

REVIEW ARTICLE | JUNE 11 2024

## Diffractive deep neural networks: Theories, optimization, and applications

Haijia Chen ; Shaozhen Lou; Quan Wang ; Peifeng Huang; Huigao Duan; Yueqiang Hu  



*Appl. Phys. Rev.* 11, 021332 (2024)  
<https://doi.org/10.1063/5.0191977>



### Articles You May Be Interested In

Compensating the distorted OAM beams with near zero time delay

*Appl. Phys. Lett.* (July 2022)

Fundamentals and recent developments of free-space optical neural networks

*J. Appl. Phys.* (July 2024)

# Diffractive deep neural networks: Theories, optimization, and applications



Cite as: Appl. Phys. Rev. 11, 021332 (2024); doi: 10.1063/5.0191977

Submitted: 18 December 2023 · Accepted: 13 May 2024 ·

Published Online: 11 June 2024



Haijia Chen,<sup>1</sup> Shaozhen Lou,<sup>1</sup> Quan Wang,<sup>1</sup> Peifeng Huang,<sup>1</sup> Huigao Duan,<sup>1,2</sup> and Yueqiang Hu<sup>1,2,3,a)</sup>

## AFFILIATIONS

<sup>1</sup>National Research Center for High-Efficiency Grinding, College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, People's Republic of China

<sup>2</sup>Greater Bay Area Institute for Innovation, Hunan University, Guangzhou 511300, People's Republic of China

<sup>3</sup>Advanced Manufacturing Laboratory of Micro-Nano Optical Devices, Shenzhen Research Institute, Hunan University, Shenzhen 518000, China

<sup>a)</sup>Author to whom correspondence should be addressed: [huyq@hnu.edu.cn](mailto:huyq@hnu.edu.cn)

## ABSTRACT

Optical neural networks (ONN) are experiencing a renaissance, driven by the transformative impact of artificial intelligence, as arithmetic pressures are progressively increasing the demand for optical computation. Diffractive deep neural networks ( $D^2NN$ ) are the important subclass of ONN, providing a novel architecture for computation with trained diffractive layers. Given that  $D^2NN$  directly process light waves, they inherently parallelize multiple tasks and reduce data processing latency, positioning them as a promising technology for future optical computing applications. This paper begins with a brief review of the evolution of ONN and a concept of  $D^2NN$ , followed by a detailed discussion of the theoretical foundations, model optimizations, and application scenarios of  $D^2NN$ . Furthermore, by analyzing current application scenarios and technical limitations, this paper provides an evidence-based prediction of the future trajectory of  $D^2NN$  and outlines a roadmap of research and development efforts to unlock its full potential.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0191977>

## TABLE OF CONTENTS

I. INTRODUCTION.....	1	D. Applications of $D^2NN$ on image reconstruction .....	17
A. The evolution of ONN.....	1	IV. SUMMARY AND OUTLOOK.....	18
B. The concept of $D^2NN$ .....	2	A. Nonlinear technology .....	18
II. THEORETICAL ARCHITECTURE AND		B. Multiplexing .....	19
OPTIMIZATION OF $D^2NN$ .....	3	C. On-chip integration .....	19
A. The theoretical architecture of $D^2NN$ .....	3		
B. Optimization of the performance and			
functionality of the $D^2NN$ .....			
1. Optimization of reasoning and generalization	5		
capabilities .....	5		
2. Nonlinear activation function.....	6		
3. Multiplexing.....	8		
4. Dynamic reconfiguration.....	10		
5. On-chip integration .....	10		
6. Incoherent light source .....	11		
III. APPLICATIONS OF $D^2NN$ .....	13		
A. Applications of $D^2NN$ on beam shaping .....	13		
B. Applications of $D^2NN$ on privacy encryption .....	14		
C. Applications of $D^2NN$ on communication			
engineering .....	15		

## I. INTRODUCTION

As deep learning<sup>1</sup> is increasingly applied in fields such as computer vision,<sup>2,3</sup> speech recognition,<sup>4</sup> natural language processing,<sup>5</sup> and autonomous driving,<sup>6</sup> the amount of information that computers need to process will further increase, posing important challenges to large-scale computing due to the exponential growth of power consumption in digital computers. However, as light is boson that does not interact with each other, it can achieve interconnection and high parallel computing at the speed of light, and using photons as a medium for information transmission can theoretically achieve zero-power transport.<sup>7–9</sup> Inspired by these, researchers have been vigorously promoting the research and development of optical computing for decades.<sup>6,10–12</sup>

### A. The evolution of ONN

ONN are neural network systems that enable information processing and computation based on optical components. Unlike traditional electronic-based neural networks, ONN facilitate information transmission and computation through optical propagation. The evolution of ONN is depicted in Fig. 1. Research on ONN began in the 1980s with Psaltis *et al.*<sup>10,13</sup> conducting the first ONN experiment. In 1995, Psaltis *et al.* implemented a neural network model using conjugate mirrors during the holographic reading process.<sup>14</sup> In 2004, Yeh *et al.* developed a system for target recognition using gratings and spatial light modulators.<sup>15</sup> Deng *et al.*<sup>16</sup> in 2005 employed diffractive optics to produce holograms for optically weighted interconnections. In 2007, Romariz and Wagner implemented an optoelectronic approach to develop a dynamic system akin to the FitzHugh–Nagumo (FN) neural model.<sup>17</sup> Slow convergence in learning and training was attributed to hardware development limitations and computational response speeds of that era. Initial micro-nano-processing techniques could not overcome neuron size limitations, adversely affecting ONN performance and hindering its development. However, advancements in semiconductor processes and deep learning algorithms have led to resurgence of ONN in general-purpose computing, leveraging their high parallelism, low energy consumption, and rapid computation. ONN are categorized by neuron interconnection forms into interference, diffraction, coupling, and scattering types.

The basic computational unit of the interferometric ONN<sup>4,18,19</sup> is the Mach–Zehnder Interferometer (MZI), which utilizes the principle of light interference to interconnect neurons. These types of ONN can perform arbitrary matrix multiplication without fundamental loss and are easily configurable and controllable. The diffractive ONN primarily employ the principle of light diffraction to facilitate neuron interconnections. By embedding the trained network parameters into physical devices through micro-nano-processing, these ONN<sup>19–21</sup> can perform optical computations at the speed of light without requiring an external power supply. The core function of the coupled ONN<sup>22–24</sup> involves linear operations facilitated by wavelength division multiplexing (WDM), which significantly enhances transmittance and enables the separation and synthesis of light. The scattering ONN<sup>25–27</sup> utilize the scattering properties of both linear and nonlinear materials to

interconnect the layers of each network, enabling both linear and non-linear operations.

D<sup>2</sup>NN utilize the diffraction of light through the surface of its design to perform calculations, providing a significant advantage in large-scale computing over other forms of ONN. In recent years, the development of D<sup>2</sup>NN has advanced dramatically, and Fig. 2 illustrates the breakthroughs in nonlinearity,<sup>28</sup> multiplexing,<sup>29</sup> incoherent light<sup>30</sup> and on-chip integration,<sup>31</sup> along with their extensive applications in beam shaping,<sup>32,33</sup> privacy encryption,<sup>34,35</sup> communication engineering<sup>36,37</sup> and image reconstruction.<sup>38,39</sup> Further details will be provided in Secs. II B 2, II B 3, II B 5, II B 6, and III A–D. These examples highlight not only the transformative potential of D<sup>2</sup>NN in optical computing but also establish new paradigms for computing architectures where light plays a central role.

### B. The concept of D<sup>2</sup>NN

D<sup>2</sup>NN represent a type of optical computing architecture that leverages the wave properties of light to perform computational tasks associated with deep learning. These networks are designed as a sequence of diffractive layers capable of controlling the phase and amplitude of the transmitted or reflected light. The working principle of D<sup>2</sup>NN, illustrated in Fig. 3(a), involves multiple layers of diffractive surfaces, each with a pattern capable of diffracting the incident light onto the next layer. As the light passes through these layers, the diffractive patterns cumulatively form the layers of an artificial neural network, depicted in Fig. 3(b). The pattern on each layer represents the weight matrix of that layer in a traditional neural network, with the diffractive effects modulating the light's intensity, analogous to the activation of artificial neurons.

The integration of optical computing and deep learning, though recent, has achieved remarkable successes in various aspects. Therefore, this paper will first introduce the evolution of ONN and then the concept of D<sup>2</sup>NN. Subsequently, this paper will focus on analyzing optical diffraction-based ONN, also referred to as D<sup>2</sup>NN. The first part introduces the theoretical model of D<sup>2</sup>NN and summarizes recent research findings on its internal mechanism. The second part describes the mainstream application scenarios of D<sup>2</sup>NN. The third part summarizes the development of D<sup>2</sup>NN and provides an outlook on their future development.

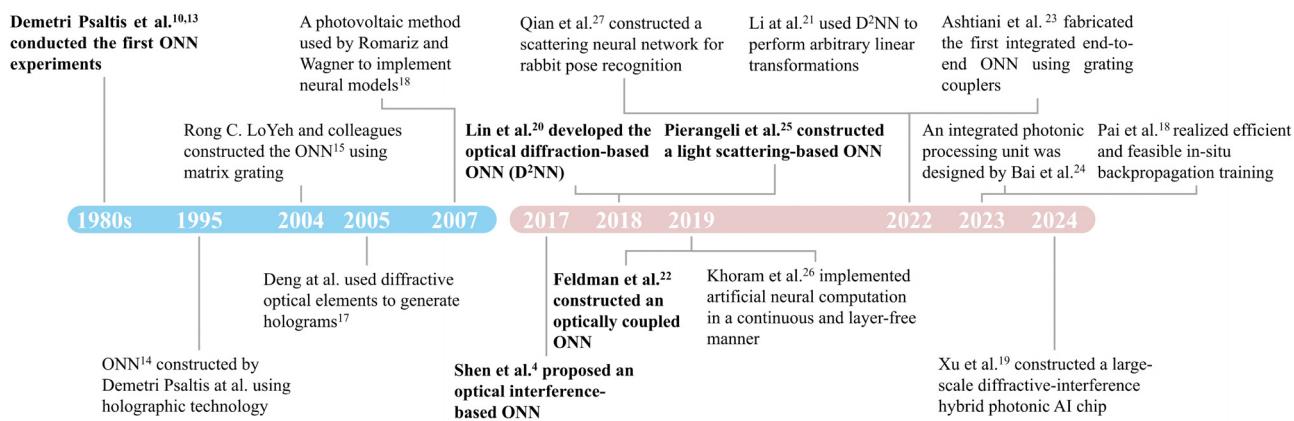
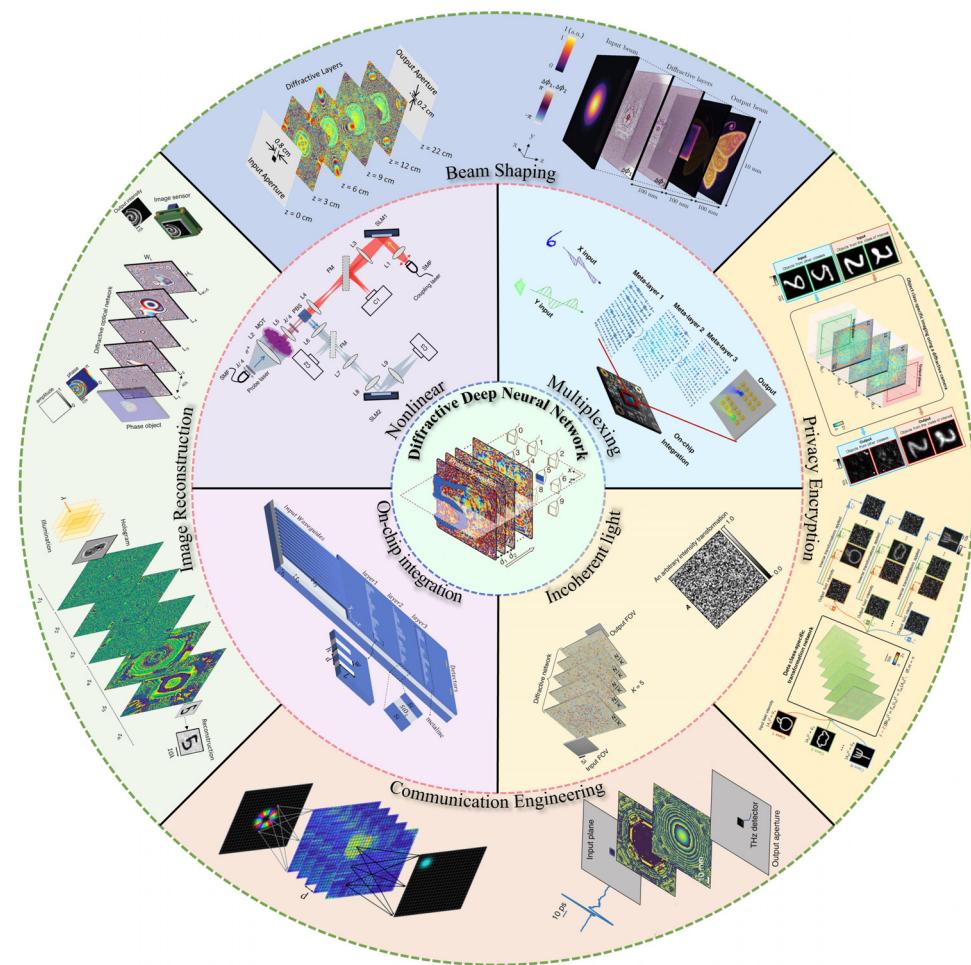


FIG. 1. The evolution of ONN.



**FIG. 2.** Overview of various D<sup>2</sup>NN. Reproduced with permission from Lin *et al.*, *Science* **361**, 1004 (2018). Copyright 2018 Authors.<sup>20</sup> Reproduced with permission from Zuo *et al.*, *Optica* **6**, 1132 (2019); Copyright 2019 Authors, licensed under a Creative Commons Attribution (CC BY) License.<sup>28</sup> Reproduced with permission from Luo *et al.*, *Light* **11**, 158 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) License.<sup>29</sup> Reproduced with permission from Rahman *et al.*, *Light* **12**, 195 (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) License.<sup>30</sup> Reproduced with permission from Fu *et al.*, *Opt. Express* **29**, 31924 (2021); Copyright 2021 Authors, licensed under a Creative Commons Attribution (CC BY) License.<sup>31</sup> Reproduced with permission from Veli *et al.*, *Nat. Commun.* **12**, 37 (2021); Copyright 2021 Authors, licensed under a Creative Commons Attribution (CC BY) License.<sup>32</sup> Reproduced with permission Buske *et al.*, *Opt. Express* **30**, 22798 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) License.<sup>33</sup> Reproduced with permission Bai *et al.*, *eLight* **2**, 14 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>34</sup> Reproduced with permission Bai *et al.*, *Adv. Mater.* **35**, 2212091 (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>35</sup> Reproduced with permission Luo *et al.*, *Light* **8**, 112 (2019); Copyright 2019 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>36</sup> Reproduced with permission from Huang *et al.*, *Phys. Rev. Appl.* **15**, 014037 (2021). Copyright 2021 American Physical Society.<sup>37</sup> Reproduced with permission from Sakib Rahman and Ozcan, *ACS Photonics* **8**, 3375 (2021). Copyright 2021 American Chemical Society.<sup>38</sup> Reproduced with permission from Mengu and Ozcan, *Opt. Mater.* **10**, 2200281 (2022). Copyright 2022 Wiley-VCH GmbH.<sup>39</sup>

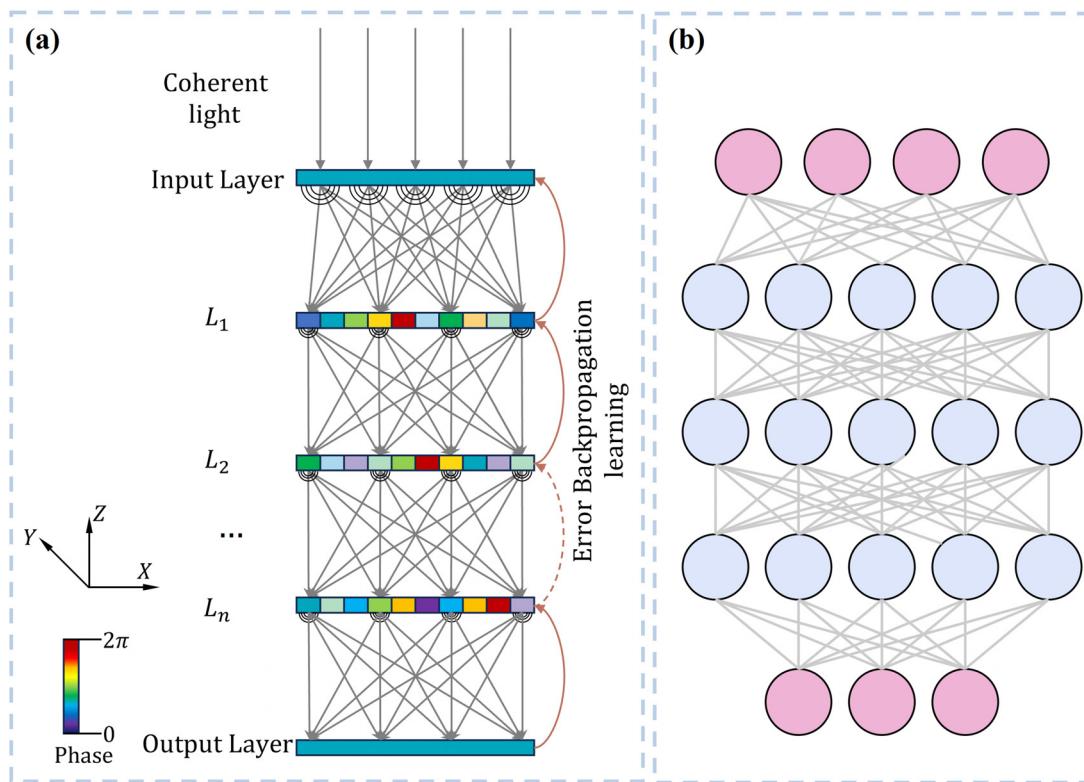
## II. THEORETICAL ARCHITECTURE AND OPTIMIZATION OF D<sup>2</sup>NN

Exploiting the isomorphism between the Huygens–Fresnel principle and the dense neural network architecture, Lin *et al.*<sup>20</sup> leveraged the computational capabilities of deep neural networks to simulate wave propagation, resulting in the creation of D<sup>2</sup>NN. D<sup>2</sup>NN consist of multiple transmission and/or reflection diffractive layers that use the interaction between light and matter to jointly modulate the input light signal, generating the desired output field. They can directly obtain the

2D or 3D input signal of the object and process the optical information encoded by the amplitude, phase, spectrum, and polarization of the input light.

