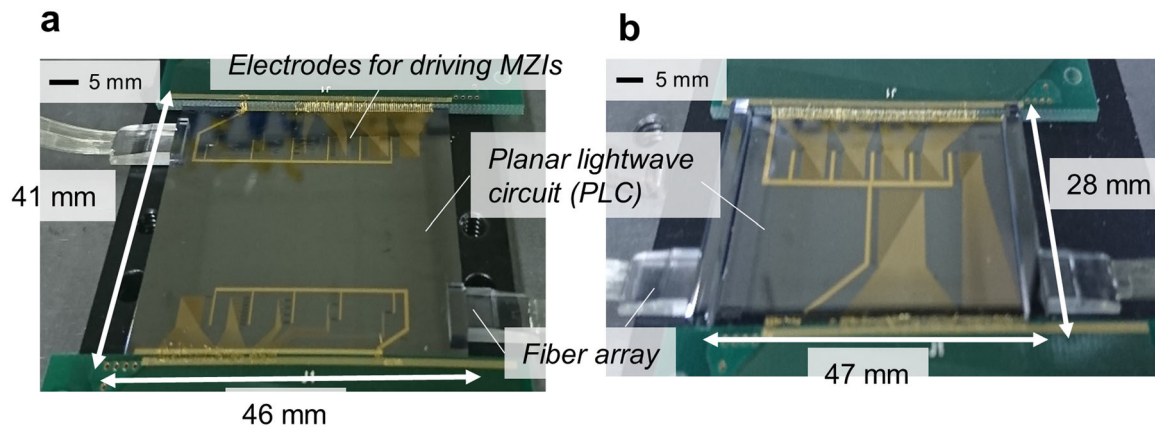


Fig. 2 Fabricated optical circuit for reservoir computing. **a, b** Appearance of fabricated optical circuit for **a** input mask and **b** reservoir circuit. Silica-based waveguide platform technology, called the planar lightwave circuit (PLC), was employed to integrate a complex optical system on a compact chip.

array. The footprint of the input and reservoir circuit were 41×46 and $28 \times 47 \text{ mm}^2$, respectively.

Here, we performed experiments to examine the fabricated chip's performance. First, we input the ASE light to the PLC for input mask and observed the transmittance spectrum to confirm the filtering performance. In this setup, two neighboring port were coupled to the center output port with the same optical power. Figure 3a shows the experimentally obtained filtering shape of the fabricated input mask circuit for the center output port (number 16) over the wavelength range of 1530–1565 nm. A magnified view is shown in Fig. 3b. For this experiment, the delay difference was set to θ as shown in Fig. 3e. As can be seen in Fig. 3a, flat and repeated wavelength dependence is obtained, suggesting the circuit can be used for an ultra-wideband optical processor. The observed extinction ratio of 20 dB suggests the good performance of the fabricated MZIs. The observed wavelength duration of 120 GHz suggests accurate implementation of θ . By changing the interfering port as shown in Fig. 3f, we observed a change in the wavelength duration of 60 GHz as expected [Fig. 3c]. When we randomly set the input mask parameter [Fig. 3g], we can set the random filtering shape as shown in Fig. 3d. Thus, the fabricated circuits can be used for the optical mask circuit. Next, we examined the time domain response and the stability of the fabricated chip. For this experiment, we input the masked signal into the center coherent cavity (number 8) to check the overall characteristics of the RC system as shown in Fig. 3k. Figure 3i shows the reservoir response of cavity number 8 before and after the 3-h operation. The input signal [Fig. 3h] and the output error [Fig. 3j] are also plotted. In this experiment, we randomly generated the mask function and α and φ were set to 0.5 and 0.1π , respectively. A continuous coherent pulse with repeating time of 512 ps [top of Fig. 3e] was input to the system. The response of the center output port was observed with an oscilloscope. As shown in this figure, we can successfully generate the time domain response along the time domain thanks to the input mask and reservoir circuit [Fig. 3i]. As is well known, a coherent system is highly sensitive to the changes in the optical phase due to changes in the surrounding temperature or vibration. Thus, the previous coherent cavity-based RC requires a complex feedback system to stabilize the RC outputs¹³. On the other hand, thanks to the solid-state circuit technology, the reservoir response for the fabricated circuit is stable over a period of three hours without any feedback operation, which is enough time to execute most RC tasks.

Chaotic time-series expectation. We used the Santa-Fe time-series prediction task⁴⁸ to evaluate the RC performance. The aim

of this task is to perform one-step-ahead prediction of chaotic data. The chaotic laser data were generated from a far-infrared laser. We used 2000 steps for training and 1000 steps for testing. For the training, we utilized the same teacher signal $d(n) = u(n + 1)$. The training was done by standard Tikhonov regression. The amplitude of the Santa Fe time series was normalized so that the input signal $u(n)$ of the Santa Fe time series ranged from 0 to 1.

First, we compared the bandwidth of input signal $u(t)$ and reservoir output $|x(t)|^2$ to confirm our concept described above. We used the center cavity (lane 8) for this experiment. The relative feedback phase was set to $\Delta\phi_l = 0.1\pi$. Figure 4a shows the measured bandwidth of input data $u(t)$ of Santa-Fe chaotic time-series and reservoir output $|x_l(t)|^2$. As can be seen in this figure, the bandwidth of the RF input is elongated thanks to the optical pulse modulation. Thus, we can execute RC processing beyond the input RF bandwidth. The observed maximum bandwidth was limited by the bandwidth of the oscilloscope (~ 20 GHz) in this experiment. By utilizing an electrical system with higher bandwidth, the signals can be up-converted to a much higher frequency region with the same optical circuits.

Next, we examined the performance on the chaotic-time-series expectation of our RC for a single cavity. Figure 4b shows the normalized mean square error (NMSE) as a function of feedback gain. For comparison, the simulation performance for the passive cavity and optoelectric reservoir with $N = 16$ are also plotted in the figure. For this simulation, we estimated the NMSE using Eqs. (3) and (5) with $q = 1$. The impulse response was set randomly to keep the following constraint: $|\sum h_i|_i|^2 = 1$. The feedback gain was swept from zero to one, and the relative feedback phase was set to $\Delta\phi_l = 0.1\pi$, the same as the physical setup. The experimental results well agree with the simulation, suggesting the good performance of our constructed optical circuit. As shown in the figure, the NMSE highly depended on the spectral radius of the reservoir connection (i.e., feedback gain in this case). This behavior was also observed in previous RC systems, and it implies the enhanced memory capacity in the large feedback condition. The observed NMSE was comparable to the optoelectric RC¹¹ in spite of the lack of the nonlinear element in the cavity.

