

Advanced Image Classification Using a Differential Diffractive Network with “Learned” Structured Illumination

Jiajun Zhang, Shuyan Zhang, Weijie Shi, Yong Hu, Zheng-Gao Dong, Jiaqi Li,* and Weibing Lu



Cite This: *ACS Photonics* 2024, 11, 5289–5298



Read Online

ACCESS |

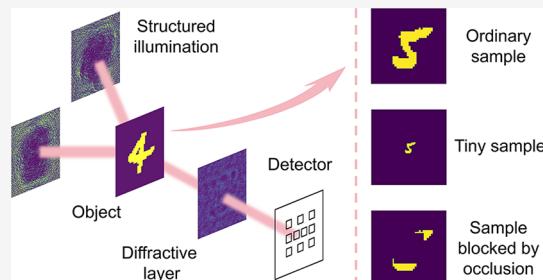
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: As a new optical machine learning framework, the diffractive deep neural network (D^2NN) has attracted much attention due to its advantages such as low power consumption, parallel computing, and fast execution speed. Here, we demonstrate a new optical neural network design of a differential D^2NN with structured illumination. In this scheme, the illumination patterns participate in the training process of the network and are optimized by an end-to-end technique. With the application of differential detection, the non-negativity constraint in a diffractive neural network can be alleviated. The test results show that this network architecture can achieve 97.63 and 88.10% classification accuracies on the MNIST and Fashion-MNIST data sets using only one diffractive layer, which exceeds the effect achieved by the five-layer traditional D^2NN . Moreover, this network architecture can achieve a comprehensive improvement over a traditional D^2NN in the challenging classification problems of tiny samples and samples blocked by occlusions. Compared with the traditional D^2NN , this scheme innovatively uses the illumination patterns as new degrees of freedom in system design, which can effectively improve classification ability and reduce the space complexity of the optical neural network.

KEYWORDS: optical computing, optical machine learning, optical neural network, diffractive deep neural network, optical information processing



INTRODUCTION

Photons are ideal carriers for information transmission and processing, which enable optical computing to be more efficient.^{1–5} The diffractive deep neural network (D^2NN) based on photonic technology is considered to be one of the effective ways to solve the limited development of traditional electronic artificial neural networks in the post-Moore era. As an emerging optical computing platform, the D^2NN can use optical devices such as metasurfaces to achieve physical construction and perform machine learning tasks in an all-optical manner.^{6,7} Based on its ability in optical information processing, the D^2NN has been applied to computational imaging,^{8–13} wavefront processing,^{14–16} optical logic operation,^{17–21} image recognition,^{22–27} etc. The D^2NN cascades multiple diffractive layers to ensure that the network can normally perform machine learning tasks. However, the cascading of multiple diffractive layers will bring about alignment and processing difficulties in practical applications.²⁸ Especially when the system uses visible light, the error accumulation caused by misalignment of multiple layers is more prominent because the size of each diffractive layer is generally small. Due to this misalignment, the performance of the D^2NN will be lower than expected.

When multiple diffractive layers are cascaded, the D^2NN will have larger space requirements and alignment challenges. In a sense, the alignment problems can only be mitigated as much

as possible, not avoided. To alleviate the alignment problem, one feasible solution is to introduce random errors during the training process, which can make the D^2NN work normally even with a certain misalignment.^{29,30} But, this method does not reduce the spatial complexity of the D^2NN . Another way is to use optical processing layers as few as possible while keeping its inference capability. In the specific implementation, this effect can be achieved by realizing system polarization multiplexing,³¹ optimizing the system's Fresnel number^{32–34} or utilizing knowledge distillation^{35,36} to compress the network. Since fewer optical processing layers are used, the error accumulation caused by misalignment of multiple layers can be alleviated, and the system takes up less space.

In this work, we demonstrate a differential D^2NN system with structured illumination. Structured light is used to illuminate the sample to form the input of the differential D^2NN , and the optical network processes the input light field, containing sample information. When there is only one diffractive layer, this network architecture can achieve 97.63

Received: August 13, 2024

Revised: November 7, 2024

Accepted: November 13, 2024

Published: November 20, 2024



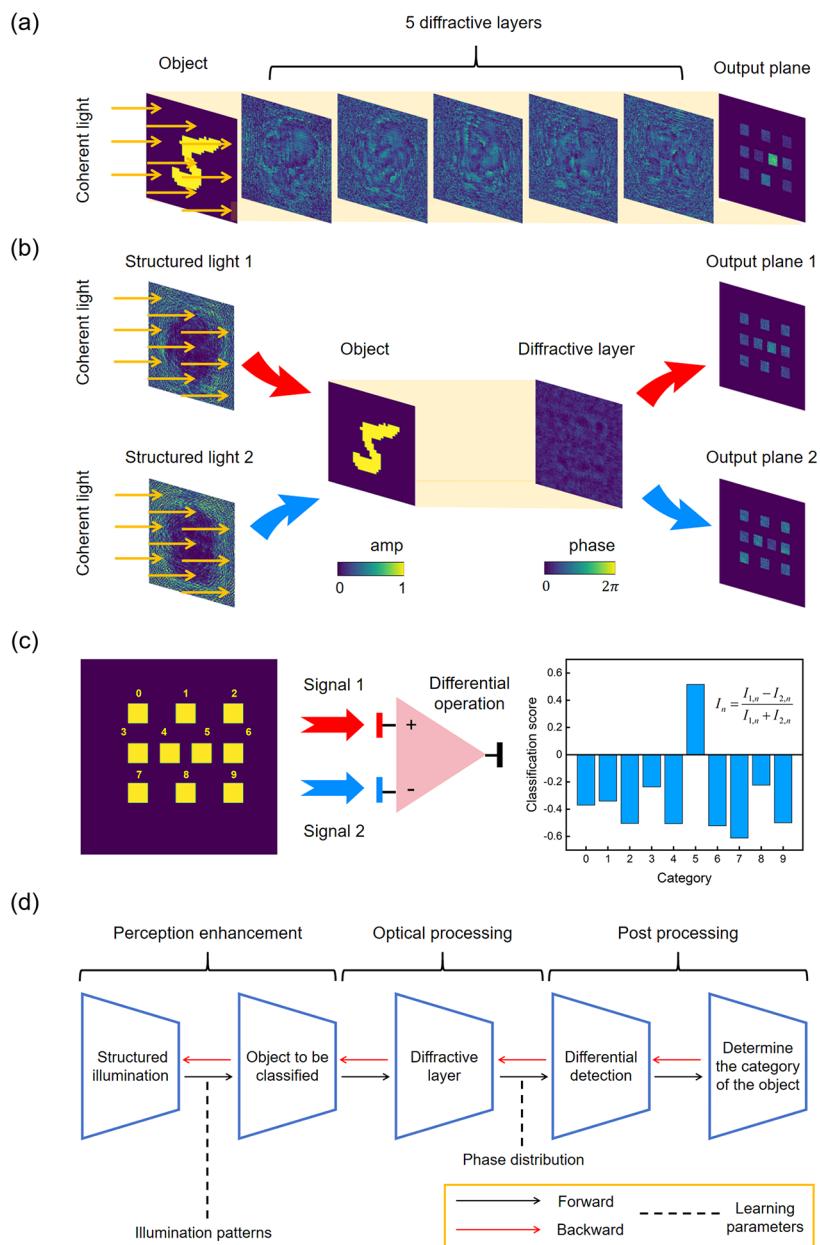


Figure 1. (a) Five-layer traditional D²NN. (b) Single-layer differential D²NN with structured illumination. (c) Schematic diagram of differential operation. (d) Optimization process of the entire diffractive neural network.

and 88.10% classification accuracies on the MNIST and Fashion-MNIST data sets. The classification accuracy exceeds the performance achieved by the five-layer traditional D²NN. Furthermore, the classification performance of this network architecture on tiny samples and samples blocked by occlusions is also discussed. The results show that this network architecture can be applied to more complex application scenarios compared with the traditional D²NN. This can effectively improve the reliance of the diffractive neural network on the cascade of multiple diffractive layers when performing complex tasks so that the network can achieve better task execution results and keep lower spatial complexity.