### A. The theoretical architecture of D<sup>2</sup>NN

This section presents the theoretical architecture of D<sup>2</sup>NN, exemplifying the modulation of amplitude and phase of an input optical signal. By simulating the structure of an electronic neural



**FIG. 3.** Schematic comparison between D<sup>2</sup>NN and electronic neural networks: (a) Schematic of D<sup>2</sup>NN and (b) schematic of an electronic neural network.

network (ENN), the modulation targets (amplitude, phase) of the optical field on the input optical signal are treated as a learnable parameter, and light diffraction is employed to connect the layers instead of electrical interconnections. Therefore, constructing a mathematical model for the propagation of light waves between diffractive layers is essential for building D<sup>2</sup>NN.

The diffractive operation is central to D<sup>2</sup>NN computation, and its accuracy directly impacts the performance of the entire neural network. Diffraction, a fundamental property of light propagation, manifests its wave nature. As light is an electromagnetic wave with vector properties, accurate calculation of the diffraction of light waves must take these vector properties into account. However, solving diffractive problems using vector waves undoubtedly complicates the design of the entire network. Based on actual design and experience, the characteristic size of the diffractive layer in the neural network is much larger than half the wavelength of the light signal. Therefore, using scalar diffraction theory to solve the diffraction problem is sufficient to meet the design requirements, ensuring that the imaged image will not undergo significant aberration and that the intensity can be maintained with good uniformity. Commonly used scalar diffraction theories include Kirchhoff diffraction theory and its two approximations, namely, Fresnel diffraction and Fraunhofer diffraction, as well as the Rayleigh-Sommerfeld diffraction integral.<sup>40</sup> However, Kirchhoff diffraction theory faces issues with mathematical self-consistency due to the nature of its boundary conditions. Therefore, the more popular Rayleigh-Sommerfeld diffraction integral is usually chosen, first, because it is more “manifestly covariant,”<sup>41</sup> and second, because it is applied to the

diffraction on the plane screen aperture, which aligns with the setting of the diffractive layer in the neural network.

According to the Rayleigh-Sommerfeld diffraction integral equation, the propagation of a light wave between diffractive layers can be reduced to the solution of the optical complex amplitude at an arbitrary position in the next layer after the diffraction of the light wave in the previous layer,

$$\mathbf{U}^l(P) = \int \int \mathbf{U}^{l-1} \cdot h(x, y, z) dx dy. \quad (1)$$

Among them,  $\mathbf{U}^l(P)$  represents the complex amplitude value at any position P (coordinates  $(x, y, z)$ ) in the  $l$ th layer of the diffractive layer;  $\mathbf{U}^{l-1}$  represents the sum of the output optical complex amplitude in the  $(l-1)$ th layer of the diffractive layer;  $h(x, y, z)$  represents the response function in the Rayleigh-Sommerfeld diffraction integral equation, and its specific expression is

$$h(x, y, z) = \frac{1}{2\pi} \cdot \frac{z - z_i}{R^2} \cdot \left( \frac{1}{2\pi R} + \frac{1}{j\lambda} \right) \cdot \exp\left(\frac{j2\pi R}{\lambda}\right). \quad (2)$$

In the equation,  $(x_i, y_i, z_i)$  in  $R = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}$  represents the coordinates of the  $i$ th neuron in the  $(l-1)$ th layer of the diffractive layer, and  $\lambda$  represents the wavelength of the incident wave.

#### Forward Propagation:

- (1) For the input layer ( $l = 0$ ): Under the interaction of the planar wave of the loaded signal and the  $i$ th neuron in the input layer, the wave function generated by diffraction is

$$\mathbf{u}_i^0 = I \cdot h(x, y, z). \quad (3)$$

In the equation,  $I$  represents the planar wave carrying information in the phase and amplitude channel, which is usually a complex value.

- (2) For the hidden layer ( $1 \leq l \leq M$ ): The wave function generated by diffraction for the  $i$ th neuron in the  $l$ th layer of the hidden layer is

$$\mathbf{u}_i^l = \sum_k [\mathbf{u}_i^{l-1} \cdot t_i^{l-1}(x_i, y_i, z_i)] \cdot h(x, y, z). \quad (4)$$

In the equation,  $k$  represents the total number of neurons in each layer;  $\mathbf{u}_i^{l-1}$  represents the complex amplitude value of the output wave of the  $i$ th neuron in the  $(l-1)$ th layer of the diffractive layer; and  $t_i^{l-1}(x_i, y_i, z_i)$  represents the transmission coefficient of the  $i$ th neuron in the  $(l-1)$ th layer of the diffractive layer, and its specific expression is as follows:

$$t_i^{l-1}(x_i, y_i, z_i) = a_i^{l-1}(x_i, y_i, z_i) \cdot \exp[j\phi_i^{l-1}(x_i, y_i, z_i)]. \quad (5)$$

In the equation,  $a_i^{l-1}(x_i, y_i, z_i)$  represents the modulation amplitude function of the  $i$ th neuron, and  $\phi_i^{l-1}(x_i, y_i, z_i)$  represents the modulation phase function of the  $i$ th neuron.

- (3) For the output layer ( $L = M + 1$ ),

$$S_i^{M+1} = \left| \sum_k \mathbf{u}_i^{M+1} \right|^2. \quad (6)$$

In the equation,  $S_i^{M+1}$  represents the optical intensity detected on the output plane.

Error Backpropagation:

$$\text{Loss} = \frac{1}{k} \sum_i^k \text{abs}(S_i^{l+1} - G_i^{l+1}). \quad (7)$$

In the equation,  $G_i^{l+1}$  represents the expected light intensity in the output plane.

The forward propagation ends for the reverse optimization process, and the loss function is defined as the mean square error between the output light field of the last layer and the theoretical target light field, and the gradientization of the output light field is processed and simplified to obtain the  $L$ th layer as

$$\frac{\partial \mathbf{u}_k^{M+1}}{\partial \phi_i^{M-L}} = j \cdot \mathbf{u}_i^{M-L} \cdot t_i^{M-L} \cdot \sum_{k_1} h \cdot t_{k_1}^M \cdots \sum_{k_L} h \cdot t_{k_L}^{M-L+1} \cdot h. \quad (8)$$

## B. Optimization of the performance and functionality of the D<sup>2</sup>NN

Once trained, D<sup>2</sup>NN can be fabricated using micro-nano-processing techniques, which offer the advantages of compact structure, reduced processing complexity, and enhanced computational capability compared to traditional ONN, thereby becoming a focal point in the realm of optical computing. Among these challenges, the key concerns of this research field include enhancing the performance for larger and more complex target tasks and achieving micro-integration of D<sup>2</sup>NN. Consequently, researchers have embarked on theoretical studies of D<sup>2</sup>NN, yielding numerous significant findings. This section will delineate the optimization of reasoning and generalization

capabilities, nonlinear activation function, multiplexing, dynamic reconfiguration, on-chip integration, and incoherent light source.

### 1. Optimization of reasoning and generalization capabilities

Since the concept of D<sup>2</sup>NN was proposed, its vast potential has been recognized. To enable D<sup>2</sup>NN to supplement or potentially supplant ENN in practical engineering applications, enhancing the network's reasoning and generalization capabilities is essential.

Consequently, a significant research effort has been directed toward the optimization of neural network training algorithms. This includes the adoption of ReLU to mitigate the gradient vanishing problem inherent to the sigmoid function when used as a loss function, and the substitution of mean square error (MSE) with softmax-cross entropy loss,<sup>42</sup> which demonstrates higher suitability for image classification tasks. These adjustments aim to enhance the computational accuracy of D<sup>2</sup>NN, as detailed in Table I. This approach is apt for D<sup>2</sup>NN configurations with a limited number of diffractive layers. However, with an increase in the number of diffractive layers, challenges such as gradient explosion and vanishing persist, engendering training instability. In response, Dou *et al.* introduced a concept termed residual D<sup>2</sup>NN (Res-D<sup>2</sup>NN),<sup>43</sup> drawing on the residual structural framework utilized in ENN.<sup>44</sup> This was accomplished through the construction of novel trainable optical shortcut connections that bypass multiple nonlinear diffractive layers, thereby effectively mitigating gradient-related challenges in deep neural networks, as illustrated in Table I. However, this methodology primarily addresses networks comprising a higher count of diffractive layers and is yet to be experimentally validated, remaining theoretical. Furthermore, the introduction of nonlinearity into D<sup>2</sup>NN poses significant challenges, necessitating the use of high-intensity input light and specialized materials. Considering these factors, Fang *et al.*, inspired by foundational previous research, developed an innovative optoelectronic D<sup>2</sup>NN.<sup>45</sup> This model utilizes knowledge distillation alongside stochastic gradient descent  $\beta$ -Lasso (SGD- $\beta$ -Lasso), aimed at augmenting knowledge transfer efficacy, to address the shortfall in nonlinearity and to refine neural network classification precision (refer to Table I). Employing knowledge distillation, they harnessed and transferred characteristics from AllConvNet,<sup>46</sup> specifically local connectivity and weight sharing, which are better suited for image processing, to their neural networks, thus attaining inference capabilities on par with electronic neural networks.

Beyond the algorithmic optimization of individual D<sup>2</sup>NN, numerous researchers have adapted concepts such as Bagging<sup>47</sup> and ensemble learning<sup>48,49</sup> from the domain of machine learning, amalgamating multiple models to surpass the prediction performance achievable by a singular model. Illustrated in Fig. 4(a), a differential detection network system, as delineated in referenced studies,<sup>50,51</sup> integrates differential measurement techniques and synergistic benefits to independently train multiple D<sup>2</sup>NN. This is achieved by generating diverse training sets through the introduction of varied filters between the incident signal light and the D<sup>2</sup>NN, enabling the extraction of distinct features from the incident signal light. These features are then synthesized into a unified model via a pruning algorithm. Consequently, this methodology markedly enhanced the computational prowess of D<sup>2</sup>NN (refer to Table I for computational accuracy details). Significantly, this research exploited D<sup>2</sup>NN's inherent high

**TABLE I.** Classification accuracy of various neural networks.

Dataset	Model	Test accuracy
MNIST	All-optical (5 layers, Phase only, Sigmoid-MSE)	91.75% (Ref. 42)
	All-Optical (5 layers, Phase only, ReLU-Cross Entropy)	97.18% (Ref. 42)
	All-Optical (5 layers, Complex modulation, ReLU-Cross Entropy)	97.81% (Ref. 42)
	Res-D <sup>2</sup> NN (5 layers, Phase only)	97.5% (Ref. 43)
	Res-D <sup>2</sup> NN (20 layers, Phase only)	98.4% (Ref. 43)
	D <sup>2</sup> NN (20 layers, Phase only)	96.0% (Ref. 43)
	Class-specific D <sup>2</sup> NN (5 layers)	98.52% (Ref. 50)
	EL-MDUNet	99.06% (Ref. 54)
	All-Optical (5 layers, Phase only, Sigmoid-MSE)	81.13% (Ref. 42)
	All-Optical (5 layers, Phase only, ReLU-Cross Entropy)	85.40% (Ref. 42)
Fashion-MNIST	All-Optical (5 layers, Complex modulation, Sigmoid-MSE)	86.33% (Ref. 42)
	All-Optical (5 layers, Complex modulation, ReLU-Cross Entropy)	86.68% (Ref. 42)
	Res-D <sup>2</sup> NN (5 layers, Phase only)	85.1% (Ref. 43)
	Res-D <sup>2</sup> NN (20 layers, Phase only)	88.4% (Ref. 43)
	D <sup>2</sup> NN (20 layers, Phase only)	83.0% (Ref. 43)
	Class-specific D <sup>2</sup> NN (5 layers)	91.48% (Ref. 50)
	All-Optical BS-D <sup>2</sup> NN (4 layers, C = 3)	85.38% (Ref. 53)
	All-Optical BS-D <sup>2</sup> NN (4 layers, C = 5)	84.73% (Ref. 53)
	Hybrid BS-D <sup>2</sup> NN (4 layers, C = 3)	88.92% (Ref. 53)
	Hybrid BS-D <sup>2</sup> NN (4 layers, C = 5)	88.96% (Ref. 53)
Cifar-10(gray)	EL-MDUNet	89.86% (Ref. 54)
	D <sup>2</sup> NN	44.27% (Ref. 45)
	Hybrid D <sup>2</sup> NN	52.38% (Ref. 45)
	Hybrid D <sup>2</sup> NN+KD+SGD- $\beta$ -Lasso	70.19% (Ref. 45)
	AllConvNet	88.43% (Ref. 45)
	Class-specific D <sup>2</sup> NN (5 layers)	50.82% (Ref. 50)
	Ensemble D <sup>2</sup> NN (N = 14)	61.14% (Ref. 51)
Cats vs. Dogs (gray)	Ensemble D <sup>2</sup> NN (N = 30)	62.13% (Ref. 51)
	D <sup>2</sup> NN	60.45% (Ref. 45)
	Hybrid D <sup>2</sup> NN	69.64% (Ref. 45)
	Hybrid D <sup>2</sup> NN+KD+SGD- $\beta$ -Lasso	85.17% (Ref. 45)
	AllConvNet	90.69% (Ref. 45)

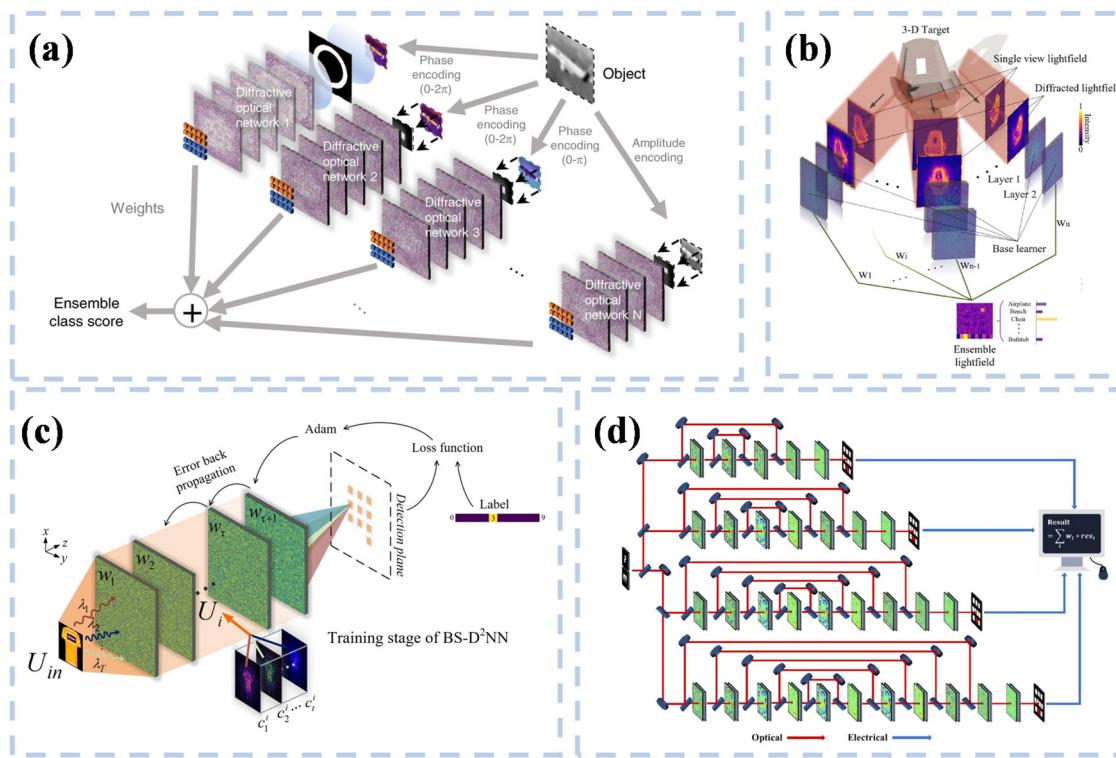
**TABLE I. (Continued.)**

Dataset	Model	Test accuracy
ModelNet 3D shape database	3-view D <sup>2</sup> NN	83.4% (Ref. 52)
	5-view D <sup>2</sup> NN	86.4% (Ref. 52)
	7-view D <sup>2</sup> NN	85.6% (Ref. 52)
	D <sup>2</sup> NN (5 layers)	73.9% (Ref. 52)
	D <sup>2</sup> NN (7 layers)	78.2% (Ref. 52)
	D <sup>2</sup> NN (9 layers)	78.3% (Ref. 52)
	D <sup>2</sup> NN (11 layers)	78.6% (Ref. 52)
	Res-D <sup>2</sup> NN (20 layers)	79.8% (Ref. 52)

parallelism to broaden the spectrum of attainable complex optical field values from solely non-negative real numbers to encompass the negative domain. This effectively surmounted the intrinsic limitation of optical systems' sensitivity to incident optical power alone, and not to complex optical fields, thereby attaining elevated test precision and enhancing the inferential capacity of neural networks. Motivated by these findings, Shi *et al.*<sup>52</sup> proposed a multi-view array-based D<sup>2</sup>NN framework [see Fig. 4(b)], addressing the limitation where prior D<sup>2</sup>NN models could only process planar image data, neglecting substantial light field information from alternate viewpoints. By leveraging light's parallelism and harnessing collective advantages for comprehensive object recognition, they positioned trainable diffractive deep neural networks at various viewpoints, substantially enhancing the computational accuracy of these networks (see Table I). A study<sup>53</sup> introduced a broad-spectrum D<sup>2</sup>NN (BS-D<sup>2</sup>NN) architecture, capitalizing on dispersion advantages and a collective strategy [see Fig. 4(c)]. Joint optimization of training across different band sub-networks via multiple channels was employed to elevate the network's computational accuracy (refer to Table I). The incident optical signal in this study comprised multiple discrete optical wave frequencies, offering a viable approach for advancing D<sup>2</sup>NN technology toward multispectral smart chip applications. However, the challenge of diminishing computational accuracy with an increasing number of channels persists, necessitating further investigation. Li *et al.*<sup>54</sup> developed a hybrid optoelectronic neural network (EL-MDUNet) utilizing an integrated learning methodology, as depicted in Fig. 4(d). Weighted voting applied to the EL-MDUNet model, featuring layers of varying depths, consequently augmented the network's blind measurement precision (detailed data provided in Table I). Furthermore, they employed an adaptation module to adjust the pixel size and resolution of the scattering layer, thereby fortifying the model's robustness and employing jump connections to facilitate deep scaling of the network.

