# Supplementary Information:
# Light-driven molecular reorientation for large-scale photonic in-memory computing

Sijia Liu[1†], Xiaonan Hu[2,3†], Peng Chen[1†*], Xinyuan Fang[2,3*], Fangtianyu Chen[1], Kenong Li[1], Rui Sun[1], Shihui Ding[1], Yiheng Zhang[1], Keyao Li[2,3], Baoli Li[2,3], Shi-Jun Ge[1], Feng Wang[1], Min Gu[2,3*] & Yan-Qing Lu[1*]

[1]National Laboratory of Solid State Microstructures, Key Laboratory of Intelligent Optical Sensing and Manipulation, College of Engineering and Applied Sciences, and Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, China

[2]School of Artificial Intelligence Science and Technology, University of Shanghai for Science and Technology, Shanghai 200093, China

[3]Institute of Photonic Chips, University of Shanghai for Science and Technology, Shanghai 200093, China

*E-mail: chenpeng@nju.edu.cn, xinyuan.fang@usst.edu.cn, gumin@usst.edu.cn, yqlu@nju.edu.cn

[†]These authors contributed equally to this work.

# Table of Contents

**Supplementary Note 1: Theoretical analysis of polarization conversions via LCs**

Liquid crystals (LCs) are intermediate materials between liquids and solid crystals. As natural birefringent materials, LCs can induce phase retardation between extraordinary light and ordinary light, which can be expressed as

$$\Gamma = \frac{2\pi\left(n_{\mathrm{eff}} - n_{\mathrm{o}}\right)d}{\lambda}, \tag{S1}$$

where $d$ is the thickness of the LC layer, $\lambda$ is the wavelength of incident light, $n_{\mathrm{o}}$ and $n_{\mathrm{eff}}$ are the ordinary and effective extraordinary refractive indices of LCs, respectively. Among them, $n_{\mathrm{eff}}$ is dependent on LCs' tilt angle $\theta$ which can be electrically tuned. Thanks to their optical anisotropy, LCs can modulate vectorial light fields variously[1]. Considering LCs with an orientation angle $\alpha$ and a phase retardation $\Gamma$, their modulation effects on vectorial light can be represented by the following Jones matrix:

$$\begin{aligned}\mathbf{J} &= \mathbf{R}(-\alpha)\cdot\begin{bmatrix}\exp\left(-i\Gamma/2\right) & 0\\ 0 & \exp\left(i\Gamma/2\right)\end{bmatrix}\cdot\mathbf{R}(\alpha)\\ &= \cos\frac{\Gamma}{2}\mathbf{I} - i\sin\frac{\Gamma}{2}\begin{bmatrix}\cos\left(2\alpha\right) & \sin\left(2\alpha\right)\\ \sin\left(2\alpha\right) & -\cos\left(2\alpha\right)\end{bmatrix}\end{aligned}, \tag{S2}$$

where $\mathbf{R}$ is the rotation matrix, $\mathbf{I}$ is the identity matrix, and $i^2 = -1$.

An arbitrarily polarized incident beam can be expressed by a Jones vector with the following expression:

$$\boldsymbol{E}_{\mathrm{in}} = \begin{bmatrix}E_x\\ E_y\end{bmatrix} = \begin{bmatrix}\cos\psi\\ \mathrm{e}^{i\delta}\sin\psi\end{bmatrix}. \tag{S3}$$

$E_x$ and $E_y$ refer to the horizontal and vertical polarization components, respectively. $\psi$ is the orientation angle of the polarization ellipse and $\delta$ is the phase difference between $E_y$ and $E_x$. Using the Jones matrix calculation method, the vectorial light field after modulation can be expressed by the Jones vector below:

$$\boldsymbol{E}_{\mathrm{out}} = \mathbf{J}\boldsymbol{E}_{\mathrm{in}} = \begin{bmatrix}\cos\frac{\Gamma}{2}\cos\psi - i\sin\frac{\Gamma}{2}\left(\cos 2\alpha\cos\psi + \sin 2\alpha\mathrm{e}^{i\delta}\sin\psi\right)\\ \cos\frac{\Gamma}{2}\mathrm{e}^{i\delta}\sin\psi - i\sin\frac{\Gamma}{2}\left(\sin 2\alpha\cos\psi - \cos 2\alpha\mathrm{e}^{i\delta}\sin\psi\right)\end{bmatrix}. \tag{S4}$$

Apart from the Jones vector, polarization can also be represented by the so-called Stokes vector, which contains the four Stokes parameters of light[2]. Normalized Stokes parameters are exactly the coordinates on the Poincaré sphere, offering a prominent geometric representation of polarization states. The relation between the Stokes parameters and the vectorial light field (elements of the Jones vector) is as follows:

$$\begin{cases} S_0 = |E_x|^2 + |E_y|^2 \\ S_1 = |E_x|^2 - |E_y|^2 \\ S_2 = 2\,\mathrm{Re}\left(E_x E_y^*\right) \\ S_3 = -2\,\mathrm{Im}\left(E_x E_y^*\right) \end{cases}. \tag{S5}$$

To explore the polarization conversion trajectory on the Poincaré sphere, the Mueller matrix, which can directly transform the Stokes vector, is introduced to describe the vectorial light modulation. The Mueller matrix $\mathbf{M}$ of a specific polarization converter can be derived from its Jones matrix[3]:

$$\mathbf{M} = \mathbf{U} \cdot (\mathbf{J} \otimes \mathbf{J}^*) \cdot \mathbf{U}^{-1}, \tag{S6}$$

where $\otimes$ denotes the Kronecker product, $\mathbf{J}^*$ is the complex conjugate of $\mathbf{J}$, and $\mathbf{U}$ is the transformation matrix defined below:

$$\mathbf{U} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix}. \tag{S7}$$

Accordingly, the polarization conversions via LCs can be expressed by the following Mueller matrix:

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2(2\alpha) + \sin^2(2\alpha)\cos\Gamma & \sin(2\alpha)\cos(2\alpha)(1-\cos\Gamma) & \sin(2\alpha)\sin\Gamma \\ 0 & \sin(2\alpha)\cos(2\alpha)(1-\cos\Gamma) & \sin^2(2\alpha) + \cos^2(2\alpha)\cos\Gamma & -\cos(2\alpha)\sin\Gamma \\ 0 & -\sin(2\alpha)\sin\Gamma & \cos(2\alpha)\sin\Gamma & \cos\Gamma \end{bmatrix}. \tag{S8}$$

Interestingly, this matrix corresponds to the rotational transformation of the Stokes parameters on the Poincaré sphere. The rotation axis is determined by the LC orientation $\alpha$, whose direction vector can be expressed as

$$\boldsymbol{n} = \left(\cos 2\alpha, \sin 2\alpha, 0\right). \tag{S9}$$

The rotation angle is equal to the phase retardation $\Gamma$. Therefore, the polarization conversion trajectory on the Poincaré sphere should be a circular arc rotating around the rotation axis $\boldsymbol{n}$. Under the full-wave condition (e.g., $\Gamma = 0$), the Mueller matrix becomes an identity matrix, keeping the polarization state unchanged. Under the half-wave condition (e.g., $\Gamma = \pi$), the trajectory spans half of a circle around axis $\boldsymbol{n}$ on the Poincaré sphere, bringing about specific polarization conversions depending on both $\alpha$ and the incident polarization state.

In the proposed large-scale photonic in-memory computing platform, the LC orientation $\alpha$ can be dynamically programmed to change the Mueller matrix (also regarded as changing the rotation axis $\boldsymbol{n}$), leading to time-varying polarization conversions. Such variation can persist without external power, thereby acting as the nonvolatile photonic memory. Furthermore, a greater variety of reconfigurable polarization converters can be introduced in future works through more flexible manipulation of LC structures. For example, by fixing the LC orientation on one substrate and rewrite the orientation on the other, twisted LC structures can be formed to induce unique polarization effects such as the adiabatic following. Utilizing the LC guest-host effect, the spatial programming of polarizer-like vectorial neurons can even be achieved. These new strategies are expected to open up new possibilities for reconfigurable vectorial DNNs and photonic in-memory computing.

**Supplementary Note 2: Definitions of loss functions for different vectorial DNNs**

This section aims to provide the loss functions for the three major functionalities reported in the manuscript, including intelligent polarization encoding, polarization-perceptive computing, and parallel image classification. Loss functions are the keys to optimizing vectorial DNN models. For different target functionalities, various loss functions can be defined to quantify the difference between the output vectorial light field and the ground truth. According to the defined loss function, error backpropagation learning can be performed based on gradient descent to minimize the loss function and thus optimize the computing performance.

Considering multiple layers of diffraction and modulation, the forward propagation model of vectorial DNNs is formulated as Equation (1) in the manuscript. In addition, the polarization detection process can also be incorporated into this model. The intensities filtered out at different polarization directions can be variously designed as the criterions for inference. Assuming the polarizer orientation angle is $\theta_{\mathrm{p}}$, the filtered light field $\begin{bmatrix}\boldsymbol{E}_x^{\mathrm{p}} & \boldsymbol{E}_y^{\mathrm{p}}\end{bmatrix}^{\mathrm{T}}$ (vectorized) can be calculated through multiplying the output vectorial light field $\begin{bmatrix}\boldsymbol{E}_x^{N+1} & \boldsymbol{E}_y^{N+1}\end{bmatrix}^{\mathrm{T}}$ (vectorized; see its expression in Equation (1) in the manuscript) by the Jones matrix of the polarizer:

$$\begin{aligned}\begin{bmatrix}\boldsymbol{E}_x^{\mathrm{p}} \\ \boldsymbol{E}_y^{\mathrm{p}}\end{bmatrix} &= \begin{bmatrix}\cos^2\theta_{\mathrm{p}}\mathbf{I} & \sin\theta_{\mathrm{p}}\cos\theta_{\mathrm{p}}\mathbf{I} \\ \sin\theta_{\mathrm{p}}\cos\theta_{\mathrm{p}}\mathbf{I} & \sin^2\theta_{\mathrm{p}}\mathbf{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{E}_x^{N+1} \\ \boldsymbol{E}_y^{N+1}\end{bmatrix} \\ &= \begin{bmatrix}\cos^2\theta_{\mathrm{p}}\boldsymbol{E}_x^{N+1} + \sin\theta_{\mathrm{p}}\cos\theta_{\mathrm{p}}\boldsymbol{E}_y^{N+1} \\ \sin\theta_{\mathrm{p}}\cos\theta_{\mathrm{p}}\boldsymbol{E}_x^{N+1} + \sin^2\theta_{\mathrm{p}}\boldsymbol{E}_y^{N+1}\end{bmatrix}\end{aligned}, \tag{S10}$$

where $\mathbf{I}$ is the identity matrix. The filtered optical intensity can be expressed as

$$\begin{aligned}\boldsymbol{O}(\theta_{\mathrm{p}}) &= \left|\boldsymbol{E}_x^{\mathrm{p}}\right|^2 + \left|\boldsymbol{E}_y^{\mathrm{p}}\right|^2 \\ &= \left|\cos^2\theta_{\mathrm{p}}\boldsymbol{E}_x^{N+1} + \sin\theta_{\mathrm{p}}\cos\theta_{\mathrm{p}}\boldsymbol{E}_y^{N+1}\right|^2 + \left|\sin\theta_{\mathrm{p}}\cos\theta_{\mathrm{p}}\boldsymbol{E}_x^{N+1} + \sin^2\theta_{\mathrm{p}}\boldsymbol{E}_y^{N+1}\right|^2\end{aligned}, \tag{S11}$$

which varies with the polarizer direction $\theta_{\mathrm{p}}$.

