

Modular Diffractive Neural Networks Using Cascaded Metasurfaces

Guannan Wang, Xiaofei Zang,* Zhiyu Tan, Teng Zhang, Zhe Gao, Yuanbo Wang, Deng Zhang, Alexander P. Shkurinov, Yiming Zhu,* and Songlin Zhuang

An all-optical diffractive neural network (DNN), as the fusion of optics and artificial intelligence, utilizes light-based computational architecture, providing potential beyond Moore's law limitations due to their low energy consumption and high parallel processing speed. However, existing DNN frameworks face limitations in realizing spatially assembled architecture (i.e., by combining two or more independent physical layers together) to create a Lego-like reconfigurable DNN that can generate additional functionalities. In this work, the modular programming concept is introduced to develop the modular diffractive neural networks (MDNNs) using the cascaded metasurfaces where each metasurface enables respective functionalities while the cascaded metasurfaces possess additional functionalities. The MDNNs showcase great advantages in flexibility and reconfigurability for multitask functionalities. When working independently, two metasurface-based modules can function as classifiers of handwritten letters and fashion products, respectively. When the two modules are assembled with different spatial orders, the additional functions of the handwritten digits classifier and imager can be obtained. Moreover, the MDNNs can be designed to mimic an architecture for high-capacity encrypted communication. This flexible and multiplexed MDNNs framework shows great potential and paves the new way for all-optical computation with multifunctional integration, massively parallel processing, and all-optical information security.

1. Introduction

Artificial intelligence (AI), with its ability to emulate human cognitive functions, such as information storage and analysis,^[1] has emerged as a pivotal force in accelerating the progress of natural sciences.^[2,3] Its applications span a diverse range of fields, including machine learning,^[4,5] pathology image diagnosis,^[6] the internet of things,^[7] and natural language processing.^[8] However, as AI progresses, the demand for high-speed and high-capacity data processing intensifies. Conventional electronic computing faces limitations due to the constraints of Moore's law and Dennard scaling.^[9] To address these challenges, optical neural networks (ONNs) have been proposed as a potential solution.^[10–17] ONNs utilize photons as carriers for information propagation and process, capitalizing on the inherent advantages of light-speed processing, parallel computation, and low energy consumption. Recent advancement in ONNs have demonstrated their potential in various fields, such as optical accelerators,^[18] machine vision,^[19] polarimetry,^[20] spectral detection,^[21] and topological photonics.^[22]

All-optical diffractive deep neural networks (D²NNs),^[23] as an important branch of ONNs, provide an all-optical computational framework for deep learning, which has significant advantages in high parallel and high-throughput arithmetic superior to other forms of ONNs.^[9] The neurons of the hidden layer in D²NNs can be constructed by the diffractive units (3D-printed subwavelength pillars) that modulate the phase or amplitude of the light field. Meanwhile, the interlayer diffraction mimics the convolution operation. Therefore, by training the neurons (i.e., modulation response of each diffractive unit) with an all-optical deep learning algorithm^[24–26] and detecting the output light field, the D²NNs are capable of neural information processing in the free space with light-speed, which has been widely employed in a plethora of applications, including object classification,^[27,28] optical mirage,^[29] imaging,^[30–34] photonic computing,^[35–37] information security,^[38,39] and polarization transformations.^[40] However, traditional D²NNs are designed with fixed functionalities, and the combination of two or

G. Wang, X. Zang, Z. Tan, T. Zhang, Z. Gao, Y. Wang, D. Zhang, Y. Zhu, S. Zhuang
Terahertz Technology Innovation Research Institute, and Shanghai Key Lab of Modern Optical System
University of Shanghai for Science and Technology
No. 516, JunGong Road, Shanghai 200093, China
E-mail: xfzang@usst.edu.cn; ymzhu@usst.edu.cn

Y. Wang, D. Zhang
Xi'an Institute of Applied Optics
Xi'an 710065, China

A. P. Shkurinov
Department of Physics
Lomonosov Moscow State University
Moscow 119992, Russia

X. Zang, Y. Zhu
Shanghai Institute of Intelligent Science and Technology
Tongji University
Shanghai 200092, China

The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/lpor.202500923>

DOI: 10.1002/lpor.202500923

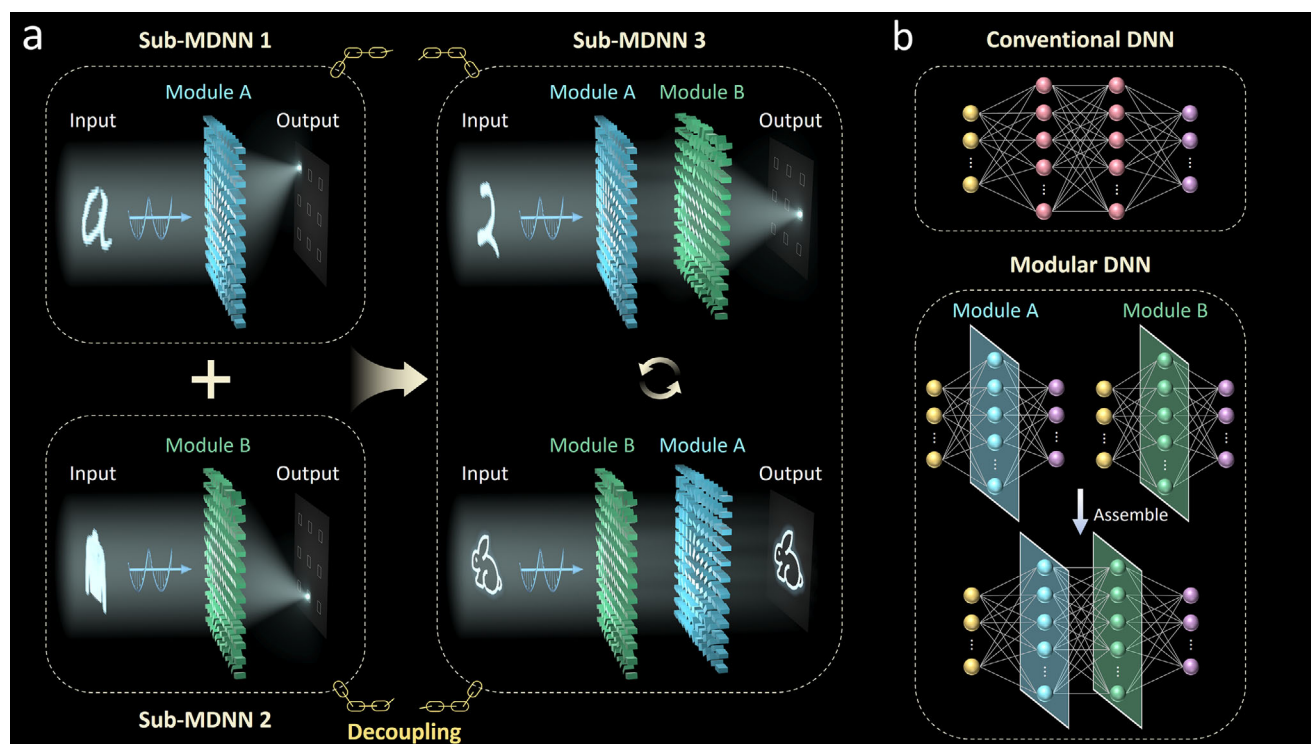


Figure 1. Schematic of modular DNNs with spatially assembled capability for processing multitask. a) Under the incidence of LCP THz waves, Sub-MDNN 1 and Sub-MDNN 2 composed of individual modules (Module A and Module B) perform the classification of handwritten letters and fashion products, Sub-MDNN 3 consisting of two modules in different assemble orders implement the classification of handwritten digits and the function of imaging. b) Comparison between the conventional DNN with rigid interlayer coupling and modular DNN with Lego-like modular assembly capability.