## 2. Nonlinear activation function

Nonlinearity originates from the spatial variation in light intensity that induces free charge carriers via photoionization. These carriers subsequently alter the local electric field distribution, effectuating a concomitant change in the refractive index and phase modulation. D<sup>2</sup>NN architectures, predicated on free-space diffraction, operate on two fundamental principles. (1) The photorefractive effect<sup>55</sup> causes localized changes in the electric field via the distribution of



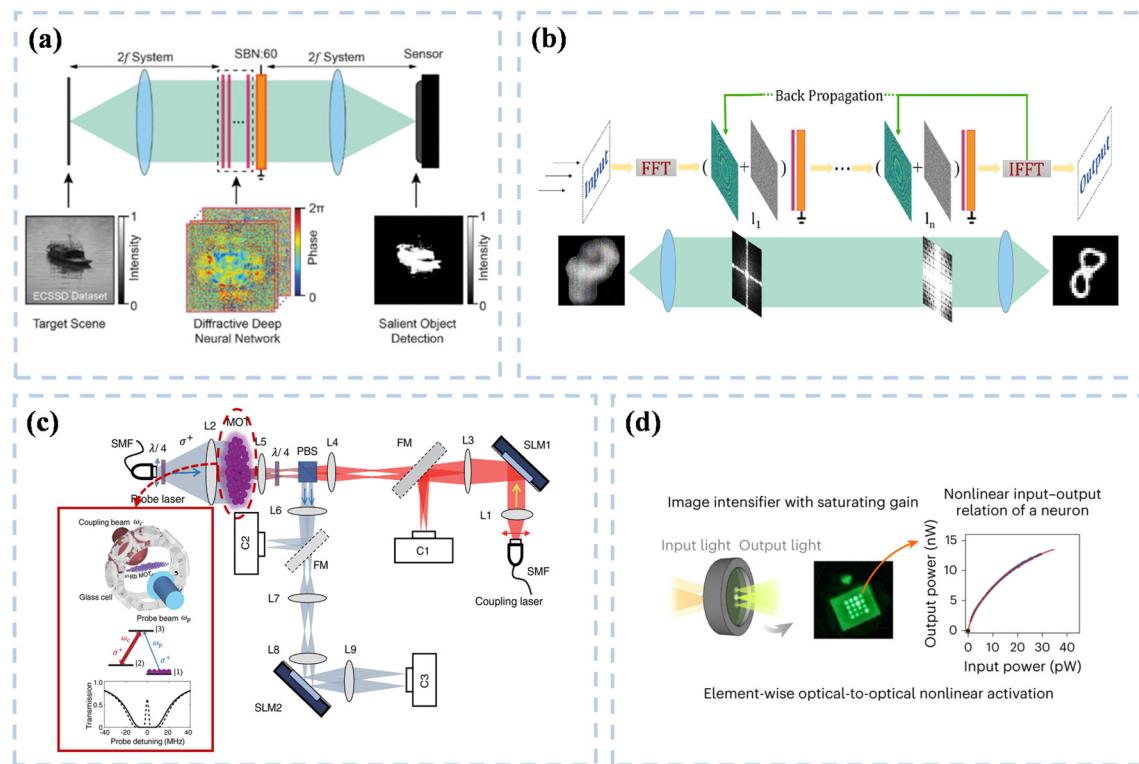
**FIG. 4.** D<sup>2</sup>NN via ensemble learning: (a) Differential detection D<sup>2</sup>NN system combining differential measurement technology, ensemble efficiency design. Reproduced with permission from Rahman *et al.*, Light **10**, 14 (2021); Copyright 2021 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>51</sup> (b) Multi-view array based D<sup>2</sup>NN. Reproduced with permission from Shi *et al.*, Opt. Lett. **46**, 3388 (2021); Copyright 2021 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>52</sup> (c) Wide spectrum D<sup>2</sup>NN architecture. Reproduced with permission from Shi *et al.*, Opt. Lett. **47**, 605 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>53</sup> (d) Hybrid optoelectronic neural network constructed by integrated learning approach. Reproduced with permission from Li *et al.*, Opt. Express **30**, 36700 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>54</sup>

12 June 2025 09:04:57

photogenerated charges in crystals, thereby triggering nonlinear optical effects. (2) The saturable absorption effect<sup>56</sup> exhibits a nonlinear response at high light intensities, attributable to the delayed degradation of excited state energy. References 53, 57, and 58 have harnessed the photorefractive effect to facilitate nonlinear operation. Reference 57 employed a photorefractive crystal (SBN:60)<sup>59,60</sup> to implement a nonlinear activation function through phase modulation, enabling intensity variations for complex mapping relationships. D<sup>2</sup>NN were positioned within a Fourier space derived from a 4f system [shown in Fig. 5(a)], affording the network enhanced computational accuracy and a more compact structure at visible wavelengths compared to the approach of Ref. 20. Song *et al.*<sup>58</sup> utilized potassium sodium strontium barium niobite (KNSBN) to effectuate the nonlinear activation function. To mitigate the issue of significant diffraction dispersion during the forward propagation of the incident beam within the diffractive layer, they adapted the network architecture from Ref. 57, integrating the computational field of the D<sup>2</sup>NN into the frequency domain through a dual-lens system. This facilitated the optimization of sub-network training across different bands via multi-channel joint optimization, enhancing the network's computational accuracy, as illustrated in Fig. 5(b). The incident optical signal in this study encompassed multiple discrete optical wave frequencies, offering a pragmatic pathway for the evolution of D<sup>2</sup>NN toward multispectral smart chip

applications. Nonetheless, the issue of diminishing computational accuracy with an increase in the number of channels persists, warranting additional investigation. The employment of photorefractive crystals for inducing nonlinearity in the aforementioned studies presents challenges, including intricate preparation processes, reduced environmental stability (e.g., temperature, humidity), and complex system integration requirements.

Studies indexed as Refs. 28 and 61–63 utilized the saturable absorption effect to incorporate a nonlinear activation function into D<sup>2</sup>NN. Specifically, Ref. 28 implemented a nonlinear activation function inspired by electromagnetically induced transparency.<sup>64,65</sup> In the D<sup>2</sup>NN configuration [refer to Fig. 5(c)], spatial light modulators and Fourier lenses facilitated linear operations, successfully classifying various stages of the Ising model. Both linear and nonlinear operations within the network were independently programmable, offering insights into the construction of large-scale D<sup>2</sup>NN. Sun *et al.*<sup>61,62</sup> through numerical simulations, deduced that ReLU activation functions, specifically Leaky-ReLU, PReLU, and RReLU, could bolster D<sup>2</sup>NN's inference capabilities, with the RReLU function showing superiority in training duration, computation speed, and recognition precision. Following this, the properties of nonlinear optical materials, such as ten-layer graphene<sup>66</sup> and zinc tetraselenide composites,<sup>67,68</sup> were harnessed to closely match the corresponding nonlinear activation



**FIG. 5.** Exploration of nonlinearity in D<sup>2</sup>NN: (a) Nonlinearity achieved through optical-bistability crystal. Reproduced with permission from Yan *et al.*, Phys. Rev. Lett. **123**, 023901 (2019). Copyright 2019 American Physical Society.<sup>57</sup> (b) Nonlinearity achieved through KNSBN. Reproduced with permission from Song *et al.*, Appl. Opt. **62**, 1082 (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>58</sup> (c) Nonlinearity achieved through electromagnetic-induced transparency. Reproduced with permission from Zuo *et al.*, Optica **6**, 1132 (2019); Copyright 2019 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>28</sup> (d) Nonlinearity achieved through LCD. Reproduced with permission from Wang *et al.*, Nat. Photonics **17**, 408 (2023). Copyright 2023 Springer Nature.<sup>63</sup>

1 June 2023 094571

functions, thereby enhancing the network's computational accuracy. However, these mentioned studies have yet to progress beyond the analog simulation phase to construct tangible physical architectures for experimental validation. Wang *et al.*<sup>63</sup> employed liquid crystal displays (LCD) to achieve nonlinear modulation of light intensity [shown in Fig. 5(d)]. This approach enabled optical neural network image sensors to circumvent the bandwidth constraints of high-resolution cameras through optical domain image compression, fostering swifter, more responsive, and efficient machine vision systems. This research illustrates D<sup>2</sup>NN's future trajectory: integration with ENN preserves D<sup>2</sup>NN's spatial parallelism while diminishing time and energy expenditures, facilitating the accomplishment of more complex tasks. Nevertheless, leveraging the saturable absorption effect for nonlinear functionality might necessitate more elaborate apparatus and experimental setups, along with heightened requirements for activating nonlinear materials, as compared to standard D<sup>2</sup>NN approaches.

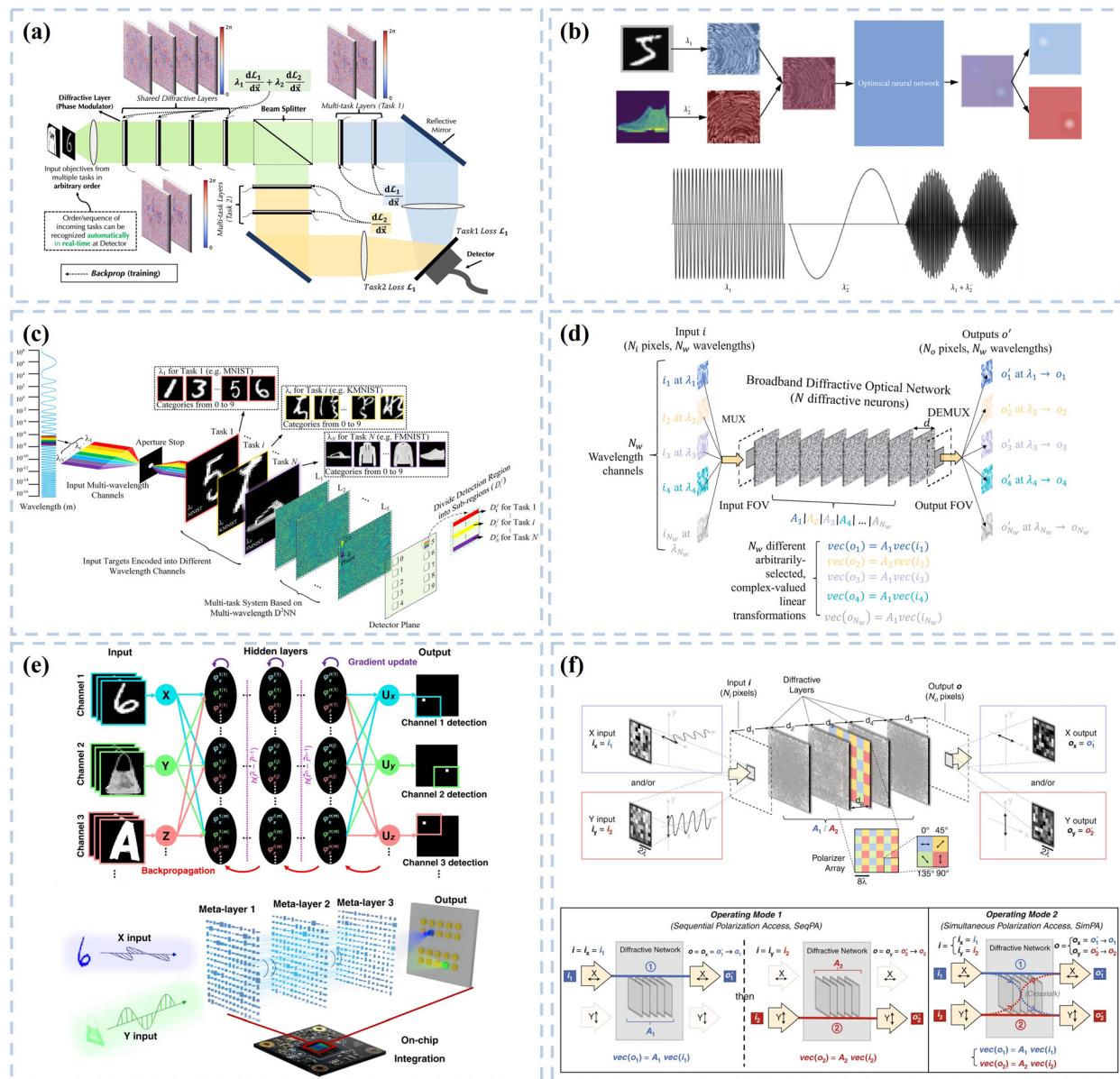
Contrary to the previously described studies, Refs. 69 and 70 employed optimization algorithms to augment the performance of nonlinear D<sup>2</sup>NN. Their research leveraged recent advancements in backpropagation<sup>71,72</sup> for training optical neural networks and successfully applied it to D<sup>2</sup>NN. Significantly, Ref. 70 highlighted the use of a quality factor to direct the backpropagation process. Diverging from the conventional method that necessitates engaging all neurons for gradient calculation, this strategy necessitated merely a single

invocation of the forward algorithm to acquire comprehensive gradient data, thereby substantially diminishing the computational burden. This work serves as a valuable reference for developing network algorithms that minimize training duration.

### 3. Multiplexing

D<sup>2</sup>NN are techniques that employ optical components or systems to implement the neural network's arithmetic functions, utilizing the characteristics of light such as wavelength, phase, amplitude, or polarization to represent information, and to enable parallel processing of data through optical components. Utilizing this characteristic, D<sup>2</sup>NN can implement multiplexing techniques, which enable the simultaneous transmission and processing of multiple signals, including spatial, wavelength, and polarization multiplexing.

**Spatial multiplexing:** This technique involves using physical space to transmit and process multiple signals simultaneously. Li *et al.*<sup>73</sup> achieved the goal of multitasking by using a multiplexer to load information from different tasks into different paths of the same wavelength, as shown in Fig. 6(a). Utilizing a hardware-software co-design scheme, they achieved real-time and multitasking processing by using multiplexers to guide the hidden layers of multiple related classification tasks, ultimately employing the same detector for detection. This approach reduced the number of neurons from 480 000 to 320 000 and



**FIG. 6.** Multiplexing of D<sup>2</sup>NN: (a) Spatial multiplexing. Reproduced with permission from Li *et al.*, Sci. Rep. **11**, 11013 (2021); Copyright 2021 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>73</sup> (b)-(d) Wavelength multiplexing. (e)-(f) Polarization multiplexing. (b) Reproduced with permission from Su *et al.*, Math. Probl. Eng. **2020**, 9748380; Copyright 2020 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>74</sup> (c) Reproduced with permission from Duan *et al.*, Nanophotonics **12**, 893 (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>75</sup> (e) Reproduced with permission from Luo *et al.*, Light Sci. Appl. **11**, 158 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>29</sup> (f) Reproduced with permission from Li *et al.*, Light Sci. Appl. **11**, 153 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>21</sup>

doubled the hardware utilization. In terms of computational accuracy loss, the accuracy drops were only 0.4% for the MNIST task and 0.3% for the MNIST-Fashion task.