In the case of a 256 × 256 vectorial DNN unit, $\boldsymbol{O}(\theta_{\mathrm{p}})$ should be a 65536-element column vector with each element representing the polarization-filtered intensity of a specific pixel. Since the detection area is not necessarily the same size as the pixel, we can extract the overall intensity within a certain detection area by multiplying a selection vector $\boldsymbol{S}$ which is the same size as $\boldsymbol{O}(\theta_\mathrm{p})$. Its element equals 1 if it is inside the detection area and 0 if it is outside. In this work, the detection area is set to 10 × 10 pixels (pixel pitch: 8 μm), so the selection vector $\boldsymbol{S}$ has 100 non-zero elements. Accordingly, the sum of intensity within the detection area can be expressed as

$$o_\mathrm{sum} = \boldsymbol{S}^\mathrm{T}\boldsymbol{O}\left(\theta_\mathrm{p}\right) = \sum_i S_i O_i, \tag{S12}$$

where $S_i$ and $O_i$ are the $i$-th elements of $\boldsymbol{S}$ and $\boldsymbol{O}(\theta_\mathrm{p})$, respectively.

## (1) Intelligent polarization encoding

To achieve image classification in a single detection area via polarization encoding, we used the softmax cross-entropy (SCE) loss as our loss function for training, which can be written as

$$L_\mathrm{SCE} = -\sum_{j=1}^{M_\mathrm{c}} y_j \log p_j \,. \tag{S13}$$

$M_\mathrm{c}$ is the number of categories, $y_j$ is the one-hot encoded ground truth, and $p_j$ is the predicted probability for class $j$. The expression of $p_j$ is as follows:

$$p_j = \mathrm{Softmax}(o_j) = \frac{\mathrm{e}^{o_j}}{\sum_{i=1}^{M_\mathrm{c}} \mathrm{e}^{o_i}}, \tag{S14}$$

where $o_j$ is the output intensity for class $j$. For the handwritten digit classification task demonstrated in Fig. 2d-h, $o_j$ is defined as

$$o_j = \boldsymbol{S}^\mathrm{T}\boldsymbol{O}\left(\frac{j-1}{M_\mathrm{c}}\pi\right), \tag{S15}$$

where $M_\mathrm{c} = 4$ and $\boldsymbol{S}$ is the selection vector for the single detection area. By minimizing the loss function in Equation (S13), digits 1 to 4 can be encoded into the vicinity of four preset linear polarization states.

6000 training images and 2000 test images were used in the training process, and a random sub-dataset of 200 test images was used in the experiment. The batch size was set as 8 and the learning rate was set as 0.01. The test accuracy in simulation reaches 96.45% after 6 epochs of training (Fig. S3).

## (2) Polarization-perceptive computing

To achieve the polarization-perceptive classification in Fig. 4, nine categories of images in three polarization states (0°, 60°, and 120°) are allocated to three detection areas (Fig. 4a). A fixed horizontal analyser is added to ensure the polarization sensing ability. For both polarization multiplexing and polarized object perception, the final category is inferred by identifying the detection area with the highest intensity.

We still adopted the SCE loss shown in Equation (S13) and (S14) for training, but $o_j$ was revised as

$$o_j = \boldsymbol{S}_j^{\mathrm{T}} \boldsymbol{O}\left(\theta_{\mathrm{p}}\right), \tag{S16}$$

where $\boldsymbol{S}_j$ is the selection vector for the detection area corresponding to class $j$ ($M_{\mathrm{c}} = 3$, so $j \in \{1, 2, 3\}$), and $\theta_{\mathrm{p}} = 0$ is the fixed analyser direction. The one-hot encoded ground truth $y_j$ in Equation (S13) is now determined by the detection area to which the training image is allocated. Notably, the input Jones vector of the training image should be carefully adjusted to its preset polarization state. By alternately feeding the model with differently polarized training data and minimizing the loss function, the vectorial DNN for both polarization multiplexing and polarized object perception can be optimized.

A total of 12000 training images and 3597 test images were used in the training process, with 600 test images (200 for each task in polarization multiplexing) used in the experiment. The batch size was set as 8 and the learning rate was set as 0.01. For polarization multiplexing, the test accuracy in simulation reaches 96.4%, 97.8%, and 97.9% for letter, fashion item, and digit classification tasks, respectively, after 6 epochs of training. For polarized object perception, the test accuracy for all 600 polarization features reaches 97.3% in simulation after 6 epochs of training.

## (3) Parallel image classification

To classify letters more efficiently, a vectorial DNN was trained to encode letters "a", "b", "c", or "d" into four polarized light spots, and was then arrayed as 12 × 12 to perform large-scale parallel computing. Only the spot corresponding to the true label is optimized as vertically polarized, whereas all the other spots tend to be horizontally polarized (Fig. S12a). The loss function was defined as the weighted combination of two SCE losses:

$$L = L_{\mathrm{SCE},1} + \gamma L_{\mathrm{SCE},2} = -\sum_{j=1}^{M_\mathrm{p}} y_{1,j} \log p_{1,j} - \gamma \sum_{k=1}^{4M_\mathrm{p}} y_{2,k} \log p_{2,k} , \tag{S17}$$

where the first term $L_{\mathrm{SCE},1}$ is aimed at optimizing the spot corresponding to the true label as vertically polarized, and the second term $L_{\mathrm{SCE},2}$ is aimed at optimizing the polarization of all four spots simultaneously. $M_\mathrm{p} = 4$ represents the four polarization directions involved, and $\gamma$ is the weight parameter. $L_{\mathrm{SCE},1}$ has the same expression as Equation (S13), but $y_{1,j}$ becomes a fixed one-hot vector [0, 0, 1, 0] (indicating vertical polarization), and $o_{1,j}$ should be revised as

$$o_{1,j} = \boldsymbol{S}_y^\mathrm{T} \boldsymbol{O}\left( \frac{j-1}{M_\mathrm{p}} \pi \right), \tag{S18}$$

where $\boldsymbol{S}_y$ is the selection vector for the detection area corresponding to the true label $y$ ($y$ equals 1, 2, 3, 4 for “a”, “b”, “c”, and “d”, respectively). As for $L_{\mathrm{SCE},2}$, $y_{2,k}$ is the combination of four one-hot encoded vectors for the four detection areas. For each area, the vector is set to [0, 0, 1, 0] (indicating vertical polarization) if the area aligns with the true label; otherwise, it is set to [1, 0, 0, 0] (indicating horizontal polarization). $o_{2,k}$ is the combination of four intensity vectors $o_{2,1,j}$, $o_{2,2,j}$, $o_{2,3,j}$, and $o_{2,4,j}$ for the four detection areas, respectively, which can be expressed as follows:

$$o_{2,m,j} = \boldsymbol{S}_m^\mathrm{T} \boldsymbol{O}\left( \frac{j-1}{M_\mathrm{p}} \pi \right), \tag{S19}$$

where $m, j \in \{1, 2, 3, 4\}$ and $\boldsymbol{S}_m$ is the selection vector for the $m$-th detection area.

By adjusting the weight $\gamma$, this model can strike a balance between optimizing the polarization of the true-label spot and that of the other three spots. In this work, $\gamma$ was selected as 0.2 to give priority to ensuring the vertical polarization of the true-label spot. 6000 training images and 1000 test images were used in the training process.

**Supplementary Note 3: Training algorithm of time-domain ensemble learning**

Ensemble learning refers to training multiple models and combining their predictions to improve the inference capability of a system[4]. Different models can learn from different attributes of the data, thus enhancing the collective performance.

In this work, we propose a training algorithm for ensemble learning in the time domain, which makes joint inference based on multiple sequential models (i.e., temporal layers). Taking intelligent polarization encoding as an example, a vectorial DNN is trained to classify eight digits in a single detection area by converting them into eight separate polarization states (Fig. S2b). If the task is implemented by only one temporal layer, the loss function can be written as

$$L_{\mathrm{SCE},1} = -\sum_{j=1}^{M_{\mathrm{c}}} y_j \log p_{j,1} \, . \tag{S20}$$

$M_{\mathrm{c}} = 8$ is the number of categories, $y_j$ is the one-hot encoded ground truth, and $p_{j,1}$ is the probability for class $j$ predicted by the single-layer vectorial DNN. Here, $p_{j,1}$ is calculated from the output intensity $o_{j,1}$ for class $j$ using the softmax function. The test accuracy reaches only 69.7% after 10 epochs of training (Fig. S2d).

As shown in Fig. S2a, additional temporal layers can be trained jointly to perform the same task. The final inference is made by adding the predictions of all temporal layers. The loss function for training the $k$-th temporal layers is expressed as

$$L_{\mathrm{SCE},k} = -\sum_{j=1}^{M_{\mathrm{c}}} y_j \log p_{j,\mathrm{joint}} - \mu \sum_{j=1}^{M_{\mathrm{c}}} y_j \log p_{j,k} \, . \tag{S21}$$

The first term aims to optimize the joint computing performance, where $p_{j,\mathrm{joint}}$ is calculated from the sum of predictions from all the existing temporal layers. The second term aims to optimize the performance of the individual $k$-th temporal layer, which can lead to more concentrated output energy. The weight parameter $\mu$ can be adjusted to balance these two effects, and was set to 0.001 in this work. As illustrated in Fig. S2d, the classification accuracy improves substantially as the number of temporal layers increases, reaching 86.4% in the case of five temporal layers, which is as effective as adding more spatial layers (Fig. S2c). This result highlights the potential for performance enhancement without causing multi-layer alignment errors.