more independent physical layers together cannot generate new functionalities. Metasurfaces have come to the fore as a revolutionary optical technology,^[41–43] offering unprecedented capability that 3D-printed diffractive structures cannot achieve, such as tunable^[44] and polarization multiplexing,^[45–48] which provides great support for D²NNs to break through the stumbling block. Notably, spin-decoupled metasurfaces^[49] enable the spin-delinked functions in two orthogonal helical channels, providing a flexible platform to construct MDNNs and generate additional functionalities in different spin channels by assembling two or more spin-decoupled metasurfaces with different spatial sequence.

In this work, inspired by the modular programming design concept, we propose MDNNs constructed by cascaded spin-decoupled metasurfaces. These metasurfaces are optimized into a Lego-like architecture, enabling spatially assembled modules. Each metasurface in MDNNs functions as an independent module that can process distinct neural network tasks. Furthermore, these modules can be assembled to form new neural networks with additional functionalities that are different from that of each individual module. Meanwhile, the assembled network is order-dependent, enabling task switching based on the assembly order of modules, which is optimized through joint training. In an experiment, we fabricated the MDNNs composed of two metasurface-based modules to demonstrate its capabilities. When working independently, these metasurface-based modules are capable of classification of the handwritten letters and fashion products, respectively. By assembling these modules, distinct

network functions, the handwritten digits classifier, and the imager can be obtained through different spatial orders. All the classification tasks achieve recognition accuracy exceeding 90% in an experiment. In addition, we numerically demonstrate that the MDNNs can be applied to the high-capacity information encryption application by transforming four categories of input cipher-texts into various letters, digits, symbols, and images as the output plaintexts. The proposed modular design paradigm for multifunctional integration in DNNs provides a new horizon toward the advancement of all-optical computation, multimodal learning, information security, and resource conservation.

2. Results

2.1. MDNNs Design Paradigm

Figure 1a showcases the schematic of the proposed MDNNs, which consist of two spin-decoupled metasurface-based modules labeled Module A (blue) and Module B (green). Each module can realize the neural networks' functionality independently with an input layer, a hidden layer (i.e., a metasurface-based module), and an output layer located at the pre-designed positions. The input plane is composed of a metallic mask plate with a specifically fabricated hollowed pattern, and details about the fabricated masks are exhibited in Note S1 (Supporting Information). For Module A (labeled Sub-MDNN 1), and upon the illumination of left circularly polarized (LCP) terahertz (THz) waves, the metallic mask generates spatially distributed light in the shape of handwritten

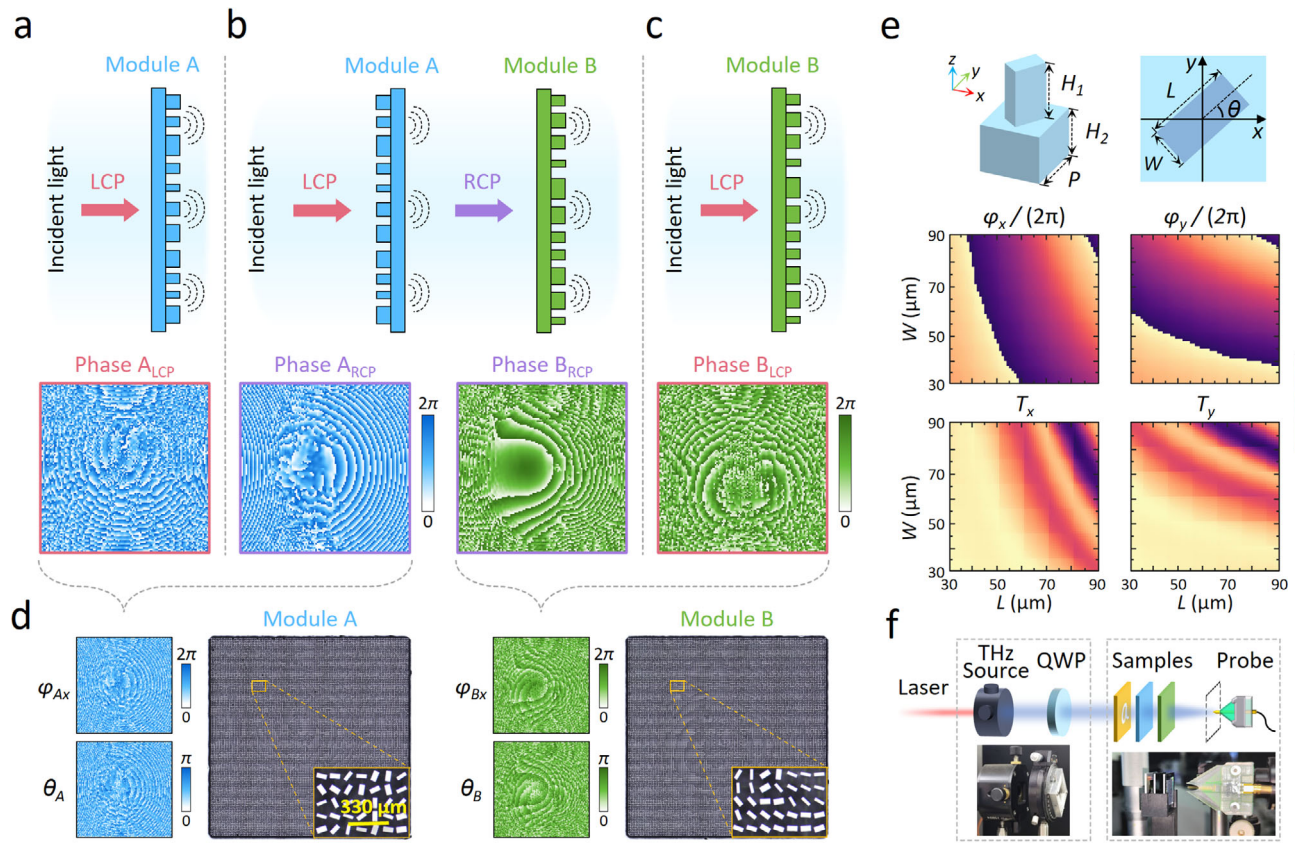


Figure 2. Design principle of MDNNs and experimental setup. a,b) The working mechanism of individual Module A and assembled-Module (Module A + Module B) under the LCP THz incidence. c) The working mechanism of individual Module B. d) Design methodology and fabricated experimental samples. The yellow mark indicates a length of 330 μm . e) Schematic diagram of a meta-unit in a spin-decoupled metasurface, and the simulated phase shifts (φ_x , φ_y) and transmission coefficients (T_x , T_y) under the illumination of x- and y-polarized THz waves. f) Experimental setup.

