Supplementary Material for

**Integrated photonic metasystem for image classifications at telecommunication wavelength**

Zi Wang, Lorry Chang, Feifan Wang, Tiantian Li and Tingyi Gu\*

Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19711, USA

\*Email: tingyigu@udel.edu

S1: Integrated photonic frameworks for matrix operation

S2: PyTorch-based training of diffractive optical network

S3: Broadband operation of the integrated spatial pattern classifier

S4: Metasystem scalability and accuracy

**Supplementary note S1: Integrated photonic frameworks for matrix operation**

Vector-by-matrix multiplication (VMM) is one of the fundamental operations in the accelerator hardware [S1]. Table S1 compares the VMM power efficiency, throughput, and footprint of integrated photonic circuits based on microring resonators (WDM), Mach Zehnder interferometer (MZI) and metasystem (this work).

Table SI: Integrated photonic frameworks for VMM.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **WDM** [S2-4] | **MZI** [S5] | **This work** |
| Method for matrix operation | Multi-wavelength modulation and summation | Singular value decomposition | Diffraction equation |
| Signal format | Temporal | Temporal | Spatial |
| Coherency | Incoherent | Coherent | Coherent |
| Device architecture | Micro-ring resonator weight banks | Mach-Zehnder interferometer | Cascaded metasurfaces |
| Matrix-vector operation | Reconfigurable | Reconfigurable | Pre-trained |
| Electronic layer | Required | Required | No |
| Weight matrix | 16×16 [S3] | 4×4 [S5] | 450×2 |
| Footprint (mm2) | 16 | 0.75 | 0.045 |
| Throughput (Tb/s) | 1 | 1 | 1000 |
| Operational power | 17fJ per MAC | 1pJ per FLOP | 0 |

**Supplementary note S2: PyTorch-based training of diffractive optical network**

The metasystem is designed by the PyTorch framework (Facebook, Inc.) [S6]. A spatial pattern classifier is used to illustrate the design process (Fig. S1a). Additional random phase noise was added in the input array and each hidden layer to represent the fabrication variation (orange blocks in Fig. S1) and coupling (yellow block in Fig. S1). The metasystem was trained to be robust against experimental variations [S7].

In each pixel of the letter image, the input information was encoded as the amplitude of the input electromagnetic (EM) field with their default phase values set at zero. Examples of the training and testing data of the inputs are illustrated in Fig. S1b. As shown in Fig. S1a, the input electric fields are multiplied with the complex-valued modulation generated by random phase noise (uniform distribution on the interval ) and the desired phase shift of each layer, and after the 2D free-space propagation, the output field of each layer is fed to the next layer as the input field. For the last layer, we calculated the cross-entropy loss between the intensity of the output field and the ground truth of the input data.

We used the Adam algorithm [S8] in back-propagation and update the neurons of the network to minimize the loss function (Fig. S1c). The whole forward/ backward propagation process can be calculated by matrix multiplication. The electric field of the layer can be calculated by

where is the input electric field of the -th neuron in the layer , is the complex transmission coefficient of the -th neuron in the layer , and is the propagation coefficient from the -th neuron in the layer to the -th neuron in the layer which can be derived from Eqs. (2) in the main text, as shown in Fig. S1c.



**Figure S1. Framework of designing a metasystem classifier robust against fabrication variations.** (a)The complex electromagnetic (EM) field propagates freely in the x-y plane. The complex EM field of each layer’s output is multiplied with the phase modulation by the metasurface layer (orange) and random phase noise (blue). The multiplied signal is propagated and transferred to the next layer (blue arrows). (b) Example training and testing input data of “Y” with different amplitude noise**.** (c)Details of the in-plane free propagation module (grey shaded area in a). The inter-layer connectivity is described by diffraction equations.

**Inter-layer connectivity by wave propagation:** As shown in Eqs. (2) in the main text, the propagation coefficient is the function of the distance and the distance along direction. Because the layers in our system are in the same length and parallel to each other, the matrix became a Toeplitz matrix [S9], as shown in the following equation:

Here we create a circulant matrix using a vector which is composed of the elements in the matrix as following:

and

where matrix is a Toeplitz matrix as shown in Eqs. (S-2). We need to calculate the multiplication of the matrix and an vector as shown in Eqs. (S-1). Here we show a fast way to do the calculation by calculating

and get the first rows of the result which is . We can calculate by following:

where *fft* and *ifft* represent Fourier and inverse Fourier transforms, respectively. is the vector in Eqs. (S-3), and is a vector composed with the vector and a zero vector with the dimension of . Instead of using the matrix with the dimension of , we used the matrix with the dimension of which reduced the requirement for both calculation and data storage. As the propagation matrix W is the most computationally expensive component, the proposed system reduces the memory requirement from O(N2) to O(N).