## Supplementary Note 4: Mapping relation between retrieved and recorded light

When exposed to linearly polarized blue/UV light, the sulfonic azo-dye SD1 molecules tend to reorient their absorption oscillators orthogonal to the local polarization direction because of their photoisomerization and dichroic absorption[5]. Via the intermolecular interaction between SD1 and LC molecules, LCs can acquire the same orientation distribution as SD1. This effect lays the foundation for the nonvolatile photonic memory demonstrated in the manuscript. Interestingly, it not only enables the optical reconfiguration of vectorial DNNs, but also allows polarization recording of the illuminating blue/UV light. Taking advantage of this vectorial light recording capability of LCs, the ingenious integration of optical computing and optical recording can be expected, uncovering LCs' potential in end-to-end all-optical DNNs. Furthermore, compared to other optical recording methods, this strategy can even optically reconstruct the information after recording, enabling the flexible retrieval of recorded vectorial light through a specific mapping relation.

Here, we discuss the mapping relation between retrieved and recorded light. Assuming that $\psi$ is the orientation angle of the polarization ellipse of the recorded vectorial light, the ultimate orientation of LC molecules not only relies on $\psi$, but also depends on the ellipticity, the exposure dosage, and the initial LC orientation. For example, the recording process exhibits a more significant response to linear polarization, while showing little response to circular polarization. With sufficient exposure dosage, the ultimate LC orientation should be orthogonal to the polarization direction of the exposure ($\alpha = \psi + 90°$). During optical retrieval, a horizontally polarized incident light can be converted to a polarization direction of $2\psi$ (i.e., $2\alpha$) under the half-wave condition, mapping the polarization information of the recorded light with a twofold relation. Notably, although recorded light with a polarization direction of $\psi$ + 90° can also be retrieved as the same polarization direction of $2\psi$, there exists a phase difference of π between these two situations, thereby preserving the information of recorded light in the phase domain.

With insufficient exposure dosage, the mapping relation will be more complex. The ultimate LC orientation will be affected by both the exposure polarization ($\psi$) and the initial LC orientation. With larger exposure dosage, the LC orientation $\alpha$ will get closer to $\psi + 90°$ from its initial orientation, thus affecting the retrieved vectorial light ($2\alpha$). To uncover this complex process, we measured the relation between retrieved and recorded light in a simplified case. The initial LC orientation was controlled as 45° by illuminating with uniformly polarized blue light (405 nm, 8.65 mW/cm$^2$, 10s). Then, the LCs were exposed to a low-power 473 nm laser with fixed 0° polarization, so the LC molecules tended to rotate to 90°. The exposure dosage was controlled by varying the power density from 0 to 0.472 mW/cm$^2$ and fixing the exposure time to 10 min. After recording, the LCs were illuminated by a 632.8 nm laser to retrieve the recorded optical information. With crossed polarizers (45° and 135°), the background light was eliminated, leaving only the more intuitive intensity information. The measured relation between retrieved and recorded intensity is shown in Fig. 5f (iii) in the manuscript, showing a nonlinear trend similar to the sigmoid function. This relation indicates that the recording effect is not significant with small exposure dosage, and tends to saturate with large exposure dosage. In a more general vectorial light recording process, this relation will also depend on the direction and ellipticity of polarization.

In collaboration with optical computing, such a relation involving both polarization and intensity holds great potential in performing nonlinear activation functions after the linear computation[6]. Furthermore, a greater variety of mapping relations can be introduced in future research by employing different photosensitive dyes. For example, dye materials with higher-sensitivity photoresponsive groups can provide a lower threshold and a steeper nonlinear transfer function. Taking advantage of the adiabatic following effect of twisted LCs or the guest-host effect of nematic LCs, the retrieved light is expected to reconstruct the same polarization direction as the recorded vectorial light. Moreover, it is worth noting that this new optical recording strategy can be utilized to record arbitrary vectorial light fields, including cylindrical vector beams[7] and vectorial holography[8], greatly facilitating optical recording technologies and polarization optics.

**Supplementary Note 5: Large-scale parallel computing, recording, and retrieval**

In the large-scale parallel photonic in-memory computing, a 256 × 256 vectorial DNN was trained and then scaled up to a 12 × 12 array, which promises to classify 144 letters simultaneously. As an example, a 3 × 2 subarray is utilized to classify six letters in parallel. For the ease of vectorial light recording, each letter ("a", "b", "c", or "d") is targeted to four light spots. Only the spot corresponding to the true label is optimized as vertically polarized, whereas all the other spots tend to be horizontally polarized (Fig. S12a; see more design details in Supplementary Note 2). In the experiment, we demonstrate this vectorial DNN array with an input image example "bad cab" (Fig. 5f (i)). The intensity of the output vectorial light field is measured without a polarizer, presenting 24 light spots in total (Fig. S12b). Its vectorial components $I_x$ and $I_y$ are measured with either a horizontal or vertical polarizer (Fig. S12d). It can be observed that the spots corresponding to the true labels (circled with white dotted lines) are relatively weaker in $I_x$ and stronger in $I_y$, which is highly consistent with our design.

Then, we used another LC device to record the output vectorial light, i.e., the 24 polarized light spots, all at once. The recording process involves a complex nonlinear mapping, which is detailed in Supplementary Note 4. After recording, the micrograph of LCs was captured with an inserted 530 nm full-wave plate (Fig. S12c), where we can read out the LC orientations from colours. Generally, the colours of the spots corresponding to the true labels (circled with white dotted lines) are closer to green, indicating the horizontal LC orientation and the vertical polarization of recorded light. This result allows us to infer the categories of the six input letters based on colours, unveiling the impressive ability of LCs in polarization recording and parallel inference.

Remarkably, we can further use the LC device after recording to retrieve the previous computing results. To achieve optical retrieval, LCs were illuminated by a 632.8 nm laser. Crossed polarizers (45° and 135°) were used to eliminate the background light, leaving a total of 24 light spots with various intensity patterns ($I_{r1}$ in Fig. S12e). As we discussed in Supplementary Note 4, the retrieved light has a complex mapping relation with the recorded light regarding both polarization and intensity. Interestingly, a morphological similarity between $I_{r1}$ and $I_x$ is observed, providing an intuitive proof for their inherent relation. Through experiments, we discover that such similarity is highly dependent on the initial orientation of the LC device. By carefully adjusting the initial LC orientation and then recording the same output vectorial light, we can also retrieve light with its intensity pattern similar to $I_y$ ($I_{r2}$ in Fig. S12e).

When performing parallel in-memory computing, we also uncover LCs' potential in parallel vectorial light recording and parallel optical retrieval, verifying the good scalability and high parallelism of the proposed photonic in-memory computing platform. Moreover, the retrieved light containing the previous computing results can be ingeniously used for the following optical computing steps (Fig. S8), enabling network extension, power regeneration, repeated data access, and even temporal sequence processing in deep DNNs.

**Supplementary Note 6: Performance assessment of the LC vectorial DNN**

The proposed LC vectorial DNN establishes the highly-desired nonvolatile photonic memory compatible with large-scale free-space interconnections, greatly enhancing the scalability of conventional photonic in-memory computing. Moreover, compared to common scalar DNNs, its reconfigurable and vectorial nature enables higher computing performance, specialized polarization-related tasks, adaptability to variable application demands, and single-shot vectorial light recording. It also offers the advantages of dispensable driving power and high energy efficiency during static operation. The metrics of the proposed LC vectorial DNN are listed in Supplementary Table S2 and are compared with those of other state-of-the-art technologies in Supplementary Tables S1, S3, and S4. A more detailed analysis is provided below.

**(1) Functionality**

The proposed large-scale photonic in-memory computing platform uniquely exploits free-space wave propagation to physically manifest matrix transformations, thus allowing higher-scalability matrix multiplication and much richer functionalities than conventional photonic in-memory computing. The so-called vectorial DNN also theoretically covers all the functionalities of scalar DNNs. With higher-dimensional network connections, it can achieve higher computing performance than conventional scalar DNNs and undertake polarization-related tasks that cannot be executed by them. For example, we demonstrate the intelligent polarization encoding that can encode different images into different polarization states within a single pixel. We also demonstrate the 3-channel (Fig. S4) and 10-channel (Fig. S6) polarization-multiplexed computing in the experiment and simulation, respectively, significantly enhancing the computing parallelism of light. More interestingly, this platform can directly sense and record vectorial light in addition to computing it. Taking polarized object perception for example, objects with different polarization features can be sensed and classified by the vectorial DNN. By constructing a convolutional vectorial DNN, real-time analysis of real-world biological samples is achieved (Fig. S7). Utilizing the photonic memory via LCs, the computing results of the vectorial DNN can be all-optically recorded in another LC device, suggesting an ingenious way to combine optical computing and recording into a single system.

**(2) Computing performance**

Leveraging the vectorial nature of light, the vectorial DNN can reasonably exceed the computing performance of scalar DNNs. As an example, we compare the digit classification accuracies of the MNIST dataset using different DNN models (Fig. S6a). The same training parameters were used for a scalar DNN, a vectorial DNN without an analyser, and a vectorial DNN with a fixed analyser. From the simulation results, the vectorial DNNs demonstrate higher classification accuracies than the scalar DNN, and the vectorial DNN with an analyser has the best performance. Intriguingly, the DNN computing performance can be further enhanced through polarization multiplexing. As shown in Fig. S6b, we compare the letter case (uppercase/lowercase) classification accuracies of 20 letters from the EMNIST dataset using different 5-layer DNN models, including a scalar DNN, a vectorial DNN without polarization multiplexing, and a vectorial DNN with 10-channel polarization multiplexing. In simulation, the polarization-multiplexed DNN exhibits the highest average classification accuracy of 93.3%. Furthermore, leveraging the nonvolatile photonic memory via LCs, the time-domain ensemble learning strategy can be adopted to enhance the inference performance of the vectorial DNN by combining the predictions from multiple temporal models (see details in Supplementary Note 3). By using large-scale neurons as a single-core processor, introducing additional diffractive layers (Fig. S13), incorporating nonlinear activations (Fig. S14), or constructing hybrid optoelectronic neural networks (Fig. S15), more complex tasks can be solved with high performance.

