

# Diffraction optical computing in free space

---

Received: 23 September 2022

---

Accepted: 9 February 2024

---

Published online: 20 February 2024

---

 Check for updates

---

Jingtian Hu<sup>1,2,3</sup>, Deniz Mengü<sup>1,2,3</sup>, Dimitrios C. Tzarouchis<sup>4,5</sup>, Brian Edwards<sup>4</sup>,  
Nader Engheta<sup>4</sup> & Aydogan Ozcan<sup>1,2,3</sup> ✉

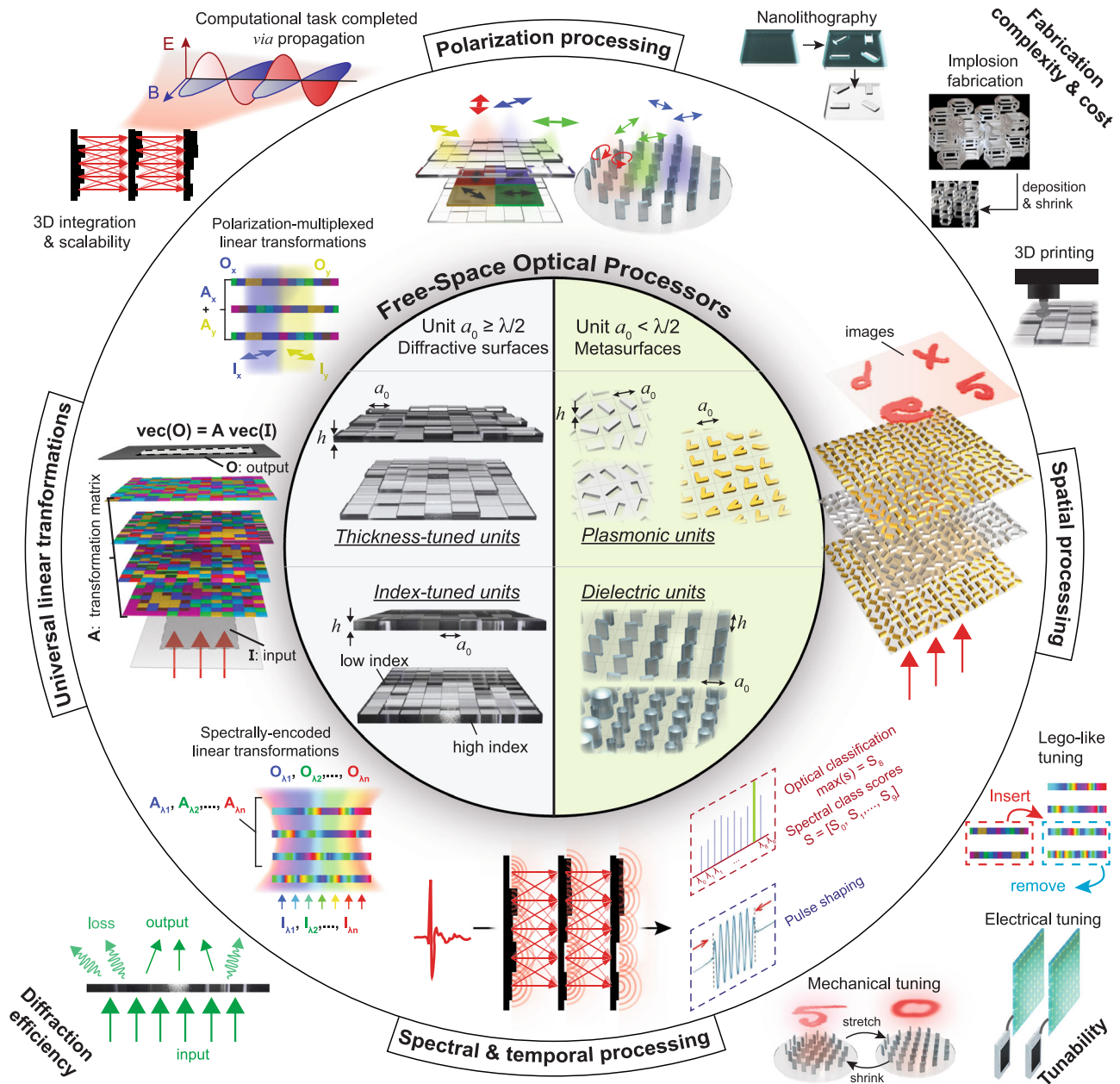
Structured optical materials create new computing paradigms using photons, with transformative impact on various fields, including machine learning, computer vision, imaging, telecommunications, and sensing. This Perspective sheds light on the potential of free-space optical systems based on engineered surfaces for advancing optical computing. Manipulating light in unprecedented ways, emerging structured surfaces enable all-optical implementation of various mathematical functions and machine learning tasks. Diffractive networks, in particular, bring deep-learning principles into the design and operation of free-space optical systems to create new functionalities. Metasurfaces consisting of deeply subwavelength units are achieving exotic optical responses that provide independent control over different properties of light and can bring major advances in computational throughput and data-transfer bandwidth of free-space optical processors. Unlike integrated photonics-based optoelectronic systems that demand preprocessed inputs, free-space optical processors have direct access to all the optical degrees of freedom that carry information about an input scene/object without needing digital recovery or preprocessing of information. To realize the full potential of free-space optical computing architectures, diffractive surfaces and metasurfaces need to advance symbiotically and co-evolve in their designs, 3D fabrication/integration, cascading, and computing accuracy to serve the needs of next-generation machine vision, computational imaging, mathematical computing, and telecommunication technologies.

The past decade has witnessed an increasing interest in optical computing platforms because of their potential to realize fast, massively parallel computation at low power consumption<sup>1–4</sup>. For example, integrated photonics-based processors<sup>5–9</sup> with on-chip interferometers and waveguide-embedded light sources were designed to replace or empower their electronic counterparts. These planar optoelectronic architectures can implement various linear and non-linear operations and can rapidly self-tune and reconfigure, making them powerful and versatile for optoelectronic computing<sup>10,11</sup>. Therefore, integrated photonics-based processors have advantages in their reconfigurability and ease of integration with electronics, making them ideal for task-specific on-chip accelerators. However, integrated

photonics devices often require preprocessing of input information, for example, to retrieve the lost phase information of an input scene due to intensity-only detection at the optoelectronic sensor array or simply to vectorize the multi-dimensional image of an input object into planar signals. In contrast, free-space optical processors (Fig. 1) can perform a given computation or inference task via light propagation and diffraction through a series of structured surfaces. Without any optoelectronic conversion or preprocessing of input information, diffractive optical processors perform computational tasks with direct access to all the encoded optical information in 3D, including the spatial phase and amplitude, polarization, spectrum, and orbital angular momentum (OAM) information of the input wave, which

---

<sup>1</sup>Electrical and Computer Engineering Department, University of California, Los Angeles, CA 90095, USA. <sup>2</sup>Bioengineering Department, University of California, Los Angeles, CA 90095, USA. <sup>3</sup>California NanoSystems Institute (CNSI), University of California, Los Angeles, CA 90095, USA. <sup>4</sup>Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA. <sup>5</sup>Meta Materials Inc., Athens 15123, Greece. ✉e-mail: [ozcan@g.ucla.edu](mailto:ozcan@g.ucla.edu)



**Fig. 1 | Overview of diffractive and metasurface systems for optical computing.** Diffractive surfaces consist of thickness- and/or index-tuned units ( $\lambda/2$  or larger) while metasurfaces consist of substructured metallic and/or dielectric units

below  $\lambda/2$ . These free-space processors can perform polarization processing, spatial processing, universal linear transformations, and spectral & temporal processing of waves.

represents an object or a scene. This direct access to all the degrees of freedom carried by the input waves provides significant advantages to free-space optical processors for various applications such as all-optical statistical inference, wireless telecommunications, and computational imaging and sensing<sup>12</sup>, where the information is created/represented and transmitted through electromagnetic waves. Therefore, free-space optical processors offer advantages in visual information processing and related applications due to their direct access to 2D or 3D optical information. Driven by some of these fundamental advantages, free-space optical computing systems (Fig. 1) have undergone a paradigm shift in their designs facilitated by emerging deep-learning methods and unique fabrication technologies.

This Perspective focuses on data-driven design and fabrication of (1) diffractive surfaces that are structured at the wavelength scale (with a unit size of  $\lambda/2$  or larger), and (2) metasurfaces with deeply subwavelength dielectric or metallic units (smaller than  $\lambda/2$ ).

Considering only the propagating electromagnetic waves and ignoring the evanescent fields, the lateral footprint of photons in free-space is  $\sim \lambda/2$ ; the optical processor design spaces covered by diffractive surfaces and metasurfaces intersect at this  $\lambda/2$  boundary, giving us a plethora of transformative opportunities to control and process light at different scales. For example, diffractive optical processors constructed based on a series of densely-packed surfaces engineered through deep-learning can achieve universal linear transformations within a compact diffractive volume by approximating any arbitrary complex-valued matrix operation with an arbitrarily small error bound<sup>13,14</sup>; such diffractive processors also enable all-optical image classification using light diffraction through structured surfaces. In addition to directly processing the spatial phase and amplitude information of an input scene, diffractive optical processors can also be trained using deep-learning to exploit other information channels, including polarization<sup>15</sup> and spectrum<sup>16–18</sup>, massively increasing the

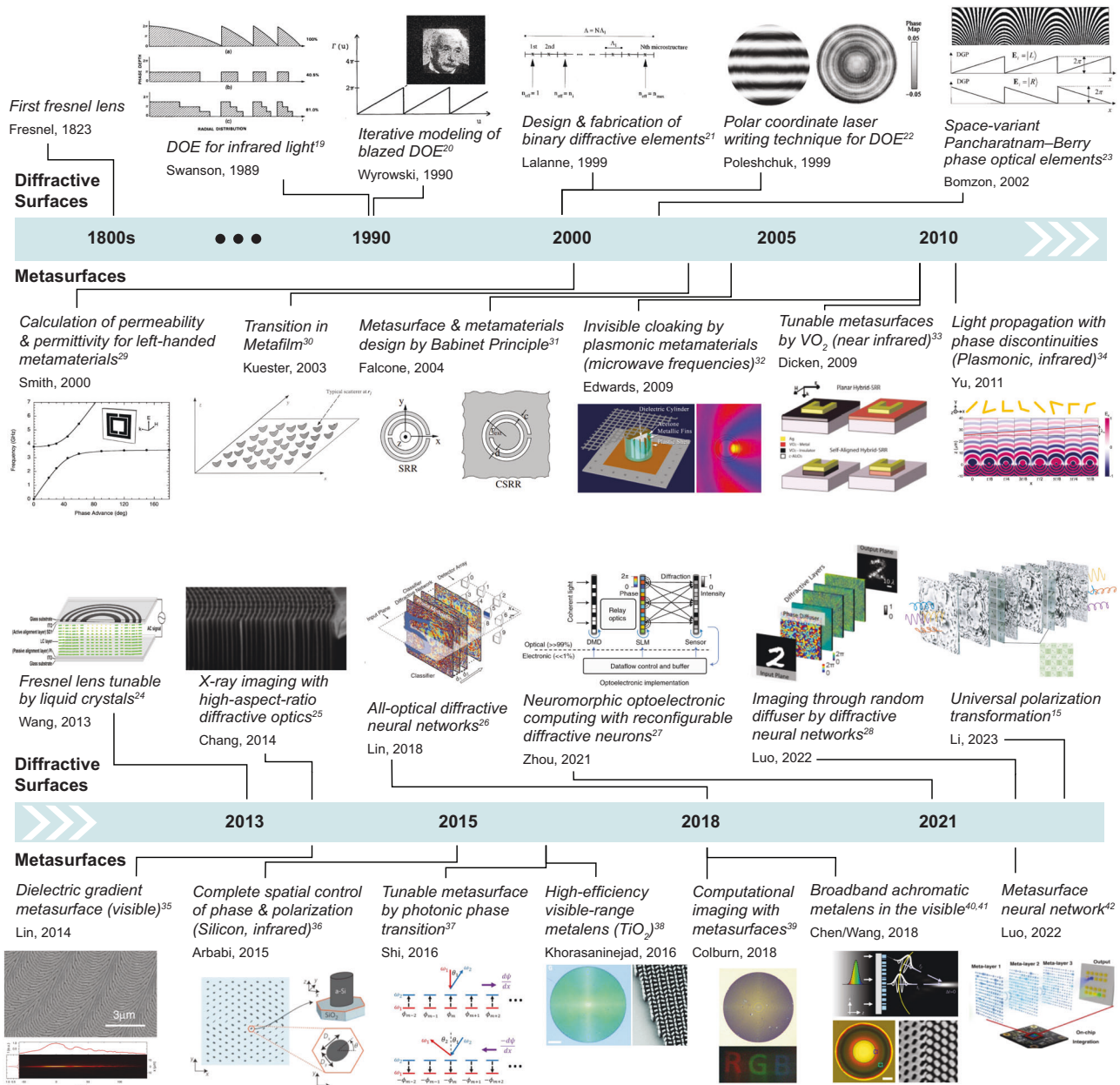
parallelism and bandwidth of free-space optical computing. Replacing or augmenting dielectric-based diffractive surfaces with metasurfaces can provide some additional degrees of freedom, including the engineering of dispersion, polarization, spin, and OAM, to further enhance the inference capacity and the computational power of free-space optical processors.