RESULTS AND DISCUSSION

Network Architecture Design. The traditional D²NN design is shown in Figure 1a. The D²NN composed of multiple

diffractive layers can modulate the incident light field so that the output light field has a specific light intensity energy distribution form. By determining the final light intensity energy distribution, deep learning tasks can be performed at the speed of light. However, the traditional D²NN requires enough diffractive layers to ensure that machine learning tasks can be completed normally. Therefore, when the number of diffractive layers is insufficient, the inference ability of the D²NN will be significantly limited. However, increasing the number of diffractive layers in the network structure will bring about problems in alignment and processing, which is a contradiction that exists in the traditional D²NN.

Structured light fields can carry abundant spatial amplitude information and have the potential to greatly increase information capacity.^{37–39} When structured light is used, the object is illuminated by the varying patterns with specific coding structures. This allows sample features to be naturally

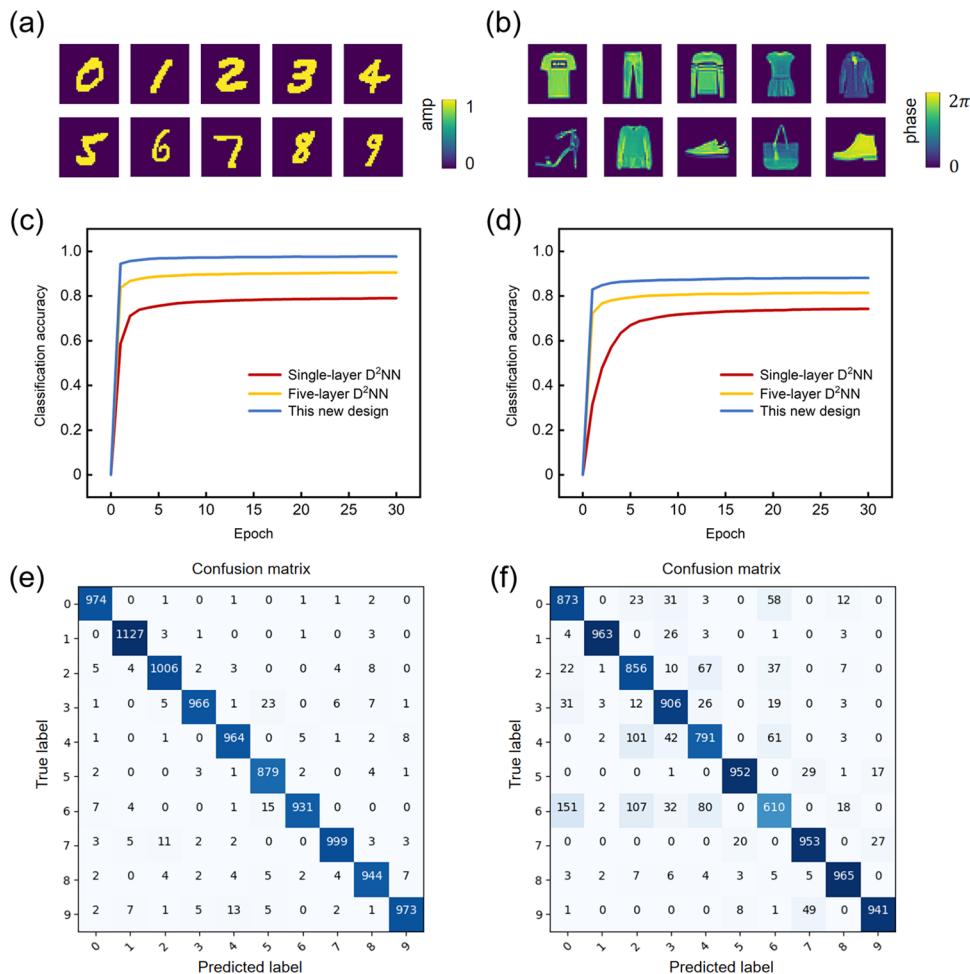


Figure 2. (a) Samples from the MNIST data set using amplitude encoding. (b) Samples of the Fashion-MNIST data set using phase encoding. (c) and (d) Classification accuracy achieved by different network architectures on the MNIST and Fashion-MNIST data sets. (e) and (f) Classification confusion matrices of the single-layer differential D²NN with structured illumination on 10,000 test images from the MNIST and Fashion-MNIST data sets.

included in the input light field, which can enable enhanced perception. Here, we propose a diffractive network framework of the differential D²NN with structured illumination. A single-layer differential D²NN is taken as an example (Figure 1b). In this network framework, the light source illuminating the sample to be classified is structured light with specific encoding instead of a simple coherent light. The illumination patterns can participate in the training and optimization process of the entire D²NN as new degrees of freedom. After training, the input light field can be encoded so that the illumination pattern contains features of the sample. Therefore, when structured illumination is used, the category differences of the samples will be amplified. This means that fewer diffractive layers can be used to complete machine learning tasks.

As another method to improve network performance, we choose differential detection to obtain classification scores for each category (Figure 1c). For the traditional D²NN, the category of the sample to be classified is mainly determined based on the light intensity of each predefined area. Thus, the non-negativity of the light intensity will limit the performance of the network.⁴⁰ In the scheme proposed in this paper, the D²NN will generate two output light fields under the two structured illuminations. By recording the light intensity of each area under the structured illumination, the final

classification score of each category can be obtained according to the corresponding difference calculation method (the training method part in the Methods Section). In this way, the final classification score of each category is freed from the non-negative constraint of light intensity.

It should be pointed out that both the traditional D²NN and the new framework that we proposed are based on the complex amplitude modulation capability of the coherent optical system. However, this is still a challenge for an incoherent optical system to maintain such wavefront processing capability. Therefore, the use of coherent light source is still the first choice for the D²NN. The optimization process of this new network framework is shown in Figure 1d. In this framework, the application of structured light can enhance the perception capabilities of this optical network. Through optical processing and post processing, the final category of the sample can be determined. With end-to-end technology, the illumination patterns and phase distribution of the diffractive layer will be optimized simultaneously.