**(3) Reconfigurability**

Through the light-driven reorientation of LC molecules, the polarization conversion trajectories on the Poincaré spheres in the proposed Poincaré-sphere-connected vectorial DNN can be dynamically programmed in a nonvolatile manner, leading to the reconfigurable photonic memory. Consequently, the LC vectorial DNN can realize flexible function switching, time-domain ensemble learning, and retrievable polarization recording, exhibiting higher adaptability and richer functionalities than fixed architectures. If properly stored (sealed and kept at room temperature), the retention time of the LC orientations can be longer than 1 year. This retention time is competitive with other nonvolatile platforms, such as phase-change materials (years), $BaTiO_3$ (hours), and memresonators (hours). In this work, up to 9.44 million reconfigurable vectorial neurons are built with 256 possible states (8 bits). The memory rewriting speed is highly dependent on the writing light's wavelength and power density. In our case, during each optical reconfiguration process, a 365 nm light-emitting diode (LED; 42.5 mW/cm$^2$, 10 s) was used to erase the original LC orientations, which can accelerate the subsequent rewriting process, and a 450 nm laser (2.05 mW/cm$^2$, 60 s) was used to rewrite new orientations. By using a higher laser power ($1.07 \times 10^4$ mW/cm$^2$) and a more sensitive wavelength (405 nm), this time can be drastically shortened to below 100 ms in the absence of LED erasing (Fig. S16). Furthermore, by using a femtosecond laser with a much higher peak power (400 nm, $1.24 \times 10^{13}$ mW/cm$^2$), the LC orientation can be rewritten with a total illumination time of only 3 ns (Fig. S17), which opens up new possibilities for nanosecond-level reconfigurable optical devices.

**(4) Scalability**

In this work, up to 2.07 million photonic memory cells (1920 × 1080) can be optically reconfigured at once using the SLM-based polarization projection system. Thanks to the nonvolatile nature of photonic memory, the LC vectorial DNN can retain the reconfigured LC orientations even when the writing laser is turned off. Therefore, the vectorial DNN can be easily scaled up (almost infinitely in theory) by translating the LC device for tiling operations. To demonstrate this ultrahigh scalability, we experimentally integrate 12 × 12 vectorial DNN units into a large matrix of reconfigurable photonic processors, which enables massively parallel photonic in-memory computing (Fig. 5e,f). Each unit was trained to classify a letter by encoding the polarization states of four light spots, leading to in-situ optical computing, recording, and retrieval of a total of 576 parallel polarization encoding channels. In this case, 9.44 million (256 × 256 × 12 × 12) reconfigurable vectorial neurons are utilized in total, which is a number much larger than the typical sizes of SLMs[9,10] and programmable metasurfaces[11]. This provides a memory capacity of 151.04 Mbit across the two LC layers (9.44 million × 8 bits × 2 layers), which is more than 100,000 times larger than that of state-of-the-art integrated photonic in-memory computing[12,13]. According to the commonly used calculation method[14], and considering the additional operations enabled by the vectorial light field, approximately $2.72 \times 10^9$ real-valued multiply-accumulate operations (MACs) can be performed in memory at a time, which can lead to an effective computing speed of $5.44 \times 10^{16}$ FLOPS, assuming real-world input and 10 MHz photodetectors are used. Moreover, the proposed photo-rewritable LC device exhibits a high transmission efficiency of 92.5% and a low insertion loss of 0.34 dB for 9.44 million parallel optical neurons, offering significant benefits for low-light computing and multi-layer scaling.

### (5) Power consumption

The total power consumption for optical computing is measured to be 96 W, including a spatial light modulator (8 W), a read laser (15 W), a signal generator (21 W), and a laptop-driven camera (52 W). This power consumption can lead to a high energy efficiency of hundreds of TOPS/W with 10 MHz photodetectors. The LC vectorial DNN has the advantage of dispensable driving power during static operation. It can retain the reconfigured LC orientations when the writing laser is turned off, and function properly under the LC half-wave condition. Although this half-wave condition is typically satisfied by applying a square-wave electrical signal (1 kHz, ~1 V), it can also be satisfied by controlling the thickness of the LC cell, thus omitting the external driving power from the signal generator, just like LC waveplates. Compared to volatile memory that requires continuous power supply, the proposed platform allows much lower power consumption during static execution of a certain task, which might be the more common scenario in some practical applications. In the case of DNN reconfiguration, a light energy of 351 nJ per neuron (42.5 mW/cm² × 64 μm²/neuron ×

10 s (LED) + 2.05 mW/cm² × 64 μm²/neuron × 60 s (450 nm laser)) is directly required to rewrite the photonic memory, which can be further reduced to 0.17 nJ/bit by using the minimal pixel size (~0.25 μm²) of LCs. The total system energy during reconfiguration (laptop computer + SLM + blue laser for LC reorientation + signal generator) is considered to be relatively higher than the scalar SLM alone (laptop computer + SLM), but the low static power consumption can effectively compensate for this. Especially, for scenarios involving large-scale data loading, computation and storage, the proposed photonic in-memory computing scheme promises to circumvent the costly data conversions and movement by integrating optical sensing, memory and computing in situ, thus significantly increasing the overall parallelism and energy efficiency.

## Supplementary Tables S1-S4

**Table S1 Comparison between diffraction-based optical computing platforms.**

| Aspect | Photo-rewritable LC | LC-SLM | Tunable metasurface (Neurophos[15]) |
|---|---|---|---|
| **Light dimension** | Vectorial | Scalar | Scalar |
| **Pixel scale** | 3072 × 3072 | 1920 × 1080 | 1024 × 1024 |
| **Feature size** | 8 μm (potentially 500 nm) | ~8 μm | Nanoscale |
| **Functionality** | Diffractive neural network (DNN) | DNN & Matrix multiplication | Matrix multiplication |
| **Transmittance/ Reflectance** | 92.5% | ~80% | <80% |
| **Tunability** | Non-contact optical tuning | Electrical tuning | Electrical tuning |
| **Volatility** | Nonvolatile (retention time: years) | Volatile | Volatile |
| **Switching time** | ~60 s (potentially 3 ns) | ~60 ms | 1 μs |
| **Manufacturing** | Free of etching and deposition | CMOS-compatible | CMOS-compatible |
| **Cost** | Low <US$100 | High ~US$10,000 | High |

**Table S2 Metrics of the proposed optical computing platform.**

| Aspect | This work (achieved with 34.8 Hz CMOS) | This work (achievable with 10 MHz photodetectors) |
|---|---|---|
| **Effective computing speed**[a] | $1.89 \times 10^{11}$ FLOPS | $5.44 \times 10^{16}$ FLOPS |
| **Power consumption** | 96 W | - |
| **Pixel size**[b] | 8 μm | 500 nm |
| **Pixel scale** | 3072 × 3072 | - |
| **Tunable neuron number** | 9.44 million | - |
| **Footprint** | 6.04 $cm^2$ | 2.36 $mm^2$ |
| **Bit precision** | 8 bits | - |
| **Insertion loss** | -0.34 dB | - |
| **Memory capacity** | 151.04 Mbit | - |
| **Computation per pass**[c] | $2.72 \times 10^9$ MACs | - |
| **Write laser energy** | 43.88 nJ/bit | 0.17 nJ/bit |
| **Read laser power** | 33 μW | - |
| **Storage retention time** | >1 year | - |

[a]The effective computing speed in the experiment is calculated to be $1.89 \times 10^{11}$ FLOPS, limited by the slow 34.8 Hz CMOS camera. If using commercially available 10 MHz photodetectors, much higher effective computing speed of $5.44 \times 10^{16}$ FLOPS (54400 TFLOPS) can be achieved.

[b]The pixel size can be reduced to 500 nm by reducing the thickness of LC cells and increasing the resolution of the LC rewriting system. Then, the footprint can be reduced to 2.36 $mm^2$ and the write laser energy can be reduced to 0.17 nJ/bit.

[c]We assume 1 multiply-accumulate operation (MAC) = 2 floating point operations (FLOPs).

**Table S3 Advantages compared to state-of-the-art optical computing technologies.**

| Source | Technology | Non-volatility | Light dimension | Tunable neuron | Bit precision | Insertion loss[a] | Memory capacity | Computation per pass[b] |
|---|---|---|---|---|---|---|---|---|
| **This work** | **Photo-rewritable LC** | **Yes** | **3D Vectorial** | **9.44 million** | **8 bits** | **-0.34 dB** | **151.04 Mbit** | **$2.72 \times 10^9$ MACs** |
| *Nature 589, 52 (2021)*[16] | PCM | Yes | 2D Scalar | 256 | 5 bits | -0.73 dB | 1280 bits | 144 MACs |
| *Nature 640, 361 (2025)*[17] | MZI | No | 2D Scalar | 4096 | 8 bits | -0.62 dB | Not supported | 4096 MACs |
| *Nat. Electron. 5, 113 (2022)*[11] | Programmable metasurface | No | 3D Scalar | 320 | 4 bits | 13 dB (amplifier integrated) | Not supported | 66816 MACs |
| *Nat. Photon. 15, 367 (2021)*[9] | DMD + LC-SLM | No | 3D Scalar | 0.49 million | 8 bits | -0.56 dB | Not supported | $7.8 \times 10^7$ MACs |

PCM, phase-change material; MZI, Mach-Zehnder interferometer; DMD, digital micromirror device; SLM, spatial light modulator.

[a]Loss accumulates with the number of layers in 3D architectures, and with the number of modulation elements in 2D architectures.

[b]Optical operations in a single pass are calculated following Ref. [14] or cited from the literatures, and are expressed in real-valued multiply-accumulate operations (MACs).

**Table S4 Comparison with state-of-the-art optical computing technologies.**

| Source | Technology | Read laser power | Write energy | Storage retention time | Footprint | Power consumption | Effective computing speed |
|---|---|---|---|---|---|---|---|
| **This work** | **Photo-rewritable LC** | **0.033 mW (wall plug efficiency: 0.022%)** | **43.88 nJ/bit (wall plug efficiency: 0.65%)** | **>1 year** | **604 mm$^2$** | **96 W** | **54400 TFLOPS*** |
| *Nature 589, 52 (2021)*[16] | PCM | 0.5 mW | 0.16 nJ/bit | >25 days | 3.63 mm$^2$ | 10 W | 4 TFLOPS |
| *Nat. Commun. 14, 2887 (2023)*[18] | PCM | 3.55 mW | 3.77 nJ/bit | >2 days | 56.02 mm$^2$ | 40.81 W | 410 TFLOPS |
| *Nat. Photon. 15, 367 (2021)*[9] | DMD + LC-SLM | 0.0014 mW | - | Volatile | - | 152.1 W | 240 TFLOPS |
| *Science 384, 202 (2024)*[19] | MZI + Diffractive units | 50 mW (wall plug efficiency: 0.14%) | - | Volatile | 28.66 mm$^2$ | 313.27 W | 50380 TFLOPS |
| *Nature 640, 361 (2025)*[17] | MZI | ~160 mW (wall plug efficiency: ~10-20%) | - | Volatile | - | 3.44 W | 8.19 TFLOPS |
| *Science 390, 1259 (2025)*[20] | SLM + Optical latent space | 0.249 mW | - | Volatile | 136.5 mm$^2$ | 53.85 W | 35700 TFLOPS |
| *Optica 11, 932 (2024)*[21] | OFC + SLM | 1 mW (wall plug efficiency: 10%) | - | Volatile | - | 32.2 W | 20.48 TFLOPS |
| *Light Sci. Appl. 15, 115 (2026)*[22] | VCSEL + SLM | 0.9 mW (wall plug efficiency: 25%) | - | Volatile | 100 mm$^2$ | 0.015 W | 0.045 TFLOPS |

PCM, phase-change material; MZI, Mach-Zehnder interferometer; DMD, digital micromirror device; SLM, spatial light modulator; OFC, optical frequency comb; VCSEL, vertical-cavity surface-emitting lasers.