Next, we investigated the performance using multiple cavities. The reservoir parameters were randomly set for each cavity. The measured reservoir responses of all the cavities are shown in Fig. 4c. Various responses were observed from each cavity. From these responses, we estimated the NMSE as a function of the number of parallel cavities as shown in Fig. 4d. The NMSE is monotonically reduced by adding the cavities, and NMSE of 0.06 is achieved, which is superior to that of a previous on-chip RC¹⁵ thanks to the increment of virtual nodes. This result indicates that

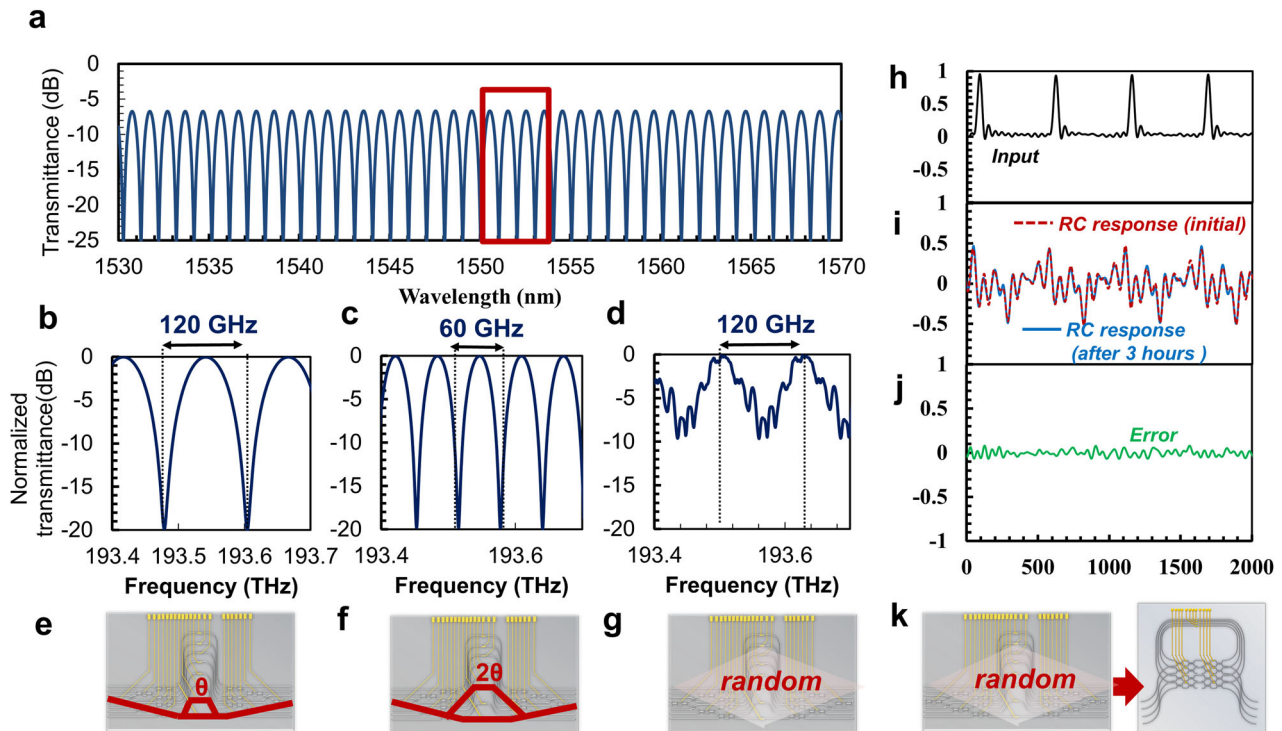


Fig. 3 Frequency and time response of fabricated circuit. **a** Experimentally obtained wavelength (frequency) filtering shape of the fabricated input mask circuit for the center output port (number 16) over the wavelength range of 1530–1565 nm. The delay difference was set to θ . **b** Magnified view for **a**. **c**, **d** Filtering shape with **c** delay difference of 2θ and **d** randomly set input mask. **e–g** Setting state of input mask circuit with **e** delay difference of θ [for **a** and **b**], **f** 2θ [for **c**], and **g** random state [for **d**]. The observed frequency response suggests the accurate implementation of input mask circuit. We also examined the time domain response and the stability of the fabricated chip by inputting the masked signal into the center coherent cavity (number 8). **h–j**, **h** Input pulse and **i** time responses of reservoir computer before and after 3-h operation and **j** their error. **k** Setting state of input mask and reservoir circuit [for **h–j**]. Thanks to the solid-state circuit technology, the reservoir response for the fabricated circuit is stable over a period of three hours without any feedback operation, which is enough time to execute most RC tasks.

the virtual nodes obtained from the parallel cavities are more effective for the chaotic time-series prediction task than those obtained from a single temporal sequence of a single cavity.

Image classification. We tested the image classification task to confirm the multidimensional data classification using multiple cavities. The dataset comprised hand-written-digits in the Modified National Institute of Standards and Technology (MNIST) database⁴⁹. The processing procedure for multidimensional inputs is shown in Fig. 5a. For preprocessing, the original 28×28 images were resized to 16×16 ones. To convert the temporal data, the 2×2 time-sliding window was used for data input. Different digit data are multiplexed in different time slots. The sliding step was set to 2. Thus, the input image data were converted to a four-dimensional time series with 64 time steps. These signals were converted to optical data, and they were processed by the input mask and reservoir optical circuits. The detailed operation for the multidimensional input is described in the “Methods” section. The output signals from each reservoir cavity were acquired by multiple photodetectors. The signals were weighted and summed by using Eq. (6). For the training, we used 60,000 images from the training dataset of MNIST. As this requires large-scale memory for Tikhonov regression, we optimized the readout weights in real time using the gradient descent method with the minibatch size of 200 and the learning rate of 5×10^{-3} . The training was executed in only one epoch. For the test phase, we validated the 10,000 data from the test dataset.