Both diffractive surfaces and metasurfaces have received major research interest in the past decades (Fig. 2)<sup>15,19–42</sup>. Despite this tremendous progress, the synergy between the two frameworks in the context of free-space optical computing has not been explored in depth. This Perspective focuses on this emerging opportunity for diffractive surfaces and metasurfaces to symbiotically shape the future designs of free-space optical computing devices and technologies that will impact various applications such as optical machine learning, statistical inference, computational camera and microscope design, and telecommunications, among many others. As part of this

Perspective’s comparative analysis of free-space diffractive computing (Fig. 1), we will also discuss four grand challenges ahead of this field of research:

- (1) Limits of computational accuracy and statistical inference capability of diffractive optical processors.
- (2) Dynamic reconfigurability for on-demand tuning of the optical processor function.
- (3) Speed, ease, and scalability of diffractive processor device fabrication, alignment, and integration, important for cost-effective and large-scale manufacturing and adoption of free-space optical processors.
- (4) Diffraction efficiency and energy consumption of the optical processor.

This Perspective focuses on these emerging opportunities and challenges centered around optical computing in free-space. We expect the comparatively analyzed methodologies in this Perspective



**Fig. 2 | Timeline of diffractive<sup>15,19–28</sup> and metasurface<sup>29–42</sup> based optical devices and systems.** Images are adapted from refs. 26,39,91 with permissions from AAAS, refs. 20–24,33,95,162 with permissions from © The Optical Society,

refs. 25,27,36,40,41 with permissions from Springer Nature, refs. 28,42 under CC BY 4.0, ref. 29 with permissions from AIP publishing, and refs. 31,32,37 with permissions from APS.

to profoundly impact the next-generation wave-based free-space computing technologies ranging from all-optical statistical inference and computational imaging to wireless communications, edge computing, and others.

## Design principles of spatially-structured surfaces for computing

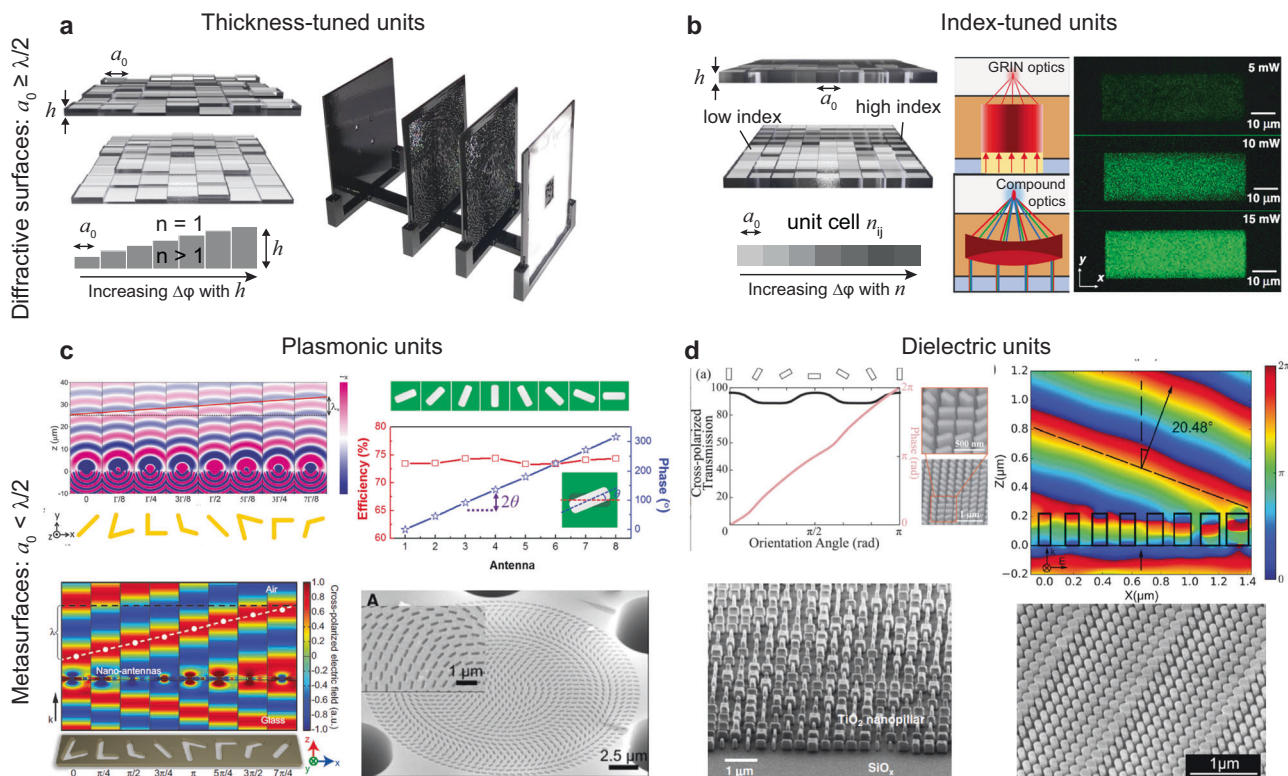
### Diffractive surfaces

A diffractive surface-based free-space optical processor is spatially engineered at the wavelength scale ( $\lambda/2$  or larger) to generate a desired spatial distribution of the complex-valued light transmission/reflection coefficients (Fig. 3a, b, Supplementary Fig. S1a). Unlike standard refractive optical components, e.g., lenses, such diffractive surfaces are pixelated in 2D, and their structure contains physical thickness and/or transmission discontinuities with a lateral period of  $\geq \lambda/2$ ; also see Supplementary Discussion 1. The structure of these variations along the optical axis is typically on the order of the wavelength of the propagating light, and therefore, they can, in general, be modeled as 2D thin optical elements. Accordingly, the relationship between the incident monochromatic complex light field,  $U_{in}(x,y)$ , and the outgoing wavefront,  $U_{out}(x,y)$ , transmitted by a diffractive surface can be written as  $U_{out}(x,y) = T(x,y)U_{in}(x,y)$  with  $T(x,y)$  denoting the 2D complex-valued transmission function of the diffractive surface at  $\lambda$ . The mathematical form of  $T(x,y)$  depends on the physical mechanism of the optical light modulation and the related fabrication techniques.

In one simple form, light modulation can be achieved via amplitude modulation by partially reflecting and/or absorbing the light incident on certain pixels while allowing transmission on others. In this case, the modulation function  $T(x,y)$  is real-valued, for example, binary (0 vs. 1 transmission). While some early research on diffractive optics heavily relied on such amplitude-only surface designs, they present

limited diffraction efficiencies. Another commonly used diffractive surface design approach relies on the difference between the refractive indices of the fabrication material (substrate),  $n_{surface}$  and the light propagation medium,  $n_{medium}$ , e.g.,  $n_{medium} = 1$  in air. In general, the index contrast of a diffractive feature (with a height of  $h$ ) with respect to the background induces an optical path delay in the form of  $\Delta\varphi = \frac{2\pi h}{\lambda}(n_{surface} - n_{medium})$ . Typically, diffractive surfaces fabricated using a single dielectric material have a constant  $n_{surface}$  over their pixels, which requires the thickness of the fabrication material to be a function of space, i.e.,  $h(x,y)$ , leading to a 2D transmission function,  $T(x,y) = \exp(j\frac{2\pi h(x,y)}{\lambda}(n_{surface} - n_{medium}))$  (Fig. 3a). With the wide availability of additive manufacturing and nanofabrication techniques<sup>43–46</sup>, this approach has become one of the most commonly employed methods in implementing diffractive surfaces. An alternative method relies on the spatial engineering of the refractive index, i.e.,  $n_{surface}(x,y)$ , while keeping  $h$  constant (shown in Fig. 3b)<sup>47,48</sup>. This approach leads to a 2D transmission function in the form of  $T(x,y) = \exp(j\frac{2\pi h}{\lambda}(n_{surface}(x,y) - n_{medium}))$ .

When the optical attenuation induced by the diffractive material is not negligible, i.e.,  $n_{surface}(x,y) = \kappa(x,y) + j\tau(x,y)$  is complex-valued, then  $T(x,y)$  applies both an amplitude and a phase modulation at every unit cell of its design. For instance, assuming  $\kappa(x,y) = \kappa$ ,  $\tau(x,y) = \tau$  and  $n_{medium} = 1$ ,  $T(x,y)$  can be written as  $T(x,y) = \exp(j\frac{2\pi h(x,y)}{\lambda}(n_{surface} - 1)) = \exp(-\frac{2\pi h(x,y)}{\lambda}(\tau)) \exp(j\frac{2\pi h(x,y)}{\lambda}(\kappa - 1))$ . This strategy for the modulation of the phase and amplitude of each diffractive unit cell, however, does *not* tailor the two variables independently because they are both controlled by  $h(x,y)$ . Independent control over both the amplitude and phase channels of each transmission function might be realized, for example, by spatial engineering of the material thickness and the complex-valued refractive index at the wavelength scale, covering thousands of individual unit cells over a diffractive surface. In fact, this is



**Fig. 3 | Design of diffractive surfaces and metasurfaces.** Diffractive unit cell designs and their principles of wavefront modulation based on (a) thickness-tuning<sup>26</sup> and (b) index-tuning<sup>47,48</sup>. Metasurface unit cell designs using (c) plasmonic<sup>34,57,112</sup> and (d) dielectric materials<sup>42,61,62</sup>. **b** This is adapted with permission

from ref. 48 by CC BY 4.0. **c** This is adapted with permission from ref. 34 by AAAS, refs. 34,57 by CC BY 4.0 and ref. 112 by ACS. **d** This is adapted with permission from ref. 42 by CC BY 4.0, ref. 62 by John Wiley and Sons and ref. 61 by ACS.

an area where metasurfaces can offer some additional degrees of freedom, which will be discussed next.

### Metasurfaces

Optical metasurfaces<sup>23,30,38,49–52</sup> are planar structured surfaces consisting of unit cells with subwavelength structures that tailor the optical wavefront by their material properties, size, and shape<sup>53,54</sup>. Based on the target wavelength range of interest, the meta-units can be made of plasmonic metals (i.e., Au, Ag, and Al) or dielectric materials (i.e., Si, GaN, and TiO<sub>2</sub>). Typically, Al and Ag are the choices of plasmonic materials for ultraviolet and visible wavelengths where Au exhibits large non-radiative losses due to interband transitions<sup>55,56</sup>. Au exhibits lower losses at longer wavelengths ranging from the near-infrared to the terahertz range<sup>55</sup>. In the visible and near-infrared parts of the spectrum, plasmonic meta-units are typically formed by metallic nanoparticles (NPs) with various sizes and shapes and their inverted nanohole structures based on Babinet's principle (Fig. 3c, Supplementary Fig. S1b)<sup>31</sup>. V-shaped meta-units tailor the wavefront by simultaneously exciting a symmetric resonant mode along the diagonal axis and an antisymmetric mode. The scattered light from each meta-unit exhibits a field amplitude and phase determined by the length  $l$  of each rod and the angle  $\Delta$  between them (Fig. 3c, left)<sup>34</sup>. In contrast, a rectangular plasmonic NP<sup>34</sup> or nanohole<sup>57</sup> oriented at an arbitrary angle  $\theta$  can manipulate the scattered phase by the Pancharatnam–Berry (P-B) principle<sup>58</sup>, where the incident light with left- or right-hand circular polarization is converted to its orthogonal circular polarization state with an additional phase delay of  $2\theta$  on the transmitted beam (Fig. 3c, right). For operation in the gigahertz to terahertz range, other meta-units with more sophisticated geometries, such as C- and H-shaped elements, were also used to achieve higher cross-polarization conversion efficiencies<sup>59</sup>. Compared to dielectric materials, metallic meta-units, in general, allow an easier fabrication process based on lithography and lift-off processes<sup>53</sup>. One of the main limitations of plasmonic materials for building metasurface-based free-space optical processors is that their large losses can degrade the output diffraction efficiency drastically, also limiting the cascading of such surfaces to enhance their function approximation power.