**Wavelength multiplexing:** This technique utilizes different wavelengths of light to convey different information. Figure 6(b) illustrates how the wavelength multiplexing technique is applied to perform classification tasks across various datasets.<sup>74</sup> This study demonstrates that

the multi-wave diffractive network (MWDN) enhances accuracy in the MNIST and MNIST-Fashion tasks by 1.15% and 3.2%, respectively, over the traditional deep neural network (DNN). This improvement is attributed to MWDN's ability to adjust learning rates by assigning specific wavelengths to different tasks, thereby optimizing task accuracy. The multi-wavelength D<sup>2</sup>NN (MD<sup>2</sup>NN) proposed by Duan *et al.*<sup>75</sup> use a joint optimization strategy that aligns specific tasks

with corresponding wavelengths, capitalizing on the adaptive learning capabilities of D<sup>2</sup>NN, as illustrated in Fig. 6(c). MD<sup>2</sup>NN significantly surpasses the single-wavelength D<sup>2</sup>NN (SD<sup>2</sup>NN) in both dual-task and quad-task classifications, achieving accuracies of 95.6% and 86.8% in dual-task scenarios, in contrast to 92.4% and 83.1% achieved by SD<sup>2</sup>NN. In quad-task scenarios, MD<sup>2</sup>NN recorded accuracies of 92.8%, 83.0%, 81.0%, and 90.4%, markedly higher than SD<sup>2</sup>NN's 64.6%, 68.7%, 52.5%, and 55.3%. This performance demonstrates MD<sup>2</sup>NN's enhanced capabilities in multi-task optical learning as task numbers increase, suggesting that its scalable design could enhance parallelism, accuracy, and generalization across various ONN architectures. Li *et al.*<sup>76</sup> developed D<sup>2</sup>NN capable of massively parallel bandwidth, achieving up to 2000 unique complex linear transformations by performing various transformations simultaneously, significantly enhancing the throughput of all-optical computing [see Fig. 6(d)]. It was determined that for the wavelength-multiplexed D<sup>2</sup>NN to successfully perform  $N_w$  different complex linear transformations, it must meet the following condition:  $N \geq 2N_w N_i N_o$ , where  $N$  represents the total number of neurons, and  $N_i$  and  $N_o$  are the numbers of input and output field-of-view pixels, respectively.

Polarization multiplexing: This technique employs various polarization states of light waves to encode distinct signals. Luo *et al.*<sup>29</sup> utilized the exceptional polarization modulation capabilities of the metasurface to load the information of different tasks onto different polarization states of the input signal light and realized the recognition of two different tasks, handwritten digits and fashion products, by using the D<sup>2</sup>NN, as shown in Fig. 6(e). Figure 6(f) shows the polarization multiplexed D<sup>2</sup>NN developed by Li *et al.*<sup>21</sup> capable of executing multiple arbitrarily chosen all-optical linear transformations. Two different physical models realized two and four linear transformations, respectively. Wang *et al.*<sup>77</sup> developed three polarization multiplexed D<sup>2</sup>NN with advanced functionalities: high-task-capacity integral classification, efficient non-interleaved Jones matrix eight-channel modulation, and customized polarization information crypto-multiplexing. These were achieved by co-training tasks across different polarization channels.

#### 4. Dynamic reconfiguration

Despite using a multiplexing scheme, the solid-state nature of neural networks cannot overcome the disadvantage of not being able to assign new functions to neural networks once they are manufactured. Consequently, some researchers have focused on dynamically reconfigurable D<sup>2</sup>NN to address the aforementioned issue.

Zhou *et al.* proposed an optoelectronic fusion computing architecture<sup>78</sup> based on the form of optical diffraction [see Fig. 7(a)], which utilized reconfigurable diffraction processing units as optoelectronic neural processors. It included large-scale diffractive neurons and weighted optical interconnections that could be programmed to establish various types of ANN with high model complexity and accuracy. To solve the problem of alignment errors and device manufacturing errors, they developed an adaptive training method, which efficiently addressed these issues and enables the network to exhibit enhanced performance. However, neuronal interconnections based on electronic circuits inevitably lead to delay and could not fully utilize the advantages of D<sup>2</sup>NN light-speed computation. Therefore, Liu *et al.* developed an active and reprogrammable D<sup>2</sup>NN architecture,<sup>79</sup> as shown in Fig. 7(b), to address the aforementioned issue. These D<sup>2</sup>NN utilized

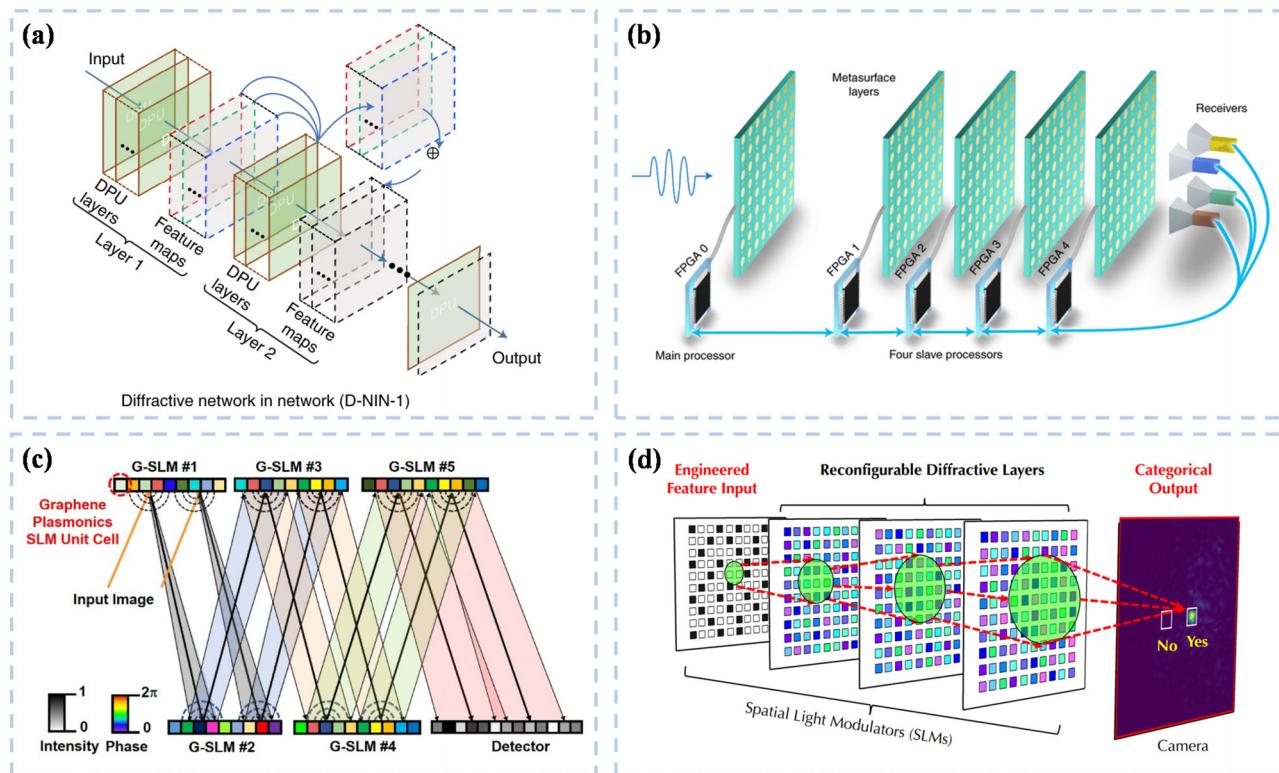
programmable metasurface<sup>80</sup> and informative metasurface<sup>81</sup> to facilitate weight reprogramming, and was capable of handling a variety of deep learning tasks. Beyond image classification, they demonstrated encoding and decoding for mobile communications, and real-time multibeam focusing. However, the system remained linear in the complex domain, with plans for incorporating nonlinear modulation to enhance the system's fitting quality in the future. In addition to utilizing programmable metasurface for multifunctional D<sup>2</sup>NN tasks, Fig. 7(c) depicts a reflective spatial light modulator composed of graphene nanoribbons designed by Zeng *et al.*<sup>82</sup> Compared with the D<sup>2</sup>NN constructed in Ref. 20, the computational accuracy of the D<sup>2</sup>NN utilizing the spatial light modulator proved equal or superior. This work further explored the relationship between the spatial optical modulators' quantity in a network and a single device's phase modulation range, easing the phase modulation requirements for spatial optical modulators and broadening their application scope in D<sup>2</sup>NN.

Likewise, Chen *et al.* developed a universal feature engineering method that transformed the features of input optical signals into binary images and employed a cascade reconfigurable liquid crystal spatial light modulator as the D<sup>2</sup>NN's physical architecture,<sup>83</sup> adaptable to various machine learning tasks [see Fig. 7(d)]. All input images and diffractive layers in this research were controllable by external signals, facilitating real-time adaptation of the D<sup>2</sup>NN for diverse tasks.

#### 5. On-chip integration

To eventually enable D<sup>2</sup>NN to partially or even fully replace ENN in computation, achieving on-chip integration constitutes a critical step. On-chip integration enhances the network's performance and computational efficiency, minimizes energy consumption and resource use, facilitates real-time operation and low latency, and promotes tight integration and scalability.

In 2021, Zarei *et al.*<sup>84</sup> constructed 5-layer on-chip D<sup>2</sup>NN using silicon 1D metasurface as shown in Fig. 8(a), and after training, these D<sup>2</sup>NN achieved 88% accuracy in classifying handwritten digits at a wavelength of 1.55  $\mu\text{m}$ . The proposed chip-scale fully passive, all-optical superneural network was notably compact and operates at the speed of light, featuring exceptionally low energy consumption. In the same year, Fu *et al.*<sup>31</sup> engineered integrated D<sup>2</sup>NN, as shown in Fig. 8(b), founded on one-dimensional metasurface atop a standard silicon-on-insulator substrate,<sup>84</sup> to execute a machine learning task for predicting coronary heart disease in a fully passive, optical manner, attaining an accuracy on par with existing techniques. In 2023, Fu *et al.*<sup>85</sup> expanded upon this foundation and implemented the D<sup>2</sup>NN on a chip using silicon photonic integration technology, see Fig. 8(c). They configured 1-layer and 3-layer D<sup>2</sup>NN hidden layers, respectively, conducted tests on, and validated them against the Iris dataset classification task, securing accuracies of 86.7% and 90% in the two D<sup>2</sup>NN, respectively. Figure 8(d) shows a miniature image classifier on a silicon photonic platform<sup>86</sup> that employs diffraction and dispersion engineering for light propagation and selective wavelength image recognition. The researchers assembled a system comprising multiple metasurface layers to achieve high-throughput vector-matrix multiplication operations via nearly 103 nanoscale phase modulators. The system demonstrated three functions: a 15-pixel spatial pattern classifier, a multichannel wavelength division multiplexer, and a hyperspectral image classifier. Beyond the use of one-dimensional metasurface for



**FIG. 7.** Dynamic reconfigurable D<sup>2</sup>NN: (a) Dynamic reconstruction based on reconfigurable diffraction processing unit. Reproduced with permission from Zhou *et al.*, Nat. Photonics **15**, 367 (2021). Copyright 2021 Springer Nature.<sup>78</sup> (b) Dynamic programming based on programmable metasurface and information metasurface. Reproduced with permission from Nat. Electron. **5**, 113 (2022). Copyright 2022 Springer Nature.<sup>79</sup> (c) Dynamic reconstruction based on tunable spatial light modulator. Reproduced with permission from Zeng *et al.*, Opt. Express **30**, 12712 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>82</sup> (d) Dynamic reconstruction based on reconfigurable liquid crystal spatial light modulator. Reproduced with permission from Chen *et al.*, Adv. Intell. Syst. **5**, 2300536 (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>83</sup>

2 June 2025 09:04:57

on-chip integration, two-dimensional metasurface also demonstrate significant application potential in this domain. Luo *et al.*<sup>29</sup> employed metasurface<sup>87–92</sup> as the physical medium of D<sup>2</sup>NN, achieving integration with complementary metal-oxide semiconductor (CMOS) technology. This integration facilitates sense-computing and offers a practical approach for D<sup>2</sup>NN application. Notably, they encoded information corresponding to distinct tasks into the input signal by using different polarization states, thereby enabling recognition of two distinct tasks, see Fig. 8(e).

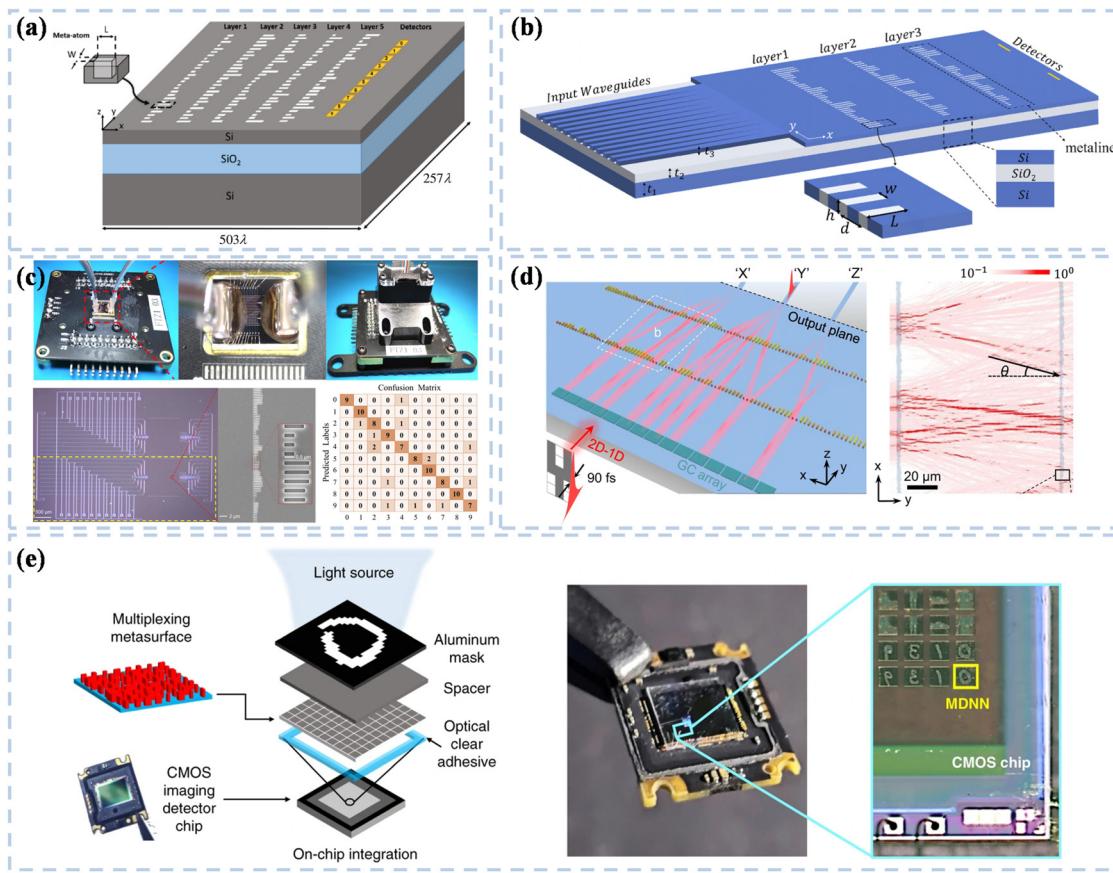
To gain a deeper understanding of D<sup>2</sup>NN's advantages and disadvantages in integration development, this study compares the aforementioned work with other types of ONN integrated on-chip, focusing on integration, energy consumption, and computational throughput. The results are presented in Table II. Here, NBU/mm<sup>2</sup> represents the total number of basic computing units that can be integrated per square millimeter, and J/operaion indicates the energy consumption required to complete each matrix operation.

## 6. Incoherent light source

Light sources in D<sup>2</sup>NN are primarily coherent. Due to the wave-like properties of coherent light, during diffraction, its wavefronts can

curve and overlap, causing image blur and distortion. This wave-like behavior may negatively impact the training and inference processes of D<sup>2</sup>NN. Furthermore, given that light sources in practical applications are largely incoherent, integrating incoherent light sources into D<sup>2</sup>NN can considerably broaden their application spectrum.