*We assume real-world input and 10 MHz photodetectors are used.

## Supplementary Figures S1-S18

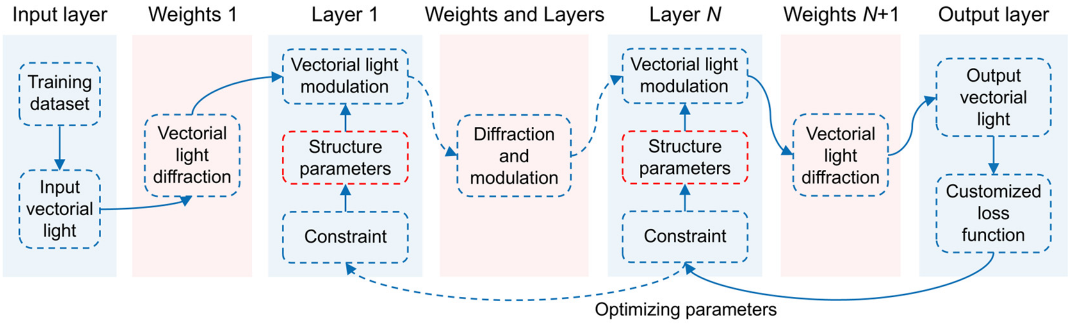


**Fig. S1 Training flowchart of the vectorial DNN.**

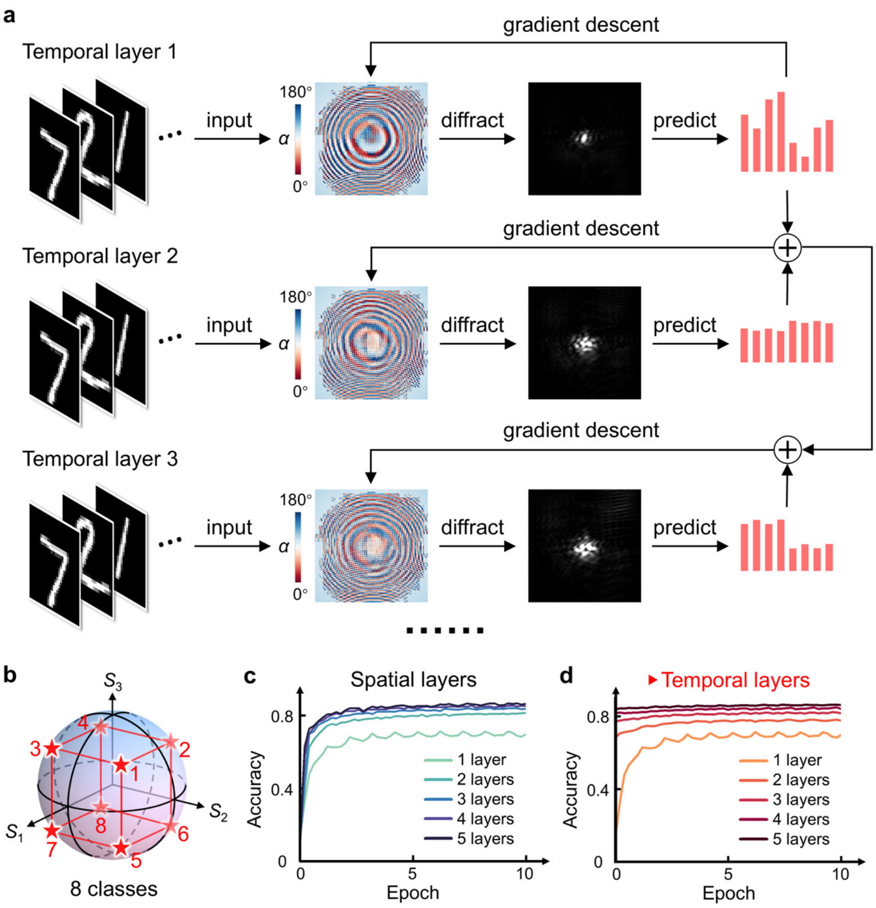


**Fig. S2 Training algorithm and simulation results of the ensemble learning. a**, Joint training of multiple temporal layers (sequential models) for the same polarization encoding task. **b**, Eight target polarization states within the single detection area for classifying eight digits. **c**,**d**, Simulated accuracy curves over epochs with different numbers of (**c**) spatial layers and (**d**) temporal layers.

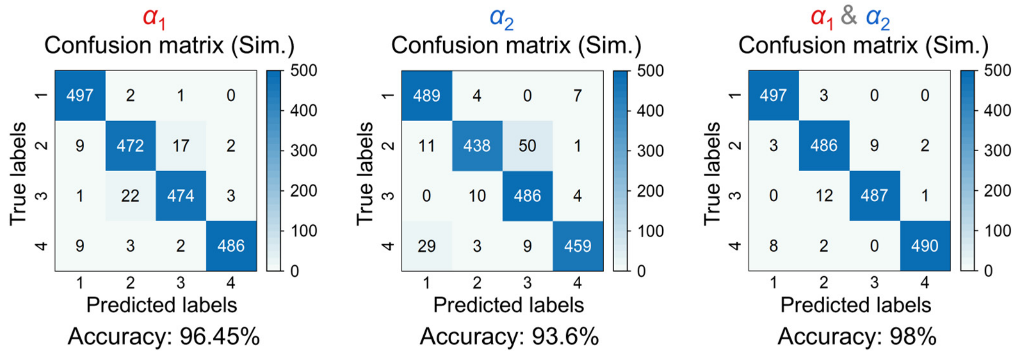


**Fig. S3 Simulated confusion matrices of polarization encoding with/without ensemble learning.** The simulated test accuracies with model 1 ($\alpha_1$), model 2 ($\alpha_2$), and both sequential models ($\alpha_1$ & $\alpha_2$) are listed below their respective confusion matrices.

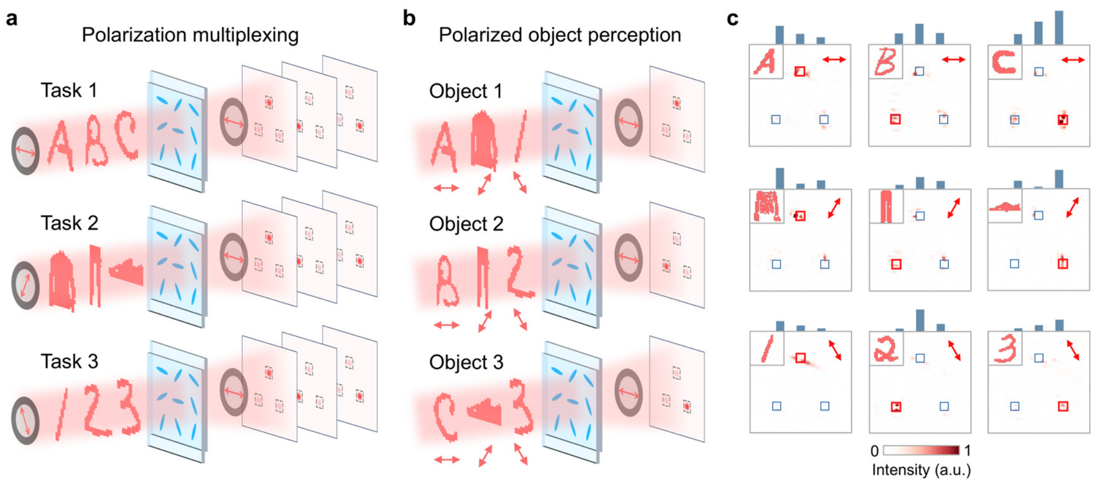


**Fig. S4 Schematic illustration and experimental results of Poincaré-sphere-perceptive computing.** **a**, Schematic of performing three independent image classification tasks in three nonorthogonal polarization channels. **b**, Schematic of conceptually classifying three categories of polarized objects by identifying their polarization features. **c**, Experimental results of Poincaré-sphere-perceptive computing. The insets represent the input images with their incident polarization marked by red arrows. The intensity patterns are their respective output fields. The sums of intensities within three detection areas are illustrated by the blue bars. The detection areas with the highest overall intensities are highlighted by red boxes, which indicate the predicted labels.

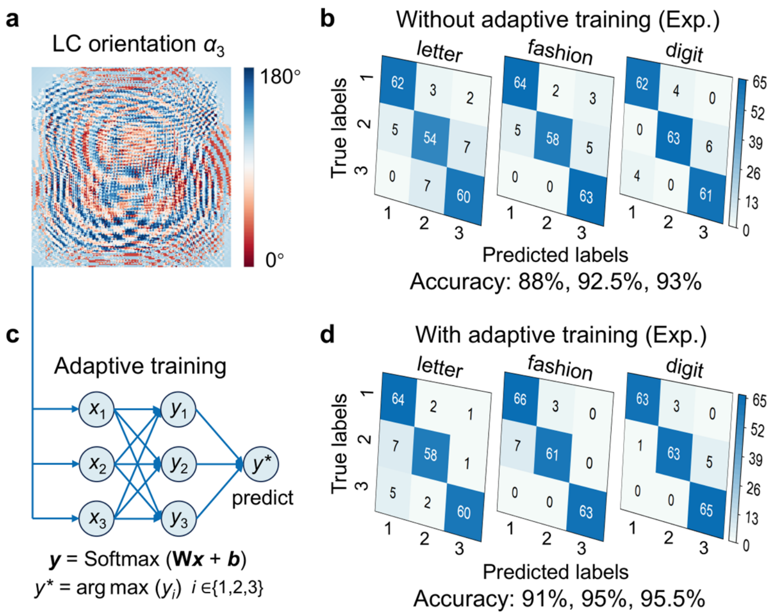


**Fig. S5 Electrical adaptive training for polarization-multiplexed computing. a**, LC orientation distribution $\alpha_3$ trained for polarization-multiplexed computing (optical neurons). **b**, Experimental confusion matrices and accuracies of the three multiplexed image classification tasks without adaptive training. **c**, Single-layer electronic neural network for adaptive training (electronic neurons). $\boldsymbol{x}$ is the optical computing result, which can be converted to $\boldsymbol{y}$ by a single fully connected layer. The final inference $y^*$ is made based on $\boldsymbol{y}$. **d**, Experimental confusion matrices and accuracies of the three tasks after combining the electrical adaptive training.