letters (e.g., “a”), and then, modulated by Module A, leading to the recognition of handwritten letters by diffracting the THz waves into specific regions in the output plane. Similarly, Module B constructs Sub-MDNN 2, and recognizes fashion products (e.g., “coat”) under the illumination of LCP THz waves. Furthermore, by assembling these two modules in different spatial orders but with the same interlayer distance, Sub-MDNN 3 is formed to realize distinct functions under the incidence of LCP THz waves. For instance, the assembly order A-to-B in Sub-MDNN 3 enables the classification of handwritten digits (e.g., “2”), while the order B-to-A facilitates the imaging operation. Notably, the functions of Sub-MDNN 1, Sub-MDNN 2, and Sub-MDNN 3 are independent, without coupling relationships and crosstalk. In addition, Figure 1b provides a schematic comparison of conventional DNN and modular DNN, illustrating the Lego-like modular assembly capability of MDNNs, whereas the rigid interlayer coupling of conventional DNN.

2.2. Principle and Methodology

To realize the modular operations, Figure 2a–c shows the inherent working mechanism of the MDNNs, which are composed of two spin-decoupled metasurfaces, served as Module A (blue) and

Module B (green). Each module consists of spin-related meta-units that enable the independent manipulation of two orthogonal circularly polarized components (i.e., LCP and RCP) of EM waves. Therefore, two spin-dependent phases can be encoded into each module for achieve distinct task requirements under the illumination of LCP and RCP THz wave. For the LCP wave incidence, Module A and Module B can modulate the light with Phase A_{LCP} and Phase B_{LCP}, generating Sub-MDNN 1 and Sub-MDNN 2 with distinct functionalities, respectively (Figure 2a,c). For Sub-MDNN 3 (Figure 2b), these two modules are assembled in a certain spatial order, such as A-to-B (or B-to-A) with a predesigned interlayer distance. Besides, the module located on the front side is rotated 180 degrees horizontally, forming a back-to-back spatial structure. Therefore, due to the geometric phase characteristic, under the illumination of LCP THz waves, the first module of Sub-MDNN 3 will modulate the light field with Phase A_{RCP} and transform the polarization state to RCP. Then, the conjugated electric field component (i.e., RCP component), diffracts to Module B and is modulated by the response of Phase B_{RCP}. Therefore, Sub-MDNN 3 can achieve additional task deployments decoupling from Sub-MDNN 1 and Sub-MDNN 2.

According to the Huygens-Fresnel principle,^[50] the wavefront generated by meta-units is composed of secondary spherical sources. Therefore, each spin-sensitive meta-unit within hidden

layers functions as an optical neuron under specific circularly polarized illumination, fully connected with neurons in the next layer through the diffraction process. For the n^{th} Sub-MDNN, according to the Rayleigh-Sommerfeld diffraction theory,^[51] the diffractive field of the $(l+1)^{\text{th}}$ layer can be expressed as:

$$E(\mathbf{r}^{l+1}) = \iint_S E(\mathbf{r}^l) \cdot e^{i\phi_n^l} \cdot h(\mathbf{r}^{l+1} - \mathbf{r}^l) dx dy, \quad (1)$$

where ϕ_n^l is the spin-decoupled phase of modules, and $h(\mathbf{r}^{l+1} - \mathbf{r}^l)$ is the impulse response that can be defined as:

$$h(\mathbf{r}^{l+1} - \mathbf{r}^l) = \frac{z^{l+1} - z^l}{R^2} \left(\frac{1}{2\pi R} - \frac{1}{j\lambda} \right) e^{j\left(\frac{2\pi R}{\lambda}\right)}, \quad (2)$$

where $R = \sqrt{(x^{l+1} - x^l)^2 + (y^{l+1} - y^l)^2 + (z^{l+1} - z^l)^2}$, λ is the working wavelength, and j is the imaginary number $\sqrt{-1}$. During training, we independently optimize the parameter ϕ_n^l of each Sub-MDNN by minimizing the mean squared error (MSE) loss function via the Adam stochastic gradient descent algorithm and ultimately obtain the Phase A/B_{LCP/RCP}. More details about the network training are supplied in Note S2 (Supporting Information).

After training, we fabricate each module with a 120×120 array of anisotropic meta-units, enabling spin-decoupled function. Each meta-unit has a period of $P = 110 \mu\text{m}$, and a height of $H_1 = H_2 = 500 \mu\text{m}$, as depicted in Figure 2e. With varying geometric parameters (L , W) and rotation angles (θ), the meta-units have different responses for the propagation phase and geometric phase, respectively. The relationship between the spin-decoupled phase ϕ_{LCP} and ϕ_{RCP} , the x -propagation phase ϕ_x , and the geometric phase θ is given by:^[52]

$$\begin{cases} \phi_{\text{LCP}} = \phi_x + 2\theta \\ \phi_{\text{RCP}} = \phi_x - 2\theta \end{cases} \quad (3)$$

Utilizing this equation and the trained Phase A/B_{LCP/RCP}, we calculate ϕ_x and θ as ϕ_{Ax} and θ_A for Module A, and ϕ_{Bx} , θ_B for Module B, as illustrated in Figure 2d. Subsequently, by leveraging the transmission coefficients and phase shifts under x - and y -polarized illumination at 0.7 THz (Figure 2e), and considering the conditions $\phi_x - \phi_x \approx \pi$ and $T_x \approx T_y \geq 70\%$, which are essential for creating a quasi-half-waveplate, we optimize and select the structural parameters of meta-units to construct the modules showcased in Figure 2d. In order to experimentally characterize the MDNNs, Figure 2f gives the schematic of the optical setup. A femtosecond laser beam impinges onto a photoconductive antenna to generate THz radiation, which is then modulated by a quarter-waveplate (QWP) and the samples and subsequently diffracted on a detection plane. A THz probe is located at the detection plane to collect the electric field distribution in the xy plane. Detailed information about the experimental design is provided in the Experimental Section. The holder for optical path alignment of modules and mask is shown in Note S3 (Supporting Information).