An example of the time domain reservoir responses (cavity number 8) is plotted in Fig. 5b. These responses indicate that we

can process the input images at a speed of 17.1 ns per image, which is much faster than the standard RC using electrical hardware; RC with same architecture takes ~ 1 ms using an 8-core 3.1-GHz CPU. The training and test results are shown in Fig. 5c, d. As shown in Fig. 5b, the training accuracy almost monotonically increased, suggesting the success of the training. The test accuracies were improved by increasing the parallelism of the RC, and we achieved maximum test accuracy of 91.3% for the 16 parallel RC case. To the best of our knowledge, this is first experimental demonstration of image classification using on-chip photonic RC. However, this value is still inferior to the accuracy for standard deep neural network models. As discussed in²¹, around 16,000 reservoir nodes are required to achieve state-of-the-art performance for the MNIST benchmark. As the node number of our device is 512, we need to construct 32 times larger circuits: e.g., the proposed optical circuit with $N = L = 128$ array is required, which needs the 128×128 optical cross-connects for the input mask and 128 array of 32-cm delay lines for the reservoir circuit. The fabrication of these components are challenging. However, it could be possible by using state-of-the-art manufacturing technology^{50–52}.

Parallel data processing using WDM. As described above, we can operate the photonic signals beyond the input RF limitation by using our device. The circuit also has potential for ultra-wideband operation thanks to the flat passband over 1530–1570 nm (~ 5 THz) as shown in Fig. 3a. However, the remaining bottleneck from the output still limits the processing speed. Here, we demonstrate parallel processing based on WDM

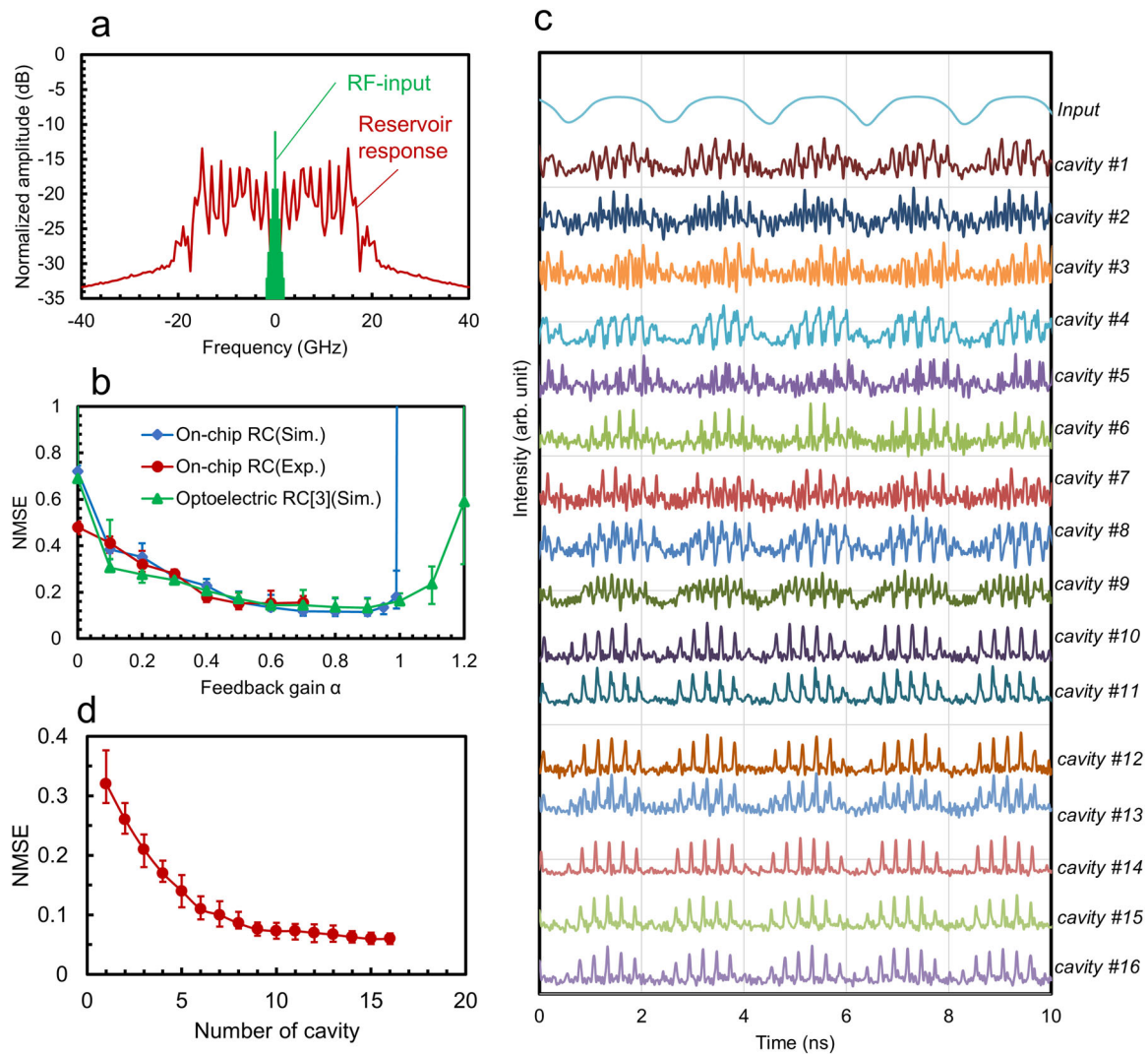


Fig. 4 Chaotic time-series expectation. **a** Observed bandwidth for radio-frequency (RF) input chaotic time-series and reservoir response. **b** Normalized mean square error (NMSE) as a function of feedback gain for single cavity experiment. **c** Reservoir response of all cavity. **d** NMSE as a function of cavity parallelism. Error bars in **b** and **d** represent one standard deviation.

to boost the computation efficiency. Figure 6a shows the WDM processing with the proposed chip. In this scheme, the optical inputs with wavelength channel spacing of $\Delta\lambda$ are multiplexed into a single optical transmission line and then input to the optical circuit. The input mask operation of our circuit is equilibrium to the frequency filtering of the up-converted signals. As shown in this figure, the filtering shape is repeated along the λ -axis due to the circulating nature of the optical phase. This feature was experimentally confirmed as shown in Fig. 3a–d. As the repeating frequency is determined by the minimum delay difference in the filter, we can generate the same optical mask function when we set the channel spacing $\Delta\lambda = c\theta$. Thus, we can use the full optical bandwidth of our device by multiplexing the multiple wavelength inputs. The feedback α_1 can be kept constant by setting the optical paths length in the MZMs to the same length. Unfortunately, the relative phase and delay length are changed due to the group delay of the circuit. However, the impact on the performance is relatively smaller than that of the other hyper parameters in RC. The output signals from the optical circuits are demultiplexed and detected by the detectors. The reservoir responses for each wavelength channel are weighted

and summed individually, which is considered as the wavelength parallel RC outputs. Thus, we can realize parallel RC on the same optical chip simultaneously. As a result, we can use the full optical bandwidth of our circuit (~ 5 THz) beyond the remaining bottleneck from the outputs bandwidth.