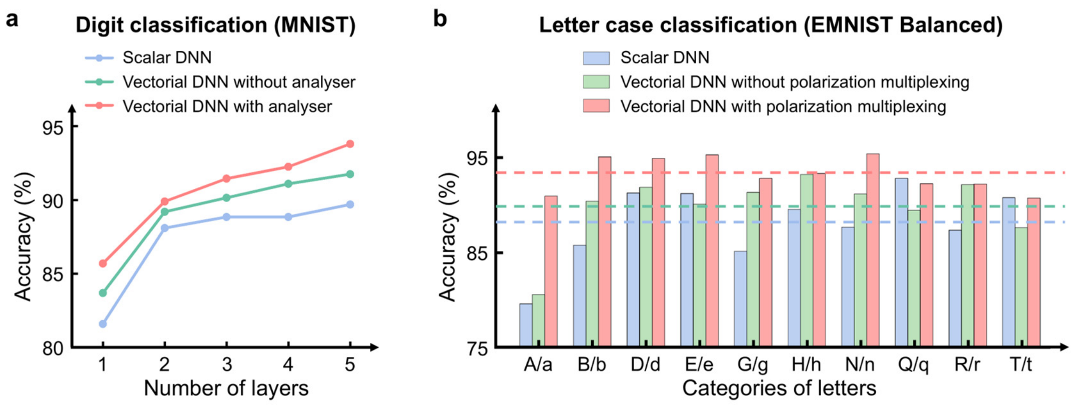


**Fig. S6 Simulated computing performance of different DNN models. a**, Digit classification accuracies of the MNIST dataset using different DNN models (with the same training parameters) in relation to their respective layer numbers. **b**, Letter case (uppercase/lowercase) classification accuracies of 20 letters from the EMNIST Balanced dataset using different DNN models. The average classification accuracies for the 20 letters are marked with dashed lines, where the 10-channel polarization-multiplexed DNN has the best performance.

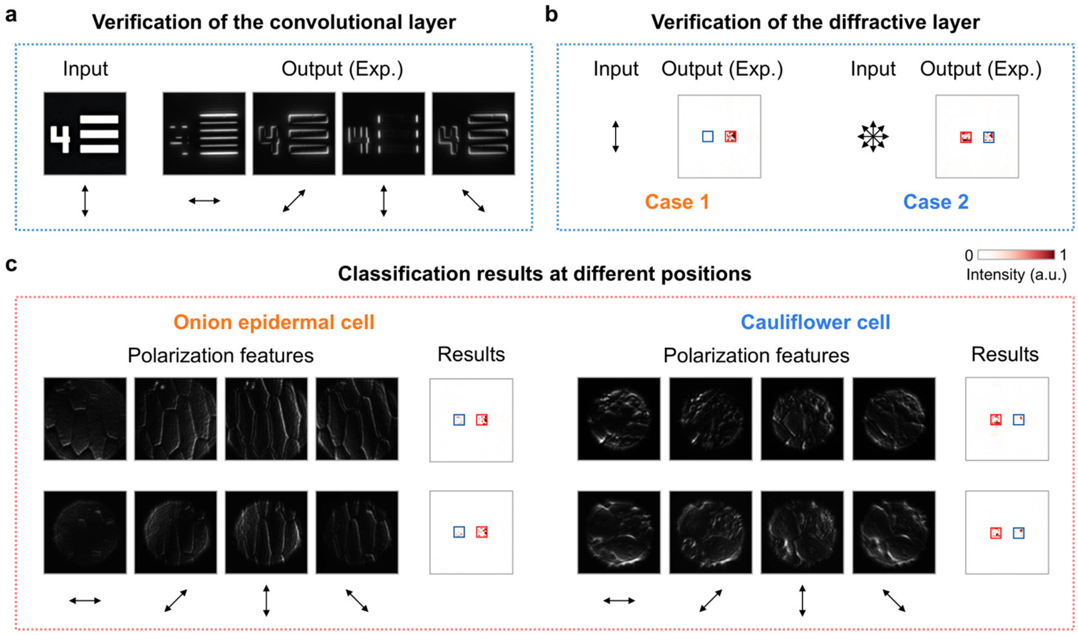


**Fig. S7 Experimental results of the convolutional vectorial DNN based on photo-rewritable LCs. a**, Verification of the LC convolutional layer, which can convert different edge directions to different polarization directions. The black arrows represent the input and output polarization states. **b**, Verification of the trained LC diffractive layer, which can classify vertically polarized light and polarization-unbiased light. **c**, Experimental classification results of onion epidermal cells and cauliflower cells at different cell positions.

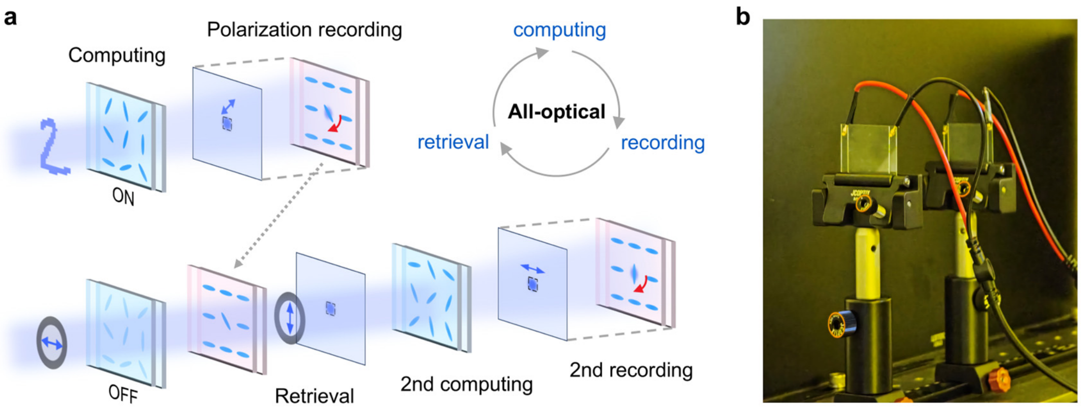


**Fig. S8 Optical recording and network extension in LC vectorial DNNs. a**, Schematic of LC-based all-optical loops, including optical computing, recording, and retrieval. The first LC device (blue) is used to perform computing tasks with blue light, such as encoding the input digit into a desirably polarized light spot. The second LC device (pink) can record the computing result through reorienting LCs orthogonal to the polarization of the output vectorial light. By switching off the first device (full-wave condition) and illuminating uniformly, the information recorded in the second LC device can be retrieved, which can be further used as the input for the next all-optical loop (the second computing and recording processes). **b**, Experimental setup for optical recording and LC device cascading.

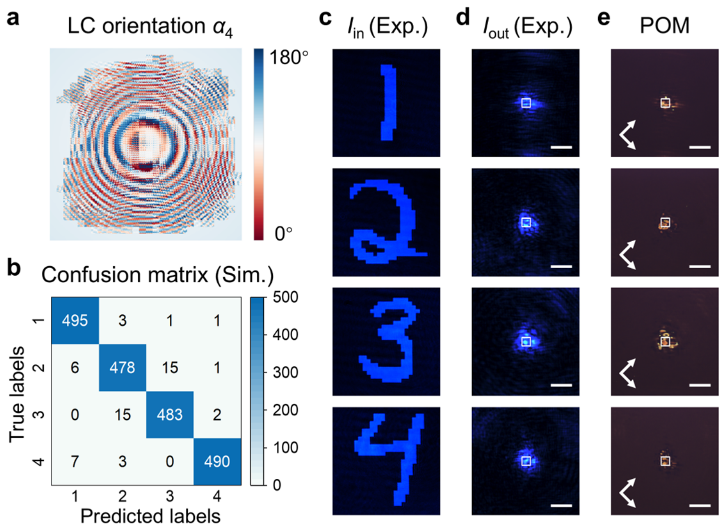


**Fig. S9 Blue-light polarization encoding and vectorial light recording. a**, LC orientation $\alpha_4$ for the blue-light polarization encoding. **b**, Simulated confusion matrix with 2000 test images. **c**,**d**, Experimental intensity patterns of (**c**) four different input images and (**d**) their respective output. **e**, Polarized optical micrographs (POM) of LCs after recording the output vectorial light. White arrows depict the crossed polarizer and analyser. All scale bars represent 200 μm.

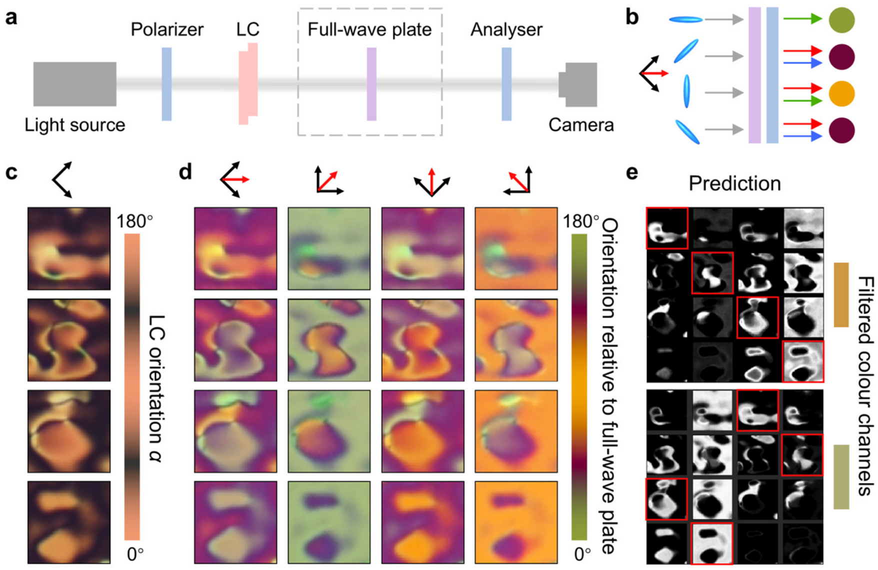


**Fig. S10 Verification of the recorded LC orientations via colours using a full-wave plate. a**, Optical setup for reading LC orientations, which combines a polarized optical microscope with an inserted 530 nm full-wave plate (the imaging lenses are not shown). **b**, Origin of colour variation for different LC orientations: different wavelengths (colourful arrows) can be filtered out through the full-wave plate (purple block) and crossed polarizers (blue block) due to the polarization interference. **c**, Enlarged micrographs of the central areas of Fig. S9e. **d**, Enlarged micrographs with variant directions of the full-wave plate (red arrows) and polarizers (black arrows), where LC orientations can be read out from colours. **e**, Two special colour channels filtered from (**d**). The brightness in red boxes is relatively higher, which is consistent with our design.