2.3. Experimental Demonstration of MDNNs for Four-Channel Multitask Processing

Figure 3 demonstrates the simulation and experiment results of the proposed MDNNs for three categories of classification tasks. For Sub-MDNN 1, the classical dataset from the Extension of Modified National Institute of Standards and Technology (EM-NIST) is utilized to train the recognition task of handwritten letters (letters “a”, “b”, and “c”). The convergence curves during training for classification tasks are provided in Note S4 (Supporting Information). The simulation and experimental validation are illustrated in Figure 3a. Here, the Sub-MDNN 1 consists of an input plane, a hidden layer (Module A), and an output plane, with the diffraction distance (i.e., propagation distance of light in free space) between the input plane and Module A, and between Module A and output plane are both set to $7500 \mu\text{m}$. To mimic the input of handwritten letter images, the input plane is constructed using a series of metallic masks featuring the shapes of each letter. The output plane is specifically designed with nine discrete recognition regions to accommodate the implementation of three types of classification tasks, where the top three detection regions are dedicated to displaying the classification results of handwritten letters. As shown in Figure 3a, Sub-MDNN 1 composed of the spin-decoupled metasurface-based Module A can work as the handwritten letter classifier. For the incidence of the letter “a”, the LCP diffractive fields will be modulated by Module A with Phase A_{LCP}, leading to the diffractive fields located at the top left corner in the output plane. The calculated energy distributions demonstrate that Sub-MDNN 1 can correctly recognize the handwritten “a.” Detailed information about the calculation method of electric-field energy distribution is provided in Note S5 (Supporting Information). A slight difference in the energy distribution between the simulation and experiment is observed due to experimental errors. Similarly, the letters “b” and “c” are recognized by Sub-MDNN 1, the diffractive fields are focused onto the top-middle and top-right recognition regions, respectively. The corresponding energy distributions confirm that Sub-MDNN 1 accurately identifies the handwritten “b” and “c.” To further characterize the classification performance, confusion matrices are provided to visualize the test results of correct and incorrect identifications, exhibiting a recognition accuracy of 90.2% in simulation and 93.3% in measurement. Here, the confusion matrices are characterized through the test dataset (which is excluded from the training dataset). For Sub-MDNN 2, Figure 3b illustrates the recognition results of the fashion products “Trouser”, “Coat”, and “Sneaker” from the Fashion-MNIST dataset. The diffraction distances set-up is identical to those in Sub-MDNN 1. For the incidence fashion products “Trouser”, “Coat”, and “Sneaker”, the LCP diffractive fields are modulated by the phase B_{LCP} of Module B and recognized with the diffractive fields focused on the bottom-left, bottom-middle, and bottom-right corner, respectively. The experiment and simulated results of the electric-field intensity distribution at the output plane and the corresponding energy distribution of the detection regions demonstrate the classification capability of Sub-MDNN 2. The calculated confusion matrix of the Sub-MDNN 2 showcased a simulated accuracy exceeding 97.6%, while the experiment accuracy is 100.0% (all 15 cases recognized). By assembling Module A and Module B in spatial order A-to-B, with Module A horizontally flipped by 180

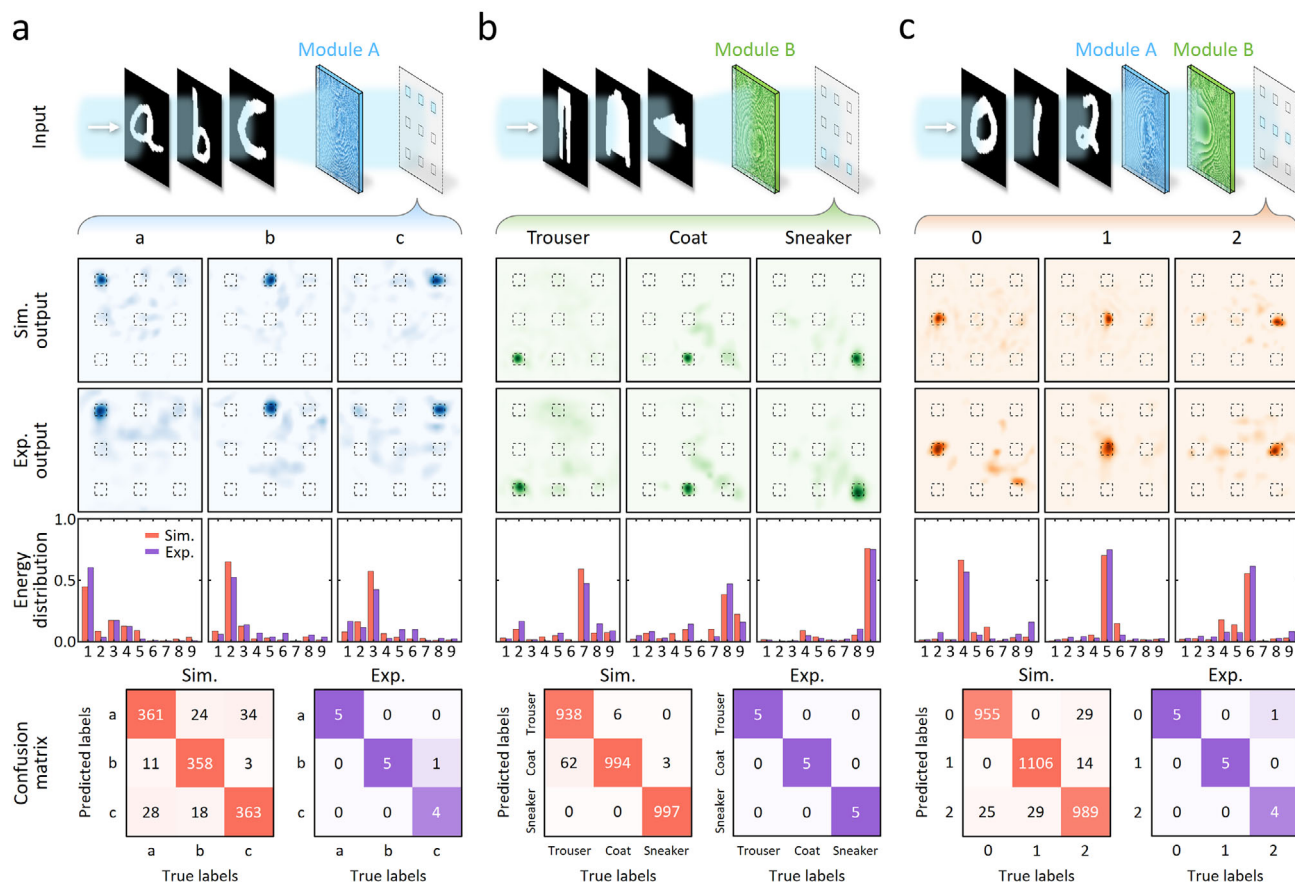


Figure 3. MDNNs for multitask processing through modular operations under the LCP incidence at 0.7 THz. a) Module A for the classification of handwritten letters “a”, “b”, “c”, and its output electric-field intensity distributions, analytic energy distributions (bar graphs) and confusion matrix in both simulation and experiment. b) Module B for the classification of fashion products “Trouser”, “Coat”, “Sneaker” and its output simulated and experimental results. c) Cascaded modules (Module A + Module B) in assemble order A-to-B for the classification of handwritten digits “0”, “1”, “2” and its output simulated and experimental results.