Partially motivated to solve this limitation, dielectric metasurfaces were developed to achieve better diffraction efficiencies in the near-infrared and visible wavelengths<sup>38,60</sup>. Typically, high aspect ratio meta-units of TiO<sub>2</sub><sup>38,42</sup> and GaN<sup>40</sup> are used in the visible range and Si meta-units<sup>61</sup> are used for near-infrared applications. Similar to their plasmonic counterparts, dielectric meta-units with rectangular or elliptical shapes can tailor the wavefront phase of circularly-polarized input light by tuning their in-plane orientations (Fig. 3d, left)<sup>36,62</sup>. This polarization selectivity of dielectric metasurfaces is also an effective tool for manipulating the spin states of light and has enabled spin to OAM conversion with high efficiency<sup>63</sup>. Polarization-insensitive phase responses can be achieved by using high-aspect-ratio dielectric posts by adjusting their diameters (Fig. 3d, right). To cover the  $0-2\pi$  phase range by diameter tuning, the dielectric phase elements need to be tall, typically larger than 400 nm<sup>64</sup>. The processing of dielectric metasurfaces generally relies on conformal deposition techniques such as atomic layer deposition (ALD)<sup>65</sup> and chemical vapor deposition (CVD)<sup>36</sup>, which are slow, expensive, and incompatible with lift-off-based fabrication processes. Instead, anisotropic dry etching steps are often required to produce the high aspect ratio nanoscale meta-units. These fabrication challenges and others will be discussed in Section “Fabrication complexity and 3D alignment requirements”.

To provide a comprehensive overview of existing free-space optical devices, we summarized in Table 1 the material compositions and fabrication methods of diffractive surfaces and metasurfaces along with their optical properties, including efficiency, spectral bandwidth, reconfigurability, and polarization responsivity. These performance metrics are also critical for designing free-space optical

computing systems and are often the determining factors for practical applications. Strategies to tackle the existing challenges and future opportunities to further improve diffractive optical processors and metasurfaces will be discussed in Section “Grand Challenges in Free-space Optical Computing and Performance Limitations”.

### Computing capabilities of free-space optical processors using structured surfaces

#### Statistical inference and data classification

Deep-learning enabled optical networks of diffractive<sup>26,66,67</sup> and meta-units<sup>42,62</sup> over a series of passive surfaces have emerged as all-optical machine learning platforms demonstrating promising capabilities in various visual inference tasks, e.g., all-optical object detection and classification referred to as diffractive deep neural networks (D<sup>2</sup>NN). Similar to electronic neural networks, diffractive optical networks also serve as parameterized function approximators. However, they process analog optical waves of a scene propagating through structured/engineered surfaces to perform statistical inference, instead of digitized signals inside a computer. While the trainable parameters of this optical diffractive black-box are the transmittance (and/or reflectance) coefficients,  $\alpha_i \exp(j\beta_i)$  of the diffractive unit cells in a predetermined coordinate,  $(x_i, y_i, z_i)$ , within the computing volume, the complex-valued weights of the connections between these diffractive features are not individually trainable, directly dictated by the light diffraction and the axial spacing between the diffractive layers. Currently, the optimization of diffractive optical networks for a given inference task uses deep-learning-based training implemented in a digital computer. The transmittance coefficients of the diffractive units within the free-space optical processor are updated using the associated gradients with respect to a penalty term (training loss function) that is specifically devised for the targeted machine learning task and the associated detector configuration at the output plane. Following the training phase, the resulting diffractive surfaces are fabricated and assembled to form the physical diffractive optical processor, which performs the target statistical inference task without any external computing power, except for the illumination light.

The early experimental demonstrations of diffractive optical networks were performed based on 3D-printed diffractive surfaces operating at THz wavelengths with e.g., 0.2 million phase-only diffractive features distributed over five dielectric diffractive surfaces<sup>26</sup>. Since these initial proof-of-concept demonstrations, various design advances have been reported leading to >98% and >90% all-optical blind classification accuracies on benchmark datasets of handwritten digits (MNIST) and fashion products (Fashion-MNIST), respectively<sup>66,68</sup>. The inference and generalization capabilities of diffractive classifiers can further be improved by exploiting machine learning methods that evoke collaboration among multiple classification networks. Optical classification systems relying on multiple class-specific diffractive optical networks exemplify such a design strategy<sup>68</sup>. As an alternative to these jointly-optimized class-specific diffractive network systems, one can also use ensemble learning techniques<sup>69,70</sup> to advance the optical inference performance of diffractive processors; these approaches will be further discussed in Section “Grand Challenges in Free-space Optical Computing and Performance Limitations”.

Beyond these all-optical image classifiers comprised of dielectric diffractive surfaces, it is also possible to extend the D<sup>2</sup>NN framework to design metasurface-enabled diffractive optical image classification networks<sup>42</sup>. An on-chip metasurface-based diffractive object classifier operating at visible wavelengths was also demonstrated<sup>42</sup>. With meta-atoms containing TiO<sub>2</sub> nanopillars placed on top of a SiO<sub>2</sub> substrate, the physical (trainable) parameters of this diffractive network were set as the width of the TiO<sub>2</sub> nanopillars in  $x$  and  $y$  directions, which controls the effective refractive index seen by the  $x$ - and  $y$ -polarized light waves. The experimentally demonstrated system was restricted to a single metasurface with 280×280 meta-atoms, optically classifying 4

**Table 1 | Survey of various diffractive surface- and metasurface-enabled designs**

Operation Wavelength/Task		Diffractive Surfaces		Metasurfaces		
		Thickness-tuned units	Index-tuned units	Plasmonic/Metallic	Dielectric	
Fabrication Methods	GHz-THz	3D printing (Stereo-lithography, Polyjet, etc.)		Photolithography <sup>191</sup> or Laser writing <sup>192</sup>	Photolithography <sup>193</sup>	
	Visible-NIR	Nanoscribe <sup>136</sup> , etc.		EBL <sup>38</sup> or DUV <sup>194</sup>	EBL/DUV/Nanoscribe <sup>195</sup>	
Materials	GHz-THz	Resin (Veroblack <sup>26</sup> , etc.)		Au, IST <sup>92</sup> , etc.	Si <sup>197</sup> , etc.	
	Visible-NIR	Resin (IP-162 <sup>136</sup> , etc.)		Hydrogel <sup>137</sup> or SiO <sub>2</sub> -TiO <sub>2</sub> <sup>198</sup>	TiO <sub>2</sub> <sup>38</sup> , GaN <sup>40</sup> , Si <sup>36</sup> , etc.	
Power Efficiency	GHz-THz	Single wavelength	26% (classification) <sup>196</sup>	31-72% (lens) <sup>199</sup>	50% (beam steering) <sup>200</sup>	68% (lens) <sup>197</sup>
		Broadband	76-91% (unidirectional imager) <sup>201</sup>	N/A	N/A	30-68% (lens) <sup>202</sup>
	Visible-NIR	Single wavelength	99% <sup>203</sup>	75% (lens) <sup>204</sup>	80% (reflection, holography) <sup>98</sup>	86% (lens) <sup>38</sup>
		Broadband	95% (lens) <sup>205</sup>	N/A	8.4-11% <sup>206</sup>	40% <sup>40</sup>
Reconfigurability	GHz-THz	Layer swapping <sup>18</sup>		Electrically-tuned <sup>121,207</sup> (single-unit-level 0-2π phase programmability)	Optical pumping tuned by DMD <sup>208</sup>	
	Visible-NIR	Spatial light modulator <sup>27</sup>		Substrate stretching <sup>103</sup>	1D liquid crystal supercells <sup>100</sup>	
Spectral Bandwidth	GHz-THz	Lens	0.3-1.5 THz <sup>209</sup>	0.4-0.6 THz <sup>210</sup>	4.2-4.5 THz <sup>211</sup>	0.3-0.8 THz <sup>202</sup>
		Imaging, Beam steering, Linear transformations	0.37-0.44 THz (unidirectional imaging) <sup>201</sup>	N/A	0.4-1 THz (beam steering) <sup>212</sup>	0.6-1 THz (beam steering) <sup>193</sup>
			0.34-0.41 THz (multi-channel linear transformation) <sup>17</sup>			
	Visible-NIR	Lens	425-700 nm <sup>205</sup>	488-1680 nm <sup>213</sup>	1200-1680 nm <sup>206</sup>	400-660 nm <sup>40</sup>
Holography		532-1550 nm <sup>214</sup>	N/A	630-1050 nm <sup>98</sup>	1.1-1.4 μm <sup>215</sup>	
Beam steering		450-950 nm <sup>216</sup>	N/A	480-660 nm <sup>217</sup>	1.4-1.8 μm <sup>218</sup>	
Polarization Responses	GHz-THz	Multi-pixel universal polarization transformations <sup>15</sup>		Single-pixel polarization conversion <sup>219</sup>	N/A	
	Visible-NIR	N/A		Versatile single-pixel polarization generation <sup>112</sup>	Subwavelength single-pixel polarization conversion <sup>36</sup> Color holography <sup>220</sup>	

data classes from each of MNIST and Fashion-MNIST datasets, encoded into two orthogonal polarization states<sup>42</sup>, with deeper architectures involving more metasurfaces, one following another, further advances in the inference accuracy spanning a larger number of data classes can be achieved.

The function approximation capabilities of free-space optical processors are not limited to statistical inference or classification tasks and can be extended to all-optically performing other general-purpose computational tasks, including e.g., logic operations and diffractive NAND gates that can be optically-cascaded<sup>71,72</sup>.

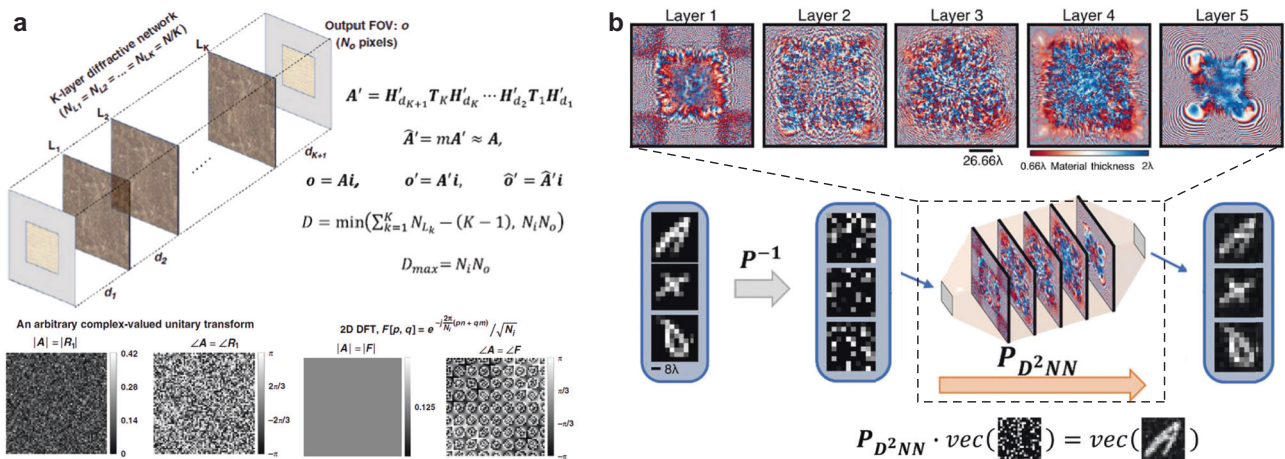
In the examples discussed so far, the diffractive optical networks, whether comprised of dielectric diffractive layers or metasurfaces, employed linear materials and lacked any form of nonlinearity except for the detector plane. In general, the statistical inference accuracy and function approximation capability of diffractive processors would benefit from power efficient and scalable integration of nonlinear optical processes as part of the free-space diffractive processor volume (see e.g., ref. 73), and this grand challenge will be discussed under Section “Grand Challenges in Free-space Optical Computing and Performance Limitations”.