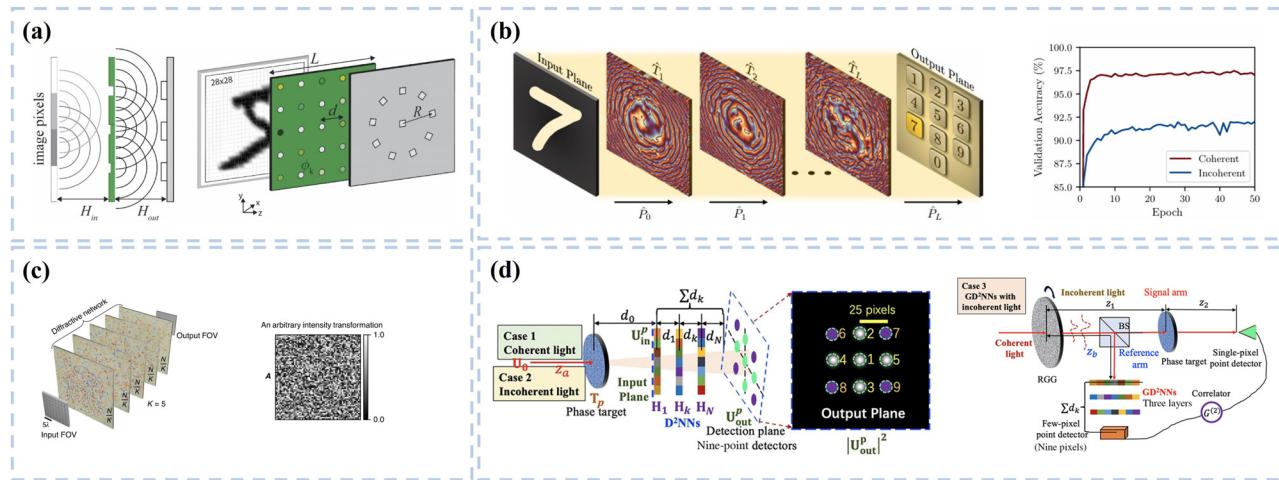
In 2022, Léonard *et al.*<sup>95</sup> conducted extensive simulations on single-layer linear diffractive metamaterials using spatially incoherent light sources, as illustrated in Fig. 9(a). Co-optimizing the training algorithm, diffractive layer properties, and system architecture resulted in recognition accuracies exceeding 94% and 85% for the MNIST and Fashion-MNIST datasets, respectively. This study offers valuable insights into developing D<sup>2</sup>NN for incoherent light, potentially enabling compact and rapid neuromorphic inference across widespread light fields. Reference 96 presented a framework for efficiently modeling and training D<sup>2</sup>NN with light of any spatial coherence, evaluating their performance on the MNIST dataset [see Fig. 9(b)]. Simulations showed D<sup>2</sup>NN achieved 97.54% and 91.17% accuracy under coherent and incoherent illumination, respectively, on the MNIST dataset. This superiority under coherent light is attributed to the nonlinear detection mechanism at the output layer, involving square amplitude relationships of the electric field. This subtle nonlinear property endows the D<sup>2</sup>NN's input signal with the capacity to



**FIG. 8.** On-chip D<sup>2</sup>NN: (a) A 5-layer on-chip D<sup>2</sup>NN using metalines. Reproduced with permission from Zarei *et al.*, Opt. Express **28**, 36668 (2020); Copyright 2020 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>84</sup> (b) Design of the on-chip D<sup>2</sup>NN based on a 1D metasurface. Reproduced with permission from Fu *et al.*, Opt. Express **29**, 31924 (2021); Copyright 2021 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>31</sup> (c) On-chip integrated D<sup>2</sup>NN. Reproduced with permission from Fu *et al.*, Nat. Commun. **14**, 70 (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>85</sup> (d) A miniature image classifier based on a silicon photonic platform. Reproduced with permission from Wang *et al.*, Nat. Commun. **13**, 2131 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>86</sup> (e) Metasurface-based on-chip integration of D<sup>2</sup>NN. Reproduced with permission from Luo *et al.*, Light **11**, 158 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>29</sup>

**TABLE II.** Selected performance comparison of on-chip integrated ONN.

Basic unit	Theoretical integration / (NBU/mm <sup>2</sup> )	Operation power consumption / (J/operation)	Throughput / Tops	Verification method
MZI (Ref. 4)	<10	$7.66 \times 10^{-14}$	6.4	Experiment
MZI (Ref. 93)	<10	$2.14 \times 10^{-13}$	21.6	Experiment
MMI (Ref. 94)	<25	$3.07 \times 10^{-14}$	30	Experiment
MMR with PCM (Ref. 9)	<5	$5.9 \times 10^{-15}$	28.8	Experiment
SWU (Ref. 84)	$\sim 3.75 \times 10^4$	N/A	$1.54 \times 10^4$	Simulation
SWU (Ref. 31)	$\sim 9.5 \times 10^3$	N/A	$1.96 \times 10^3$	Simulation
SWU (Ref. 85)	$\sim 2 \times 10^3$	$1.1 \times 10^{-17}$	$1.38 \times 10^4$	Experiment
SWU (Ref. 86)	$\sim 6.7 \times 10^3$	$4.2 \times 10^{-19}$	$4.05 \times 10^4$	Experiment
Meta-unit (Ref. 29)	$\sim 6.25 \times 10^6$	~0	N/A	Experiment



**FIG. 9.** Incoherent light as a light source: (a) D<sup>2</sup>NN for spatially incoherent light. Reproduced with permission from Léonard *et al.*, Opt. Express **30**, 12510 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>95</sup> (b) D<sup>2</sup>NN with arbitrary spatial coherence. Reproduced with permission from Filipovich *et al.*, Machine Learning with New Compute Paradigms (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>96</sup> (c) Generalized linear intensity transformation using D<sup>2</sup>NN. Reproduced with permission from Rahman *et al.*, Light **12**, 195 (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>97</sup> (d) Pattern recognition for OAM using D<sup>2</sup>NN. Reproduced with permission from Ye *et al.*, Phys. Rev. Appl. **20**, 054012 (2023). Copyright 2023 American Physical Society.

sensitively capture complex features, surpassing traditional linear processing models. Under incoherent illumination, the lack of coherence prevents the formation of stable interference patterns, reverting the input–output relationship to linear. Consequently, the D<sup>2</sup>NN behave no differently from a typical linear system, and the unique processing potential of phase information fails to manifest. Bai *et al.*<sup>30</sup> explored spatially incoherent D<sup>2</sup>NN for generalized linear intensity transformations, as depicted in Fig. 9(c). Their findings, through numerical analysis and experiments, indicate that D<sup>2</sup>NN can accurately perform linear intensity transformations given  $N \geq 2NiNo$ , where  $N$  is the number of phase-type diffractive features, and  $Ni$  and  $No$  are the numbers of features in the input and output spaces, respectively. These results are significant for designing all-optical vision processors compatible with natural incoherent light, with applications in computational microscopy and incoherent imaging via point spread function engineering. Figure 9(d) shows Ye *et al.*<sup>97</sup> combining ghost diffraction with D<sup>2</sup>NN to create D<sup>2</sup>NN that uses pseudo-thermal (incoherent) light for pattern recognition tasks involving orbital angular momentum (OAM). Despite an average signal-to-noise ratio of only 1.18, the experimental and numerical demonstrations highlighted the potential of using incoherent light for phase target classification. The discussed research suggests that incoherent light's use implies a shift away from relying solely on light interference for functions like phase modulation. In some instances, combining incoherent with coherent light can harness both advantages for various processing tasks. For instance, applications necessitating high-resolution feature detection employ coherent light, whereas those requiring enhanced image quality and system stability utilize incoherent light.

### III. APPLICATIONS OF D<sup>2</sup>NN

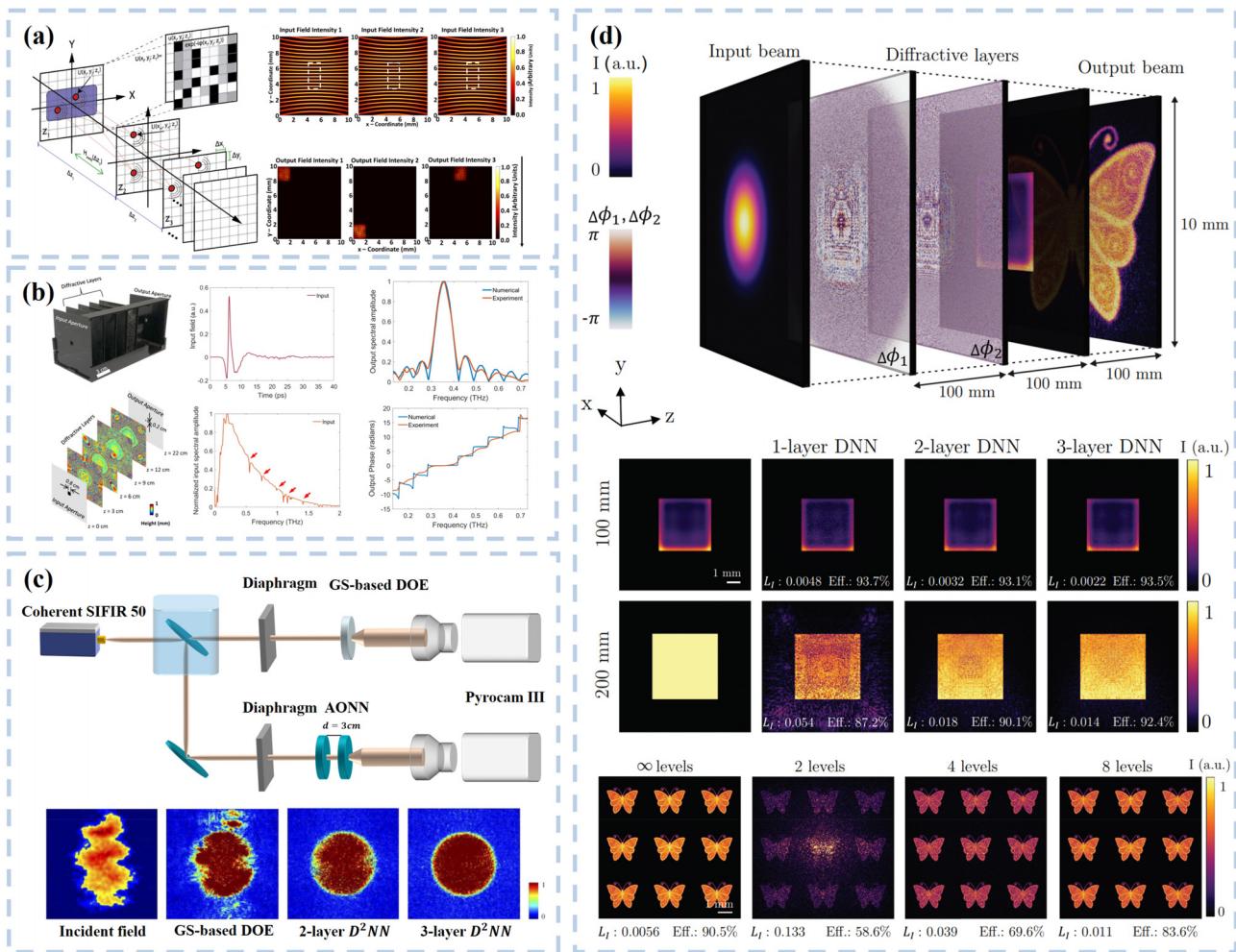
Recent advancements in AI offer innovative solutions to numerous inverse problems in optics. At the confluence of machine learning and optics, D<sup>2</sup>NN merge adaptive optics and deep learning to execute

diverse optical tasks through the design of task-specific elements. D<sup>2</sup>NN have shown impressive capabilities in fields including image recognition, beam shaping, logic operations, communications engineering, and image reconstruction. In Sec. II B, the vast majority of the work centered on image recognition tasks. This section concentrates on four key domains: beam shaping, privacy encryption, communications engineering, and image reconstruction.

#### A. Applications of D<sup>2</sup>NN on beam shaping

Beam shaping technology was developed to create beams with specific spatial intensity distributions for applications involving special wavefronts, defined optical field shapes, and patterned light intensities. Based on the wave properties of light, beam shaping technology employs reflective elements to alter light's phase of propagation for beam shaping. Concurrently, significant advancements have been achieved in beam shaping within the interdisciplinary realm of optics and deep learning.

Figure 10(a) illustrates Idehenre *et al.* utilizing D<sup>2</sup>NN to direct light to a specified area and controlling the beam steering through a light modulator.<sup>98</sup> They employed an array of photomodulators to generate the desired light source, training the network with a specialized set to efficiently steer the beam. However, this study was limited to single-wavelength beam shaping, excluding simultaneous consideration of steering capabilities and states. Veli and others used D<sup>2</sup>NN for passive all-optical shaping of arbitrary broadband pulses,<sup>32</sup> synthesizing desired optical waveforms. Building on this, they proposed a physical transfer learning approach, showing that replacing specific diffractive layers in the network allows for modular pulse shaping and tunability, as depicted in Fig. 10(b). This neural network simultaneously controlled phase and amplitude across a wideband for pulse shaping, offering a compact, energy-efficient alternative to traditional methods. Shi and others developed all-optical D<sup>2</sup>NN for efficient terahertz laser beam shaping<sup>99</sup> [see Fig. 10(c)]. In beam shaping



**FIG. 10.** Applications of D<sup>2</sup>NN on beam shaping field: (a) D<sup>2</sup>NN model for beam steering. Reproduced with permission from Idehenre and Mills, Opt. Express **28**, 36668 (2020); Copyright 2020 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>84</sup> (b) D<sup>2</sup>NN for pulse shaping. Reproduced with permission from Veli *et al.*, Nat. Commun. **12**, 37 (2021); Copyright 2021 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>32</sup> (c) D<sup>2</sup>NN for terahertz beam shaping. Reproduced with permission from Shi *et al.*, Opt. Express **29**, 7084 (2021); Copyright 2021 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>39</sup> (d) D<sup>2</sup>NN for advanced laser beam shaping. Reproduced with permission from Buske *et al.*, Opt. Express **30**, 22798 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>33</sup>

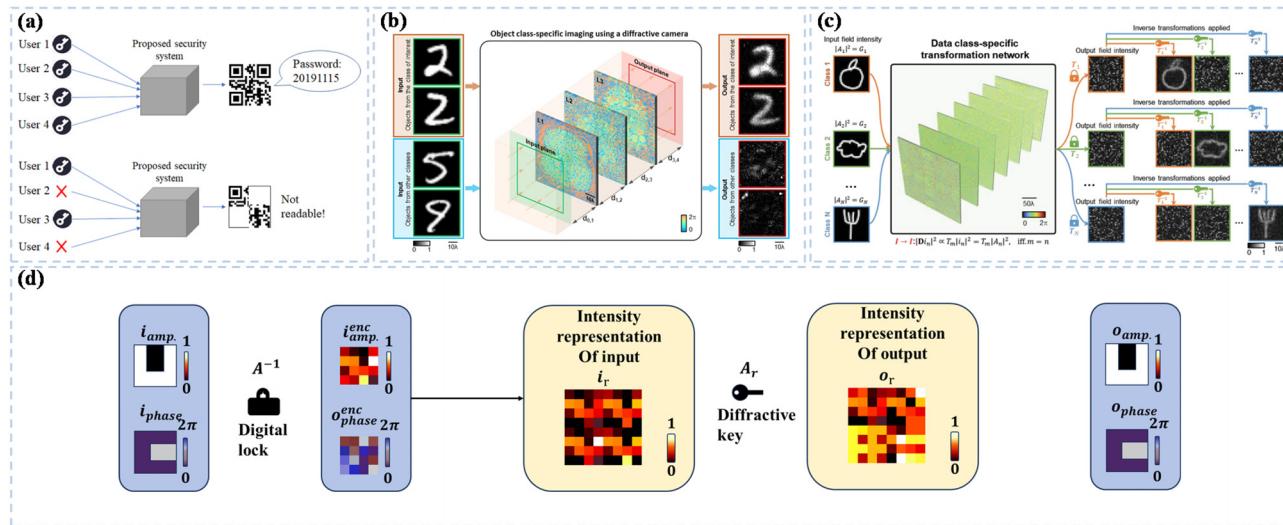
simulations, they found D<sup>2</sup>NN outperformed conventional diffractive elements, especially for beams with optical wave defects and aberrations. However, while neural networks with additional diffractive layers may enhance beam shaping performance, addressing the resultant energy loss and error accumulation poses a significant challenge in D<sup>2</sup>NN-based beam shaping. Figure 10(d) depicts Buske *et al.* utilizing D<sup>2</sup>NN to shape the beam for three targets that have different intensity distributions at different distances.<sup>33</sup> This method enabled efficient system design with cascaded diffractive optical elements and concurrent optimization of complex target field distributions across multiple planes.

## B. Applications of D<sup>2</sup>NN on privacy encryption

During the collection, transmission, and processing of diverse data types, such as images, information security emerges as crucial.

Optical cryptography,<sup>100</sup> a promising area in optical information processing, offers robust parallel processing capabilities and crucial encryption during data transmission.<sup>101</sup> Unlike traditional photonic, electronic, or optoelectronic hybrid computing platforms, D<sup>2</sup>NN can directly process 2D and 3D visual information without the need for pre-processing. This information is encoded through parameters like amplitude, phase, polarization, and the spectrum of the input light.

In 2020, Gao *et al.*<sup>102</sup> introduced a D<sup>2</sup>NN-based system for multi-image encryption and concealment. This system achieved simultaneous encryption and concealment of multiple images by directing them into separate optical channels, as illustrated in Fig. 11(a). Each input image was subjected to a distinct optical transformation, resulting in non-overlapping output images on the imaging plane. A wave-front matching algorithm optimized the phase mask for effective image encryption and concealment. The study outlines a generalized



**FIG. 11.** Applications of D<sup>2</sup>NN on privacy encryption: (a) D<sup>2</sup>NN applied to multi-image encryption and hiding. Reproduced with permission from Gao *et al.*, Opt. Commun. **463**, 125476 (2020). Copyright 2020 Elsevier B.V.<sup>102</sup> (b) D<sup>2</sup>NN for selective imaging. Reproduced with permission from Bai *et al.*, eLight **2**, 14 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>34</sup> (c) D<sup>2</sup>NN for specific data category conversion. (d) Two D<sup>2</sup>NN schemes for all-optical image encryption using incoherent illumination. Reproduced with permission from Bai *et al.*, Adv. Mater. **35**, 2212091 (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>35</sup>

framework, simulation outcomes, and potential applications, underscoring its significance in optical security and information processing. Bai *et al.*<sup>34</sup> excelled in optical pattern filtering and selective imaging via deep learning on structured transmissive surfaces. Employing the same framework, they demonstrated class-specific permutations and linear transformations for all-optical encryption, selectively altering or erasing objects based on class-specific criteria. This method enabled irreversible encryption, safeguarding the privacy of the targeted data class, as shown in Fig. 11(b). This D<sup>2</sup>NN design affords speed and efficiency by negating the need for extra digital computing or external power, ideal for remote systems requiring continuous operation with limited power. Figure 11(c) illustrates the extension of this work, where different data classes are assigned unique linear transformation matrices, and the D<sup>2</sup>NN are trained for class-specific transformations, allowing only users with the correct inverse matrix (decryption key) to recover the original information.<sup>35</sup> This framework enhances security and privacy by encoding input data into amplitude, phase, or intensity, employing class-specific encryption with predefined matrices for distinct data classes. This design introduces an additional security layer during transmission, ensuring that compromising one encryption key does not expose all original data. Furthermore, this design allows parallel, secure data distribution to multiple users, with each user decrypting only their assigned data class, facilitating novel applications. In 2024, Yang *et al.*<sup>103</sup> utilized spatially disjoint D<sup>2</sup>NN for complex-valued linear transformations and image encryption. They optimized diffractive features based on the relationship between the point spread function (PSF) and target complex-valued transformations. By training nine models with varying numbers of optimized diffractive features (N) and evaluating the MSE between the numerically measured  $\hat{A}_r$  and the target all-optical linear transformation  $A_r$ , they found that when the number of optimized diffractive features exceeds a threshold determined by the product of the input and output spatial bandwidths, the

spatially incoherent diffraction vision processor can approximate any complex-valued linear transformation and can be used for all-optical image encryption.