degrees, Sub-MDNN 3 can be constructed as shown in Figure 3c, enabling additional task deployments. Meanwhile, we demonstrate that the functions of Sub-MDNN 1, Sub-MDNN 2, and Sub-MDNN 3 are decoupled from each other, a detailed analysis is given in Note S6 (Supporting Information). The interlayer (input layer-Module A-Module B-output layer) diffraction distances are both set to 5000 μm . In the simulation and experiment, for the incidence handwritten digits “0”, “1”, and “2”, the diffractive LCP field is modulated by Module A with Phase A_{RCP} and converted to the RCP light. Then, the diffractive RCP fields are modulated by phase B_{RCP} of Module B, leading to the diffractive field focused into the middle-left, middle-middle, and middle-right recognition regions at the output plane, respectively. The measured electric-field intensity distribution and energy distribution demonstrate the correct classification of handwritten digits “0”, “1”, and “2”. The confusion matrices in the simulation showcase a recognition accuracy exceeding 96.9%, while in the experiment, it is 93.3%. To analyze the robustness in operating bandwidth/oblique incidence for Sub-MDNN 1, Sub-MDNN 2, and Sub-MDNN 3, detailed information is provided in Note S7/Note S8 (Supporting Information). The THz insertion loss of MDNNs induced by the metasurface module is discussed in Note S9 (Sup-

porting Information). Furthermore, we prove that the MDNNs can be promoted to performing more complex recognition tasks such as 10-class classification per category (for letters, fashion products, and digits), which is discussed in Note S10 (Supporting Information).

By reassembling these modules from the initial spatial order A-to-B (Figure 3c) to an inverted order B-to-A (with module B horizontally flipped by 180 degrees), the switched Sub-MDNN 3 enables the functionality of imaging, as shown in Figure 4. This capability is realized through a joint training method, depicted in Figure 4a, where the classification loss from the A-to-B network and the imaging loss from the B-to-A network are optimized simultaneously in Sub-MDNN 3, using the handwritten digits from the MNIST dataset for training. In the simulated and experimental validations, under the illumination of LCP THz waves and the input digits “0”, “1”, and “2” (Figure 4b), the imaging operation can be performed. As illustrated on the left side of Figure 4c, the measured diffractive electric-field intensity distributions in the output plane form the images of corresponding digits “0”, “1”, “2”, demonstrating its capability to implement imaging. In addition, due to the external generalization capability of the neural networks, we demonstrate the Sub-MDNN 3-based

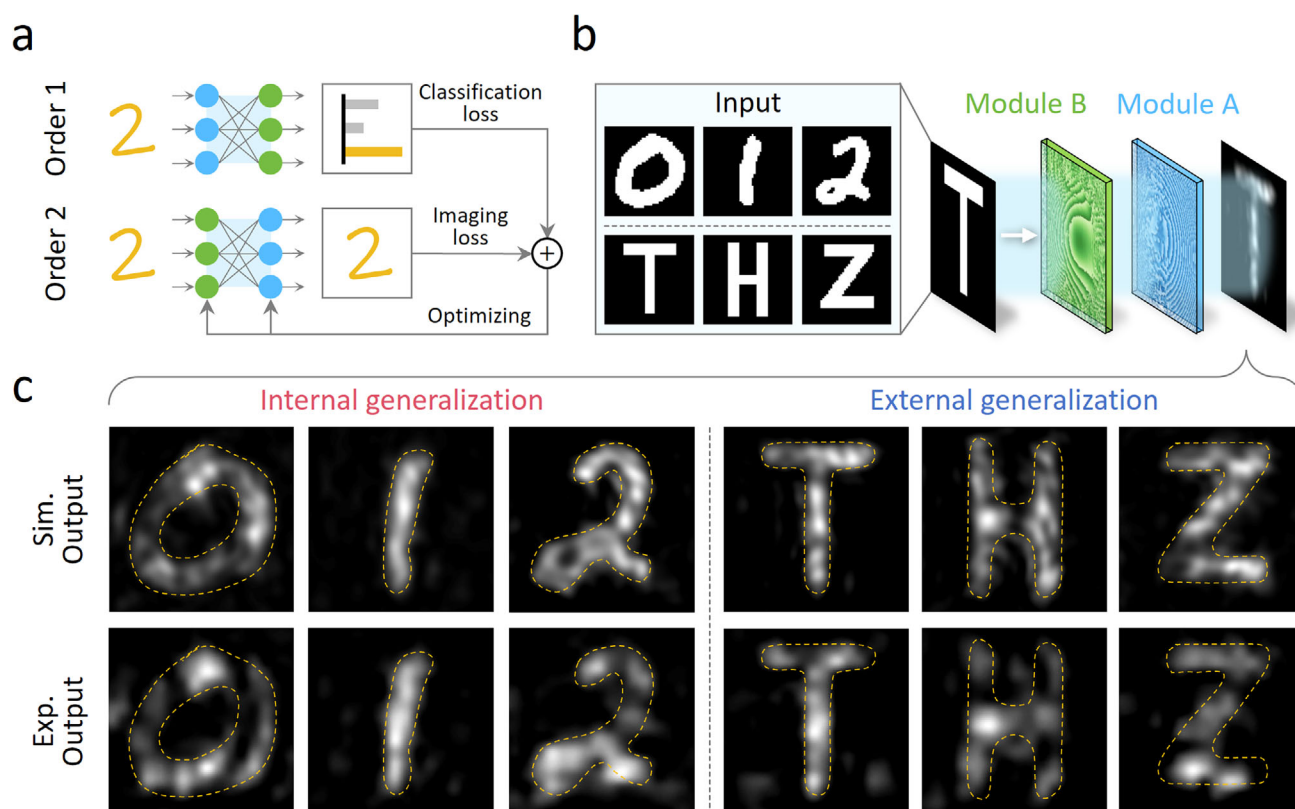


Figure 4. Order-dependent function of MDNNs. a) Schematic for the joint training of classification and imaging in two assemble orders. b) Cascaded modules (Module A + Module B) in assemble order B-to-A for the imaging, and the input images of “0”, “1”, “2” and “T”, “H”, “Z”. c) Output electric-field intensity distributions of simulation and experiment for the validation of internal generalization (i.e., imaging of “0”, “1”, “2”) and external generalization (i.e., imaging of “T”, “H”, “Z”).

imager for other untrained categories. For instance, we create custom binary images (i.e., “T”, “H”, and “Z”) as input to evaluate its external generalization. The results, presented on the right side of Figure 4c, showcase its validity for imaging the patterns “T”, “H”, and “Z”. To evaluate the impact of optical misalignment on MDNNs, detailed information is given in Note S11 (Supporting Information). The proposed MDNNs exhibit scalability that can integrate more functional modules (e.g., comprising Module A, Module B, and Module C), which is discussed in Note S12 (Supporting Information).