### Universal linear transformations

Optical networks have been broadly used to perform matrix operations, e.g.,  $y = Ax$ , where A is an arbitrary complex-valued matrix to be approximated by the optical network. Such universal linear transformations have been realized by integrated photonic circuits consisting of waveguides connected with e.g., Mach-Zehnder interferometers<sup>5,8</sup> as well as by free-space diffractive networks<sup>13,14</sup>. An integrated photonic

waveguide-based crossbar array also realized universal matrix multiplications, achieving trillions of multiply-accumulate operations per second<sup>74</sup>. One difference between these two main-stream approaches (integrated photonic circuits vs. free-space) is that the latter can directly act on the 2D or 3D optical information of an object, without the need to vectorize or pre-process the input optical information, and therefore provides a better fit for direct analog processing of visual information contained in a scene.

The all-optical transformation by a diffractive free-space processor can be modeled as a complex-valued linear matrix multiplication operation between the input and output wavefront fields. It has been shown that diffractive optical processors are universal analog computing platforms for approximating arbitrary linear transformations<sup>13,14,75,76</sup> which form the fundamental building blocks in a plethora of computational applications from communications<sup>77,78</sup> and image processing<sup>79,80</sup> to machine learning<sup>81</sup> and beyond. The universality of deep-learning-based diffractive processor designs in all-optical implementations of complex-valued linear transformations has been demonstrated through numerical studies conducted on various types of linear operators including, (1) arbitrarily chosen complex-valued unitary, nonunitary, and noninvertible linear matrices (Fig. 4a)<sup>13</sup>, (2) 2D discrete Fourier transformation, (3) arbitrary 2D permutation operations (Fig. 4b)<sup>82</sup>, and (4) high-pass filtered coherent imaging<sup>13</sup>. While these results involved spatially and temporally coherent illumination, free-space-based diffractive processors were also designed to implement arbitrary linear transformations in optical intensity using spatially incoherent light<sup>83</sup>; furthermore, wavelength-multiplexed and massively parallel universal linear transformations



**Fig. 4 | Diffraction optical networks can perform universal linear transformations.** Diffraction surfaces trained to perform (a) linear operations including an arbitrary complex-valued transform, discrete Fourier transform (DFT)<sup>13</sup> and (b) permutation operation<sup>82</sup>. Extensions of this framework to perform universal linear

transformations under spatially and/or temporally incoherent illumination were also demonstrated<sup>17,83</sup>. a, b These are adapted with permission from ref. 13 and ref. 82, respectively, by CC BY 4.0.

were also demonstrated using deep-learning-designed diffractive processors, simultaneously covering hundreds of wavelength channels<sup>17</sup>.

In general, deeper diffractive architectures where a given number of diffractive degrees of freedom ( $N$ ) is distributed across one diffractive surface following another present better optical approximation accuracies with improved output diffraction efficiencies compared to shallower architectures where the same  $N$  degrees of freedom are distributed to a smaller number of diffractive layers. Before the demonstration of this depth feature for universal linear transformations<sup>75</sup>, its first empirical evidence was presented for image classification tasks<sup>26</sup>, where deeper diffractive processors generalized to classify unknown object images better than shallower diffractive architectures.

An experimental demonstration of all-optical linear transformations e.g., an arbitrary permutation matrix with 625 input-output connections, was reported based on diffractive free-space processors comprised of  $K = 3$  dielectric modulation surfaces operating at THz wavelengths<sup>82</sup>. Deep learning-designed diffractive networks were also used to experimentally realize data class-specific all-optical transformations and image encryption<sup>84</sup>. In these diffractive designs, the visual information that is encoded into the intensity of the optical field was all-optically encrypted by a data class-specific diffractive network that implemented a separate transformation matrix for each data class—i.e., a different linear transformation for each distinct class of objects, performed by the same diffractive optical processor<sup>84</sup>. This all-optical data-class-specific transformation system was experimentally demonstrated in different parts of the electromagnetic spectrum, i.e., 1550 nm and 0.75 mm wavelengths, and can enable fast, energy-efficient data encryption using a broad range of optical sources.

As another demonstration in the visible spectrum, a single-layer Si metasurface that could represent a linear complex-valued  $S$ -matrix was reported for solving the Fredholm integral equations<sup>85</sup>. This meta-grating structure can approximate a Neumann series based on successive reflection between the metasurface and a semitransparent mirror. This analog computing platform that operates in the visible spectrum enables a highly compact and thin device that can achieve high processing speeds, with the possibility of on-chip integration.

### Optical processing of spatial information and computational imaging

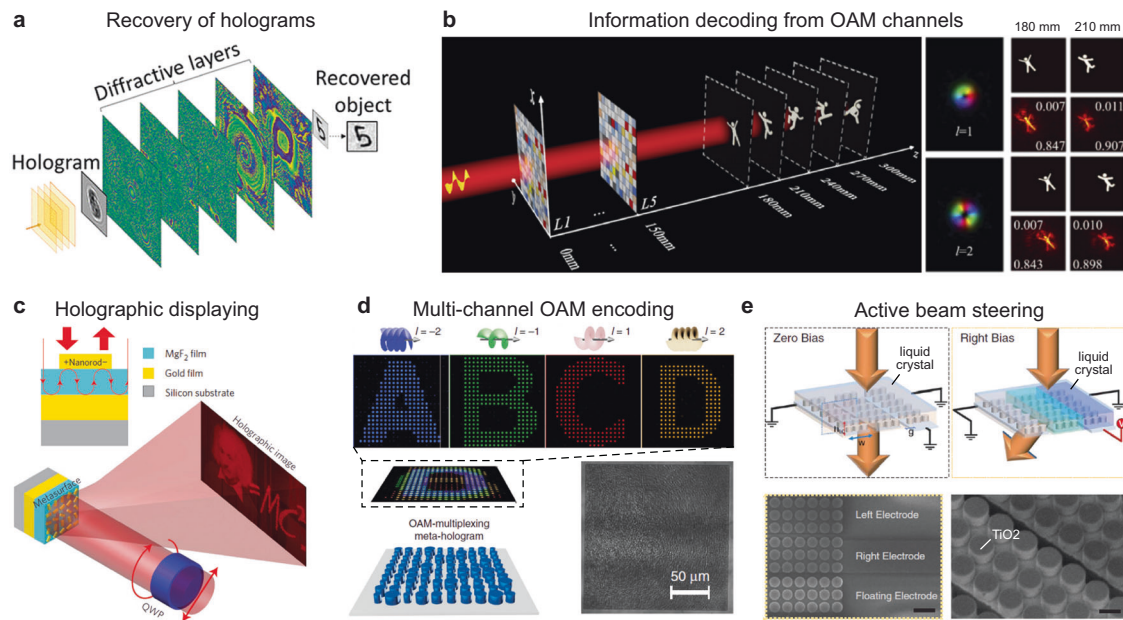
Diffractive networks that process spatially encoded information of light present some unique opportunities for computational imaging and display systems. For example, imaging through unknown, random

scattering, and diffusive media is a challenging problem of major significance for various fields ranging from biomedical imaging<sup>86</sup> to autonomous robotic systems<sup>87–89</sup>. Diffractive networks provide a reliable all-optical solution to this challenge of imaging through unknown random diffusers by learning from the scattering process of images propagating through a large number of diffusers in a deep-learning-based training process<sup>28</sup>. Remarkably, this technique shows strong generalizability to unknown diffusers with a comparable correlation length to the diffusers used during the training. Diffractive surfaces are also capable of performing selective imaging, where only the desired classes of objects are recorded while other information is all-optically erased by the diffractive system<sup>90</sup>.

Diffractive networks also provide a powerful tool to achieve pixel super-resolution that can contribute to the development of high-resolution displays using spatial light modulators (SLMs) with limited pixel density and space-bandwidth products (SBP)<sup>91</sup>. The resolution of the displayed images can be enhanced up to 4-fold by (1) encoding the high-resolution images into low-resolution SLM patterns using a convolutional neural network (CNN) and (2) decoding the low-resolution images into the high-resolution images by a jointly-trained diffractive network<sup>91</sup>. This all-optical pixel super-resolution system can surpass the SBP restrictions of the wavefront modulator while maintaining its image field-of-view. The method can also significantly reduce the data transmission and storage burden by encoding the high-resolution input images into compact, low-resolution representations.

The ability of diffractive networks to process spatially encoded optical information can also enable hologram reconstruction and all-optical recovery of quantitative phase information (Fig. 5a)<sup>67,92</sup>. Quantitative phase imaging (QPI) is a powerful label-free imaging method for imaging low-contrast samples with weak scattering by visualizing the optical path length information of specimens<sup>93</sup>. By converting the input phase information of a scene into intensity variations at the output plane, diffractive networks have achieved QPI in a compact diffractive design that axially spans  $\sim 200\text{--}300 \lambda$ <sup>93</sup>. In contrast to digital approaches that retrieve holographic images from captured intensity images using digital phase recovery algorithms, these passive free-space processors enable all-optical reconstruction of holographic images, achieving superb image quality and high diffraction efficiency that can be generalized to classes of objects different from the training datasets<sup>92,93</sup>.

Diffractive networks also allow the separation of spatially multiplexed information encoded in different OAM channels that profoundly impact future telecommunication technologies<sup>94</sup>. Upon illumination by



**Fig. 5 | Diffraction networks and metasurfaces for all-optical spatial processing of light.** All optical image recovery by diffractive networks from (a) computer-generated holograms<sup>92</sup> and (b) OAM information channels<sup>95</sup>. Metasurface-based processors for manipulating OAM with polarization and phase modulation have also been demonstrated<sup>96</sup>. Although existing demonstrations focus largely on terahertz waves and longer wavelengths, expanding to the visible ranges is expected with the developments in nanofabrication methods. We expect this technology to enable advances in holography and storage with larger scalability and information density. Metasurface spatial processors have been demonstrated as single-layer devices, for example, to generate holograms based on computer-generated phase and amplitude profiles (Fig. 5c)<sup>97,98</sup>. By sampling with a 2D Dirac comb function in the spatial frequency-domain, a metasurface allows the multiplexing of multiple holographic images in different OAM channels by the same design (Fig. 5d)<sup>99</sup>. With integration with actively tunable optical materials, metasurfaces also allow beam steering and focusing of free-space light tunable by external stimuli such as mechanical deformation and electrical biases<sup>100–103</sup>. In particular, a one-dimensional SLM has been realized by liquid crystals (Fig. 5e)<sup>100</sup>. The ability to reconfigure metasurfaces has contributed to adaptive holograms where the image can be tailored by chemical<sup>104</sup> and mechanical stimuli<sup>105</sup>. Some of these wavefront modulation and beam steering applications shown in, e.g., Fig. 5c–e are relatively limited in their computational capabilities since they perform a restricted form of wave transformation with limited input-output representations as opposed to implementing a function that can accurately transform infinitely many optical inputs to the desired representations at the output aperture (such as, e.g., Fig. 5a). With further advances in 3D nanofabrication techniques, we expect the realization of multi-layer metasurface networks to enable powerful adaptive free-space spatial processing and function representation capabilities critical for next-generation information processing and VR/AR applications.

adapted with permission from ref. 92 by ACS. **b** This is adapted with permission from ref. 95 by © The Optical Society. **c** This is adapted with permission from ref. 98 by Springer Nature. **d** This is adapted with permission from ref. 99 by Springer Nature. **e** This is adapted with permission from ref. 100 by AAAS.

vortex beams that carry OAM, the engineered surfaces produce a selective set of images at designed spatial positions tailored by the orbital quantum number ( $l$ ) (Fig. 5b)<sup>95</sup>. Metasurface-based processors for manipulating OAM with polarization and phase modulation have also been demonstrated<sup>96</sup>. Although existing demonstrations focus largely on terahertz waves and longer wavelengths, expanding to the visible ranges is expected with the developments in nanofabrication methods. We expect this technology to enable advances in holography and storage with larger scalability and information density.

Metasurface spatial processors have been demonstrated as single-layer devices, for example, to generate holograms based on computer-generated phase and amplitude profiles (Fig. 5c)<sup>97,98</sup>. By sampling with a 2D Dirac comb function in the spatial frequency-domain, a metasurface allows the multiplexing of multiple holographic images in different OAM channels by the same design (Fig. 5d)<sup>99</sup>. With integration with actively tunable optical materials, metasurfaces also allow beam steering and focusing of free-space light tunable by external stimuli such as mechanical deformation and electrical biases<sup>100–103</sup>. In particular, a one-dimensional SLM has been realized by liquid crystals (Fig. 5e)<sup>100</sup>. The ability to reconfigure metasurfaces has contributed to adaptive holograms where the image can be tailored by chemical<sup>104</sup> and mechanical stimuli<sup>105</sup>. Some of these wavefront modulation and beam steering applications shown in, e.g., Fig. 5c–e are relatively limited in their computational capabilities since they perform a restricted form of wave transformation with limited input-output representations as opposed to implementing a function that can accurately transform infinitely many optical inputs to the desired representations at the output aperture (such as, e.g., Fig. 5a). With further advances in 3D nanofabrication techniques, we expect the realization of multi-layer metasurface networks to enable powerful adaptive free-space spatial processing and function representation capabilities critical for next-generation information processing and VR/AR applications.