To make the training process match the real physical process, the size of the diffraction neuron will depend on the size of the unit structure on the diffractive layer. Generally, a spatial light modulator (SLM) or metasurface is used to physically realize the function of the diffractive layer. Taking

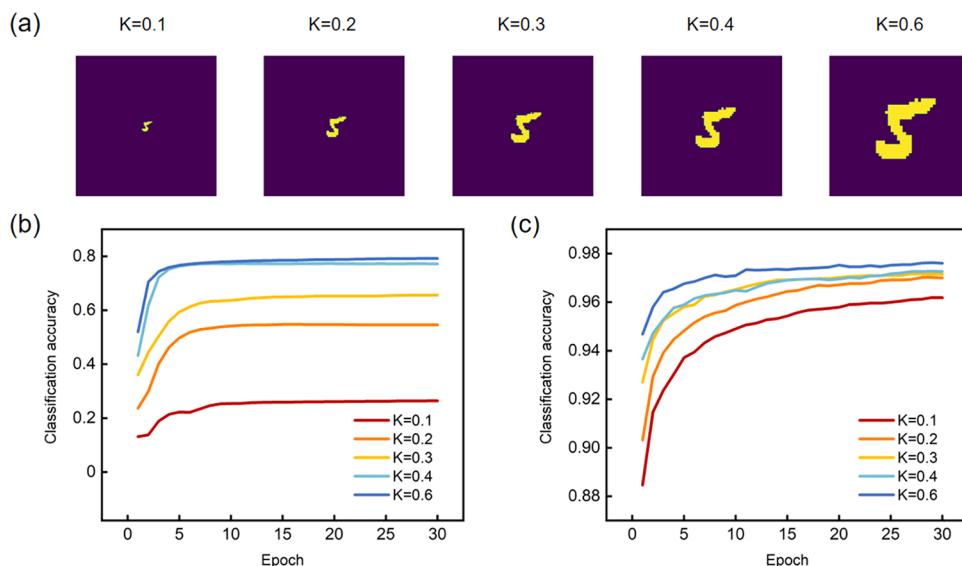


Figure 3. (a) Input light field with different K values. (b) Classification performance of the single-layer traditional D^2 NN on MNIST data set under different K values. (c) Classification performance of the single-layer differential D^2 NN with structured illumination on MNIST data set under different K values.

the PLUTO-NIR-015-C SLM (HOLOEYE) as an example, the resolution is 1920×1080 , the pixel unit size is $8 \mu\text{m}$, the operating wavelength is $650\text{--}1100 \text{ nm}$, and it can well realize the phase manipulation of the incident light field. Based on the modified maximum half-cone diffraction angle formula (*Supporting Section S1*), we choose the system parameters as the light source wavelength $\lambda = 650 \text{ nm}$, the size of diffraction neurons $d_f = 8 \mu\text{m}$, the number of neurons in the diffractive layer $M \times M = 200 \times 200$, and the spacing between layers $D = 5 \text{ cm}$. After calculation, it is found that such system parameter selection can perfectly realize fully connected diffraction neurons between layers.

Here, we load the images of the MNIST data set into the amplitude channel and the Fashion-MNIST data set into the phase channel (*Figure 2a,b*). Since the size of each sample in the MNIST and Fashion-MNIST data sets is 28×28 , up-sampling is required. As shown in *Figure 2c,d*, it can be seen that the traditional single-layer D^2 NN can only achieve 79.04 and 74.21% classification accuracies. By increasing the number of diffractive layers, the five-layer traditional D^2 NN system can achieve 90.44 and 81.43% classification accuracies. Under the same system parameters, the single layer differential D^2 NN with structured illumination can achieve 97.63 and 88.10% classification accuracies on 10,000 test images of the MNIST and Fashion-MNIST data set (the confusion matrices are given in *Figure 2e,f*). It should be noted that different model effects can be achieved by using different loss functions. Here, the MSE loss function is used to train the traditional D^2 NN model mentioned above. Although using the CE loss function can achieve a higher classification accuracy, the signal-to-noise ratio will decrease (*Supporting Section S2*). With only a single diffractive layer, the new framework can achieve results that exceed those achieved by the five-layer traditional D^2 NN. It effectively alleviates the D^2 NN's dependence on the number of diffractive layers and the alignment problems caused by multilayer cascades.

Performance on the Challenging Application Scenarios. Due to the loss of feature information, the classification of tiny samples and samples blocked by occlusions are

challenging issues for the traditional D^2 NN. Here, we use a single-layer differential D^2 NN with structured illumination and test the network's ability in these complex application scenarios. As mentioned above, the number of neurons in the diffractive layer is $M \times M = 200 \times 200$. Since the size of each sample to be classified for the MNIST data is set to 28×28 , up-sampling is required to make it consistent with the expected size. We assume that the sample images are first expanded from 28×28 to $m \times m$ ($m \leq M$) and then expanded to $M \times M$ by zero padding. Here, we define the light field ratio factor $K = m/M$. Different proportion factors K mean different sizes of light field input (*Figure 3a*). This is used to compare the perception and classification capabilities of the different D^2 NN frameworks for tiny samples.

In the traditional framework, we find that the classification accuracy of the single-layer D^2 NN is affected by the change of the size of the input light field (*Figure 3b*). As the K value changes from 0.1 to 0.4, the classification accuracy of the single-layer traditional D^2 NN will gradually increase. Then, if the K value continues to increase, the performance of the network will not change significantly. This is because the information is naturally contained within the light wavefront in the D^2 NN. When the K value is small, the input light field contains less information, and the performance of the entire network is limited by the input light field. Therefore, increasing the input light field within a certain range will bring more optical information and can effectively improve the performance of the network. However, the entire network begins to be limited by the number of diffractive layers when the input light field carries enough information. The system performance will basically not be significantly improved with the K value being further increased. When $K = 0.1$, only 26.41% classification accuracy can be achieved. Although a certain degree of improvement can be achieved by adding diffractive layers (when $K = 0.1$, the five-layer D^2 NN can achieve a classification accuracy of 66.98%), the effect is not obvious. Using the same system parameters, the single-layer differential D^2 NN with structured illumination can always maintain a classification accuracy of more than 94% (*Figure 3c*). This fully

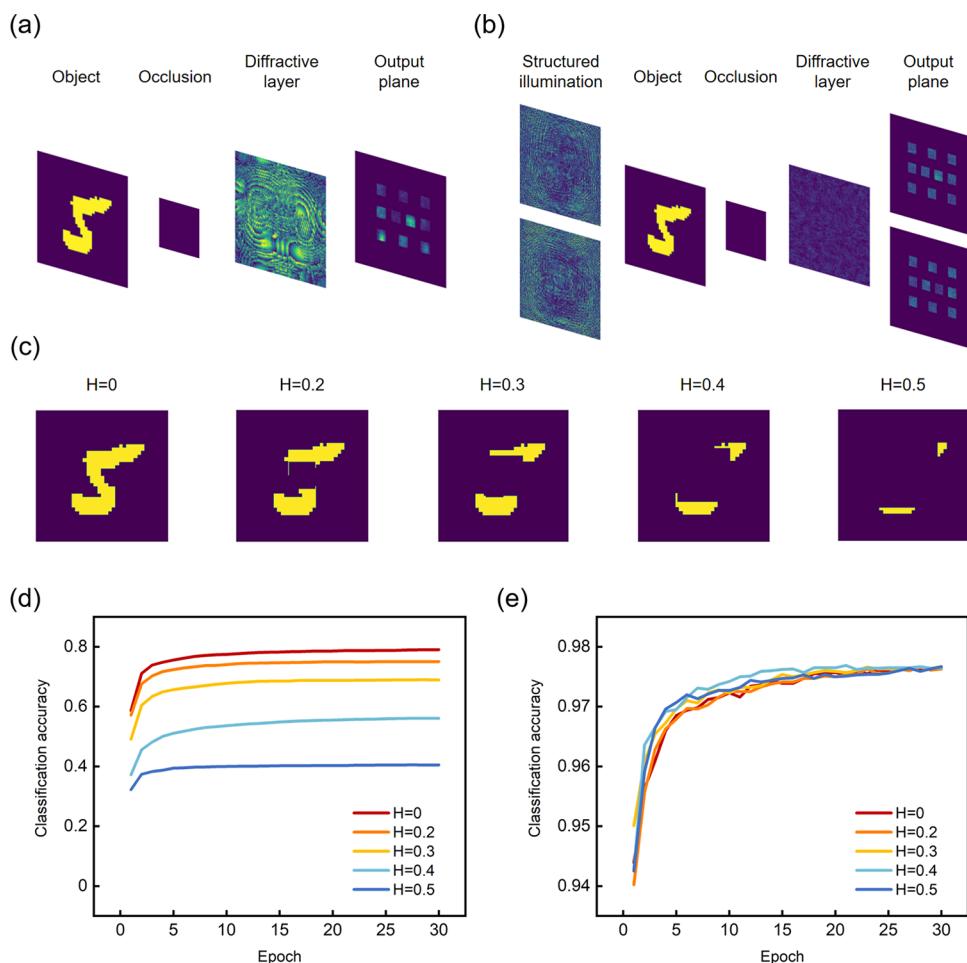


Figure 4. (a) and (b) Add occlusions in different network architectures, and the distance from the occlusions to the object and the diffractive layer is set to 2.5 cm. (c) Effect of the occlusions on the light field under different H values. (d) Classification performance of the single-layer traditional D²NN on MNIST data set under different H values. (e) Classification performance of the single-layer differential D²NN with structured illumination on the MNIST data set under different H values.

demonstrates that the network architecture proposed in this paper is not sensitive to the input light field and has the ability to perceive and classify the tiny samples.