Two implementations for image encryption proposed in this work are shown in Fig. 11(d). The first approach encoded the message into a complex image, using amplitude and phase encoding or real and imaginary part encoding. The image was then encrypted using a digital lock by applying a linear transformation ( $A^{-1}$ ) to hide the original message. At the optical receiver, the encrypted message was decrypted by optimized incoherent D<sup>2</sup>NN to optically realize the inverse transformation  $A$ . The second approach swapped the roles of encryption and decryption, i.e., encrypted the message using spatially incoherent D<sup>2</sup>NN and decrypted it using  $A^{-1}$  through digital inversion. This study highlights the potential of diffractive optical networks for all-optical information processing under various lighting conditions.

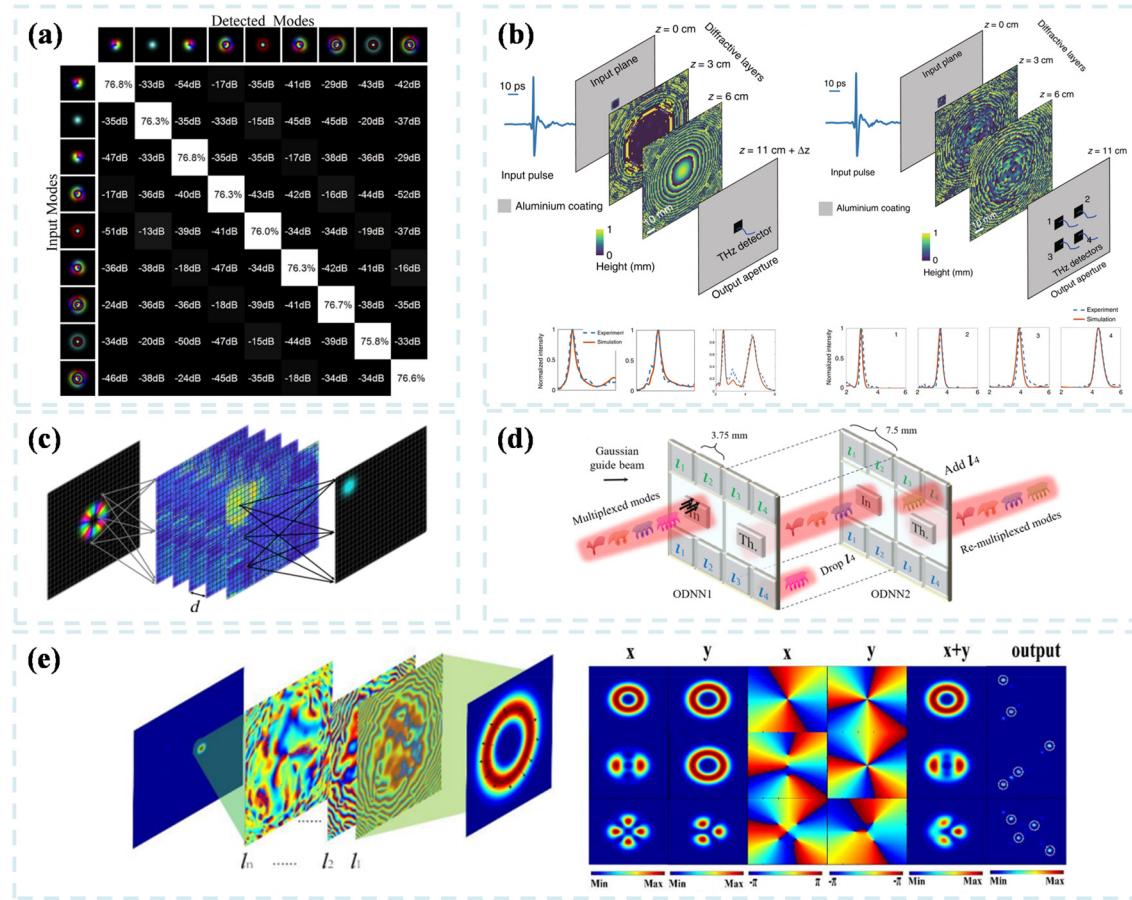
### C. Applications of D<sup>2</sup>NN on communication engineering

WDM technology<sup>104</sup> is a communication technique that combines a series of optical signals, each carrying information at different wavelengths, into a single bundle for transmission along a single optical fiber. At the receiving end, these optical signals are separated based on their wavelengths. In an era of increasing mobile data, utilizing WDM technology to enhance communication capacity is a crucial strategy in communication engineering.

The introduction of D<sup>2</sup>NN for WDM applications was pioneered by Zheng *et al.*<sup>105</sup> Their study on the phase modulation and layer-to-layer propagation of input optical signals by D<sup>2</sup>NN led to key conclusions: the inner product of any two optical fields remains unchanged after modulation by D<sup>2</sup>NN, and for complete spatial separation of the output signal light, the input optical fields must be orthogonal. This

suggests that D<sup>2</sup>NN are naturally suited for handling optically orthogonal modes. Figure 12(a) shows applications in mode conversion, mode multiplexing/demultiplexing, and pattern recognition, verified by simulation. This work provides a theoretical basis for the application of D<sup>2</sup>NN in communication engineering and further confirms their robust compatibility. Luo *et al.* constructed broadband D<sup>2</sup>NN<sup>36</sup> using time-coherent broadband terahertz pulses as the incident signal [see Fig. 12(b)]. They demonstrated adjustable single-band and dual-band spectral filters, along with spatially controlled wavelength demultiplexing, both showing good fitting results. Additionally, they designed corresponding broadband network training algorithms and incorporated material dispersion into neural network training. The orbital angular momentum has infinitely many orthogonal modes, providing unlimited channels for signal multiplexing and greatly improving communication capacity density. Consequently, the combination of D<sup>2</sup>NN and orbital angular momentum for WDM technology has become a key

research direction. Huang *et al.* established an exact mapping relationship between input and output beams by specific training of D<sup>2</sup>NN,<sup>37</sup> enabling vortex beam orbital angular momentum multiplexing, demultiplexing, mode conversion, and OAM-SK communication techniques, as shown in Fig. 12(c). The constructed D<sup>2</sup>NN could manipulate multiple vortex beams by controlling the phase and amplitude, enabling interconversion between single or mixed orbital angular momentum modes and Gaussian modes at different positions, as well as between different orbital angular momentum modes. In optical communication engineering, the orbital angular momentum interpolator multiplexer enhances dynamic interactivity in the communication network by enabling selective downloading and uploading. However, its realization is challenging due to the lack of effective mode-selective coupling and separation techniques. To address this, Xiong *et al.* designed a D<sup>2</sup>NN-based OAM multiplexer and demultiplexer<sup>106</sup> that accurately controlled optical waveforms through hybrid amplitude and



**FIG. 12.** Applications of D<sup>2</sup>NN on communication engineering: (a) Efficiency and crosstalk of a wavelength division multiplexer based on D<sup>2</sup>NN. Reproduced with permission from Zheng *et al.*, Opt. Lett. **47**, 605 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>53</sup> (b) Tunable single-pass and double-pass spectral filters, as well as spatially controlled wavelength demultiplexer. Reproduced with permission from Luo *et al.*, Light **8**, 112 (2019); Copyright 2019 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>36</sup> (c) Wave division multiplexing based on D<sup>2</sup>NN for orbital angular momentum. Reproduced with permission from Huang *et al.*, Phys. Rev. Appl. **15**, 014037 (2021). Copyright 2021 American Physical Society.<sup>37</sup> (d) Orbital angular momentum multiplexing and demultiplexing. Reproduced with permission from Xiong *et al.*, Opt. Express **29**, 31924 (2021); Copyright 2021 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>31</sup> (e) Polarization optical D<sup>2</sup>NN based on metasurface and its classification, generation, and multiplexing of polarization orbital angular momentum beams. Reproduced with permission from Zhang *et al.*, Opt. Express **30**, 22798 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>33</sup>

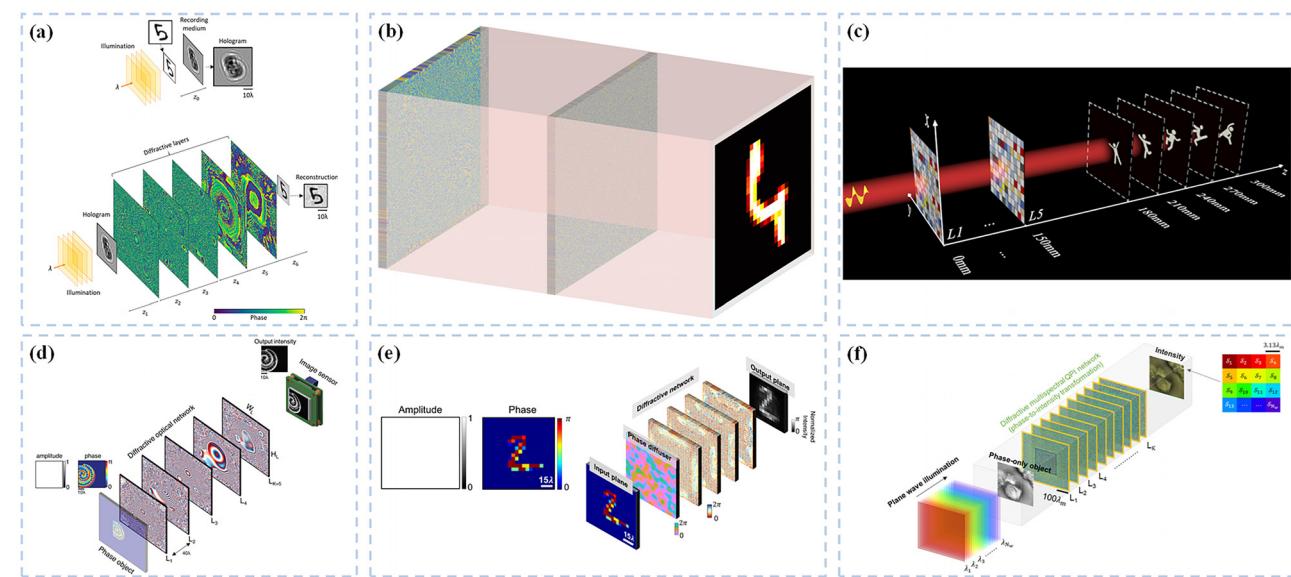
phase modulation, achieving diffraction efficiency and mode purity higher than 95%. Figure 12(d) shows an OAM multiplexer and demultiplexer formed by cascading two D<sup>2</sup>NN with selectively coupled and separated OAM modes. This work also analyzed factors affecting the communication performance of D<sup>2</sup>NN, including atmospheric turbulence, layer spacing, and diffraction screen size, and proposed an improved optimization scheme. The above work struggles with beams having different spatial orbital angular momentum modes with polarization direction distributions. For this reason, Zhang *et al.*<sup>107</sup> proposed metasurface-based D<sup>2</sup>NN for classifying, generating, multiplexing, and demultiplexing polarized OAM beams, as shown in Fig. 12(e). They designed a rectangular cell structure of an anisotropic metasurface using a wavefront matching algorithm to independently modulate the phase of each incoming polarized signal light beam. The researchers successfully classified and demultiplexed 14 orthogonally polarized vortex beams and demonstrated through simulation that the 10-layer network could achieve high-quality generation and demultiplexing of cylindrical vector beams.

#### D. Applications of D<sup>2</sup>NN on image reconstruction

The image reconstruction technique,<sup>108</sup> utilizing external measurements to detail the shape of three-dimensional objects, extends beyond its traditional use in radiological medical equipment. With technological advancements, the application of this reconstruction technique has expanded into various fields, such as industrial inspection, bio-imaging, security, and art restoration. Integrating D<sup>2</sup>NN into image reconstruction has notably expanded its application scope and

capabilities, particularly in hologram reconstruction and quantitative phase imaging.

Holographic imaging<sup>109</sup> grounded in wavefront recording principles, captures both amplitude and phase information of light waves in a hologram. During hologram reconstruction, the hologram is illuminated by a reference wave, typically a laser beam of the recording's original wavelength. This process enables the reproduction of the original object's complete three-dimensional wavefront image. The D<sup>2</sup>NN-based all-optical hologram reconstruction method,<sup>38</sup> demonstrated in Fig. 13(a), effectively removes double-image artifacts, enabling high-quality reconstruction of rotated and scaled images with robust performance. By adjusting network training parameters, the method also enhanced diffraction efficiency and expanded the hologram recording's depth of field. However, this study is limited to intensity-directed holographic reconstruction; future research should validate the technique on complex images with uniform intensity distributions. Liao *et al.* developed a D<sup>2</sup>NN-based terahertz hologram reconstructor,<sup>110</sup> with simulations and experiments confirming its capability to dynamically reconstruct multiple digital holograms. Figure 13(b) illustrates how modifying the loss function enabled the generation of uniform amplitude surface holograms. Nevertheless, challenges remain in resolving inter-structure coupling, enhancing transmittance, and controlling fabrication alignment errors. Huang *et al.*<sup>111</sup> utilized orbital angular momentum multiplexing<sup>12,113</sup> with five-layer D<sup>2</sup>NN to enrich holographic imaging's information content and achieve spatial depth in holograms. Figure 13(c) shows independent modulation of multiple orbital angular momentum modes and spatial depths using the Missy transform and linear modulation. Furthermore, training with varied



**FIG. 13.** Applications of D<sup>2</sup>NN on image reconstruction: (a) All-optical holographic reconstruction based on D<sup>2</sup>NN. Reproduced with permission from Sakib Rahman and Ozcan, ACS Photonics **8**, 3375 (2021). Copyright 2021 American Chemical Society.<sup>38</sup> (b) Terahertz holographic dynamic reconstruction based on D<sup>2</sup>NN. (c) Image reconstruction using orbital angular momentum multiplexing. Reproduced with permission from Huang *et al.*, Opt. Express **30**, 22798 (2022); Copyright 2022 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>33</sup> (d) D<sup>2</sup>NN for QPI. Reproduced with permission from Opt. Mater. **10**, 2200281 (2022). Copyright 2022 Wiley-VCH GmbH.<sup>39</sup> (e) D<sup>2</sup>NN for QPI via random diffractors. Reproduced with permission from Li *et al.*, Light **4**, 17 (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>115</sup> (f) Multispectral QPI. Reproduced with permission from Shen *et al.*, Adv. Intell. Syst. **5**, 2300300 (2023); Copyright 2023 Authors, licensed under a Creative Commons Attribution (CC BY) license.<sup>116</sup>

evaluation indexes enhanced the D<sup>2</sup>NN's noise resistance and robustness, introducing a depth-controllable imaging technique that achieved holographic imaging with spatial depth through multiplanar light conversion and *in situ* propagation.

Quantitative phase imaging (QPI)<sup>114</sup> employs the modulating effect of a sample on light wavefronts to measure the optical path length, capturing phase information for constructing 2D or 3D refractive index distribution maps. Mengu and Ozcan<sup>39</sup> introduced a five-layer D<sup>2</sup>NN-based quantized phase imaging system that directly transforms input phase information into an intensity image, facilitating all-optical phase imaging sans digital reconstruction algorithms, as shown in Fig. 13(d). Testing across various datasets (Iny-Imagenet, CIFAR-10, Fashion-MNIST) yielded SSIM and PSNR values of  $0.824 \pm 0.050$  ( $26.43 \text{ dB} \pm 2.69$ ),  $0.917 \pm 0.041$  ( $31.98 \text{ dB} \pm 3.15$ ), and  $0.596 \pm 0.116$  ( $26.94 \text{ dB} \pm 1.5$ ) respectively, demonstrating the system's accuracy and reproducibility. This D<sup>2</sup>NN system, boasting a compact axial range of 200–300 wavelengths, shows promise for achieving high efficiency and frame rates. In Fig. 13(e), Li *et al.*<sup>115</sup> developed D<sup>2</sup>NN for QPI utilizing random diffraction. The team assessed the network's performance under known and unknown random lenses, including a lensless state, analyzing imaging effects through mean correlation coefficient (PCC) and absolute phase error, and examined the impact of correcting aberrations caused by random lenses on image reconstruction quality. This highlights D<sup>2</sup>NN's potential to revolutionize markerless microscopy and sensing technologies, particularly in biomedical applications. Shen *et al.*<sup>116</sup> employed deep learning-optimized spatial engineering techniques in constructing D<sup>2</sup>NN for multispectral QPI of transparent objects, depicted in Fig. 13(f). Two D<sup>2</sup>NN were developed for multispectral QPI in 9 and 16 spectral bands, tested over 10 diffractive layers. The 9-band and 16-band designs yielded SSIM values of  $0.770 \pm 0.015$  and  $0.726 \pm 0.031$ , and PSNR values of  $17.04 \pm 0.33 \text{ dB}$  and  $16.67 \pm 0.43 \text{ dB}$ , respectively. Variations in SSIM and PSNR likely arise from the differential QPI performance of the two designs across spectral bands. These studies provide a foundation for developing compact, efficient high-throughput solutions for quantitative phase microscopy and spectroscopy, tunable to operating wavelengths, opening new avenues in QPI and spectroscopic device design.