For the demonstration, we examined the parallel RC processing using two wavelength inputs. We used different datasets (MNIST and Fashion-MNIST datasets) for the classification task. The input spectrum is shown in Fig. 6b. The Fashion-MNIST data were modulated onto the wavelength of 1550 nm, and the MNIST data were modulated onto that of 1550.966 nm, corresponding to channel spacing of 120-GHz to set the same optical mask shape. We employed only one cavity in this demonstration. The training results are shown in Fig. 6c. The training accuracy increased almost monotonically, which suggests the success of the proposed WDM approach. The observed test accuracies were 79.2% for the MNIST and 70.1% for the Fashion-MNIST datasets. The results for the MNIST are almost comparable to those for the single-wavelength experiment [the result for the single cavity in Fig. 5d] despite the change of wavelength. This suggests that the performance degradation in the WDM approach is small. Thus,

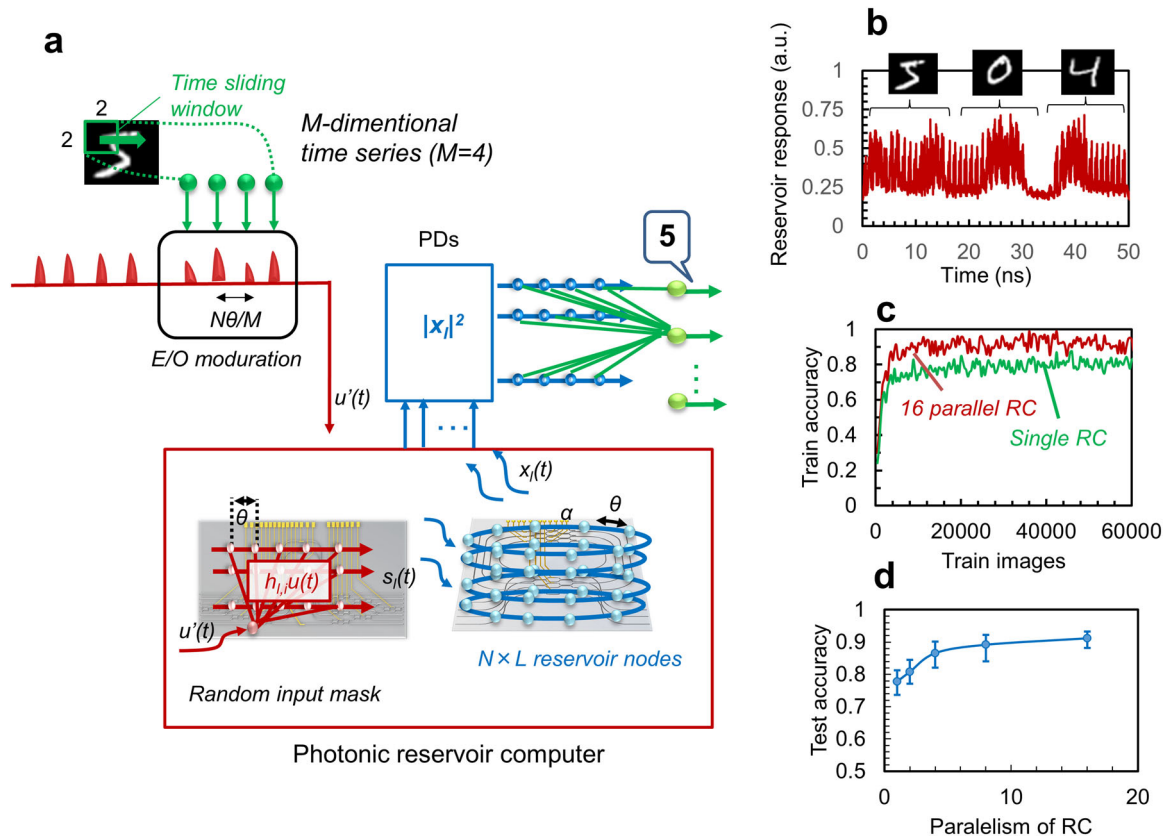


Fig. 5 Hand-written-digit recognition. **a** Processing scheme for Modified National Institute of Standards and Technology (MNIST) database. Original digit images were converted to optical time-series data, and they were processed by the input mask and reservoir optical circuits. The signals were detected by photodetectors (PDs) and classified by digital output layer. **b** Measured time response of reservoir for the center cavity ($l = 8$). **c** Training accuracy as a function of learned images. **d** Test accuracy as a function of cavity parallelism. Error bars represent one standard deviation.

we think that we can improve the accuracy to 91.3% by using the multiple cavities as shown in Fig. 5d. Although we only confirmed the feasibility for two-wavelength input due to the limitations of the WDM experimental setup, the fabricated circuit can support 40 wavelength channels over the telecom C-band at least. The results confirm the potential of ultra-wideband processing through WDM.

Discussion

Here, we discuss the computational efficiency of the proposed circuit. Although there are many indices for expressing the performance of a computing device, the multiply-accumulate per second speed ($\text{MAC}\cdot\text{s}^{-1}$) is now widely considered to be a milestone in the photonic neuromorphic computation region^{9,53,54}. Thus, we discuss this index for the reservoir computer. Our photonic circuit can compute the input and reservoir layer propagation as described in Eqs. (3) and (5). The equilibrium algorithm for our photonic reservoir computer with multiple ring cavities for the case of one-dimensional input ($M = 1$) is shown in Algorithm 1. The second term is input mask operation as described in Eq. (3). The first term and the summation of first and second terms are the reservoir operation as described as Eq. (5). Note that Ω_{ij} in the Algorithm 1 is an inter-reservoir connection, which depends on the q value (delay length of each cavity). For ease of understanding, we describe the case for $q = 1$. The computational complexity of the photonic circuit does not depend on the q value. This algorithm requires the

following MAC operations for each time step: 6 NL times multiplication for the reservoir operation (first term), $6N$ $L\sigma$ times multiplication for the mask operation, where σ is the non-zero density of h_{ij} (second term), and $(2NL)$ times for summation of the first and second terms. Note that the processing is executed on the complex space, which requires two (six times) calculations for the sum (multiplex) operation. The average σ value for the FIR filter is considered as 0.5. Thus, the total MAC for each time-step in our device is considered as 11 NL. As the operation time of each time step for the optical circuit is determined by modulation time interval T , the MAC/s can be expressed as $11L/\theta$. For the M -dimensional case, input mask operation becomes more complex, and it requires $0.2(4M - 1)\sigma NL$ times operations (see Supplementary Note 1). Thus, the $\text{MAC}\cdot\text{s}^{-1}$ for our photonic RC can be described as follows:

$$\text{MAC} \cdot \text{s}^{-1} = (4M + 7)NL/T = (4M + 7)L/\theta. \quad (8)$$

As M is an input-data-oriented value, L/θ is the important value for the device. The L value is limited to the size and device components of the optical circuit. The $T (=N\theta)$ is determined by the optical pulse width. Note that the typical single delay-line RC approach gives $\text{MAC}\cdot\text{s}^{-1}$ of only $2N/\theta$ because it does not have an optical input mask and $L = 1$ with real-valued processing.