For the classification task of samples blocked by occlusions, the detection scenes are constructed to test the performance of different network architectures (Figure 4a,b). We also set that the number of neurons in the diffractive layer is $M \times M = 200 \times 200$, and the light field ratio factor $K = 0.75$. The occlusion is placed between the object and the diffractive layer to block the transmission of the light wavefront in this area. Assuming that the size of the opaque occlusion is $h \times h$ ($h \leq M$), we define the occlusion ratio factor as $H = h/M$ (Figure 4c). Different H values mean that the obstruction of the input light field is different. This is used to compare the perception and classification capabilities of the different D²NN frameworks for samples blocked by occlusions.

As shown in Figure 4d, the classification accuracy of the single-layer traditional D²NN for samples blocked by occlusions will gradually decrease as the H value increases. When $H = 0.5$, the classification accuracy of the diffractive neural network on the MNIST data set is only 40.49%. The reason is that light waves carrying the information on samples cannot be sent to the D²NN for processing completely due to the presence of occlusions. It can be improved by increasing the number of diffractive layers (when $H = 0.5$, the five-layer

D²NN can achieve a classification accuracy of 58.96%), but the improvement effect is not obvious. In contrast, when the K value changes from 0.2 to 0.5, the classification accuracy of the single-layer differential D²NN with structured illumination on the MNIST data set remains above 97% (Figure 4e). The test results show that the scheme proposed in this paper can effectively realize the classification of samples blocked by occlusions and can be applied to more complex application scenarios.

Physical Construction and Simulation. Through preliminary training, the ideal structured illumination pattern that the system needs to load can be obtained. It should be noted that the value of each pixel unit on the structured illumination pattern is limited to the range of 0–1 during the training process. 0 means that there is no light output at this pixel unit, and 1 means that the light at this pixel unit is completely output to the next layer. However, in real experimental scenarios, structured illumination patterns often need to be loaded onto amplitude-modulated SLM (spatial light modulators) or DMD (digital micromirror device) or implemented using metasurfaces. Taking an 8-bit amplitude-modulated SLM as an example, it can realize 256-level manipulation of each pixel unit between completely blocking light (value 0) and fully transmitting light (value 1). Therefore, the value of each pixel unit can be only 256 specific numbers

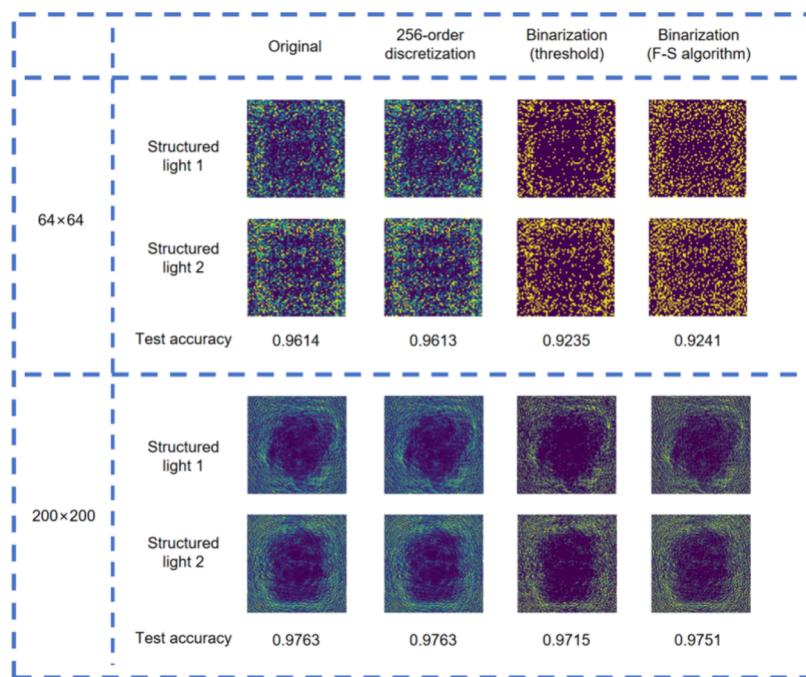


Figure 5. When the number of diffraction neurons on each layer of the network is 64×64 and 200×200 , the impact of different discretization schemes on the network performance. All tests are performed on 10,000 test images of the MNIST data set.

between 0 and 1. Similarly, for DMD, each pixel unit has only two states: completely blocked and completely transparent, so the value of each pixel unit can be only 0 or 1. When structured illumination is actually used, the value of each pixel unit on the trained illumination pattern needs to be discretized according to the corresponding requirements.

Here, we discuss the performance of the differential D²NN with structured illumination using different discretization methods, as shown in Figure 5. The 256-order discretization means that the value of each pixel unit is operated according to $p_{\text{new}} = \text{round}(p_{\text{old}} \cdot 255)/255$ to ensure that all values are discrete to 256 specific numbers between 0 and 1. Threshold binarization refers to discretization by setting a threshold value, and here, we set 0.5 as the threshold. Pixel values above the threshold are set to 1, and other pixel values are set to 0. Besides, the method of binarization using the Floyd–Steinberg dithering algorithm (F–S algorithm) can better preserve the structural information on the pattern.³⁹ In this algorithm, the discrete error of each pixel will be transferred to the surrounding pixels so that the details of the pattern can be retained as much as possible after binarization (Supporting Section S3).

From the results, the 256-order discretization has basically no impact on the final classification performance of the network because the division of each step is relatively small. This means that using an 8-bit amplitude-modulated SLM to load the trained structured illumination pattern will not cause too much fluctuation in the network performance. Regardless of whether it is threshold binarization or binarization using the F–S algorithm, the network performance will be reduced. Compared with the large network architecture (such as 200×200 neurons per layer), the small network architecture (such as 64×64 neurons per layer) is more sensitive to these discretization operations. Moreover, the test results show that when loading structured illumination patterns into the DMD,

using the F–S algorithm for binary discretization can alleviate the decline of the classification accuracy.

Based on the ability of DMD and SLM to modulate the wavefront, we physically constructed the single-layer differential D²NN with structured light illumination. In the optical system design, 650 nm laser is selected as the light source. DMD is used as a structured light generating device to load the illumination patterns in sequence. When using a single diffractive layer, the phase-modulated SLM is used to simulate the phase distribution on the diffractive layer obtained by training. The lens groups (lens1 and lens2, lens3 and lens4) in the middle are used for collimation and beam expansion so that the light field matches the pixel units of the DMD and SLM. The linear polarizer is applied to adjust the incident light intensity, while the stray light and high-order diffractions can be filtered out by the optical iris. The specific device setup is shown in Figure 6a. As mentioned above, binarization based on the F–S dithering algorithm can preserve the structural information on the pattern to the greatest extent. Thus, we need to perform similar processing on the trained illumination patterns. Then, the discretized patterns are loaded into the DMD and two output light fields can be obtained on the detector sequentially. By recording and comparing the light intensities of the test areas and performing differential calculations, the final classification category of the sample will be determined. Here, under three different types of input light fields, the output light fields obtained by training and the predicted output light fields after discretization are shown in Figure 6b. The results show that the system that we proposed can perform the classification task for different test samples precisely.