#### IV. SUMMARY AND OUTLOOK

This paper provides an overview of D<sup>2</sup>NN, details the evolution of ONN and the D<sup>2</sup>NN concept, explores the theoretical models and mathematical analyses of D<sup>2</sup>NN, and summarizes recent research on D<sup>2</sup>NN's internal mechanisms.

D<sup>2</sup>NN have made great progress due to the continuous efforts of researchers, but the ability of the network to handle complex tasks and the computational accuracy still need to be further improved. To this end, the following points will be analyzed:

1. Current mainstream physical architectures for D<sup>2</sup>NN include using spatial light modulators (SLM) alongside traditional diffractive elements, employing 3D printing technology, and utilizing metasurface.<sup>117</sup> The advantage of employing SLMs lies in their dynamic adjustability, offering D<sup>2</sup>NN the capability for real-time modification. The architecture can be applied to different D<sup>2</sup>NN designs, as well. However, SLM present clear disadvantages, including low integration levels and response speeds constrained by the SLM themselves. 3D printing technology,<sup>118</sup> while reducing manufacturing costs and facilitating rapid design

iterations of D<sup>2</sup>NN, is constrained by printing precision and material choice, leaving significant room for improvement in network integration and optical efficiency. Metasurface<sup>119</sup> enables precise micro-nano scale regulation of optical wavefronts, aiding in intelligent high-dimensional information identification and on-chip integration for D<sup>2</sup>NN. However, their static nature limits flexibility compared to SLM-based systems.

2. Given the complexity and potential inefficiency in training D<sup>2</sup>NN due to its current architecture, the integrated design of hardware and software becomes essential for optimizing system performance. This approach allows for the customization of physical hardware for specific optical processing tasks and software adjustments to fully leverage the benefits of optical hardware. On the other hand, incorporating advanced algorithmic strategies from electronic neural networks, like ensemble learning and knowledge distillation, into D<sup>2</sup>NN design can significantly enhance computational efficiency. Moreover, reverse engineering the optical network's structure via optimization algorithms during the diffractive network design phase represents a significant research area for D<sup>2</sup>NN. This strategy enhances both the overall design efficiency and the network's computational performance by leveraging deep learning's robust optimization capabilities.
3. While multilayer D<sup>2</sup>NN can enhance recognition accuracy, challenges such as lateral and axial deviations, surface reflections, diffraction efficiencies, and manufacturing defects, including in diffractive layers, must be addressed to realize compact D<sup>2</sup>NN.
4. The integration of neural networks' data processing power with the speed and efficiency of optical systems makes D<sup>2</sup>NN a promising technology across various fields. Presently, D<sup>2</sup>NN image processing predominantly involves simple static datasets like MNIST, Fashion-MNIST, and Cifar-10. Expanding into complex homomorphic image processing applications, such as face recognition, object tracking, and motion capture, represents a valuable direction.

In conclusion, this paper will offer insights and recommendations for the future advancement of D<sup>2</sup>NN, informed by its current state of development and focusing on three critical areas: nonlinear technology, multiplexing, and on-chip integration.

#### A. Nonlinear technology

First, activating optical nonlinearities necessitates high optical energy and specialized materials, making the pursuit of materials that bolster nonlinear effects without excessive energy a pivotal direction for future research. This includes researching nonlinear optical materials, 2D materials, and hybrid materials with high nonlinear responsiveness capable of lower power threshold activation. Second, given the sizable nature of current nonlinear systems, devising methods for their efficient integration and packaging into practical applications without compromising performance is crucial. Therefore, exploring the use of nonlinear metasurface and optical coatings presents a promising direction. Current approaches employ optoelectronic hybridization to introduce nonlinearities, though this results in repetitive conversions that affect both response time and energy consumption. Further integrating the optoelectronic conversion process into the diffractive layer's design could mitigate performance losses due to these conversions. Finally, beyond spatial structure nonlinear integration,

incorporating time-dimension modulation presents an opportunity to introduce nonlinearities. For instance, leveraging a pulsed light source to elicit a nonlinear response temporally enables the creation of time-space coupled D<sup>2</sup>NN systems. This approach facilitates overcoming existing geometrical and material limitations. Future endeavors should concentrate on achieving greater integration and miniaturization of nonlinear D<sup>2</sup>NN, ensuring high performance and energy efficiency.

## B. Multiplexing

First, implementing advanced techniques like OAM multiplexing and dense wavelength division multiplexing (DWDM) not only broadens data transmission bandwidth but also boosts D<sup>2</sup>NN's multitasking capabilities. OAM multiplexing enables simultaneous signal transmission through unique orbital angular momenta for each channel, while DWDM increases data channel independence by reducing wavelength spacing. Future research should delve into integrating these advanced multiplexing techniques within D<sup>2</sup>NN systems, a topic of considerable potential. Second, leveraging metasurface<sup>[120,121]</sup> for their modulation capabilities to unify spatial, wavelength, and polarization multiplexing, enhancing multi-information and multi-channel sensing, presents a promising avenue for synergistically boosting D<sup>2</sup>NN computational abilities. Finally, while multiplexing in D<sup>2</sup>NN complicates network training, researching methods to enhance training speed and optimize objectives remains a critical area of study.

## C. On-chip integration

First, existing research structures are limited to phase modulation. There is a demand for on-chip diffractive structures capable of multi-dimensional optical field modulation within constrained degrees of freedom to handle complex tasks. Second, adopting bottom-up approaches to identify novel diffractive materials, like 2D materials,<sup>[122]</sup> promises robust performance in compact sizes. In addition, incorporating advanced diffractive elements, including metasurface, with detector front-ends will lead to the development of rapid and intelligent sensing devices. These elements will grant detectors real-time optical signal modulation, capturing multifaceted optical information. Such intelligent sensing platforms will benefit specialized fields like image processing and broader industries, including biomedical imaging, high-speed communications, and data storage. In image processing and machine vision, the parallelism of integrated diffractive elements and preprocessing capabilities will boost processing speed and efficiency, pushing the boundaries of current image perception capabilities. Finally, on-chip networks facilitate general-purpose logic operations, significantly accelerating and transforming electronic chip functionalities. This endeavor concentrates on integrating optical computing within on-chip networks, particularly through the design and optimization of optical logic gates. Precise control over the diffractive light path, simulating classical logic operations, enables on-chip optical neural networks to implement specific algorithms and evolves into programmable optical processing units for various computational tasks.

## ACKNOWLEDGMENTS

The authors acknowledge the financial support from the National Key Research and Development Program of China (Grant No. 2023YFB2806504), the National Natural Science Foundation of China (Grant No. 62275078), the Natural Science Foundation of

Hunan Province of China (Grant No. 2022JJ20020), the Science and Technology Innovation Program of Hunan Province (Grant No. 2023RC3101), and the Shenzhen Science and Technology Program (Grant No. JCYJ20220530160405013).

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

## Author Contributions

**Haijia Chen:** Data curation (equal); Investigation (equal); Writing – original draft (equal). **Shaozhen Lou:** Investigation (equal); Writing – original draft (equal). **Quan Wang:** Data curation (equal); Validation (equal). **Peifeng Huang:** Validation (equal); Writing – original draft (equal). **Huigao Duan:** Validation (equal); Writing – original draft (equal); Writing – review & editing (equal). **Yueqiang Hu:** Supervision (equal); Writing – review & editing (equal).

## DATA AVAILABILITY

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

## REFERENCES

- <sup>1</sup>Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436 (2015).
- <sup>2</sup>K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNET classification," in *Proceedings of the IEEE International Conference on Computer Vision* (IEEE, 2015), p. 1026.
- <sup>3</sup>O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, and M. Bernstein, "ImageNET large scale visual recognition challenge," *Int. J. Comput. Vision* **115**, 211 (2015).
- <sup>4</sup>Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, and D. Englund, "Deep learning with coherent nanophotonic circuits," *Nat. Photonics* **11**, 441 (2017).
- <sup>5</sup>R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in *Proceedings of the 25th International Conference on Machine Learning* (Association for Computing Machinery, 2008), p. 160.
- <sup>6</sup>L. Cutrona, E. Leith, C. Palermo, and L. Porcello, "Optical data processing and filtering systems," *IRE Trans. Inf. Theory* **6**, 386 (1960).
- <sup>7</sup>J.-H. Han, F. Boeuf, J. Fujikata, S. Takahashi, S. Takagi, and M. Takenaka, "Efficient low-loss InGaAsP/Si hybrid MOS optical modulator," *Nat. Photonics* **11**, 486 (2017).
- <sup>8</sup>J. Liu, A. S. Raja, M. Karpov, B. Ghadiani, M. H. Pfeiffer, B. Du, N. J. Engelsen, H. Guo, M. Zervas, and T. J. Kippenberg, "Ultralow-power chip-based soliton microcombs for photonic integration," *Optica* **5**, 1347 (2018).
- <sup>9</sup>J. Feldmann, N. Youngblood, M. Karpov, H. Gehring, X. Li, M. Stappers, M. Le Gallo, X. Fu, A. Lukashchuk, and A. S. Raja, "Parallel convolutional processing using an integrated photonic tensor core," *Nature* **589**, 52 (2021).
- <sup>10</sup>D. Psaltis and N. Farhat, "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback," *Opt. Lett.* **10**, 98 (1985).
- <sup>11</sup>K. Wagner and D. Psaltis, "Multilayer optical learning networks," *Appl. Opt.* **26**, 5061 (1987).
- <sup>12</sup>K. Vandoorne, J. Dambre, D. Verstraeten, B. Schrauwen, and P. Bienstman, "Parallel reservoir computing using optical amplifiers," *IEEE Trans. Neural Networks* **22**, 1469 (2011).
- <sup>13</sup>N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, "Optical implementation of the Hopfield model," *Appl. Opt.* **24**, 1469 (1985).
- <sup>14</sup>D. Psaltis, D. Brady, X.-G. Gu, and S. Lin, "Holography in artificial neural networks," *Nature* **343**, 325 (1990).

- <sup>15</sup>S. L. Yeh, R. C. Lo, and C. Y. Shi, "Optical implementation of the Hopfield neural network with matrix gratings," *Appl. Opt.* **43**, 858 (2004).
- <sup>16</sup>Z. Deng, D.-K. Qing, P. R. Hemmer, and M. S. Zubairy, "Implementation of optical associative memory by a computer-generated hologram with a novel thresholding scheme," *Opt. Lett.* **30**, 1944–1946 (2005).
- <sup>17</sup>A. R. Romariz and K. H. Wagner, "Tunable vertical-cavity surface-emitting laser with feedback to implement a pulsed neural model. I. Principles and experimental demonstration," *Appl. Opt.* **46**, 4736 (2007).
- <sup>18</sup>S. Pai, Z. Sun, T. W. Hughes, T. Park, B. Bartlett, I. A. Williamson, M. Minkov, M. Milanizadeh, N. Abebe, and F. Morichetti, "Experimentally realized in situ backpropagation for deep learning in photonic neural networks," *Science* **380**, 398 (2023).
- <sup>19</sup>Z. Xu, T. Zhou, M. Ma, C. Deng, Q. Dai, and L. Fang, "Large-scale photonic chiplet Taichi empowers 160-TOPS/W artificial general intelligence," *Science* **384**, 202 (2024).
- <sup>20</sup>X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, "All-optical machine learning using diffractive deep neural networks," *Science* **361**, 1004 (2018).
- <sup>21</sup>J. Li, Y.-C. Hung, O. Kulce, D. Mengu, and A. Ozcan, "Polarization multiplexed diffractive computing: All-optical implementation of a group of linear transformations through a polarization-encoded diffractive network," *Light Sci. Appl.* **11**, 153 (2022).
- <sup>22</sup>J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature* **569**, 208 (2019).
- <sup>23</sup>F. Ashtiani, A. J. Geers, and F. J. N. Aflatouni, "An on-chip photonic deep neural network for image classification," *Nature* **606**, 501 (2022).
- <sup>24</sup>B. Bai, Q. Yang, H. Shu, L. Chang, F. Yang, B. Shen, Z. Tao, J. Wang, S. Xu, and W. Xie, "Microcomb-based integrated photonic processing unit," *Nat. Commun.* **14**, 66 (2023).
- <sup>25</sup>D. Pierangeli, V. Palmieri, G. Marcucci, C. Moriconi, G. Perini, M. De Spirito, M. Papi, and C. Conti, "Deep optical neural network by living tumour brain cells," *arxiv:1812.09311* (2018).
- <sup>26</sup>E. Khoram, A. Chen, D. Liu, L. Ying, Q. Wang, M. Yuan, and Z. Yu, "Nanophotonic media for artificial neural inference," *Photonics Res.* **7**, 823 (2019).
- <sup>27</sup>C. Qian, Z. Wang, H. Qian, T. Cai, B. Zheng, X. Lin, Y. Shen, I. Kaminer, E. Li, and H. Chen, "Dynamic recognition and mirage using neuro-metamaterials," *Nat. Commun.* **13**, 2694 (2022).
- <sup>28</sup>Y. Zuo, B. Li, Y. Zhao, Y. Jiang, Y.-C. Chen, P. Chen, G.-B. Jo, J. Liu, and S. Du, "All-optical neural network with nonlinear activation functions," *Optica* **6**, 1132 (2019).
- <sup>29</sup>X. Luo, Y. Hu, X. Ou, X. Li, J. Lai, N. Liu, X. Cheng, A. Pan, and H. Duan, "Metasurface-enabled on-chip multiplexed diffractive neural networks in the visible," *Light Sci. Appl.* **11**, 158 (2022).
- <sup>30</sup>M. S. S. Rahman, X. Yang, J. Li, B. Bai, and A. Ozcan, "Universal linear intensity transformations using spatially incoherent diffractive processors," *Light Sci. Appl.* **12**, 195 (2023).
- <sup>31</sup>T. Fu, Y. Zang, H. Huang, Z. Du, C. Hu, M. Chen, S. Yang, and H. Chen, "On-chip photonic diffractive optical neural network based on a spatial domain electromagnetic propagation model," *Opt. Express* **29**, 31924 (2021).
- <sup>32</sup>M. Veli, D. Mengu, N. T. Yardimci, Y. Luo, J. Li, Y. Rivenson, M. Jarrahi, and A. Ozcan, "Terahertz pulse shaping using diffractive surfaces," *Nat. Commun.* **12**, 37 (2021).
- <sup>33</sup>P. Buske, A. Völl, M. Eisebitt, J. Stollenwerk, and C. Holly, "Advanced beam shaping for laser materials processing based on diffractive neural networks," *Opt. Express* **30**, 22798 (2022).
- <sup>34</sup>B. Bai, Y. Luo, T. Gan, J. Hu, Y. Li, Y. Zhao, D. Mengu, M. Jarrahi, and A. Ozcan, "To image, or not to image: Class-specific diffractive cameras with all-optical erasure of undesired objects," *eLight* **2**(1), 14 (2022).
- <sup>35</sup>B. Bai, H. Wei, X. Yang, T. Gan, D. Mengu, M. Jarrahi, and A. Ozcan, "Data-class-specific all-optical transformations and encryption," *Adv. Mater.* **35**, 2212091 (2023).
- <sup>36</sup>Y. Luo, D. Mengu, N. T. Yardimci, Y. Rivenson, M. Veli, M. Jarrahi, and A. Ozcan, "Design of task-specific optical systems using broadband diffractive neural networks," *Light Sci. Appl.* **8**, 112 (2019).
- <sup>37</sup>Z. Huang, P. Wang, J. Liu, W. Xiong, Y. He, J. Xiao, H. Ye, Y. Li, S. Chen, and D. Fan, "All-optical signal processing of vortex beams with diffractive deep neural networks," *Phys. Rev. Appl.* **15**, 014037 (2021).
- <sup>38</sup>M. S. Sakib Rahman and A. Ozcan, "Computer-free, all-optical reconstruction of holograms using diffractive networks," *ACS Photonics* **8**, 3375 (2021).
- <sup>39</sup>D. Mengu and A. Ozcan, "All-optical phase recovery: Diffractive computing for quantitative phase imaging," *Adv. Opt. Mater.* **10**, 2200281 (2022).
- <sup>40</sup>M. Born and E. Wolf, "Principles of optics: Electromagnetic theory of propagation," in *Interference and Diffraction of Light* (Elsevier, 2013).
- <sup>41</sup>N. Mukunda, "Consistency of Rayleigh's diffraction formulas with Kirchhoff's boundary conditions," *J. Opt. Soc. Am.* **52**, 336 (1962).
- <sup>42</sup>D. Mengu, Y. Luo, Y. Rivenson, and A. Ozcan, "Analysis of diffractive optical neural networks and their integration with electronic neural networks," *IEEE J. Sel. Top. Quantum Electron.* **26**, 3700114 (2020).
- <sup>43</sup>H. Dou, Y. Deng, T. Yan, H. Wu, X. Lin, and Q. Dai, "Residual D<sup>2</sup>NN: Training diffractive deep neural networks via learnable light shortcuts," *Opt. Lett.* **45**, 2688 (2020).
- <sup>44</sup>V. Wiley and T. Lucas, "Computer vision and image processing: A paper review," *Int. J. Artif. Intell. Res.* **2**, 29 (2018).
- <sup>45</sup>T. Fang, J. Li, X. Zhang, and X. Dong, "Classification accuracy improvement of the optical diffractive deep neural network by employing a knowledge distillation and stochastic gradient descent  $\beta$ -Lasso joint training framework," *Opt. Express* **29**, 44264 (2021).
- <sup>46</sup>J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, "Striving for simplicity: The all convolutional net," *arXiv:1412.6806* (2014).
- <sup>47</sup>L. Breiman, "Bagging predictors," *Mach. Learn.* **24**, 123 (1996).
- <sup>48</sup>B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles," in *Advances in Neural Information Processing Systems* (Neural Information Processing Systems Foundation, 2017), Vol. 30.
- <sup>49</sup>C. Ju, A. Bibaut, and M. van der Laan, "The relative performance of ensemble methods with deep convolutional neural networks for image classification," *J. Appl. Stat.* **45**, 2800 (2018).
- <sup>50</sup>J. Li, D. Mengu, Y. Luo, Y. Rivenson, and A. Ozcan, "Class-specific differential detection in diffractive optical neural networks improves inference accuracy," *Adv. Photonics* **1**, 046001 (2019).
- <sup>51</sup>M. S. S. Rahman, J. Li, D. Mengu, Y. Rivenson, and A. Ozcan, "Ensemble learning of diffractive optical networks," *Light Sci. Appl.* **10**, 14 (2021).
- <sup>52</sup>J. Shi, L. Zhou, T. Liu, C. Hu, K. Liu, J. Luo, H. Wang, C. Xie, and X. Zhang, "Multiple-view D<sup>2</sup>NNs array: Realizing robust 3D object recognition," *Opt. Lett.* **46**, 3388 (2021).
- <sup>53</sup>J. Shi, Y. Chen, and X. Zhang, "Broad-spectrum diffractive network via ensemble learning," *Opt. Lett.* **47**, 605 (2022).
- <sup>54</sup>Y. Li, Z. Zheng, R. Li, Q. Chen, H. Luan, H. Yang, Q. Zhang, and M. Gu, "Multiscale diffractive U-Net: A robust all-optical deep learning framework modeled with sampling and skip connections," *Opt. Express* **30**, 36700 (2022).
- <sup>55</sup>F. Chen, J. LaMacchia, and D. Fraser, "Holographic storage in lithium niobate," *Appl. Phys. Lett.* **13**, 223 (1968).
- <sup>56</sup>L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, 1995).
- <sup>57</sup>T. Yan, J. Wu, T. Zhou, H. Xie, F. Xu, J. Fan, L. Fang, X. Lin, and Q. Dai, "Fourier-space diffractive deep neural network," *Phys. Rev. Lett.* **123**, 023901 (2019).
- <sup>58</sup>M. Song, R. Li, and J. Wang, "Only frequency domain diffractive deep neural networks," *Appl. Opt.* **62**, 1082 (2023).
- <sup>59</sup>D. N. Christodoulides, T. H. Coskun, M. Mitchell, and M. Segev, "Theory of incoherent self-focusing in biased photorefractive media," *Phys. Rev. Lett.* **78**, 646 (1997).
- <sup>60</sup>L. Waller, G. Situ, and J. W. Fleischer, "Phase-space measurement and coherence synthesis of optical beams," *Nat. Photonics* **6**, 474 (2012).
- <sup>61</sup>Y. Sun, M. Dong, M. Yu, J. Xia, X. Zhang, Y. Bai, L. Lu, and L. Zhu, "Nonlinear all-optical diffractive deep neural network with 10.6  $\mu$ m wavelength for image classification," *Int. J. Opt.* **2021**, 6667495.
- <sup>62</sup>Y. Sun, M. Dong, M. Yu, L. Lu, S. Liang, J. Xia, and L. Zhu, "Modeling and simulation of all-optical diffractive neural network based on nonlinear optical materials," *Opt. Lett.* **47**, 126 (2022).
- <sup>63</sup>T. Wang, M. M. Sohoni, L. G. Wright, M. M. Stein, S.-Y. Ma, T. Onodera, M. G. Anderson, and P. L. McMahon, "Image sensing with multilayer nonlinear optical neural networks," *Nat. Photonics* **17**, 408 (2023).