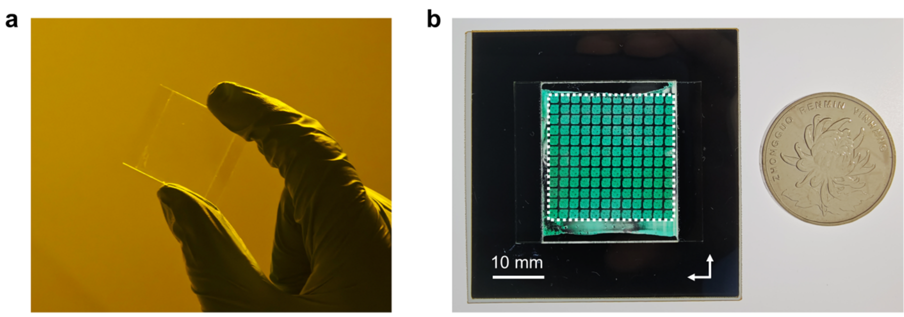


**Fig. S11 Real-world photo of the LC device for large-scale photonic in-memory computing. a**, LC device containing 12 × 12 vectorial DNN units, which is transparent without polarizers. **b**, The same LC device under crossed polarizers (white arrows). The effective area (12 × 12 processor units) is enclosed by dotted lines. A coin is used as a reference object. The scale bar represents 10 mm.

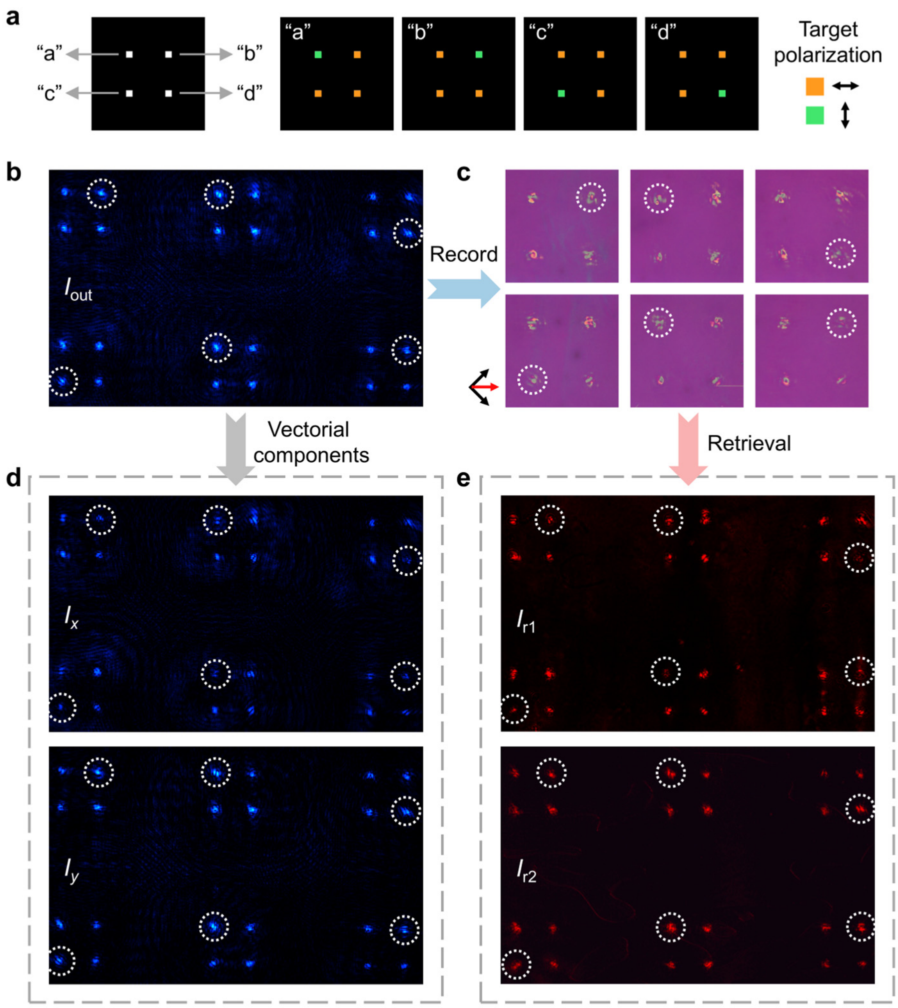


**Fig. S12 Experimental results of parallel optical recording and retrieval in large-scale photonic in-memory computing. a**, Target polarization in the four detection areas for the parallel letter classification task. **b**, Intensity pattern of the output vectorial light with an input image example "bad cab". Spots corresponding to the true labels are circled with white dotted lines. **c**, Micrographs of LCs after recording the vectorial light in (**b**). Black and red arrows represent the crossed polarizers and the full-wave plate, respectively. **d**, Horizontally (upper) and vertically (lower) polarized components of (**b**). **e**, Optically retrieved computing results after recording them with two different initial LC orientations.

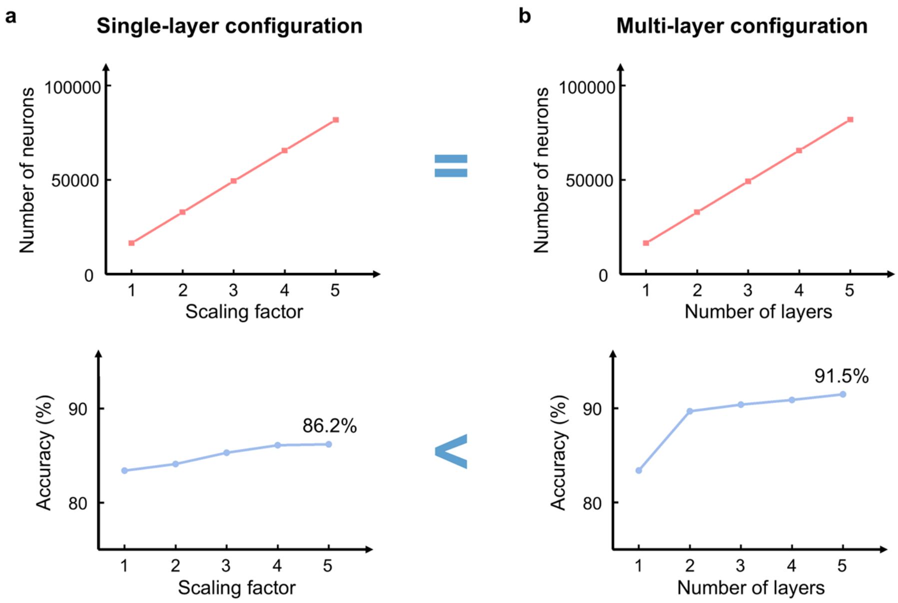


**Fig. S13 Relationship between computational performance and the number of neurons in single-layer and multi-layer configurations, respectively. a,** MNIST digit classification accuracy with different numbers of trainable neurons in a single layer. **b,** MNIST digit classification accuracy with different numbers of layers, with the same number of trainable neurons as in **a**.

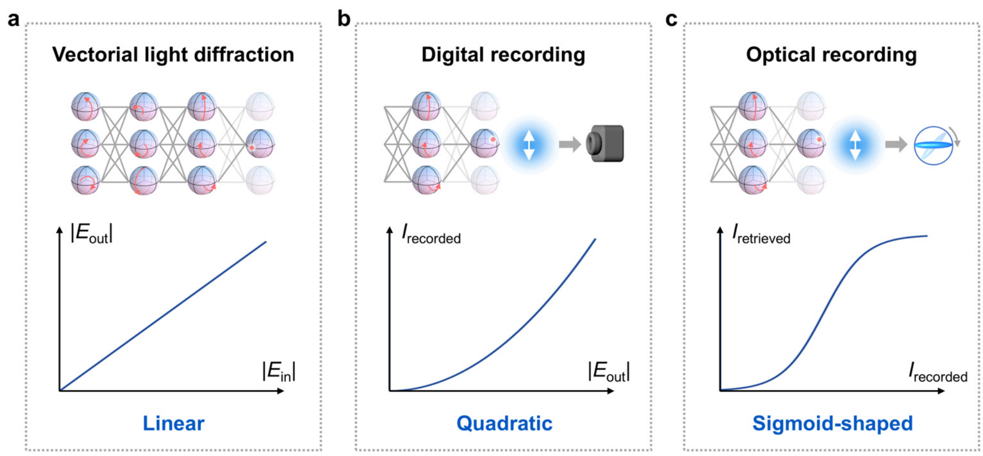


**Fig. S14 Analysis of nonlinearity intrinsically built in the proposed architecture. a**, Linearity in vectorial light modulation and diffraction. **b**, Quadratic nonlinearity in digital recording. **c**, Sigmoid-shaped nonlinearity in optical recording.