In addition, the physical construction of the diffractive network framework that we reported is not limited to the optical system mentioned above. As a novel type of planar optical element, metasurfaces are composed of subwavelength resonators for manipulating the wavefront of light^{41,42} and

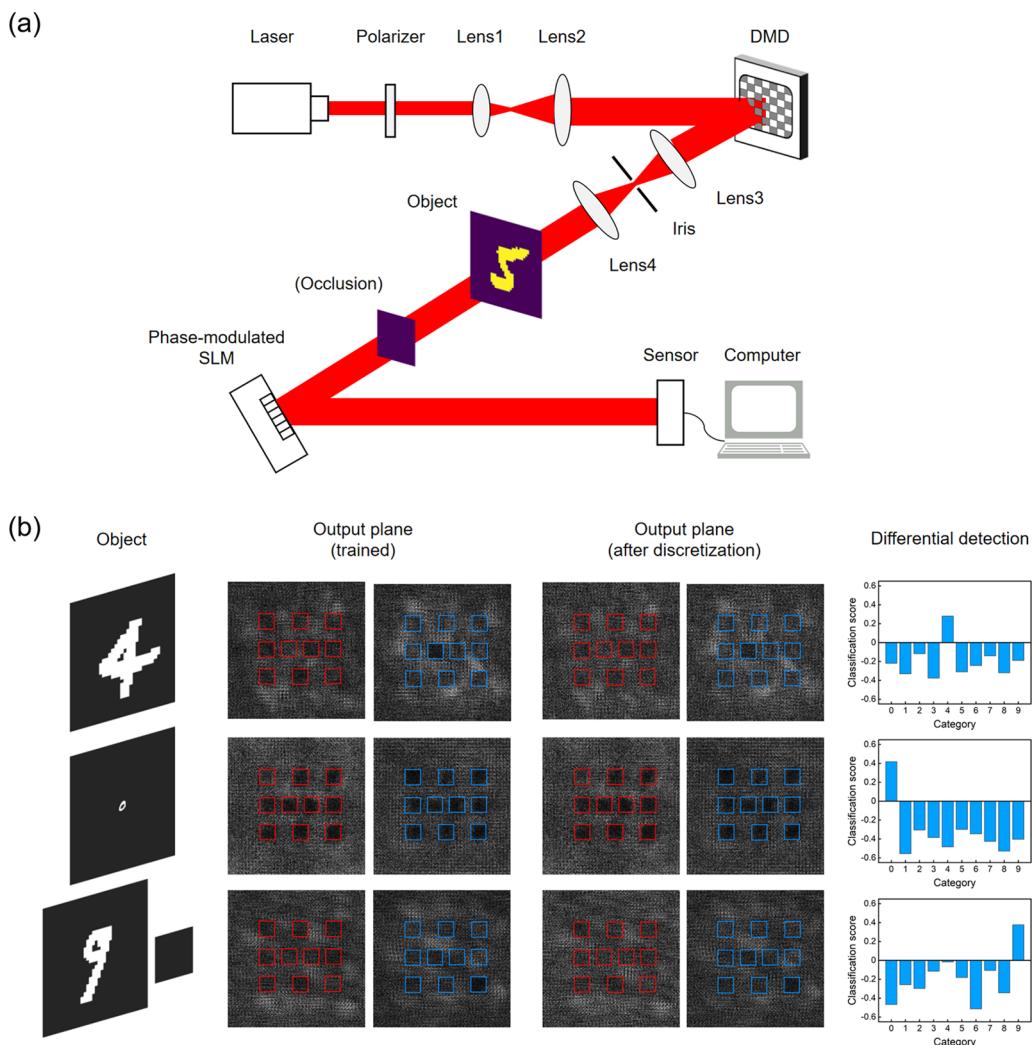


Figure 6. (a) Optical system design for the single-layer differential D²NN system with structured illumination. When testing the network's ability to classify samples obstructed by occlusions, the occlusions need to be placed between the object and the SLM. (b) Trained output light field and the predicted simulated output light field under different illumination patterns.

have a wide range of applications, such as optical holography.^{43,44} Thus, both the generation of structured light and the regulation function of the diffraction layer can also be achieved through metasurfaces.^{45–47} Combined with the polarization multiplexing and passive characteristics of metasurfaces, the entire system can achieve parallel computing, lower energy consumption, and less computing time. However, due to the reconfigurability of DMD, different illumination patterns can be loaded using only one DMD. This reuse of optical devices simplifies the entire system and enables the physical construction of the optical network to be completed in a convenient and feasible way.

METHODS

Mathematical Model. The proposed D²NN follows optical diffraction theory and a deep learning structure. In the diffractive network, each diffractive layer in the D²NN is composed of many unit structures. According to Huygens' principle, each unit structure can be regarded as a secondary subwave source. The optical propagation process between diffractive layers can be described by the Rayleigh–Sommerfeld equation. The equation is given as

$$w_i^l(x, y, z) = \frac{z - z_i}{r^2} \left(\frac{1}{2\pi r} + \frac{1}{j\lambda} \right) \exp(j \frac{2\pi r}{\lambda}) \quad (1)$$

where $w_i^l(x, y, z)$ refers to the light field distribution at the diffraction neuron with coordinates (x, y, z) in the $l+1$ 'th layer, which is generated by the i 'th diffraction neuron with coordinates (x_i, y_i, z_i) in the l 'th layer. λ refers to the wavelength of the illumination light, $r = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}$ and j refers to the imaginary unit. Therefore, the light field of the i 'th neuron in the l 'th layer of neurons can be regarded as

$$u_i^l(x_i, y_i, z_i) = \sum_j t_i^l(x_i, y_i, z_i) \cdot w_i^{l-1}(x_i, y_i, z_i) \cdot u_j^{l-1}(x_i, y_i, z_i) \quad (2)$$

where $t_i^l(x_i, y_i, z_i)$ is the transmission function of the diffraction neuron, which consists of two parts: amplitude and phase. It reflects the complex value modulation of the light field by the diffractive layer and can be expressed as

$$t_i^l(x_i, y_i, z_i) = a_i^l(x_i, y_i, z_i) \cdot \exp[j\varphi_i^l(x_i, y_i, z_i)] \quad (3)$$

where α and ϕ refer to the amplitude and phase coefficients in the transmission function of each diffraction neuron, respectively. All of the research in this paper is based on the phase-type D²NN, so α is set to a constant 1, and ϕ is a trainable parameter allowed to take values in the range of 0– 2π .

The classification performance of different network architectures for samples blocked by occlusion is also discussed in this paper. On the plane of opaque occlusion, the amplitude is modulated by the occlusion function. The formula is expressed as

$$o(x, y) = \begin{cases} 0, & (x, y) \text{ in the blocked part} \\ 1, & \text{otherwise} \end{cases} \quad (4)$$

where “0” means that the propagation of light will be blocked and “1” means that light will be completely transmitted through.