Algorithm 1 Photonic reservoir computing with multiple ring cavity.

for $n = 0:n_{\text{max}}$ (time evolution).

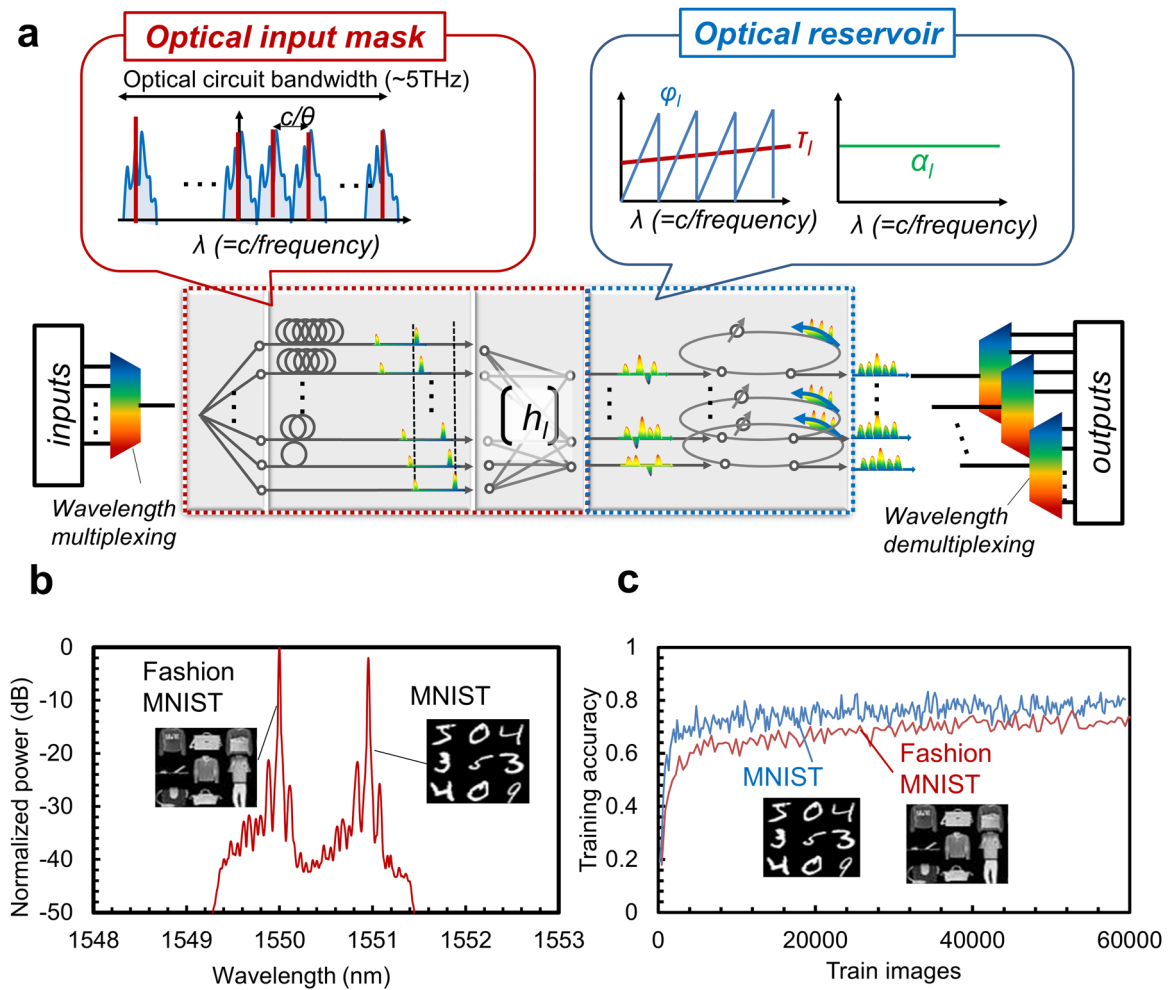


Fig. 6 Parallel processing using wavelength division multiplexing (WDM). **a** Schematic of parallel reservoir computing using WDM. As the filtering shape of input mask is repeated along the wavelength (λ) axis, we can process the WDM-input data in parallel with $c\theta$ channel spacing. **b** Spectra of WDM-inputs. **c** Training accuracy as a function of learned images.

for $l = 1:L$ (loop for multiple cavity)

$$\begin{pmatrix} x_1(n,l) \\ x_2(n,l) \\ x_3(n,l) \\ \vdots \\ x_{N-1}(n,l) \\ x_N(n,l) \end{pmatrix} = \alpha_l \exp(j\Delta\varphi_l) \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1(n-1,l) \\ x_2(n-1,l) \\ x_3(n-1,l) \\ \vdots \\ x_{N-1}(n-1,l) \\ x_N(n-2,l) \end{pmatrix} + \begin{pmatrix} h_{l,1} \\ h_{l,2} \\ \vdots \\ h_{l,N-1} \\ h_{l,N} \end{pmatrix} u'(n).$$

end
end

Figure 7 shows estimated $\text{MAC}\cdot\text{s}^{-1}$ for each wavelength ($\text{MAC}\cdot\text{s}^{-1}\cdot\lambda^{-1}$) as a function of θ for various L values. Here, we assume $M = 1$ [Fig. 7a] and $M = 4$ [Fig. 7b], which are the values we used for the chaotic series expectation and image classification tasks in this work. In our circuit, the spatially distributed delay lines, optical cavities, and optical interferometers solve Eqs. (3) and (5) in parallel, which result in much faster computation speed beyond the input RF-bandwidth. The performance reached 21.12 and 44.16 T $\text{MAC}\cdot\text{s}^{-1}$. for $M = 1$ and 4 for our fabricated circuit ($\theta = 12.5$ ps, $L = 16$). The total power consumption of our system is estimated to be 1330 W: 20 W for the optical circuit; 20 W for the detector, including the post RF amplifier; 120 W for the