**Loss function and optimizer definition:** The softmax of the output intensity is defined as:

and the cross-entropy loss is defined as

where is the ground truth of the input data, and is the output intensity. The gradient can be derived as

and

where is the phase shifts of the neurons, is the complex conjugate of .

**Supplementary note S3: Broadband operation of the integrated spatial pattern classifier**

We verified the broadband property of the classifier neural network in both simulations and experiments as shown in Fig. S2. Even though the classifier neural network is designed around 1550nm, it can also work with the input wavelength of 1520nm and 1620nm. The overall output intensity for the input wavelength at 1500nm and 1600nm is lower than the input wavelength at 1550nm.

 

**Figure S2. Broadband operation of the classifier.** FDTD simulated spatial distribution of light on the output plane (grey lines) compared with experimental results (squares with error bar), as the wavelength of the input c.w. laser set at (a) 1520 nm, (b) 1550 nm and (c) 1620 nm. The experimentally measured values are marked as blue, red, and yellow dots for the output port representing ‘X’, ‘Y’, ‘Z’ respectively. The corresponding confusion matrixes of the experiments are shown in (d), (e) and (f).

**Supplementary note S4: Metasystem scalability and accuracy**

The scalability of the design algorithm is verified by a more complicated system of a Modified National Institute of Standards and Technology (MNIST) handwritten digit database. The accuracy of the output is evaluated by Python simulations. Fig. S3 compares the loss functions and truth tables of the small scale (15-pixel inputs for ‘XYZ’) and large scale (784-pixel inputs for MNIST). The proposed system in section S1 is efficient. The accuracy of the small-scale system converges after 1 Epoch and only 3 epoch brings the accuracy of a larger system to be 96%. Compared with free-space diffractive neural network [S10-S11], our structure has fewer neurons and faster training speed.



**Figure S3. Letters and handwriting digits (MNIST) recognition training for phase-only ID2NN.** (a) Training convergence plots. Inset: example of input and (b) confusion matrix for MNIST handwriting digits with 784 pixels, and (c-d) letters images with 15 pixels.

The system accuracy not only depends on the number of the neuorn, but also the depth of the neuron network.We investigated the system accuracy versus neuron number per layer, the number of layers, and the interlayer spacing (Fig. S4). A metasystem with 5 layers and 4000 neurons per layer achieves accuracy of more than 96%. With the same number of neurons, the metasystem accuracy significantly improves in a multi-layer system (Fig. S4a). The inter-layer spacing increases the connectivity, and thus the system accuracy (Fig. S4b). At least 1000 neurons per layer and 3 layers are needed for maintaining high accuracy of 92%. A trade-off between the device accuracy and the metasystem footprint is observed.



**Figure S4. Scalability of the integrated diffractive optical network for MNIST. (a)** Simulated accuracy versus the same total number of the weight elements, with different number of layers (N). The inter-layer distance is 100 µm. **(b)** Accuracy versus layer number with inter-layer distance (D) of 50, 100, 250 and 500 µm. The number of phase shifters per layer is 1000.

**References**

[S1] Kitayama, K. et al. Novel frontier of photonics for data processing—Photonic accelerator. APL Photonics 4, 090901 (2019).

[S2] Xu, X. et al. 11 TOPS photonic convolutional accelerator for optical neural networks. Nature 589, 44–51 (2021).

[S3] Feldmann, J. et al. Parallel convolutional processing using an integrated photonic tensor core. Nature 589, 52–58 (2021).

[S4] Tait, A. N. et al. Neuromorphic photonic networks using silicon photonic weight banks. Sci. Rep. 7, 7430 (2017).

[S5] Shen, Y. et al. Deep learning with coherent nanophotonic circuits. Nat. Photonics 11, 441–446 (2017).

[S6] Paszke, A. et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. arXiv:1912.01703 (2019).

[S7] Pai, S. et al. Parallel Programming of an Arbitrary Feedforward Photonic Network. IEEE J. Select. Topics Quantum Electron. 26, 1–13 (2020).

[S8] Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 (2017).

[S9] Golub, G. H. & Van Loan, C. F. Matrix computations. (The Johns Hopkins University Press, 2013).

[S10] Backer, A. S. Computational inverse design for cascaded systems of metasurface optics. Opt. Express 27, 30308 (2019).

[S11] Zhu, D., Liu, Z., Raju, L., Kim, A. S. & Cai, W. Multifunctional Meta-Optic Systems: Inversely Designed with Artificial Intelligence. arXiv:2007.00130 [physics] (2020).