2.4. MDNNs Enable High-Capacity All-Optical Encryption for Four Types of Secret Messages

By leveraging the flexibility and reconfigurability of the proposed MDNNs, we offer a high-capacity all-optical encryption strategy for high-security information transmission. Figure 5 demonstrates three types of encryption operations and corresponding decryption scenarios. In the encryption progress, as depicted in Figure 5a, two categories of plaintext messages consisted of handwritten letters “M”, “E”, “T”, “A”, etc., and handwritten digits “2”, “0”, “2”, “5”, etc., are encrypted into the ciphertext messages comprising digits “0”, “1”, “2”, “3”, etc., and letters “a”, “b”, “c”, “d”, etc. This encryption process is achieved by training Sub-MDNN

1 and Sub-MDNN 2 with the handwritten digits and letters from MNIST and EMNIST datasets as the network input, while the specific binary images are the ground truth. After training, the encryption keys (Key1 and Key2) are generated as the network parameters phase A_{LCP} and phase B_{LCP} , respectively. The Sub-MDNN 3 is trained to handle more complex encryption contents, as shown in Figure 5b. For Sub-MDNN 3 in the assemble order A-to-B, the encryption progress compiles the element symbol “Si”, “Ti”, “Cu”, and “Au”, etc., into the double-handwritten digits “22”, “36”, “41”, and “70”, etc., through the training process. By joint training with another assemble order B-to-A, the animal images “monkey”, “bull”, “dog”, and “rabbit”, etc., are encrypted into the double-handwritten letters “ab”, “cc”, “nq”, and “rt”, etc., as shown in Figure 5c. Meanwhile, Key3 and Key4 are generated in two assembly sequences from phase A_{RCP} and phase B_{RCP} . Following this, Module A and Module B can be rapidly constructed using Phase $A/B_{LCP/RCP}$, as illustrated in Figure 5d–f. Then Meta-Key1 to Meta-Key4 are constructed by Sub-MDNN 1, Sub-MDNN 2, Sub-MDNN 3 assembled A-to-B and assembled B-to-A, respectively.

In the decryption progress, the recipients have different Meta-Keys, and only the authorized recipients can access the specific Meta-Keys. For instance, as depicted in Figure 5d, when the sender transmits encrypted information to agent “Alice” with Meta-Key1 and agent “Daniel” with Meta-Key2, they can

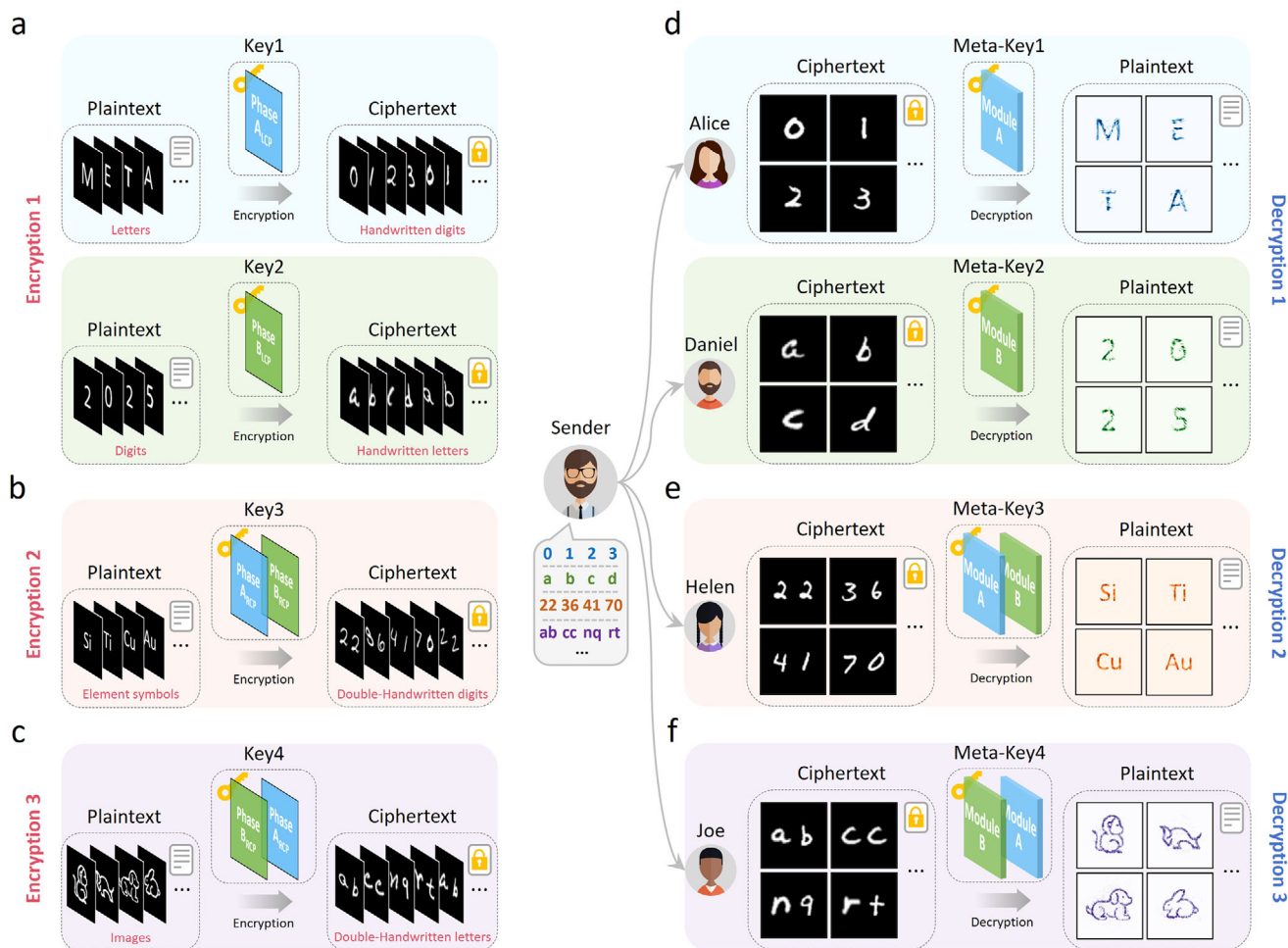


Figure 5. High-capacity all-optical information encryption strategy based on MDNNs. a–c) The encryption from four types of plaintexts into ciphertexts including handwritten “0”, “1”, “2”, “3” for Key1, “a”, “b”, “c”, “d” for Key2, “22”, “36”, “41”, “70” for Key3 and “ab”, “cc”, “nq”, “rt” for Key 4. d–f) The calculated electric-field intensity distribution at output plane, where four Meta-Keys (Meta-Key1 to Meta-Key4) are used to decrypt ciphertexts into plaintexts, including letters (“M”, “E”, “T”, “A”), digits (“2”, “0”, “2”, “5”), element symbols (“Si”, “Ti”, “Cu”, “Au”), and images (“monkey”, “bull”, “dog”, “rabbit”).