optical pulse source, including the optical modulator and RF amplifier; 700 W for the oscilloscope, 180 W for the arbitrary waveform generator, and 290 W for our desk-top PC. As the values were estimated from nominal values in the catalog spec sheet, the actual total power consumption is less than this estimation. Thus, the energy efficiency of our circuit [$\text{MAC}\cdot\text{s}^{-1}$ per watt ($\text{MAC}\cdot\text{s}^{-1}\cdot\text{W}^{-1}$)] can be estimated as 15.9 and 33.2 G $\text{MAC}\cdot\text{s}^{-1}\cdot\text{W}^{-1}$ for $M = 1$ and 4. This value is for state-of-the-art electric computational devices (the present best performance is 21.108 G $\text{MAC}\cdot\text{s}^{-1}$ for MN-3⁵⁵ in June 2020). As can be seen, most of the power consumption originates from the electric devices (oscilloscope, AWG, and desk-top PC); therefore, we can reduce the power consumption much more by constructing application-specific circuits. In addition, as demonstrated in this study, the WDM technique enables the optical circuit to share the parallel data inputs. Our circuit supports the C-band (1530–1570 nm, which has ~ 5 -THz bandwidth) with wavelength-spacing of 120 GHz. Thus, we can potentially use 40 wavelength channels in our circuit, which realizes petascale optical reservoir processing (0.845 and 1.77 P $\text{MAC}\cdot\text{s}^{-1}$ for $M = 1$ and 4). These speeds are much higher than the theoretical operation speed of recent CPUs [~ 500 G $\text{MAC}\cdot\text{s}^{-1}$ (16 $\text{MAC} \times 3$ GHz $\times 10$ core)] and GPUs [~ 6 T $\text{MAC}\cdot\text{s}^{-1}$ (2 $\text{MAC} \times 1$ GHz $\times 3000$ core)]. This value is not far

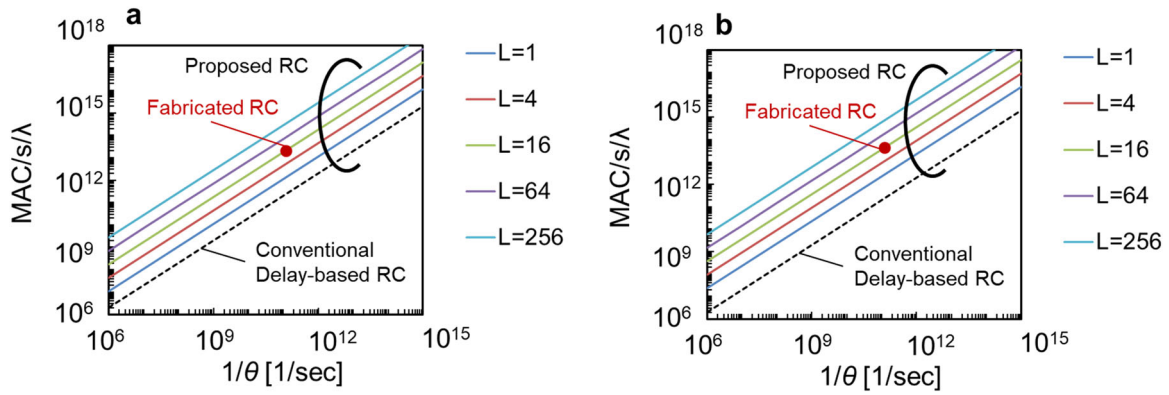


Fig. 7 Estimated processing speed. **a, b** Estimated multiply-accumulate (MAC) per second per wavelength λ ($\text{MAC s}^{-1} \lambda^{-1}$) for proposed photonic RC as a function of $1/\theta$ with **a** $M = 1$ and **b** $M = 4$.

from that for current state-of-the art of supercomputers, which ranges from 1 to 100 PMAC/s. Therefore, our approach poses a challenge to conventional Turing-von Neumann machines, and it confirms the great potential of photonic neuromorphic processing.

Conclusion

In this paper, we demonstrated photonic RC based on coherent linear photonic processors. In contrast to previous approaches, both the input signals and reservoir weights are optically encoded in the spatiotemporal domain, which enables scalable integration on a compact chip. The ultrafast optical pulse source up-converts the input signals to a higher frequency region beyond the electrical bandwidth. We also demonstrated parallel processing based on the wavelength division multiplexing (WDM), which enables the use of ultrawide optical bandwidth (>THz) as a computational resource. As a result, the computation efficiency is boosted beyond the computation bottleneck remaining from the I/O bandwidth. Experiments for the standard benchmarks showed good performance for chaotic time-series forecasting and image classifications with record-high processing speed of ~ 17.1 ns per image. The device can achieve $21.12 \text{ T MAC s}^{-1}$ for each wavelength and can reach peta-scale computation speed on a single photonic chip by using WDM.

Methods

Experimental setup. Optical pulses with a pulse width of 30 ps and repeating time T of ~ 266.7 ps ($1/120 \text{ GHz} \times 32$), corresponding to $p = 1$, were generated by using a coherent laser light source with a 1550-nm wavelength and an amplitude modulator with a 45-GHz bandwidth. The input data was generated by using an arbitrary waveform generator with a 20-GHz bandwidth and sampling rate of 60 GSa/s. The synthesizer and arbitrary waveform generator (AWG) were synchronized by using the same clock. The signal was amplified by an optical fiber amplifier to 10 dBm and input to the fabricated PLC. The mask condition and reservoir condition were randomly set by controlling the Mach-Zehnder modulators and phase shifters in the chip. The output signals were filtered by the wavelength filter. They were measure by using 50-GHz bandwidth photodetectors, and the observed RF signals were amplified by using low-noise amplifier with a 40-GHz bandwidth. The signals were acquired by using a digital storage oscilloscope with an 18-GHz bandwidth and 60 GSa/s (DSO). Thus, the measurable node count decreased to $\theta'/\theta = 1/2$ [$\theta' \sim 16.7$ ps, $N' = 16$].