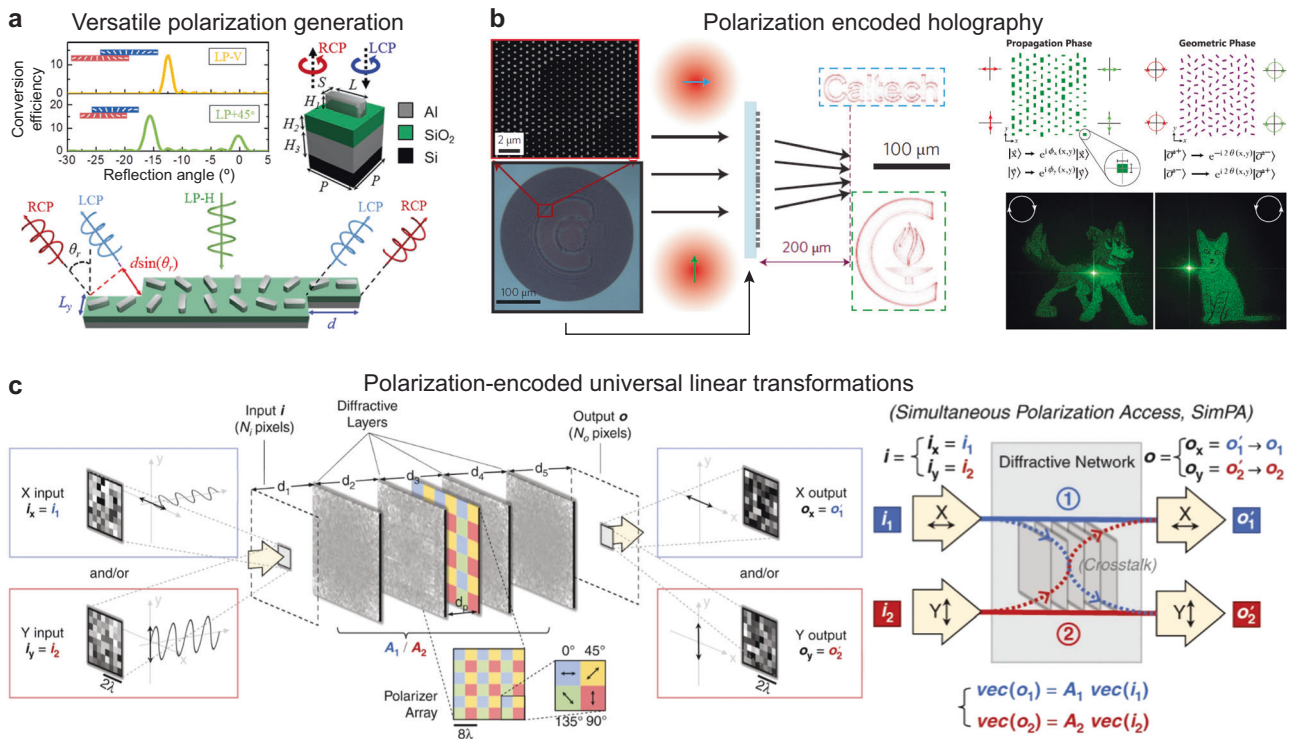
### Optical processing of spectral and temporal information

Diffractive surfaces also present powerful tools for processing the spectral and temporal information of light. One area of application

that has long exploited diffractive surfaces to manipulate the temporal profile of light is the shaping of broadband optical pulses in a vast range of applications, e.g., lightwave communications, biomedical imaging, and sensing. Typically, diffractive pulse shaping systems<sup>106</sup> consist of a dynamic/passive diffractive modulation surface sandwiched between two Fourier transform lenses (or concave mirrors)<sup>107</sup>. In this conventional setup of optical pulse shaping, using metasurfaces with advanced meta-units, e.g., Si nanopillars on a fused-silica substrate, brings additional degrees of freedom and capabilities leading to, e.g., ultra-wideband operation<sup>108</sup>. However, one particular drawback associated with this standard pulse shaping architecture is that it often results in bulky optical systems due to the use of lenses in paraxial regime (satisfying the condition for Fourier transformation) in addition to other components such as gratings and mirrors<sup>109</sup>. Moreover, while high-quality lenses, mirrors, and gratings are widely available for applications in certain parts of the electromagnetic spectrum, e.g., visible range, for optical pulses occupying some other parts of the electromagnetic spectrum, e.g., THz, the quality of these components is often inadequate. Deep learning-based design of diffractive optical processors, on the other hand, offers lens-free, data-driven, and compact alternatives for optical pulse shaping. Experimental demonstration of a compact, lensless pulse shaping platform that relies on 3D-printed dielectric diffractive surfaces manipulating an input broadband THz pulse has been reported<sup>18</sup>.

In the context of broadband light processing, task-specific dielectric diffractive surfaces structured using deep-learning can also provide solutions to various challenging inverse design problems, including single/multi-passband spectral filters and spatially-controlled wavelength-demultiplexing<sup>16</sup>. The advantages of a deep-learning-based task-specific design of diffractive optical processors can be further exploited by replacing the dielectric modulation surfaces with metamaterial-based solutions developed for applications in e.g., achromatic flat optics<sup>40,41</sup>. The additional degrees of freedom provided by the sophisticated meta-units, e.g., GaN nanopillars on Al<sub>2</sub>O<sub>3</sub> substrate, can enhance the broadband light processing capabilities of diffractive processors by providing e.g., simultaneous tuning of





**Fig. 6 | Free-space polarization processors based on diffractive networks and metasurfaces.** **a** Metasurfaces for versatile polarization generation by nanorod units<sup>112</sup>. **b** Polarization-dependent holographic images using anisotropic meta-units<sup>36,113</sup>. **c** Diffractive networks enabled multichannel universal linear

transformation by polarization multiplexing using a polarizer array<sup>14</sup>. **a** This is adapted with permission from ref. 112 by ACS. **b** This is adapted with permission from ref. 36 by Springer Nature and ref. 113 from APS. **c** This is adapted with permission from ref. 14 by CC BY 4.0.

modulation response and dispersion characteristics of individual meta-units.

**Optical processing of polarization information**

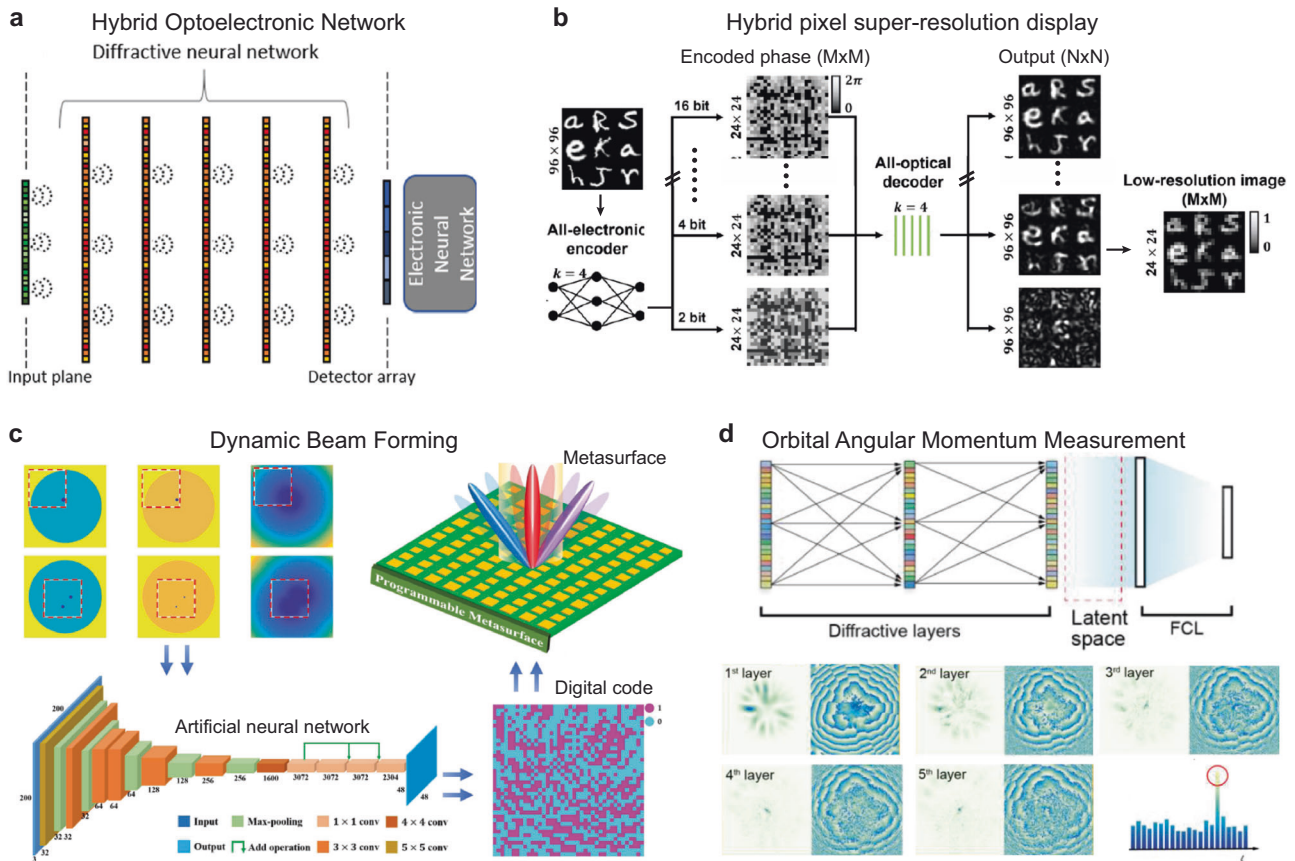
Since its invention, metasurfaces have been widely used for the manipulation of polarization states using their anisotropic subwavelength units that exhibit polarization-selective optical responses<sup>23,30,38,49–51,53,110,111</sup>. Based on their in-plane orientations ( $\varphi$ ), P-B geometric phase units such as nanorods generate opposite phase changes ( $\pm 2\varphi$ ) on left- and right-hand circularly-polarized light, respectively (Fig. 6a)<sup>112</sup>. Therefore, a metasurface with a linear phase gradient can separate a linearly polarized incident beam into two circularly-polarized beams in the opposite direction. A combination of two P-B supercells of meta-deflectors with opposite phase gradients at a subwavelength scale can generate all polarization states by tuning the shift distances (Fig. 6a). Polarization-sensitive metasurfaces can also produce holographic images encoded in different cross-polarization channels (Fig. 6b)<sup>36,113</sup>. Existing metasurfaces have demonstrated polarization processing by single-layer designs that typically accept only plane-wave inputs. Although multilayered metasurface polarization converters have been demonstrated in the literature<sup>114,115</sup>, these reports are based on periodic lattices and form single-pixel polarization processors. Metasurface-based networks for versatile polarization processing are still lacking but important for high-throughput information processing systems. This leaves significant room for improvement: a multiscale design strategy that concurrently optimizes the unit-level designs and the global structures of metasurfaces can provide unprecedented possibilities for polarization processing.

We want to point out that diffractive surfaces consisting of isotropic materials and unit cell designs can also realize polarization processing. Combined with a polarizer array, an all-optically polarization-multiplexed diffractive processor<sup>14</sup> has been demonstrated to

conduct multiple linear transformations by a series of isotropic diffractive layers (Fig. 6c). Fundamentally, this diffractive processor (composed of isotropic dielectric diffractive layers) gained its polarization-processing capability by spatial modulation of polarized beams generated by the polarizer array, which acts as a polarization seed within the free-space diffractive processor. Designed by a deep-learning training procedure, this free-space platform can perform concurrently multiple complex-valued matrix-vector operations through different polarization-encoded channels as long as the total number ( $N$ ) of diffractive neurons satisfies  $N \geq N_i N_o N_p$ , where  $N_i$ ,  $N_o$  are the number of pixels at the input and output fields, and  $N_p$  is the number of transformations encoded as combinations of input-output polarizations. This unique design strategy that achieves universal polarization transformations<sup>15</sup> with global spatial modulation at the wavelength scale can also be combined with the unit-level polarization-tuning of metasurfaces to further enhance their processing capabilities.