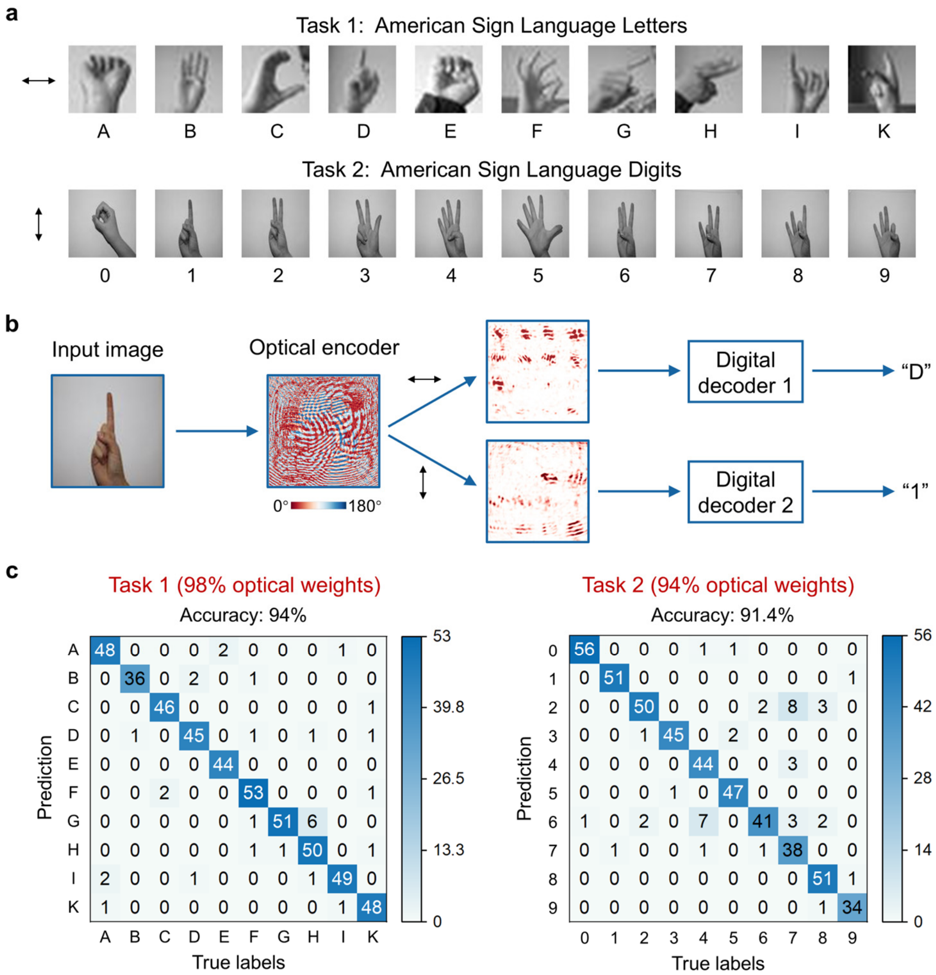


**Fig. S15 Sign language classification based on a hybrid optoelectronic neural network. a**, Polarization multiplexing of two classification tasks. **b**, Architecture of the optoelectronic neural network. The optical encoder consists of 1080 × 1080 trainable neurons, which is significantly larger than the two digital decoders (Digital decoder 1: 64 and 32 neurons in the two hidden layers; Digital decoder 2: 128 and 128 neurons in the two hidden layers). **c**, Experimental confusion matrices of the polarization-multiplexed sign language classification.

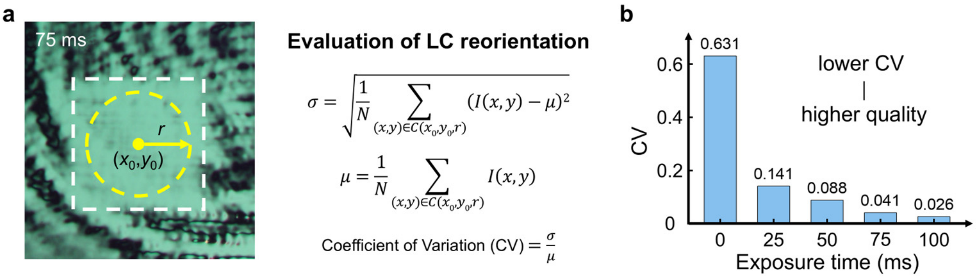


**Fig. S16 Experimental results of rewriting LCs with high-power continuous laser and the employed evaluation method. a**, Evaluation method of the LC reorientation quality based on the coefficient of variation (CV) within the reorientation area of the polarized optical micrograph. An example (exposure time: 75 ms) is presented to show the evaluation process. The originally complex LC orientations were rewritten into a uniform orientation within a square exposure region (white dashed box). The corresponding CV is calculated from the intensity deviation $\sigma$ (standard deviation of intensity) and the mean intensity value $\mu$ within the central circular region (yellow dashed circle). **b**, Experimental results of the CV using different exposure time ranging from 0 to 100 ms, wherein the lower CV indicates the higher LC reorientation quality. A 405 nm laser with a high power density of $1.07 \times 10^4$ mW/cm² was used for LC reorientation in the absence of LED erasing.

**a** Femtosecond laser parameters

| | | | |
|---|---|---|---|
| Central Wavelength | 400 nm | Repetition Rate | 1 kHz |
| Pulse Duration | 100 fs | Laser-on time | 30 s |
| Average Power | $1.24 \times 10^3$ mW/cm² | Pulse Count | 30 k |
| Peak Power | $1.24 \times 10^{13}$ mW/cm² | Illumination time | 3 ns |

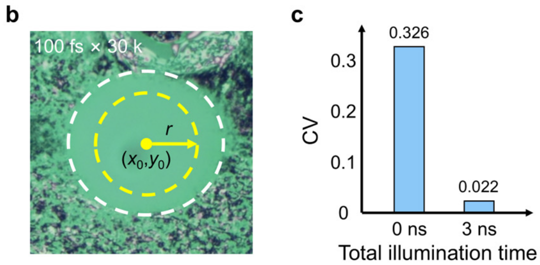


**Fig. S17 Experimental results of rewriting LCs with ultrahigh-peak-power femtosecond laser. a**, Parameters of the used femtosecond laser. **b**, Polarized optical micrograph of the LC device after illuminating it with 30,000 100-fs pulses (3 ns in total). **c**, Experimental results of the CV within the reorientation area of the micrograph before and after 3-ns illumination.

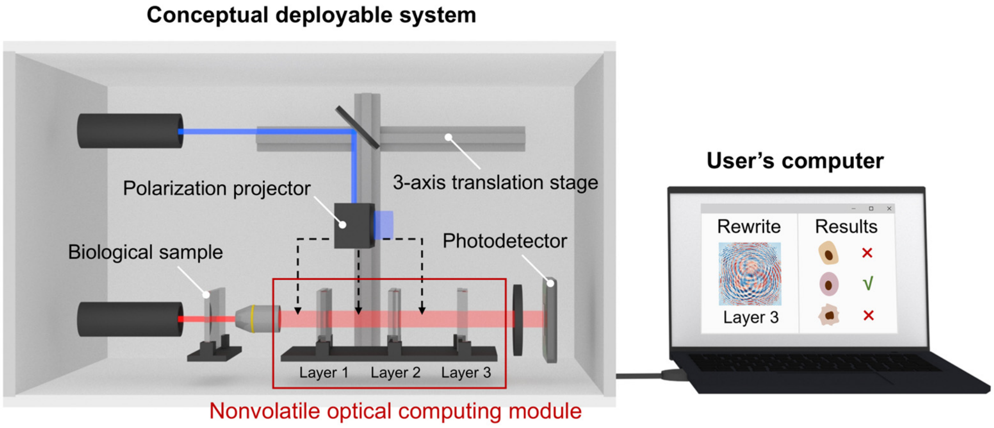


**Fig. S18 Envisioned deployable system for real-time biological sample analysis.** The system consists of a nonvolatile optical computing module and a translational rewriting module. The optical computing module, based on a multilayer LC vectorial DNN, can operate independently for static tasks with ultralow static power consumption. The LC vectorial DNN enables direct sensing and processing of optical information from biological samples in the optical domain, allowing classification of different types of tissues or cells through optical computation. The optical computing results can be captured by a photodetector and read out by software on the user's computer. For function switching, the translational rewriting module can be employed to re-program the LC vectorial DNN. Under software control on the user's computer, the polarization projector can be precisely positioned via a 3-axis translation stage to rewrite selected LC layers. This conceptual deployable system enables more efficient analysis of information-intensive biological samples, with higher computing speed and lower power consumption.

## Supplementary Information References


1. Chen, P., Wei, B. Y., Hu, W. & Lu, Y. Q. Liquid-crystal-mediated geometric phase: from transmissive to broadband reflective planar optics. *Adv. Mater.* **32**, 1903665 (2020).
2. Bass, M., Stryland, E. W. V., Williams, D. R. & Wolfe, W. L. *Handbook of Optics : Devices Measurements and Properties* Ch. 22 (New York: McGraw-Hill Inc., 1995).
3. Roy-Br´ehonnet, F. L., Jeune, B. L., Eliès, P., Cariou, J. & Lotrian, J. Optical media and target characterization by Mueller matrix decomposition. *J. Phys. D: Appl. Phys.* **29**, 34-38 (1996).
4. Rahman, M. S. S., Li, J., Mengu, D., Rivenson, Y. & Ozcan, A. Ensemble learning of diffractive optical networks. *Light Sci. Appl.* **10**, 14 (2021).
5. Akiyama, H. *et al.* Synthesis and properties of azo dye aligning layers for liquid crystal cells. *Liq. Cryst.* **29**, 1321-1327 (2002).
6. Zuo, Y. *et al.* All-optical neural network with nonlinear activation functions. *Optica* **6**, 1132-1137 (2019).
7. Zhan, Q. Cylindrical vector beams: from mathematical concepts to applications. *Adv. Opt. Photonics* **1**, 1-57 (2009).
8. Liu, S.-J. *et al.* Bi-chiral nanostructures featuring dynamic optical rotatory dispersion for polychromatic light multiplexing. *Adv. Mater.* **35**, 2301714 (2023).
9. Zhou, T. *et al.* Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit. *Nat. Photonics* **15**, 367-373 (2021).
10. Miscuglio, M. *et al.* Massively parallel amplitude-only Fourier neural network. *Optica* **7**, 1812-1819 (2020).
11. Liu, C. *et al.* A programmable diffractive deep neural network based on a digital-coding metasurface array. *Nat. Electron.* **5**, 113-122 (2022).
12. Ríos, C. *et al.* In-memory computing on a photonic platform. *Sci. Adv.* **5**, eaau5759 (2019).
13. Youngblood, N., Ríos Ocampo, C. A., Pernice, W. H. P. & Bhaskaran, H. Integrated optical memristors. *Nat. Photonics* **17**, 561-572 (2023).

14. Zhou, T. *et al.* In situ optical backpropagation training of diffractive optical neural networks. *Photonics Res.* **8**, 940-953 (2020).
15. Neurophos. *https://www.neurophos.com/*.
16. Feldmann, J. *et al.* Parallel convolutional processing using an integrated photonic tensor core. *Nature* **589**, 52-58 (2021).
17. Hua, S. *et al.* An integrated large-scale photonic accelerator with ultralow latency. *Nature* **640**, 361-367 (2025).
18. Zhou, W. *et al.* In-memory photonic dot-product engine with electrically programmable weight banks. *Nat. Commun.* **14**, 2887 (2023).
19. Xu, Z. *et al.* Large-scale photonic chiplet Taichi empowers 160-TOPS/W artificial general intelligence. *Science* **384**, 202-209 (2024).
20. Chen, Y. *et al.* All-optical synthesis chip for large-scale intelligent semantic vision generation. *Science* **390**, 1259-1265 (2025).
21. Latifpour, M. H., Park, B. J., Yamamoto, Y. & Suh, M.-G. Hyperspectral in-memory computing with optical frequency combs and programmable optical memories. *Optica* **11**, 932-939 (2024).
22. Liang, Y. *et al.* High-clockrate free-space optical in-memory computing. *Light Sci. Appl.* **15**, 115 (2026).