Multidimensional data processing. To process the multidimensional data such as image information, RC inputs often become time-dependent vector $\mathbf{u}(t) = [u_1(t), u_2(t), \dots, u_M(t)]$. To process the such type of data, $(M \times N)$ -sized random matrix are usually employed for the input mask^{20,21}. In the previous work, it is easy to realize because the input mask was processed in the electric domain. However, there are no reports for input masking for photonic implementation. Here, we consider the method for input masking by using the same photonic convolutional filter. In our device, there are two ways to process the multidimensional inputs. The first way is simply inputs the signal to multiple input ports like the conventional

on-chip ANNs⁴. However, this architecture requires multiple laser sources, which requires complex and expensive hardware update; e.g., for MNIST task, which is a standard image classification task, requires 28 inputs. It requires 28 laser inputs with optoelectric modulator array. In addition, when their relative phases and/or intensity are fluctuated, the optical output is also fluctuated due to interferometric condition changes. It leads the poor accuracy of the system. Thus, we adapt the second choice, that is time division multiplexing as shown in Fig. 1. This method does not require any hardware update. In this method, the multi-dimensional inputs are arranged in different time slot. To realize it, the radio-frequency (RF) input signal $\mathbf{u}(t)$ is set to following form.

$$\mathbf{u}(t) = [u_1(t), u_2(t), \dots, u_M(t), u_1(t+1), u_2(t+1), \dots, u_M(t+1), u_1(t+2), \dots], \tag{9}$$

where the modulation interval T' is set to $T' = T/M$. The repeating time of optical pulse is also set to T' . Then, the $\delta(t)$ is described as follows;

$$\delta(n, i) = \begin{cases} 1 & (i = N/M) \\ 0 & (i \neq N/M) \end{cases}, \tag{10}$$

Then, the modulated optical pulse $u'(n)$ for discretized time $t(n, i)$ can be described as the orthonormal basis of input vector

$$\begin{pmatrix} u'(n, 1) \\ \vdots \\ u'(n, \frac{N}{M}) \\ \vdots \\ u'(n, \frac{2N}{M}) \\ \vdots \\ u'(n, N) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ u_1(n) \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ u_2(n) \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ u_M(n) \end{pmatrix} \tag{11}$$

To satisfy the synchronization between RF-input and optical pulse, N/M should be set to integer value. The reservoir processing and post processing are same as the way for one-dimensional case. Thus we can process the multidimensional data by using the same optical circuit.

Training of readout. The parameter of RC is only readout weight ω , which forms a linear combination of the reservoir states. For the training, we collected the reservoir response $|x|^2$ from each photodetector during training period, T_{train} . Then, we obtained the $N_r \times T_{\text{train}}$ state matrix S , where N_r is the number of reservoir nodes (in our case, $N_r = NL$). The goal of the optimization is to find the ω in such a way that the actual output $\mathbf{Y} = \omega S$ matches the desired output $\mathbf{Y}_{\text{teacher}}$ as close as possible in the least-squares sense. This is a linear problem. For the offline case, the optimum ω is calculated by using the Moore-Penrose pseudoinverse S^\dagger of state matrix S :

$$S^\dagger = (S^T S)^{-1} S^T, \tag{12}$$

$$\omega = \mathbf{Y}_{\text{teacher}} (S^\dagger - \lambda S^T \mathbf{I})^{-1}, \tag{13}$$

where λ is the parameter for Tikhonov regularization to avoid overfitting, and \mathbf{I} is an N_r -dimensional unity matrix. For the online training, we divided the training data into chunks (mini-batches). The readout was updated by the following simple gradient descent method.

$$\omega \leftarrow \omega + \varepsilon (\mathbf{Y}_{\text{teacher}} - \mathbf{Y}) S, \tag{14}$$

where ε is learning rate.

Data availability

The data for the current study are available from the corresponding author upon reasonable request.

Received: 29 July 2020; Accepted: 16 December 2020;