**Hybrid systems: Integration of free-space optical processors with electronic computing**

Free-space processors that integrate both optical and digital neural networks<sup>62,66,102,116</sup> can perform certain tasks with performances not possible by purely digital or optical components alone (Fig. 7a). This hybrid strategy has also been applied to integrated photonic systems<sup>117,118</sup>. For pixel super-resolution, which we have discussed in 2.3, for example, the digital CNN that encodes the high-resolution image into low-resolution SLM phase patterns not only serves as the essential interface to wavefront modulators but also reduces the burden of data transmission and storage (Fig. 7b)<sup>91</sup>. Hybrid networks have also realized machine vision using a single-pixel detector based on spectral encoding<sup>119</sup>. The diffractive network processes the spatial features of an input object illuminated by a broadband source and encodes the information into the power spectrum of the diffracted



**Fig. 7 | Hybrid optical-electronic networks for intelligent free-space processing.** **a** Hybrid optoelectronic network for classification<sup>66</sup>. **b** Digital encoder enables pixel super-resolution displaying by diffractive networks<sup>91</sup>. **c** Deep-learning design of digital code for dynamic beamforming on a reconfigurable metasurface<sup>120</sup>.

**d** Robust, high-accuracy OAM measurement by a hybrid network<sup>123</sup>. **a** This is adapted with permission from ref. 66 by CC BY 4.0. **b** This is adapted with permission from ref. 91 by AAAS. **c** This is adapted with permission from ref. 120 by IEEE. **d** This is adapted with permission from ref. 123 by CC BY 4.0.

light that is processed by the free-space diffractive processor that is optimized using deep-learning to communicate with a single-pixel at its output. In addition to the classification of input objects based solely on the spectrum detected at the single-pixel output aperture, a fully-connected digital neural network can also reconstruct an image of the input object from the spectrally encoded classification signal detected by the single-pixel<sup>119</sup>.

We want to note that a single-step joint training of the optical and electronic parts often stagnates at a local minimum with undesirable performance. Instead, improved inference performance can be obtained from a two-step method that (1) trains the optical layers with additional virtual optical layers that act as a placeholder for the electronic network and then (2) jointly trains the untrained digital layers cascaded with the partially-trained optical front-end. This two-step training strategy achieves improved performance by providing a better initial condition for the diffractive network. Overall, this two-step transfer-learning-like training procedure can effectively augment the capabilities of a trained model and facilitates the integration of optical and electronic networks. The hybrid approach is also critical for increasing the depth of metasurface-based free-space processors<sup>62</sup>, which have been limited to a single optical layer by current limitations in nanofabrication. The depth advantages of optical and hybrid networks will be further discussed in Section “Tunability and reconfigurability”.

Digital deep-learning methods can also guide the structural design and reconfiguration of diffractive and metasurface devices<sup>120</sup>. Reconfigurable phase layers and metasurfaces can be realized by SLMs (in the visible) and programmable optoelectronic units (GHz range) (Fig. 7c)<sup>120,121</sup>. Hybrid neural networks also provide powerful tools for

the recovery and measurement of phase-encoded information. For example, phase-encoded images formed by overlapping two phase images of MNIST numbers can be classified by a diffractive processor, and each digit can be accurately reconstructed by a digital network<sup>122</sup>. The optical network compresses the size of the input phase image to reduce the number of pixels by 20–65 times, which facilitates the reconstruction of both images regardless of the phase ambiguity. The ability to process phase information also allows the hybrid systems to manipulate OAM beams, which form a promising technology for next-generation information processing and telecommunication (Fig. 7d). Remarkably, this diffractive network integrated with a single fully-connected layer can achieve high calculation speed and energy efficiency where 99.98% of computations are performed optically. The OAM reconstruction achieved an extraordinary accuracy with a mean-squared error (MSE) of  $10^{-5}$ – $10^{-3}$ <sup>123</sup>. The compact system (footprint  $\sim 200\lambda$ ) can achieve high robustness to the atmospheric turbulence and the spatial displacements of OAM beams, including transverse and angular shifts. Recently, a demonstration of electrically tuned metasurfaces can generate the OAM beam with dynamically-tunable topological charges<sup>124</sup>. We believe that the integration of digital neural networks with these reconfigurable (ideally jointly-trained) metasurfaces can contribute to a broad range of futuristic information technologies.

### Grand challenges in free-space optical computing and performance limitations

Various practical applications of free-space optical systems, such as imaging and computing, require tackling various challenges in device fabrication and performance. Diffraction efficiency, in particular, is

critical for enabling energy-efficient technologies ranging from QPI and microscopy to holographic displays. Emerging innovations in the design methods and optical materials that can improve output diffraction efficiencies will be discussed in Section “Diffraction efficiency and power requirements”. The ability to tune and reconfigure free-space devices during their operation is also critical for various applications, including telecommunications (e.g., active beamforming) and imaging. To discuss these challenges, we will cover the critical problems and emerging opportunities in realizing tunable and reconfigurable diffractive and metasurface-based networks in Section “Tunability and reconfigurability”.

Furthermore, the nanofabrication techniques needed for building these 3D optical networks will be discussed in Section “Fabrication complexity and 3D alignment requirements”. Section “Computation speed, parallelism, and scalability” will analyze the strategies to achieve fast, parallel free-space processors with the scalability that is essential in modern computing platforms for solving complex problems. Lastly, for free-space processors to become a game-changer for the next-generation of computing technologies, they need to achieve competitive computation accuracy and precision compared to their electronic counterparts in the computational tasks they are designed for. Some of the potential techniques to improve the accuracy of diffractive and metasurface-based networks for computation and inference tasks will be discussed in Section “Computation accuracy and inference capability”.

### Diffraction efficiency and power requirements

The overall diffraction efficiency of metasurface systems, defined as the percentage of the optical power detected at the targeted sensor pixels with respect to the total input power, has long been a limiting factor for achieving competitive performance in imaging and display technology<sup>125</sup>. Plasmonic metasurfaces operating in the transmission mode, in particular, often exhibit low diffraction efficiencies that prohibit their applications in imaging and spectroscopy. Plasmonic meta-units exhibit non-radiative losses due to their materials properties and suffer from radiative modes that direct ballistic photons to undesirable targets. While the search for low-loss plasmonic materials in the visible and near-infrared range is still an ongoing research effort<sup>56,126–128</sup>, designs of metasurfaces that operate in reflection mode have realized a diffraction efficiency of >80% for holography<sup>98</sup> and beam steering<sup>129</sup> by using metal-insulator-metal structures.

All-dielectric metasurfaces can realize higher power efficiencies for transmission applications using materials such as TiO<sub>2</sub> for the visible range<sup>38</sup> and Si for the near-infrared range<sup>64</sup>. By tailoring their size and shape, the meta-units can exhibit destructive interference of electrical and magnetic dipole modes<sup>130</sup> in reflection to suppress reflection losses and enhance transmission efficiency<sup>125</sup>. Ideal designs of dielectric P-B phase elements (such as nanofins) can function as perfect half-wave plates that shift the phase of the incident light polarized along the long axis of the meta-units by  $\pi$  with respect to the light polarized along their short axis. Experimentally, the focusing efficiency of such P-B devices can achieve ~86% at 405 nm<sup>38</sup>. In comparison, polarization-insensitive metasurfaces consisting of cylindrical dielectric posts that tailor the wavefront phase by their diameters have reported diffraction efficiencies higher than 90%<sup>131,132</sup>. Furthermore, broadband, polarization-insensitive metasurfaces have recently achieved a diffraction efficiency of 90% in the 450–700 nm wavelength range by dispersion-engineered meta-units<sup>133</sup>.

This relatively high efficiency of dielectric metasurfaces comes at the cost of difficult fabrication, which will be discussed in Section “Fabrication complexity and 3D alignment requirements”. Briefly, commonly used dielectric materials such as TiO<sub>2</sub> and Si are produced by slow and costly methods such as ALD and CVD. Realization of multi-layer metasurface designs by these high-temperature processes, which preclude organic layer-to-layer spacers, is also difficult. Therefore, we

believe that there are still tremendous opportunities in the discovery of low-loss, high-index optical materials that allow simpler fabrication. The dielectric nanoposts with small diameters typically have a very high aspect ratio that makes their accurate nanofabrication challenging, and therefore, the experimental diffraction efficiency can be compromised due to fabrication imperfections.

Diffractive networks and dielectric-based optical processors composed of  $\lambda/2$  lateral feature sizes (Section “Diffractive surfaces”) exhibit power efficiencies that largely depend on their training loss function and design process. For example, diffractive network training for all-optical image classification using MSE loss promotes higher signal contrast and better power efficiency (routing e.g., 25.07% of the output photons to the correct photodetector assigned to the correct label for MNIST image data) while the cross-entropy loss favors a significantly better classification performance with a compromise in the diffraction efficiency (by e.g., ~10-fold)<sup>66</sup>. In general, the output diffraction efficiency of a multilayer diffractive network shows a depth advantage where an increase in the number of diffractive surfaces/layers improves its overall power efficiency because of the additional degrees of freedom in the system<sup>66,134</sup>. In these cascaded optical processor designs, the diffractive efficiency at the output plane can be further improved by a compromise in the performance of the network, such as imaging contrast, signal-to-noise ratio, and test object classification accuracy. This trade-off between output diffraction efficiency and performance can be engineered readily by the choice of loss functions and the level of diffraction efficiency penalty<sup>18,66,92,93</sup>. For example, for a diffractive processor design that performed all-optical QPI through the conversion of phase-only signals into quantitative intensity patterns<sup>93</sup>, an increased output diffraction efficiency of >11% was achieved by increasing the weight of the diffraction efficiency related penalty term in the loss function at the cost of a slight decrease in the image quality (i.e., a structural similarity index reduction of <0.06). A similar trade-off was observed in the all-optical reconstruction of holograms<sup>67,92</sup>, where the diffraction efficiency of the reconstructed holographic images at the output plane was increased to >26% by increasing the weight of the diffraction efficiency penalty term in the training loss function, which resulted in a tolerable compromise in the reconstruction quality.

In experimental validation, however, material absorption can lead to a larger loss with an increasing number of successive layers (or thicker monolithic designs) if a low-loss material for the fabrication of the 3D diffractive processor is not available for the wavelength range of operation. In the network training process, the lossy diffractive layers can be modeled as complex modulators (Section “Diffractive surfaces”) using the measured extinction coefficient of the material so that the design parameters (e.g., the number of layers, layer-to-layer separation, etc.) can be optimized during the training. For the visible part of the electromagnetic spectrum, this poses less of a challenge since there are various low-loss optical materials (e.g., glasses, polymers) that can be engineered at the wavelength scale in 3D<sup>135–137</sup>. Materials such as TiO<sub>2</sub> and Si are also suitable for building low-loss diffractive layers in the visible and near-infrared wavelength ranges, respectively.