In the process of solving the D²NN model, using the Rayleigh–Sommerfeld formula often brings a huge computational burden. Under appropriate interlayer spacing, the calculation results using the Fresnel scalar diffraction theory are equivalent to those using the Rayleigh–Sommerfeld formula. However, the former can effectively reduce the calculation complexity.^{48,49} At this time, the complex amplitude light field distribution of the i 'th neuron in the l 'th layer of neurons can be characterized as

$$\left\{ \begin{array}{l} u_i^l(x_i, y_i) = F^{-1} \left\{ H(f_x, f_y) \cdot F \left[t_i^{l-1}(x_i, y_i) \cdot u_i^{l-1}(x_i, y_i) \right] \right\} \\ H(f_x, f_y) = \exp \left[j \frac{2\pi(z - z_i)}{\lambda} \right] \cdot \exp \left[-j\lambda\pi(z - z_i) \right] \\ \left[f_x^2 + f_y^2 \right] \end{array} \right. \quad (5)$$

where F and F^{-1} represent the Fourier transform and inverse Fourier transform, respectively, and $H(f_x, f_y)$ describes the propagation process of the light in the space.

Training Method. The network architecture proposed in this paper is trained and tested on two widely used data sets: MNIST and Fashion-MNIST. A total of 60,000 samples are used as training data, and 10,000 samples are selected for blind testing to verify the classification ability of the model. To make the sample size match the network, all samples are processed by up-sampling before use.

The differential detection method is to obtain the classification score for each category by the normalized difference of the corresponding intensity signal pair. For an n -classification problem, we assume that the light intensity distributions of the n predefined areas on the output plane are $I_{1,n}$ and $I_{2,n}$ with the two structured illuminations. 1 and 2 are used to distinguish the light intensity distribution on the same area under two consecutive structured illuminations. Then, the predicted classification score I_n after differential operation can be expressed as

$$I_n = \frac{I_{1,n} - I_{2,n}}{I_{1,n} + I_{2,n}} \quad (6)$$

Here, the soft-max cross-entropy (SCE) loss function is chosen as the loss function during training. To improve the convergence speed and performance of the network, we divide

the predicted classification score by a scaling factor T (also known as “temperature”, it is a hyperparameter in the neural network training process). The processed predicted classification score is $I'_n = I_n/T$. For the SCE loss function, the softmax function is used to process the predicted classification score for obtaining the classification probability of each category, which can be expressed as

$$p_n = \frac{\exp(I'_n)}{\sum_{n=1}^N \exp(I'_n)} \quad (7)$$

where p_n refers to the predicted classification probability of the n 'th category. Assuming that the n 'th entry of the truth label vector is y_n , the loss function can be written as

$$\text{loss} = -\frac{1}{N} \sum_{n=1}^N y_n \log(p_n) \quad (8)$$

All of the neural networks in this work were simulated using Python (v3.7.12) and Google TensorFlow (v2.4.0) frameworks. The Adam optimizer was used for optimization during the training process of all models, and the initial learning rate was set to 0.001. To train and test our models, we used a computer with an NVIDIA GeForce RTX 3090 graphics processing unit (GPU) and an Intel(R) Xeon(R) Gold 6133 CPU @ 2.50 GHz and 64 GB of RAM, running the Microsoft Windows 10 operating system. All the models were trained for 30 epochs, and each round of training took about 20 min.

CONCLUSIONS

In this work, a framework design of a differential D²NN with structured illumination is proposed. This scheme involves the illumination patterns as optimizable targets in the system design, which will effectively improve the perception ability of the D²NN. When there is only one diffractive layer, the network architecture can achieve 97.63 and 88.10% classification accuracies on the MNIST and Fashion-MNIST data sets. For the classification problems of tiny samples and samples blocked by occlusions, this new optical network framework can also achieve better performance in these complex machine learning tasks compared with the traditional D²NN. The results show that this scheme can further reduce the space complexity of the D²NN while maintaining its inference ability. On the basis of reducing the use of diffractive layers, the error accumulation caused by the misalignment of multiple layers can be alleviated.

However, compared to the electronic neural network, this network architecture lacks nonlinear activation functions. It is possible to consider using nonlinear materials to realize the construction of optical nonlinear layers^{50–53} so that the inference ability of the network can be further improved. In addition, the training process of the network architecture still needs to be implemented on the computer platform. When larger models are deployed, the demand for computing resources is still large. Once the task is changed, retraining and redesign are required. To further improve reconfigurability and reduce deployment energy consumption, new optical computing platforms represented by optoelectronic hybrid platforms should be developed.^{54–56} It is expected that more related research will be proposed in the future, which will fully demonstrate the advantages of the optical computing platform.

ASSOCIATED CONTENT

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsphtronics.4c01511>.

Modified maximum half-cone diffraction angle formula; comparison of models trained with different loss function; Floyd–Steinberg dithering algorithm; framework design in Fourier Space ([PDF](#))

AUTHOR INFORMATION

Corresponding Author

Jiaqi Li – Key Laboratory of Quantum Materials and Devices of Ministry of Education, School of Physics, Southeast University, Nanjing 211189, China;  [orcid.org/0000-0002-8868-9839](#); Email: ljq@seu.edu.cn

Authors

Jiajun Zhang – Key Laboratory of Quantum Materials and Devices of Ministry of Education, School of Physics, Southeast University, Nanjing 211189, China

Shuyan Zhang – Key Laboratory of Quantum Materials and Devices of Ministry of Education, School of Physics, Southeast University, Nanjing 211189, China

Weijie Shi – Key Laboratory of Quantum Materials and Devices of Ministry of Education, School of Physics, Southeast University, Nanjing 211189, China

Yong Hu – Key Laboratory of Quantum Materials and Devices of Ministry of Education, School of Physics, Southeast University, Nanjing 211189, China

Zheng-Gao Dong – Key Laboratory of Quantum Materials and Devices of Ministry of Education, School of Physics, Southeast University, Nanjing 211189, China;  [orcid.org/0000-0003-1655-0253](#)

Weibing Lu – Center for Flexible RF Technology, State Key Lab of Millimeter waves, School of Information Science and Engineering, Southeast University, Nanjing 210096, China;  [orcid.org/0000-0001-6337-1739](#)

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsphtronics.4c01511>

Author Contributions

J.Z. conceived the idea, executed the research, and revised and finalized the manuscript; J.L. supervised the whole project and provided critical guidance. All authors provided helpful comments during the research process and have given approval to the final version of the manuscript.

Funding

This work is supported by the National Natural Science Foundation of China (62231001), China National Funds for Distinguished Young Scientists (61925103), the Fundamental Research Funds for the Central Universities (2242022k60004), and the National Natural Science Foundation of China (12174052).

Notes

The authors declare no competing financial interest.