Published online: 10 February 2021

References

- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
- Merolla, P. A. et al. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* **345**, 668–673 (2014).
- Misra, J. & Saha, I. Artificial neural networks in hardware: a survey of two decades of progress. *Neurocomputing* **74**, 239–255 (2010).
- Shen, Y. et al. Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **11**, 441–446 (2017).
- Lin, X. et al. All-optical machine learning using diffractive deep neural networks. *Science* <https://doi.org/10.1126/science.aat8084> (2018).
- Hughes, T. W., Minkov, M., Shi, Y. & Fan, S. Training of photonic neural networks through in situ backpropagation and gradient measurement. *Optica* **5**, 864–871 (2018).
- Hamerly, R., Bernstein, L., Sludds, A., Soljačić, M. & Englund, D. Large-scale optical neural networks based on photoelectric multiplication. *Phys. Rev. X* **9**, 021032 (2019).
- Feldmann, J., Youngblood, N., Wright, C. D., Bhaskaran, H. & Pernice, W. H. P. All-optical spiking neuromorphic networks with self-learning capabilities. *Nature* **569**, 208–214 (2019).
- Feldmann, J. et al. Parallel convolution processing using an integrated photonic tensor core. *Nature* **589**, 52–58 (2021).
- Tait, A. N. et al. Neuromorphic photonic networks using silicon photonic weight banks. *Sci. Rep.* **7**, 7430 (2017).
- Larger, L. et al. Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing. *Opt. Express* **20**, 3241 (2012).
- Brunner, D., Soriano, M. C., Mirasso, C. R. & Fischer, I. Parallel photonic information processing at gigabyte per second data rates using transient states. *Nat. Commun.* **4**, 1364 (2013).
- Vinckier, Q. et al. High-performance photonic reservoir computer based on a coherently driven passive cavity. *Optica* **2**, 438 (2015).
- Vandoorne, K. et al. Experimental demonstration of reservoir computing on a silicon photonics chip. *Nat. Commun.* **5**, 3541 (2014).
- Kuriki, Y., Nakayama, J., Takano, K. & Uchida, A. Impact of input mask signals on delay-based photonic reservoir computing with semiconductor lasers. *Opt. Express* **26**, 5777–5788 (2018).
- Harkhoe, K., Verschaffelt, G., Katumba, A., Bienstman, P. & Van der Sande, G. Demonstrating delay-based reservoir computing using a compact photonic integrated chip. *Opt. Exp.* **28**, 3086 (2020).
- Bueno, J. et al. Reinforcement learning in a large-scale photonic recurrent neural network. *Optica* **5**, 756–760 (2018).
- Nakajima, M., Inubushi, M., Goh, T. & Hashimoto, T. Coherently driven ultrafast complex-valued photonic reservoir computing. In *Proceedings of Conference on Lasers and Electro-Optics (CLEO) paper SM1C.4* (Optical Society of America, Washington D.C., 2018).
- Argyris, A., Bueno, J. & Fischer, I. Photonic machine learning implementation for signal recovery in optical communications. *Sci. Rep.* **8**, art. 8487 (2018).
- Larger, L. et al. High-speed photonic reservoir computing using a time-delay based architecture: million words per second classification. *Phys. Rev. X* **7**, 11015 (2017).
- Antonik, P., Marsal, N. & Rontani, D. Large-scale spatiotemporal photonic reservoir computer for image classification. *IEEE J. Sel. Top. Quantum Electron.* **26**, 1–12 (2020).
- Goodman, J. Four decades of optical information processing. *Opt. Photon. News* **2122**, 11–15 (1991).
- Bocker, R. P. Matrix multiplication using incoherent optical techniques. *Appl. Opt.* **13**, 1670–1676 (1974).
- Psaltis, D. & Lin, S. Optoelectronic implementations of neural networks. *IEEE Commun. Mag.* **27**, 1112, 37–40 (1989).
- Tomson, D. et al. Roadmap on silicon photonics. *J. Opt.* **18**, 073003 (2016).
- Harris, N. C. et al. Linear programmable nanophotonic processors. *Optica* **5**, 1623–1631 (2018).
- Carolan, J. et al. Universal linear optics. *Science* **349**, 711–716 (2015).
- Jaeger, H. & Haas, H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* **304**, 78–79 (2004).
- Maass, W., Natschläger, T. & Markram, H. “Real-time computing without stable states: a new framework for neural computation based on perturbations.”. *Neural Comput.* **14**, 2531 (2002).
- Lukoševičius, M. & Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **3**, 127 (2009).
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. *Parallel Distributed Processing* Vol. 1 (MIT, Cambridge, 1986).
- The 2006/07 forecasting competition for neural networks & computational intelligence. <http://www.neural-forecasting-competition.com/NN3/> (2006).
- Antonik, P., Marsal, N., Brunner, D. & Rontani, D. Human action recognition with a large-scale brain-inspired photonic computer. *Nat. Mach. Intell.* **1**, 530–537 (2019).
- Tanaka, G. et al. “Recent advances in physical reservoir computing: a review. *Neural Netw.* **115**, 100–123 (2019).
- Appeltant, L. et al. Information processing using a single dynamical node as complex system. *Nat. Commun.* **2**, 468 (2011).
- Ortin, S. & Pesquera, L. Reservoir computing with an ensemble of time-delay reservoirs. *Cogn. Comput.* **9**, 327–336 (2017).
- Sugano, C., Kanno, K. & Uchida, A. Reservoir computing using multiple lasers with feedback on a photonic integrated circuit. *IEEE IEEE J. Sel. Top. Quantum Electron.* **26**, 1500409 (2020).
- Aratake, A. High reliability of silica-based 1 × 8 optical splitter modules for outside plant. *J. Lightwave Technol.* **34**, 27 (2016).
- Takahashi, H. High performance planar lightwave circuit devices for large capacity transmission. *Opt. Exp.* **19**, B173 (2011).
- Sakamoto, J., Goh, T., Katayose, S., Kasahara, R. & Hashimoto, T. Shape-optimized multi-mode interference for a wideband visible light coupler. *Opt. Commun.* **443**, 221 (2019).
- Carolan, J. et al. Universal linear optics. *Science* **349**, 711 (2015).
- Akatsuka, T. et al. Optical frequency distribution using laser repeater stations with planar lightwave circuits. *Opt. Exp.* **28**, 9186 (2020).
- Sasayama, K., Okuno, M. & Habara, K. Coherent optical transversal filter using silica-based waveguides for high-speed signal processing. *J. Lightwave Technol.* **9**, 1225 (1991).
- Clements, W. R. et al. Optimal design for universal multiport interferometers. *Optica* **3**, 1460–1465 (2016).
- Reck, M., Zeilinger, A., Bernstein, H. J. & Bertani, P. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.* **73**, 58–61 (1994).
- Miller, D. A. B. Self-configuring universal linear optical component. *Photon. Res.* **1**, 1–15 (2013).
- Miller, D. A. B. Perfect optics with imperfect components. *Optica* **2**, 747–750 (2015).
- A. S. Weigend, A. S. & Gershenfeld, N. A. Time series prediction: forecasting the future and understanding the past. <http://www-psychem.stanford.edu/~andreas/Time-Series/SantaFe.html> (1993).
- LeCun, Y., Cortes, C. & Burges, C. J. C. The MNIST database of handwritten digits (1998).
- Seok, T. J., Kwon, K., Henriksson, J., Luo, J. & Wu, M. C. Wafer-scale silicon photonic switches beyond die size limit. *Optica* **6**, 490–494 (2019).
- Kominato, T. et al. Extremely low-loss (0.3 dB/m) and long silica-based waveguides with large width and clothoid curve connection. In *Proceeding of European Conference on Optical Communication (ECOC) paper Tu1.4.3* (Optical Society of America, Washington D.C., 2004).
- Lee, H., Chen, T., Li, J., Painter, O. & Vahala, K. J. Ultra-low-loss optical delay line on a silicon chip. *Nat. Commun.* **3**, 867 (2012).
- Nahmias, M. A. et al. Photonic multiply-accumulate operations for neural networks. *IEEE J. Sel. Top. Quantum Electron.* **26**, 7701518 (2019).
- Totović, A., Dabos, G., Passalis, N., Tefas, A. & Pler, N. Femtojoule per MAC neuromorphic photonics: an energy and technology roadmap. *J. Sel. Top. Quantum Electron.* **26**, 8800115 (2020).
- Top500 Project, *Green500 June 2020* <https://www.top500.org/lists/green500/2020/06/> (2020).

Acknowledgements

The Authors are grateful to Takashi Goh, Masanobu Inubushi, Shiori Konisho, Soichi Oka, and Satoshi Shigematsu for useful discussions. We also thank Fumikatsu Sugimoto for technical assistance with the experiments.

Author contributions

M.N. conceived the basic concept of the presented on-chip reservoir computer. T.H. supervised the project. M.N. designed and characterized the optical circuits. M.N. and K.T. carried out the experiments for the reservoir computing. M.N. wrote the initial draft of the manuscript. All the author analyzed and discussed the results and contributed to write the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42005-021-00519-1>.

Correspondence and requests for materials should be addressed to M.N.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021