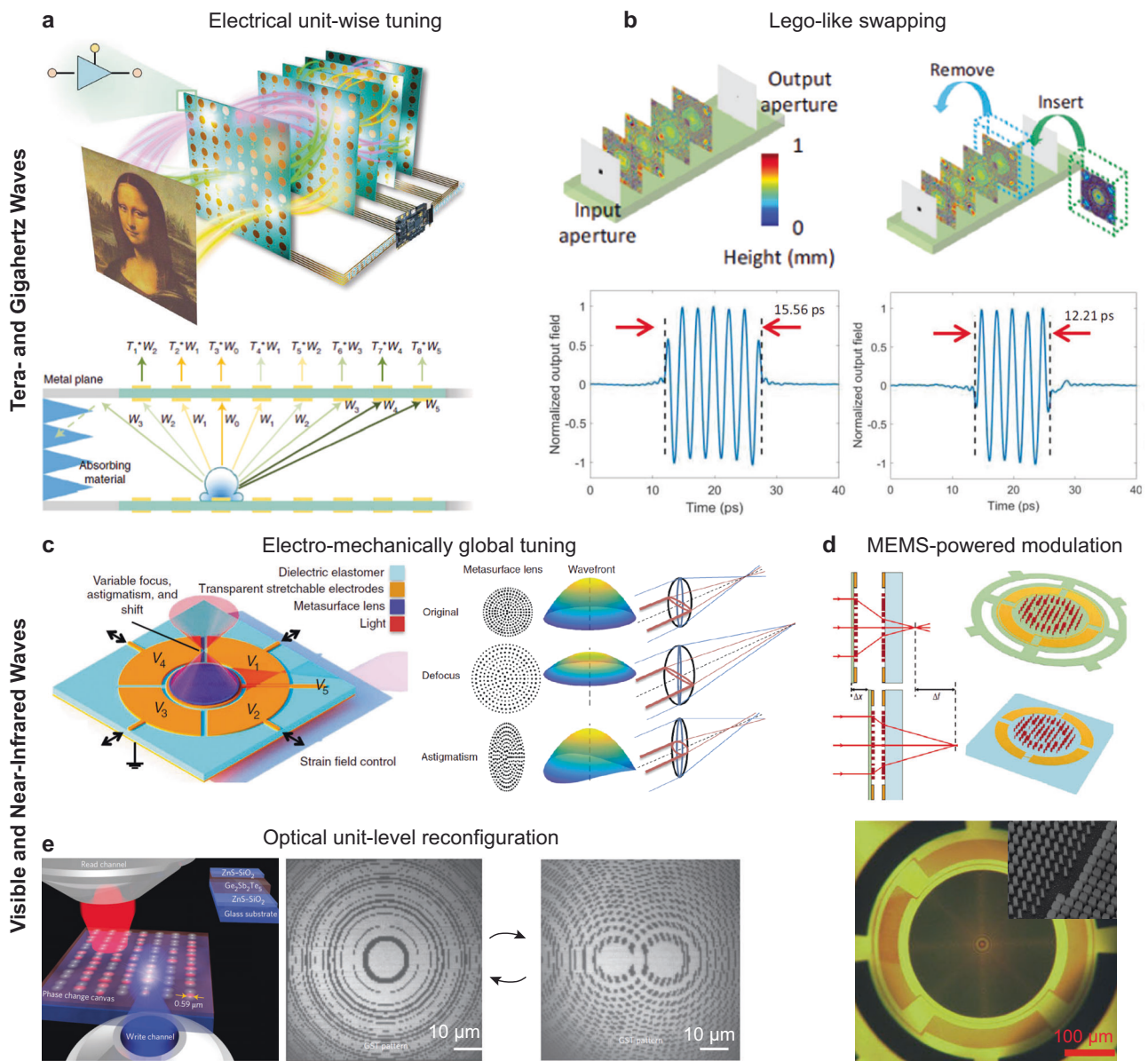
### Tunability and reconfigurability

The ability to fine-tune the design of free-space optical processors is critical for adjusting their performance based on the task, which is especially important when more than one metric needs to be optimized concurrently<sup>138,139</sup>. In a typical example of an image classification problem, the main objective to optimize is the classification accuracy, but the power efficiency and signal contrast are also important in practical applications. Therefore, on-demand tuning of the network after the fabrication is critical for optimizing the performance of these free-space processors based on the need of the specific tasks. Diffractive free-space processors consisting of transmissive layers also

allow Lego-like swapping of layers to tune their output function (Fig. 8a)<sup>18</sup>. For example, based on a transfer learning approach, some existing layers in a pre-trained diffractive network can be physically replaced with newly trained and fabricated transmissive layers to enable the on-demand synthesis of new pulses with desired pulse shapes and durations. The connectivity of a diffractive network, which can be tuned by the layer-to-layer distance  $\Delta z$ , is another important parameter that determines the output of the diffractive optical processor<sup>16</sup>. For terahertz pulse shaping, for example, the same set of diffractive layers produced a tunable center frequency from 0.349 THz to 0.399 THz by physically reducing  $\Delta z$  from 30 mm to 25 mm<sup>18</sup>. To apply this  $\Delta z$  tuning of diffractive and metasurface layers that operate at the visible part of the spectrum, we expect MEMS-based systems<sup>102,140</sup> to become useful for the precise control over  $\Delta z$  at micrometer length scales. Alternatively, the effective optical path length between the passive layers can be tailored by modulating the background refractive index without changing the physical

distances between the layers. This index-tunability can be realized by a range of materials systems, including liquid crystals<sup>100</sup>, electro-optic materials<sup>141–143</sup>, and phase-change materials<sup>144</sup>.

The tunability of metasurfaces is also important for achieving adaptive responses in applications ranging from imaging to spectroscopy. Mechanical tuning, for example, is a common approach for realizing active metasurfaces with tunable global optical responses<sup>103,105</sup>. By patterning meta-units on a flexible electroactive substrate, the focal point of a metalens can be shifted and stretched upon applying electrical biases, which adjusts the spacings ( $a_0$ ) between the meta-units (Fig. 8b)<sup>101</sup>. Microelectromechanical systems (MEMS) that adjust the distance between two metasurfaces can also tailor the output plane (Fig. 8c)<sup>102</sup>. The ability to shift and reshape the foci of metalenses is critical for building compact microscopes and miniaturized cameras. These mechanical systems based on adjusting lattice constants can also modulate the spectral responses of the metasurfaces<sup>145,146</sup>. By adjusting the spacing between the meta-units,



**Fig. 8 | Tunable and reconfigurable responses from metasurfaces and diffractive networks.** Tuning of device responses by (a) Lego-like swapping of the layer designs<sup>18</sup>, (b) electromechanical modulation that deforms the metasurface substrate<sup>101</sup>, (c) MEMS systems that adjust distances between two metasurfaces<sup>102</sup>. Unit-level metasurface reconfiguration by (d) electrical tuning of

field-programmable gate arrays<sup>121</sup> and (e) reversible laser writing on phase-change materials<sup>144</sup>. **a** This is adapted with permission from ref. 18 by Springer Nature. **b** This is adapted with permission from ref. 101 by AAAS. **c–e** These are adapted with permission from ref. 102, ref. 121, and ref. 144, respectively, by Springer Nature.

the resonance of a plasmonic metasurface can be tuned over a wide wavelength range (e.g., 500–600 nm). In general, mechanical modulation is slow (e.g., with a response time of tens of ms)<sup>101</sup> and limited to changing the global structures of metasurfaces but can not access and individually change their degrees of freedom. Integration with emerging 2D materials such as graphene<sup>147,148</sup> and black phosphorous<sup>149</sup> is also promising to enable gate-tunability within metasurface networks. In addition, metasurfaces with a silicon-on-lithium niobate architecture demonstrated wavefront shaping and modulation based on electro-optic tuning<sup>150</sup>.

In addition to tunability (fine-tuning the function of an optical processor), the reconfigurability of optical processors allows these devices to evolve and adapt to different tasks to achieve optimal performance in a dynamic environment. For diffractive free-space processors, fast reconfiguration at the single-neuron level can further enable in-situ network training, where the diffractive devices are tailored in real-time based on their measured responses<sup>27</sup>. This training strategy that uses experimental results instead of simulated responses can not only realize accurate evaluation of the optical processor performance but also reduce the training time for dynamically changing tasks whose simulation is computationally expensive. Gate-tunable plasmonic units were realized by graphene-gold resonators with a single-unit phase tunability of  $0-1.28\pi$  in the mid-infrared range and were used for achieving active beam steering<sup>147</sup>. A fully-programmable metasurface network was also demonstrated for GHz wave applications using electrical field-programmable gate arrays (Fig. 8d)<sup>12</sup>. These GHz wave processors can be reconfigured to perform a range of tasks ranging from matrix inversion<sup>151</sup> to image classification and information encoding/decoding<sup>152</sup>. The ability to conduct fully-programmable field synthesis is also valuable for beamforming and holographic image projection applications<sup>153</sup>. For visible and near-infrared ranges, however, unit-level adjustment of the diffractive surfaces and metasurfaces is more challenging because the feature sizes are significantly smaller. Unit-level reconfiguration has been realized by focused laser writing and erasing of phase-change materials<sup>144,154</sup> such as  $\text{Ge}_2\text{Sb}_2\text{Te}_5$  (Fig. 8e). This approach, however, requires sophisticated instrumentation for the active modulation and can not apply to multi-layer processors with current technology.

Overall, reconfigurable diffractive and metasurface processors are expected to have more complex structures and fewer materials choices that are harder to fabricate than their static counterparts. For visible-range applications, in particular, the development of fully-dynamic nanoscale meta-units with comparable fine features as static meta-units is extremely challenging with the current nanofabrication techniques. However, reconfigurable free-space devices with competitive performances can be realized with a new paradigm of hybrid network structure consisting of (1) fixed diffractive and metasurface layers with carefully engineered subwavelength features that provide the fine control over optical wavefront, (2) tunable optical superstructures with micron-scale features sizes for introducing tunable and reconfigurable optical responses to the system, and (3) digital networks for the pre- and post-processing of the input/output. We expect this hybrid framework to become the next-generation of high-performance reconfigurable free-space processors that can be integrated with portable devices.

### Fabrication complexity and 3D alignment requirements

Precise fabrication of diffractive free-space processors is important for ensuring the accuracy of these analog computing platforms, which do not have the same error-correction ability as digital computing. For these optical computing systems, in particular, the misalignments between the diffractive or metasurface layers will introduce potential systematic errors to the computing results and should be ideally limited to a size comparable to the illumination wavelength. The ability to mitigate these misalignment errors is essential for fully harnessing the

computing capabilities of multilayer 3D diffractive networks. Therefore, the fabrication of diffractive free-space processors that operate at visible wavelengths requires the ability to create layers of nanoscale units with accurate phase modulation using e.g., spatial engineering of the refractive index and/or thickness of the material. For instance, 3D printing methods based on two-photon polymerization of photo-sensitive materials<sup>45,46</sup> are promising techniques for achieving sub-wavelength feature resolution using tightly focused laser beams (Supplementary Fig. S2a). Commercial solutions such as the Quantum X developed by Nanoscribe<sup>135,136</sup> and MicroFAB-3D developed by Microlight3D<sup>43</sup> have demonstrated the printing of sophisticated photonic structures such as photonic crystals<sup>155</sup> with a resolution below 200 nm. Implosion lithography<sup>137</sup> can further push the resolution limit down to 50 nm by printing a magnified version of the target structure in a hydrogel and shrinking the entire structure by dehydration (Supplementary Fig. S2b). One of the major challenges in creating visible-range diffractive free-space processors consisting of multiple thickness-tuned layers is their mechanical robustness. Although structural supports are added between the diffractive layers to be printed/fabricated in the same fabrication session, the layers during the curing process can undergo deformations that can produce inaccurate layer-to-layer distances and degrade the performance of the optical processor. Therefore, monolithic architectures would be preferred for building robust, miniaturized diffractive free-space processors. For example, laser writing techniques<sup>48</sup> can generate index-tuned diffractive layers in a single dielectric volume to achieve high mechanical robustness. A current limitation of this technique is that the index-tuning by exposure time may have limited spatial resolution and index-tuning range. Therefore, the discovery of novel material platforms that allow index-tuning is critical for building robust, monolithic diffractive optical processors.

On the other hand, optical metasurfaces are typically fabricated by an electron beam lithography process consisting of steps including electron beam writing, deposition, and etching (Supplementary Fig. S2c)<sup>65</sup>. In a typical fabrication process for a visible-range metasurface made of  $\text{TiO}_2$  (a high-index, low-loss dielectric material), the metasurface is fabricated by (1) electron beam writing of the design on the resist, (2) ALD-based fabrication of the  $\text{TiO}_2$  with excess thickness, and (3) etching of the excess  $\text{TiO}_2$  and removal of the electron beam resist to obtain the final metasurface pattern<sup>38,65</sup>. Si- or GaN-metasurfaces can be fabricated by similar processes where an etch mask with the metasurface design is produced by electron beam lithography and then the pattern is transferred to the Si or GaN film by anisotropic reactive-ion etching<sup>36,40</sup>.

Current fabrication methods for free-space metasurface processors, however, can only realize single-layer designs<sup>42,62</sup> or multilayer designs that do not require accurate ( $<\lambda/2$ ) layer-to-layer alignment<sup>64,156-158</sup>. An alignment technique between different fabrication layers is still lacking and this feature is essential for achieving multilayer metasurface processors for performing e.g., universal linear transformations in a transmissive geometry between a diffraction-limited input field and output. Also, interlayer structural supports are needed to construct multilayer metasurface-based optical processors. Unlike their THz counterparts that can be assembled from free-standing layers in the air, visible-range metasurfaces must be fabricated on robust, smooth substrate materials. Therefore, dielectric filler materials with high transparency and mechanical stability would be preferred spacers that separate the metasurface layers. The major challenge is that deposition methods such as ALD and PECVD for high-quality transparent materials are slow ( $<5$  nm/min for ALD of  $\text{Al}_2\text{O}_3$ <sup>159</sup> and 167 nm/min for PECVD of  $\text{SiO}_2$  by SAMCO) and expensive. The residual stress accumulated during these high-temperature deposition processes also makes the film vulnerable to cracking and fracture at thicknesses larger than  $15\ \mu\text{m}$ <sup>160</sup>, which restricts the designs of metasurface processors that can be realized. The poor thermal conductivity

of common dielectric materials also produces cooling issues in dry etching processes and prohibits the directional etching required for producing high-aspect ratio dielectric meta-units<sup>161</sup>. We expect the development of novel dielectric materials and nanoscale 3D alignment methods to be crucial for realizing metasurface networks formed by successive layers of jointly-optimized meta-layers.