REFERENCES

- (1) Shen, Y.; Harris, N. C.; Skirlo, S.; Prabhu, M.; Baehr-Jones, T.; Hochberg, M.; Sun, X.; Zhao, S.; Larochelle, H.; Englund, D.; Soljačić, M. Deep Learning with Coherent Nanophotonic Circuits. *Nat. Photonics* **2017**, *11* (7), 441–446.
- (2) Wetzstein, G.; Ozcan, A.; Gigan, S.; Fan, S.; Englund, D.; Soljačić, M.; Denz, C.; Miller, D. A. B.; Psaltis, D. Inference in Artificial Intelligence with Deep Optics and Photonics. *Nature* **2020**, *588* (7836), 39–47.
- (3) Liu, J.; Wu, Q.; Sui, X.; Chen, Q.; Gu, G.; Wang, L.; Li, S. Research Progress in Optical Neural Networks: Theory, Applications and Developments. *Photonix* **2021**, *2* (1), 5.
- (4) Hu, J.; Mengu, D.; Tzarouchis, D. C.; Edwards, B.; Engheta, N.; Ozcan, A. Diffractive Optical Computing in Free Space. *Nat. Commun.* **2024**, *15* (1), 1525.
- (5) Hughes, T. W.; Minkov, M.; Shi, Y.; Fan, S. Training of Photonic Neural Networks through in Situ Backpropagation and Gradient Measurement. *Optica* **2018**, *5* (7), 864.
- (6) Lin, X.; Rivenson, Y.; Yardimci, N. T.; Veli, M.; Luo, Y.; Jarrahi, M.; Ozcan, A. All-Optical Machine Learning Using Diffractive Deep Neural Networks. *Science* **2018**, *361* (6406), 1004–1008.
- (7) Badloe, T.; Lee, S.; Rho, J. Computation at the Speed of Light: Metamaterials for All-Optical Calculations and Neural Networks. *Adv. Photon.* **2022**, *4* (6), No. 064002.
- (8) Li, J.; Li, Y.; Gan, T.; Shen, C.-Y.; Jarrahi, M.; Ozcan, A. All-Optical Complex Field Imaging Using Diffractive Processors. *Light Sci. Appl.* **2024**, *13* (1), 120.
- (9) Luo, Y.; Zhao, Y.; Li, J.; Çetintas, E.; Rivenson, Y.; Jarrahi, M.; Ozcan, A. Computational Imaging without a Computer: Seeing through Random Diffusers at the Speed of Light. *eLight* **2022**, *2* (1), 4.
- (10) Sakib Rahman, M. S.; Ozcan, A. Computer-Free, All-Optical Reconstruction of Holograms Using Diffractive Networks. *ACS Photonics* **2021**, *8* (11), 3375–3384.
- (11) Hu, J.; Liao, K.; Dinç, N. U.; Gigli, C.; Bai, B.; Gan, T.; Li, X.; Chen, H.; Yang, X.; Li, Y.; İlşıl, Ç.; Rahman, M. S. S.; Li, J.; Hu, X.; Jarrahi, M.; Psaltis, D.; Ozcan, A. Subwavelength Imaging Using a Solid-Immersion Diffractive Optical Processor. *eLight* **2024**, *4* (1), 8.
- (12) Liao, D.; Chan, K. F.; Chan, C. H.; Zhang, Q.; Wang, H. All-Optical Diffractive Neural Networked Terahertz Hologram. *Opt. Lett.* **2020**, *45* (10), 2906.
- (13) Bai, B.; Luo, Y.; Gan, T.; Hu, J.; Li, Y.; Zhao, Y.; Mengu, D.; Jarrahi, M.; Ozcan, A. To Image, or Not to Image: Class-Specific Diffractive Cameras with All-Optical Erasure of Undesired Objects. *eLight* **2022**, *2* (1), 14.
- (14) Shen, C.-Y.; Li, J.; Gan, T.; Li, Y.; Jarrahi, M.; Ozcan, A. All-Optical Phase Conjugation Using Diffractive Wavefront Processing. *Nat. Commun.* **2024**, *15* (1), 4989.
- (15) Zhang, J.; Ye, Z.; Yin, J.; Lang, L.; Jiao, S. Polarized Deep Diffractive Neural Network for Sorting, Generation, Multiplexing, and de-Multiplexing of Orbital Angular Momentum Modes. *Opt. Express* **2022**, *30* (15), 26728.
- (16) Veli, M.; Mengu, D.; Yardimci, N. T.; Luo, Y.; Li, J.; Rivenson, Y.; Jarrahi, M.; Ozcan, A. Terahertz Pulse Shaping Using Diffractive Surfaces. *Nat. Commun.* **2021**, *12* (1), 37.
- (17) Ding, X.; Zhao, Z.; Xie, P.; Cai, D.; Meng, F.; Wang, C.; Wu, Q.; Liu, J.; Burokur, S. N.; Hu, G. Metasurface-Based Optical Logic Operators Driven by Diffractive Neural Networks. *Adv. Mater.* **2024**, *36* (9), No. 2308993.
- (18) Wang, P.; Xiong, W.; Huang, Z.; He, Y.; Xie, Z.; Liu, J.; Ye, H.; Li, Y.; Fan, D.; Chen, S. Orbital Angular Momentum Mode Logical Operation Using Optical Diffractive Neural Network. *Photon. Res.* **2021**, *9* (10), 2116.
- (19) Qian, C.; Lin, X.; Lin, X.; Xu, J.; Sun, Y.; Li, E.; Zhang, B.; Chen, H. Performing Optical Logic Operations by a Diffractive Neural Network. *Light Sci. Appl.* **2020**, *9* (1), 59.
- (20) Luo, Y.; Mengu, D.; Ozcan, A. Cascadable All-Optical NAND Gates Using Diffractive Networks. *Sci. Rep.* **2022**, *12* (1), 7121.