- <sup>64</sup>J. P. Marangos, "Electromagnetically induced transparency," *J. Mod. Opt.* **45**, 471 (1998).
- <sup>65</sup>M. Fleischhauer, A. Imamoglu, and J. P. Marangos, "Electromagnetically induced transparency: Optics in coherent media," *Rev. Mod. Phys.* **77**, 633 (2005).
- <sup>66</sup>C. Kai, Z. Lianqing, L. Xiaoping, Y. Qifeng, and L. Fei, "All-polarization-maintaining fiber laser mode-locked by graphene," *Infrared Laser Eng.* **46**, 1005004 (2017).
- <sup>67</sup>M. Divyasree, E. Shiju, J. Francis, P. Anusha, S. V. Rao, and K. Chandrasekharan, "ZnSe/PVP nanocomposites: Synthesis, structural and nonlinear optical analysis," *Mater. Chem. Phys.* **197**, 208 (2017).
- <sup>68</sup>R. Divya, N. Manikandan, T. S. Girisun, and G. Vinitha, "Investigations on the structural, morphological, linear and third order nonlinear optical properties of manganese doped zinc selenide nanoparticles for optical limiting application," *Opt. Mater.* **100**, 109641 (2020).
- <sup>69</sup>T. Zhou, L. Fang, T. Yan, J. Wu, Y. Li, J. Fan, H. Wu, X. Lin, and Q. Dai, "In situ optical backpropagation training of diffractive optical neural networks," *Photonics Res.* **8**, 940 (2020).
- <sup>70</sup>I. U. Idehenre, E. S. Harper, and M. S. Mills, "Diffractive deep neural network adjoint assist or (DNA)<sup>2</sup>: A fast and efficient nonlinear diffractive neural network implementation," *Opt. Express* **30**, 7441 (2022).
- <sup>71</sup>M. Hermans, J. Dambre, and P. Bienstman, "Optoelectronic systems trained with backpropagation through time," *IEEE Trans. Neural Networks Learn. Syst.* **26**, 1545 (2015).
- <sup>72</sup>T. W. Hughes, M. Minkov, Y. Shi, and S. Fan, "Training of photonic neural networks through in situ backpropagation and gradient measurement," *Optica* **5**, 864 (2018).
- <sup>73</sup>Y. Li, R. Chen, B. Sensale-Rodriguez, W. Gao, and C. Yu, "Real-time multi-task diffractive deep neural networks via hardware-software co-design," *Sci. Rep.* **11**, 11013 (2021).
- <sup>74</sup>J. Su, Y. Yuan, C. Liu, and J. Li, "Multitask learning by multiwave optical diffractive network," *Math. Probl. Eng.* **2020**, 9748380.
- <sup>75</sup>Z. Duan, H. Chen, and X. Lin, "Optical multi-task learning using multi-wavelength diffractive deep neural networks," *Nanophotonics* **12**, 893 (2023).
- <sup>76</sup>J. Li, T. Gan, B. Bai, Y. Luo, M. Jarrahi, and A. Ozcan, "Massively parallel universal linear transformations using a wavelength-multiplexed diffractive optical network," *Adv. Photonics* **5**, 016003 (2023).
- <sup>77</sup>Y. Wang, A. Yu, Y. Cheng, and J. J. L. Qi, "Matrix diffractive deep neural networks merging polarization into meta-devices," *Laser Photonics Rev.* **18**, 2300903 (2024).
- <sup>78</sup>T. Zhou, X. Lin, J. Wu, Y. Chen, H. Xie, Y. Li, J. Fan, H. Wu, L. Fang, and Q. Dai, "Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit," *Nat. Photonics* **15**, 367 (2021).
- <sup>79</sup>C. Liu, Q. Ma, Z. J. Luo, Q. R. Hong, Q. Xiao, H. C. Zhang, L. Miao, W. M. Yu, Q. Cheng, and L. Li, "A programmable diffractive deep neural network based on a digital-coding metasurface array," *Nat. Electron.* **5**, 113 (2022).
- <sup>80</sup>T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light Sci. Appl.* **3**, e218 (2014).
- <sup>81</sup>T. J. Cui, S. Liu, and L. Zhang, "Information metamaterials and metasurfaces," *J. Mater. Chem. C* **5**, 3644 (2017).
- <sup>82</sup>H. Zeng, J. Fan, Y. Zhang, Y. Su, C. Qiu, and W. Gao, "Graphene plasmonic spatial light modulator for reconfigurable diffractive optical neural networks," *Opt. Express* **30**, 12712 (2022).
- <sup>83</sup>R. Chen, Y. Tang, J. Ma, and W. Gao, "Scientific computing with diffractive optical neural networks," *Adv. Intell. Syst.* **5**, 2300536 (2023).
- <sup>84</sup>S. Zarei, M. Marzban, and A. Khavasi, "Integrated photonic neural network based on silicon metalenses," *Opt. Express* **28**, 36668 (2020).
- <sup>85</sup>T. Fu, Y. Zang, Y. Huang, Z. Du, H. Huang, C. Hu, M. Chen, S. Yang, and H. Chen, "Photonic machine learning with on-chip diffractive optics," *Nat. Commun.* **14**, 70 (2023).
- <sup>86</sup>Z. Wang, L. Chang, F. Wang, T. Li, and T. Gu, "Integrated photonic metasystem for image classifications at telecommunication wavelength," *Nat. Commun.* **13**, 2131 (2022).
- <sup>87</sup>N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, "Light propagation with phase discontinuities: Generalized laws of reflection and refraction," *Science* **334**, 333 (2011).
- <sup>88</sup>S. Sun, K.-Y. Yang, C.-M. Wang, T.-K. Juan, W. T. Chen, C. Y. Liao, Q. He, S. Xiao, W.-T. Kung, and G.-Y. Guo, "High-efficiency broadband anomalous reflection by gradient meta-surfaces," *Nano Lett.* **12**, 6223 (2012).
- <sup>89</sup>Y. Hu, L. Li, R. Wang, J. Song, H. Wang, H. Duan, J. Ji, and Y. Meng, "High-speed parallel plasmonic direct-writing nanolithography using metasurface-based plasmonic lens," *Engineering* **7**, 1623 (2021).
- <sup>90</sup>Y. Hu, Y. Jiang, Y. Zhang, X. Yang, X. Ou, L. Li, X. Kong, X. Liu, C.-W. Qiu, and H. Duan, "Asymptotic dispersion engineering for ultra-broadband metasurfaces," *Nat. Commun.* **14**, 6649 (2023).
- <sup>91</sup>X. Luo, S. Dong, Z. Wei, Z. Wang, Y. Hu, H. Duan, and X. Cheng, "Full-Fourier-component tailorabile optical neural meta-transformer," *Laser Photonics Rev.* **17**, 2300272 (2023).
- <sup>92</sup>H. Yang, P. He, K. Ou, Y. Hu, Y. Jiang, X. Ou, H. Jia, Z. Xie, X. Yuan, and H. Duan, "Angular momentum holography via a minimalist metasurface for optical nested encryption," *Light Sci. Appl.* **12**, 79 (2023).
- <sup>93</sup>H. Zhang, M. Gu, X. Jiang, J. Thompson, H. Cai, S. Paesani, R. Santagati, A. Laing, Y. Zhang, and M. Yung, "An optical neural chip for implementing complex-valued neural network," *Nat. Commun.* **12**, 457 (2021).
- <sup>94</sup>X. Zhao, H. Lv, C. Chen, S. Tang, X. Liu, and Q. Qi, "On-chip reconfigurable optical neural networks," *Res. Square* (2021).
- <sup>95</sup>F. Léonard, E. J. Fuller, C. M. Teeter, and C. M. Vineyard, "High accuracy single-layer free-space diffractive neuromorphic classifiers for spatially incoherent light," *Opt. Express* **30**, 12510 (2022).
- <sup>96</sup>M. J. Filipovich, A. Malyshev, and A. Lvovsky, "Diffractive optical neural networks with arbitrary spatial coherence," in *Machine Learning with New Compute Paradigms* (2023).
- <sup>97</sup>Z. Ye, C. Zhou, C.-X. Ding, J. Zhao, S. Jiao, H.-B. Wang, and J. Xiong, "Ghost diffractive deep neural networks: optical classifications using light's second-order coherence," *Phys. Rev. Appl.* **20**, 054012 (2023).
- <sup>98</sup>I. Idehenre and M. Mills, "Multi-directional beam steering using diffractive neural networks," *Opt. Express* **28**, 25915 (2020).
- <sup>99</sup>J. Shi, D. Wei, C. Hu, M. Chen, K. Liu, J. Luo, and X. Zhang, "Robust light beam diffractive shaping based on a kind of compact all-optical neural network," *Opt. Express* **29**, 7084 (2021).
- <sup>100</sup>P. Regfregier and B. Javid, "Optical image encryption based on input plane and Fourier plane random encoding," *Opt. Lett.* **20**, 767 (1995).
- <sup>101</sup>A. Alfalou and C. Brosseau, "Optical image compression and encryption methods," *Adv. Opt. Photonics* **1**, 589 (2009).
- <sup>102</sup>Y. Gao, S. Jiao, J. Fang, T. Lei, Z. Xie, and X. Yuan, "Multiple-image encryption and hiding with an optical diffractive neural network," *Opt. Commun.* **463**, 125476 (2020).
- <sup>103</sup>X. Yang, M. S. S. Rahman, B. Bai, J. Li, and A. Ozcan, "Complex-valued universal linear transformations and image encryption using spatially incoherent diffractive networks," *Adv. Photonics Nexus* **3**, 016010 (2024).
- <sup>104</sup>O. E. DeLange, "Wide-band optical communication systems: Part II—Frequency-division multiplexing," *Proc. IEEE* **58**, 1683 (1970).
- <sup>105</sup>S. Zheng, S. Xu, and D. Fan, "Orthogonality of diffractive deep neural network," *Opt. Lett.* **47**, 1798 (2022).
- <sup>106</sup>W. Xiong, Z. Huang, P. Wang, X. Wang, Y. He, C. Wang, J. Liu, H. Ye, D. Fan, and S. Chen, "Optical diffractive deep neural network-based orbital angular momentum mode add-drop multiplexer," *Opt. Express* **29**, 36936 (2021).
- <sup>107</sup>J. Zhang, Z. Ye, J. Yin, L. Lang, and S. Jiao, "Polarized deep diffractive neural network for sorting, generation, multiplexing, and de-multiplexing of orbital angular momentum modes," *Opt. Express* **30**, 26728 (2022).
- <sup>108</sup>M. Defrise and G. T. Gullberg, "Image reconstruction," *Phys. Med. Biol.* **51**, R139 (2006).
- <sup>109</sup>R. K. Kostuk, *Holography: Principles and Applications* (CRC Press, 2019).
- <sup>110</sup>D. Liao, M. Wang, K. F. Chan, C. H. Chan, and H. Wang, "A deep-learning enabled discrete dielectric lens antenna for terahertz reconfigurable holographic imaging," *IEEE Antennas Wireless Propag. Lett.* **21**, 823 (2022).
- <sup>111</sup>Z. Huang, Y. He, P. Wang, W. Xiong, H. Wu, J. Liu, H. Ye, Y. Li, D. Fan, and S. Chen, "Orbital angular momentum deep multiplexing holography via an optical diffractive neural network," *Opt. Express* **30**, 5569 (2022).
- <sup>112</sup>L. Allen, M. W. Beijersbergen, R. Spreeuw, and J. Woerdman, "Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes," *Phys. Rev. A* **45**, 8185 (1992).

- <sup>113</sup>J. Wang, J.-Y. Yang, I. M. Fazal, N. Ahmed, Y. Yan, H. Huang, Y. Ren, Y. Yue, S. Dolinar, and M. Tur, "Terabit free-space data transmission employing orbital angular momentum multiplexing," *Nat. Photonics* **6**, 488 (2012).
- <sup>114</sup>Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in bio-medicine," *Nat. Photonics* **12**, 578 (2018).
- <sup>115</sup>Y. Li, Y. Luo, D. Mengu, B. Bai, and A. Ozcan, "Quantitative phase imaging (QPI) through random diffusers using a diffractive optical network," *Light Adv. Manuf.* **4**, 17 (2023).
- <sup>116</sup>C.-Y. Shen, J. Li, D. Mengu, and A. Ozcan, "Multispectral quantitative phase imaging using a diffractive optical network," *Adv. Intell. Syst.* **5**, 2300300 (2023).
- <sup>117</sup>L. L. Ma, C. Liu, S. B. Wu, P. Chen, Q. M. Chen, J. X. Qian, S. J. Ge, Y. H. Wu, W. Hu, and Y. Q. Lu, "Programmable self-propelling actuators enabled by a dynamic helical medium," *Sci. Adv.* **7**, eabh3505 (2021).
- <sup>118</sup>J. Cheng, S. Yu, R. Wang, and Q. Ge, "Digital light processing based multimaternal 3D printing: Challenges, solutions and perspectives," *Int. J. Extreme Manuf.* **6**, 042006 (2024).
- <sup>119</sup>M. Pan, Y. Fu, M. Zheng, H. Chen, Y. Zang, H. Duan, Q. Li, M. Qiu, and Y. Hu, "Dielectric metalens for miniaturized imaging systems: Progress and challenges," *Light Sci. Appl.* **11**, 195 (2022).
- <sup>120</sup>Y. Hu, L. Li, Y. Wang, M. Meng, L. Jin, X. Luo, Y. Chen, X. Li, S. Xiao, and H. Wang, "Trichromatic and tripolarization-channel holography with noninterleaved dielectric metasurface," *Nano Lett.* **20**, 994 (2020).
- <sup>121</sup>H. Yang, Y. Jiang, Y. Hu, K. Ou, and H. Duan, "Noninterleaved metasurface for full-polarization three-dimensional vectorial holography," *Laser Photonics Rev.* **16**, 2200351 (2022).
- <sup>122</sup>Y. Hu, X. Luo, Y. Chen, Q. Liu, X. Li, Y. Wang, N. Liu, and H. Duan, "3D-integrated metasurfaces for full-colour holography," *Light Sci. Appl.* **8**, 86 (2019).