respectively decode the received ciphertext into the actual information. The calculated electric-field intensity distributions at the output plane successfully decrypt the plaintext “M”, “E”, “T”, “A”, and “2”, “0”, “2”, “5”. Agent “Helen” and “Joe” have the authority to possess both Module A and Module B, allowing them to use Meta-Key3 and Meta-Key4, separately (Figure 5e,f). The decrypted messages “Si”, “Ti”, “Cu”, “Au” and “monkey”, “bull”, “dog”, “rabbit” are revealed by “Helen” and “Joe” through the calculated electric-field intensity distributions, respectively. Notably, if the messages transmitted from the sender are stolen by unauthorized individuals, they can only obtain misleading and useless information. More detailed training and testing methods for MDNNs-based encryption are provided in Note S13 (Supporting Information).

3. Conclusion

In summary, we propose and experimentally validate a Lego-like MDNNs framework to implement diverse tasks through flexible

spatial assembly. By cascading two spin-decoupled metasurfaces-based modules and utilizing its spin-dependent polarization channel, MDNNs are constructed to process four distinct functions. The experiment results are in good agreement with the theory. Module A and Module B can function independently to implement the classification tasks for handwritten letters and fashion products, respectively. Furthermore, by assembling these modules in different spatial orders, distinct task deployments are achieved. Specifically, the handwritten digits classification and the imaging operation are concurrently realized in this assembled network from two assemble orders (i.e., A-to-B and B-to-A). Moreover, we demonstrate that the MDNNs framework can be promoted to versatile applications, such as high-capacity and high-security all-optical information transmission and encryption. Four types of secret messages, including letters, digits, element symbols, and animal images, are encrypted into a variety of ciphertexts, which only can be decoded by the Meta-Keys from MDNNs. This flexible, multi-function integrated MDNNs framework, along with its underlying mechanism, opens up new

possibilities for intelligent photonic systems, enabling real-time multitask processing and efficient resource utilization and promoting the development of next-generation all-optical computing technologies.

4. Experimental Section

Training Methods: The MDNNs were trained using PyTorch framework (v1.10.2) on Windows 11 operating system (Microsoft) with an Intel(R) Core(TM) i9-13900H CPU @2.6 GHz, a GeForce RTX 4060 GPU (NVIDIA), and a 128GB RAM. To optimize the phase parameters of each layer of the designed MDNNs, The mean squared error (MSE) loss function was used, and the Adam optimizer, with a training batch size of 8 and a learning rate of 0.001.

Sample Fabrication: The spin-decoupled metasurfaces were fabricated on 1000- μm -thick intrinsic silicon wafers (Orient: $\langle 100 \rangle$, R: $>10\,000\,\Omega\text{-cm}$, Polish: 2sp) from HF-Kejing company. The process begins with ultrasonic cleaning, followed by AZP4620 photoresist layer (Thickness: 7–8 μm) spin-coating (6000 RPM), and baking (100 $^{\circ}\text{C}$, 1 min). After exposed using a mask aligner (IMP SF-100), the resist film with micro-pillars was baked again (100 $^{\circ}\text{C}$, 2 min) and etched using the DRIE (SF6/C4F8, 95 min). Meanwhile, the golden masks were made by spin-coating the photoresist onto one side of the silica film, then exposing, metal coating and ultrasonic stripping, finally serve as the inputs of MDNNs.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

X.Z. and Y.Z. acknowledge the National Natural Science Foundation of China (61988102, 62271320, and 61871268), the “Shuguang” Program of Shanghai Education Commission (19SG44), the Science and Technology Commission of Shanghai Municipality (21S31907400), and Shanghai Municipal Science and Technology Commission (22JC1400200).

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

X.Z. and G.W. initiated the idea. G.W. conducted the numerical simulations. G.W. performed the measurements. G.W., X.Z., and Z.T. prepared the manuscript. X.Z., Z.T. and Y.Z. supervised the project. All the authors discussed and analysed the results.

Data Availability Statement

The data that support the findings of this study are available in the Supporting Information of this article.

Keywords

encryption, imaging, metasurfaces, modular diffractive neural networks, multitasking, parallel processing, pattern recognition

Received: April 18, 2025

Revised: June 4, 2025

Published online:

- [1] A. Krogh, *Nat. Biotechnol.* **2008**, 26, 195.
- [2] Y. LeCun, Y. Bengio, G. Hinton, *Nature* **2015**, 521, 436.
- [3] Y. Xu, F. Wang, T. Zhang, *Innovation* **2024**, 5, 100725.
- [4] J. Suo, W. Zhang, J. Gong, X. Yuan, D. J. Brady, Q. Dai, *Proc. IEEE* **2023**, 111, 1607.
- [5] Y. Jia, C. Qian, Z. Fan, T. Cai, E.-P. Li, H. Chen, *Light Sci. Appl.* **2023**, 12, 82.
- [6] M. K. K. Niazi, A. V. Parwani, M. N. Gurcam, *Lancet Oncol.* **2019**, 20, 253.
- [7] Z. Sun, M. Zhu, Z. Zhang, Z. Chen, Q. Shi, X. Shan, R. C. H. Yeow, C. Lee, *Adv. Sci.* **2021**, 8, 2100230.
- [8] R. Collobert, J. Weston, *Proc. of the 25th Int. Conf. on Machine Learning*, Association for Computing Machinery, Helsinki, Finland **2008**, pp. 160–167.
- [9] H. Chen, S. Lou, Q. Wang, P. Huang, H. Duan, Y. Hu, *Appl. Phys. Rev.* **2024**, 11, 021332.
- [10] H. J. Caulfield, J. Kinser, S. K. Rogers, *Proc. IEEE* **1989**, 77, 1573.
- [11] X. Xu, M. Tan, B. Corcoran, J. Wu, A. Boes, T. G. Nguyen, S. T. Chu, B. E. Little, D. G. Hicks, R. Morandotti, A. Mitchell, D. J. Moss, *Nature* **2021**, 589, 44.
- [12] T. Zhou, X. Lin, J. Wu, Y. Chen, H. Xie, Y. Li, J. Fan, H. Wu, L. Fang, Q. Dai, *Nat. Photonics* **2021**, 15, 367.
- [13] F. Ashtiani, A. J. Geers, F. Aflatouni, *Nature* **2022**, 606, 501.
- [14] Z. Xue, T. Zhou, Z. Xu, S. Yu, Q. Dai, L. Fang, *Nature* **2024**, 632, 280.
- [15] G. Westzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljačić, C. Denz, D. A. B. Miller, D. Psaltis, *Nature* **2020**, 588, 39.
- [16] M. K. Chen, X. Liu, Y. Sun, D. P. Tsai, *Chem. Rev.* **2022**, 122, 15356.
- [17] X. Liu, M. K. Chen, D. P. Tsai, *Laser Photonics Rev.* **2024**, 18, 2300456.
- [18] A. Song, S. N. M. Kottapalli, R. Goyal, B. Schölkopf, P. Fischer, *Nat. Commun.* **2024**, 15, 10692.
- [19] H. Zheng, Q. Liu, I. I. Kravchenko, X. Zhang, Y. Huo, J. G. Valentine, *Nat. Nanotechnol.* **2024**, 19, 471.
- [20] C. Chen, X. Xiao, X. Ye, J. Sun, J. Ji, R. Yu, W. Song, S. Zhu, T. Li, *Light Sci. Appl.* **2023**, 12, 288.
- [21] H. Jiang, Y. Chen, W. Guo, Y. Zhang, R. Zhou, M. Gu, F. Zhong, Z. Ni, J. Lu, *Nat. Commun.* **2024**, 15, 8347.
- [22] Z. Xue, T. Zhou, Z. Xu, S. Yu, Q. Dai, *Nature* **2024**, 632, 280.
- [23] X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, A. Ozcan, *Science* **2018**, 361, 1004.
- [24] Z. Fan, C. Qian, Y. Jia, Y. Feng, H. Qian, E.-P. Li, R. Fleury, H. Chen, *Nat. Commun.* **2024**, 15, 9416.
- [25] H. Chi, Y. Hu, X. Ou, Y. Jiang, D. Yu, S. Lou, Q. Wang, Q. Xie, C.-W. Qiu, H. Duan, *Adv. Mater.* **2025**, 37, 2419621.
- [26] Y. Wang, A. Yu, Y. Cheng, J. Qi, *Laser Photonics Rev.* **2024**, 18, 2300903.
- [27] X. Luo, Y. Hu, X. Ou, X. Li, J. Lai, N. Liu, X. Cheng, A. Pan, H. Duan, *Light Sci. Appl.* **2022**, 11, 158.
- [28] C. He, D. Zhao, F. Fan, H. Zhou, X. Li, Y. Li, J. Li, F. Dong, Y.-X. Miao, Y. Wang, L. Huang, *Opto-Electron. Adv.* **2024**, 7, 230005.
- [29] C. Qian, Z. Wang, H. Qian, T. Cai, B. Zheng, X. Lin, Y. Shen, I. Kaminer, E. Li, H. Chen, *Nat. Commun.* **2022**, 13, 2694.
- [30] J. Li, T. Gan, Y. Zhao, B. Bai, C.-Y. Shen, S. Sun, M. Jarrahi, A. Ozcan, *Sci. Adv.* **2023**, 9, adg1505.
- [31] B. Bai, X. Yang, T. Gan, J. Li, D. Meng, M. Jarrahi, A. Ozcan, *Light Sci. Appl.* **2024**, 13, 178.
- [32] J. Li, Y. Li, T. Gan, C.-Y. Shen, M. Jarrahi, A. Ozcan, *Light Sci. Appl.* **2024**, 13, 120.
- [33] J. Hu, K. Liao, N. U. Dinc, C. Gigli, B. Bai, T. Gan, X. Li, H. Chen, X. Yang, Y. Li, Ç. Işıl, M. S. S. Rahman, J. Li, X. Hu, M. Jarrahi, D. Psaltis, A. Ozcan, *eLight* **2024**, 4, 8.
- [34] Ç. Işıl, T. Gan, F. O. Ardic, K. Montesoglu, J. Digani, H. Karaca, H. Chen, J. Li, D. Meng, M. Jarrahi, K. Akşit, A. Ozcan, *Light Sci. Appl.* **2024**, 13, 43.

- [35] Z. Zhan, H. Wang, Q. Liu, X. Fu, *Nat. Commun.* **2024**, *15*, 10643.
- [36] X. Ding, Z. Zhao, P. Xie, D. Cai, F. Meng, C. Wang, Q. Wu, J. Liu, S. N. Burokur, G. Hu, *Adv. Mater.* **2024**, *36*, 2308993.
- [37] Z. Zhao, Y. Wang, C. Guan, K. Zhang, Q. Wu, H. Li, J. Liu, S. N. Burokur, X. Ding, *PhotoniX* **2022**, *3*, 15.
- [38] B. Bai, R. Lee, Y. Li, T. Gan, Y. Wang, M. Jarrahi, A. Ozcan, *Sci. Adv.* **2024**, *10*, adn9420.
- [39] G. Ma, X. Yang, B. Bai, J. Li, Y. Li, T. Gan, C.-Y. Shen, Y. Zhang, Y. Li, Ç. Işıl, M. Jarrahi, A. Ozcan, *Laser Photonics Rev.* **2024**, *18*, 2400238.
- [40] Y. Li, J. Li, Y. Zhao, T. Gan, J. Hu, M. Jarrahi, A. Ozcan, *Adv. Mater.* **2023**, *35*, 2303395.
- [41] N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, Z. Gaburro, *Science* **2011**, *334*, 333.
- [42] L. Huang, X. Chen, H. Mühlenbernd, G. Li, B. Bai, Q. Tan, G. Jin, T. Zentgraf, S. Zhang, *Nano Lett.* **2012**, *12*, 5750.
- [43] C. Qian, I. Kaminer, H. Chen, *Nat. Commun.* **2025**, *16*, 1154.
- [44] Y. Jia, H. Lu, Z. Fan, B. Wu, F. Qu, M.-J. Zhao, C. Qian, H. Chen, *IEEE Trans. Antennas Propag.* **2024**, *72*, 4532.
- [45] Y. Wang, Y. Wang, A. Yu, M. Hu, Q. Wang, C. Pang, H. Xiong, Y. Cheng, J. Qi, *Adv. Mater.* **2025**, *37*, 2408978.
- [46] H. Li, A. Yu, C. Pang, Y. Wang, J. Qi, *Laser Photonics Rev.* **2024**, *18*, 2400300.
- [47] Y. Wang, C. Pang, J. Qi, *Laser Photonics Rev.* **2024**, *18*, 2300832.
- [48] Y. H. Wang, C. Pang, Y. Wang, J. Qi, *Ann. Phys.* **2025**, *537*, 2400158.
- [49] R. C. Devlin, A. Ambrosio, N. A. Rubin, J. P. B. Mueller, F. Capasso, *Science* **2017**, *358*, 896.
- [50] L. Mandel, E. Wolf, *J. Opt. Soc. Am.* **1961**, *51*, 815.
- [51] E. Wolf, E. W. Marchand, *J. Opt. Soc. Am.* **1964**, *54*, 587.
- [52] S. Li, X. Li, G. Wang, S. Liu, L. Zhang, C. Zeng, L. Wang, Q. Sun, W. Zhao, W. Zhang, *Adv. Opt. Mater.* **2019**, *7*, 1801365.