- (21) Zhang, Z.; Feng, F.; Gan, J.; Lin, W.; Chen, G.; Somekh, M. G.; Yuan, X. Space-Time Projection Enabled Ultrafast All-Optical Diffractive Neural Network. *Laser Photonics Rev.* **2024**, *18*, No. 2301367.
- (22) Bai, B.; Li, Y.; Luo, Y.; Li, X.; Çetintas, E.; Jarrahi, M.; Ozcan, A. All-Optical Image Classification through Unknown Random Diffusers Using a Single-Pixel Diffractive Network. *Light Sci. Appl.* **2023**, *12* (1), 69.
- (23) Rahman, M. S. S.; Li, J.; Mengu, D.; Rivenson, Y.; Ozcan, A. Ensemble Learning of Diffractive Optical Networks. *Light Sci. Appl.* **2021**, *10* (1), 14.
- (24) Duan, Z.; Chen, H.; Lin, X. Optical Multi-Task Learning Using Multi-Wavelength Diffractive Deep Neural Networks. *Nanophotonics* **2023**, *12* (5), 893–903.
- (25) Kulce, O.; Mengu, D.; Rivenson, Y.; Ozcan, A. All-Optical Information-Processing Capacity of Diffractive Surfaces. *Light Sci. Appl.* **2021**, *10* (1), 25.
- (26) Yan, T.; Wu, J.; Zhou, T.; Xie, H.; Xu, F.; Fan, J.; Fang, L.; Lin, X.; Dai, Q. Fourier-Space Diffractive Deep Neural Network. *Phys. Rev. Lett.* **2019**, *123* (2), No. 023901.
- (27) Mengu, D.; Rivenson, Y.; Ozcan, A. Scale-, Shift-, and Rotation-Invariant Diffractive Optical Networks. *ACS Photonics* **2021**, *8* (1), 324–334.
- (28) Jiao, S.; Feng, J.; Gao, Y.; Lei, T.; Xie, Z.; Yuan, X. Optical Machine Learning with Incoherent Light and a Single-Pixel Detector. *Opt. Lett.* **2019**, *44* (21), 5186.
- (29) Mengu, D.; Zhao, Y.; Yardimci, N. T.; Rivenson, Y.; Jarrahi, M.; Ozcan, A. Misalignment Resilient Diffractive Optical Networks. *Nanophotonics* **2020**, *9* (13), 4207–4219.
- (30) Rahman, M. S. S.; Gan, T.; Deger, E. A.; Işıl, Ç.; Jarrahi, M.; Ozcan, A. Learning Diffractive Optical Communication around Arbitrary Opaque Occlusions. *Nat. Commun.* **2023**, *14* (1), 6830.
- (31) Xia, R.; Wu, L.; Tao, J.; Zhao, M.; Yang, Z. Monolayer Directional Metasurface for All-Optical Image Classifier Doublet. *Opt. Lett.* **2024**, *49* (9), 2505.
- (32) Zheng, M.; Shi, L.; Zi, J. Optimize Performance of a Diffractive Neural Network by Controlling the Fresnel Number. *Photon. Res.* **2022**, *10* (11), 2667.
- (33) Zheng, M.; Liu, W.; Shi, L.; Zi, J. Diffractive Neural Networks with Improved Expressive Power for Gray-Scale Image Classification. *Photon. Res.* **2024**, *12* (6), 1159.
- (34) Wang, J.; Chai, S.; Gu, W.; Li, B.; Jiang, X.; Zhang, Y.; Liao, H.; Liu, X.; Ta, D. Screening COVID-19 from Chest X-Ray Images by an Optical Diffractive Neural Network with the Optimized F Number. *Photon. Res.* **2024**, *12* (7), 1410.
- (35) Xiang, J.; Colburn, S.; Majumdar, A.; Shlizerman, E. Knowledge Distillation Circumvents Nonlinearity for Optical Convolutional Neural Networks. *Appl. Opt.* **2022**, *61* (9), 2173–2183.
- (36) Wirth-Singh, A.; Xiang, J.; Choi, M.; Fröch, J. E.; Huang, L.; Colburn, S.; Shlizerman, E.; Majumdar, A. Compressed Meta-Optical Encoder for Image Classification, 2024. *ArXiv ArXiv:2406.06534*.
- (37) Choi, E.; Kim, G.; Yun, J.; Jeon, Y.; Rho, J.; Baek, S.-H. 360° Structured Light with Learned Metasurfaces. *Nat. Photonics* **2024**, *18* (8), 848–855.
- (38) Kim, G.; Kim, Y.; Yun, J.; Moon, S.-W.; Kim, S.; Kim, J.; Park, J.; Badloe, T.; Kim, I.; Rho, J. Metasurface-Driven Full-Space Structured Light for Three-Dimensional Imaging. *Nat. Commun.* **2022**, *13* (1), 5920.
- (39) Zhang, Z.; Li, X.; Zheng, S.; Yao, M.; Zheng, G.; Zhong, J. Image-Free Classification of Fast-Moving Objects Using “Learned” Structured Illumination and Single-Pixel Detection. *Opt. Express* **2020**, *28* (9), 13269.
- (40) Li, J.; Mengu, D.; Luo, Y.; Rivenson, Y.; Ozcan, A. Class-Specific Differential Detection in Diffractive Optical Neural Networks Improves Inference Accuracy. *Adv. Photon.* **2019**, *1* (04), 1.
- (41) Lee, S.; Park, C.; Rho, J. Mapping Information and Light: Trends of AI-Enabled Metaphotonics. *Curr. Opin. Solid State Mater. Sci.* **2024**, *29*, No. 101144.
- (42) So, S.; Mun, J.; Park, J.; Rho, J. Revisiting the Design Strategies for Metasurfaces: Fundamental Physics, Optimization, and Beyond. *Adv. Mater.* **2023**, *35* (43), No. 2206399.
- (43) Kim, J.; Im, J.-H.; So, S.; Choi, Y.; Kang, H.; Lim, B.; Lee, M.; Kim, Y.-K.; Rho, J. Dynamic Hyperspectral Holography Enabled by Inverse-Designed Metasurfaces with Oblique Helicoidal Cholesterics. *Adv. Mater.* **2024**, *36* (24), No. 2311785.
- (44) So, S.; Kim, J.; Badloe, T.; Lee, C.; Yang, Y.; Kang, H.; Rho, J. Multicolor and 3D Holography Generated by Inverse-Designed Single-Cell Metasurfaces. *Adv. Mater.* **2023**, *35* (17), No. 2208520.
- (45) Liu, C.; Ma, Q.; Luo, Z. J.; Hong, Q. R.; Xiao, Q.; Zhang, H. C.; Miao, L.; Yu, W. M.; Cheng, Q.; Li, L.; Cui, T. J. A Programmable Diffractive Deep Neural Network Based on a Digital-Coding Metasurface Array. *Nat. Electron.* **2022**, *5* (2), 113–122.
- (46) Luo, X.; Hu, Y.; Ou, X.; Li, X.; Lai, J.; Liu, N.; Cheng, X.; Pan, A.; Duan, H. Metasurface-Enabled on-Chip Multiplexed Diffractive Neural Networks in the Visible. *Light Sci. Appl.* **2022**, *11* (1), 158.
- (47) Luo, X.; Dong, S.; Wei, Z.; Wang, Z.; Hu, Y.; Duan, H.; Cheng, X. Full-Fourier-Component Tailorable Optical Neural Meta-Transformer. *Laser Photonics Rev.* **2023**, *17* (12), No. 2300272.
- (48) Zhang, K.; Liao, K.; Cheng, H.; Feng, S.; Hu, X. Advanced All-Optical Classification Using Orbital-Angular-Momentum-Encoded Diffractive Networks. *Adv. Photon. Nexus* **2023**, *2* (06), No. 066006.
- (49) Zhou, T.; Lin, X.; Wu, J.; Chen, Y.; Xie, H.; Li, Y.; Fan, J.; Wu, H.; Fang, L.; Dai, Q. Large-Scale Neuromorphic Optoelectronic Computing with a Reconfigurable Diffractive Processing Unit. *Nat. Photonics* **2021**, *15* (5), 367–373.
- (50) Schlickriede, C.; Kruk, S. S.; Wang, L.; Sain, B.; Kivshar, Y.; Zentgraf, T. Nonlinear Imaging with All-Dielectric Metasurfaces. *Nano Lett.* **2020**, *20* (6), 4370–4376.
- (51) Wang, T.; Sohoni, M. M.; Wright, L. G.; Stein, M. M.; Ma, S.-Y.; Onodera, T.; Anderson, M. G.; McMahon, P. L. Image Sensing with Multilayer Nonlinear Optical Neural Networks. *Nat. Photonics* **2023**, *17* (5), 408–415.
- (52) Zuo, Y.; Li, B.; Zhao, Y.; Jiang, Y.; Chen, Y.-C.; Chen, P.; Jo, G.-B.; Liu, J.; Du, S. All-Optical Neural Network with Nonlinear Activation Functions. *Optica* **2019**, *6* (9), 1132.
- (53) Ryou, A.; Whitehead, J.; Zhelyaznyakov, M.; Anderson, P.; Keskin, C.; Bajcsy, M.; Majumdar, A. Free-Space Optical Neural Network Based on Thermal Atomic Nonlinearity. *Photon. Res.* **2021**, *9* (4), B128.
- (54) Chen, Y.; Nazhamaiti, M.; Xu, H.; Meng, Y.; Zhou, T.; Li, G.; Fan, J.; Wei, Q.; Wu, J.; Qiao, F.; Fang, L.; Dai, Q. All-Analog Photoelectronic Chip for High-Speed Vision Tasks. *Nature* **2023**, *623* (7985), 48–57.
- (55) Cheng, J.; Huang, C.; Zhang, J.; Wu, B.; Zhang, W.; Liu, X.; Zhang, J.; Tang, Y.; Zhou, H.; Zhang, Q.; Gu, M.; Dong, J.; Zhang, X. Multimodal Deep Learning Using On-Chip Diffractive Optics with in Situ Training Capability. *Nat. Commun.* **2024**, *15* (1), 6189.
- (56) Yuan, X.; Wang, Y.; Xu, Z.; Zhou, T.; Fang, L. Training Large-Scale Optoelectronic Neural Networks with Dual-Neuron Optical-Artificial Learning. *Nat. Commun.* **2023**, *14* (1), 7110.