In general, nanofabrication methods that allow scalable, low-cost production of metasurface networks are essential for the wide-scale use of such optical processors. One promising approach to realizing large-scale parallel nanofabrication is photolithography which can replicate the metasurface designs defined on a photomask on transparent wafers (Supplementary Fig. S2d)<sup>162</sup>. Using a 193-nm light source and liquid-immersion technology, the state-of-the-art deep-ultraviolet (DUV) lithography processes can easily produce meta-units below 40 nm<sup>163</sup>. Inverse design of photomasks using sub-resolution assist features can further improve the quality of the fine features<sup>164</sup>. The main disadvantage of DUV photolithography is its extremely high instrumental cost. In contrast, soft lithography methods such as nanoimprinting<sup>165</sup> and solvent-assisted nanoscale embossing<sup>166</sup> can produce metasurfaces with features below 60 nm using benchtop (Supplementary Fig. S2e) and are suitable for usage in labs and small-scale manufacturing. In a typical soft lithography process, an elastomeric stamp is obtained by molding from a template with the metasurface design and brought into conformal contact with a photoresist-coated substrate. Via mechanical pressure or solvent interactions, the metasurface designs are formed on the substrate and subsequent pattern transfer steps such as deposition and etching are used to produce the metasurface devices.

Besides direct improvements in instrumentation and materials, the challenges in nanofabrication of diffractive and metasurface-based free-space optical processors can be resolved in the design process by considering the imperfections in fabrication. Previously, diffractive networks that are invariant to the scale-, shift, and rotation of input images have been demonstrated by introducing these variations in the training stage<sup>167</sup>. An extension of this approach can tremendously improve the robustness of the free-space optical processor toward the unavoidable misalignments between the layers or other imperfections by including them as random parameters in the training process<sup>82,90,168</sup>. We expect this ‘vaccination’ strategy to apply to a wide range of imperfections in fabrication and experimental systems (such as inaccuracy in phase and amplitude modulation) and can be customized based on the limitations of the selected fabrication method. Furthermore, in contrast to the designs that use analytic wave-optics approaches, the deep-learning-based diffractive and metasurface-based optical processors that are trained using data can be designed based on an incomplete phase coverage (less than  $2\pi$ , that can be selected on demand) to accommodate materials-related refractive index limitations and phase restrictions to enable simpler fabrication methods. All in all, deep-learning-enabled designs of diffractive and meta-material based optical processors would be able to better handle some of these experimental restrictions due to fabrication imperfections and limitations in e.g., resolution, feature size, depth, refractive index, and mechanical misalignments. There is still considerable work needed for building cascaded free-space optical processors that can operate at IR and visible wavelengths at diffraction-limited SBP, in order to fully utilize the density of information encoding in free-space.

### Computation speed, parallelism, and scalability

Free-space optical processors present a unique advantage since their computation is completed as the input light propagates through the optical processor and its volume/layers. While electronic processors and integrated photonics-based processors also utilize the speed of electromagnetic wave propagation, the free-space processors *complete* all the calculations and inference tasks with a single pass of wave propagation, without any digital storage/transmission or pre-processing of

information. Therefore, the computing speed of diffractive and metasurface systems for the same task can be increased drastically by scaling the system down to the nanoscale feature sizes, operating at visible or IR wavelengths, with a total axial thickness of e.g.,  $\sim 100\text{--}200\lambda$ . To manipulate this wavelength range, however, the diffractive features used in free-space optical processors need to have dimensions of  $>100\text{ nm}$ , significantly larger than the state-of-the-art transistors on chips made by the 3-nm process. In addition, the size of free-space optical processors that can be accurately modeled and designed (through e.g., a deep-learning-based training process) is still limited by the speed of digital computers and their memory restrictions. Unlike CPUs and GPUs that can be assembled into clusters to solve advanced computation problems, a versatile strategy to assemble diffractive networks into a large-scale cluster of free-space optical processors is still lacking, but highly desired to fully utilize their advantages in parallelism. Therefore, the ability to construct low-loss, large-scale networks of diffractive optical processors is thus critical for building fast, free-space computing platforms with significantly lower power consumption.

One promising strategy to achieve a high density of processing units in free-space optical processing is to use a *diffractive volume* where every discrete cell in the 3D space is spatially engineered to collectively process the input optical information. Such a volumetric optical processor can have a very large number of optimized diffractive features, e.g.,  $>10^9$  within a compact material volume, which would normally be challenging for the multilayered designs to achieve even with nanoscale meta-units. Such a large density of optimized diffractive features within a 3D compact topology could be another unique aspect of free-space-based visual information processors and is hard to match with planar integrated photonic circuit-based designs. Such dense 3D lattices of diffractive units also have the potential to utilize evanescent waves (containing much larger spatial frequencies) to further increase the computation capability. However, the design of a diffractive volume (that can process evanescent waves) is challenging not only because a fast and accurate forward model that can quantify the scattering between these densely-packed units is lacking, but also due to the extreme difficulties in the fabrication of such monolithic free-space processors. For example, a nanophotonic design volume based on simulations by finite-difference frequency-domain calculations and fabrication by laser-tuning of refractive index was demonstrated, but the total volume is limited to  $4\lambda \times 4\lambda \times 6\lambda$ <sup>169</sup>. This diffractive computing medium needs to scale up significantly using e.g., emerging nanofabrication<sup>137</sup> and optical simulation methods<sup>170</sup>. For scaled-up designs, the optical simulation technique used by the forward model must be fast and accurate while allowing the computation of gradients essential for the deep-learning training process. Well-established numerical methods such as the finite-elements method<sup>171</sup> and finite-difference time-domain<sup>172</sup> exhibit high-accuracy, but can not meet the speed and gradient-computation requirements. One promising simulation method is MaxwellNet<sup>170</sup>, a physics-driven neural network that can solve for the optical responses of a diffractive volume. Using the residual of Maxwell’s equations as the losses, the method can be trained to quickly predict the phase and amplitude modulation of a structure given the spatial distribution of the material property.

Despite the challenges above, the diffractive and metasurfaces still have tremendous opportunities and advantages in information processing because of their unique operation. For example, the massive parallelism of optical systems allows the same free-space optical processor to perform multiple tasks simultaneously with the multiplexing of wavelength, polarization, and spin<sup>14,17</sup>. In particular, the subwavelength structures of meta-units provide additional structural degrees-of-freedom that allow independent, pixel-level control over e.g., polarization<sup>112,113</sup> and dispersion<sup>173</sup>. We believe that rationally-designed metasurface-based networks can form multispectral, polarization-multiplexed free-space processors that can process each

channel independently while providing on-demand control of the cross-coupling among different channels—all within the diffraction limit of light and the corresponding SBP of the input and output apertures. In addition, free-space processors can be integrated directly into the light path (without the need for digitization, storage/transmission, and pre-processing of information) to perform tasks such as image reconstruction and class-selective imaging with zero time delay<sup>28,90</sup>. Therefore, we expect that free-space optics will form powerful task-specific processors to perform analog computation and imaging tasks at the speed of light.

Overall, diffractive surfaces and metasurfaces can potentially provide a scalable free-space platform for highly parallel processing of visual information that is difficult for integrated photonics-based systems to process directly, which are, by design, more suitable for processing already digitized information. For example, the 2D phase information of a scene/object can be directly processed by free-space-based diffractive processors without the need for pre-processing of information (such as phase retrieval and vectorization of the resulting phase image) which would normally be required for an integrated photonics-based processor before it can act on the phase image. Free-space optical processors, in contrast, can directly process 2D or 3D visual information and work with significantly larger inputs since their physical size linearly grows with the size of the input FOV. For inference tasks such as image recognition and classification, free-space optical processors based on diffractive and metasurface networks allow highly dense connectivity in 3D, which is harder to achieve in integrated photonics<sup>174</sup>. Although the number of trainable parameters of optical neural networks, as of today, is still significantly smaller than the state-of-the-art electronic deep neural networks<sup>175</sup> (with, e.g.,  $\sim 10^{12}$  parameters), free-space-based processors can still provide unique advantages due to their direct access to the optical information of input scenes/objects without the need for any digital cameras, image recording/transmission, or pre-processing of data, providing ultra-fast and power efficient visual information processors—potentially integrated with mobile devices.

### Computation accuracy and inference capability

For free-space diffractive and metasurface networks to play a key role in next-generation technology, they must reach a computation accuracy and precision comparable to their existing electronic counterparts in the target computational tasks. In digital electronic systems, the precision is determined by the number of discrete states implemented by binary logic units to approximate continuous variables<sup>176</sup>. Increasing the precision of analog devices, in general, requires significant advances in materials and fabrication techniques, where the cost can increase exponentially with increases in the desired precision<sup>177,178</sup>. In particular, free-space optical processors as analog optical information processing systems exhibit an explicit limit in their precision, practically limited by their design, fabrication, and measurement processes. For example, errors in diffractive computing can result from inaccuracy in the layer thicknesses or the topology of the fabricated subwavelength features due to imperfections of the 3D printing process. Misalignments between diffractive layers can also introduce nontrivial errors that lead to incorrect computational and inference outcomes<sup>168</sup>. Although statistical inference and computational imaging tasks usually do not require high precision<sup>179</sup> and can tolerate some degrees of errors, other applications such as linear transformations<sup>13</sup> can be much more sensitive. Some of these limitations in nanofabrication and misalignments can potentially be addressed by appropriate design and network training strategies that take these factors into account during the deep-learning-based optimization and training of the free-space processor (discussed in Section “Fabrication complexity and 3D alignment requirements”).

The computational performance of current deep neural networks systems can be understood by the universal approximation theorem<sup>180</sup>, which states that multi-layered networks with a sufficient number of hidden units and adequate nonlinear activation functions can approximate any continuous function mapping from closed, bounded space to another one with finite dimensions<sup>180–182</sup>. To achieve universal function approximation through free-space optical networks, however, materials and neuron designs that can achieve nonlinear optical responses would be needed<sup>73</sup>. Nonlinearity has been implemented in diffractive optical processors using a combination of photodetector/sensor arrays and dynamic electro-optic modulation devices<sup>27,183</sup>. However, this approach not only significantly increases the complexity of the devices but also introduces undesirable active power consumption to the system. The speed of operation will also be limited by the performance of the dynamic electro-optic modulation devices. Another interesting demonstration of nonlinearity using transitions between atomic states of <sup>85</sup>Rb<sup>184</sup>, but the use of vulnerable atomic systems makes this system impractical in realistic devices. A practical material system that enables large nonlinearity and easy fabrication is still the main challenge in developing the next-generation of free-space machine-learning devices<sup>185</sup>. Besides the direct implementation of nonlinear activation functions, the inference capacity can also be improved by innovations in diffractive processor architecture and training strategies, including class-specific network training<sup>68</sup> and ensemble learning<sup>69</sup> as detailed in earlier sections. A disadvantage of both of these strategies (class-specific diffractive networks and ensemble learning approach) is that multiple diffractive networks need to operate together, requiring sophisticated 3D integration and optical alignment in free-space, further complicating the optical hardware. As an alternative approach, inspired by pixel super-resolution microscopy using object shifts<sup>186</sup>, time-lapse image classification using a single diffractive network was reported that utilized random or ordered lateral shifts of an object with respect to the input aperture to achieve ~62% inference accuracy for classifying CIFAR-10 test images using a single time-lapse diffractive optical network<sup>187</sup>.

### Conclusions

In this Perspective, we discussed the recent advances in diffractive surface and metasurface designs as free-space optical processors and their implications and potential impact on future analog visual computing technologies (Fig. 9). Starting from their fundamental design and operation principles, we analyzed the capabilities and potentials of optical diffractive networks as a platform for statistical inference<sup>26</sup> and universal linear transformations<sup>13,14</sup>. As discussed earlier, we expect major advances in the design and fabrication of diffractive and metasurface-based free-space optical processors to achieve more competitive performance in inference tasks, especially if the 3D alignment and assembly challenges in the fabrication of multi-layer designs can be overcome. We also reviewed some of the emerging approaches that can potentially overcome the practical challenges in implementing nonlinear activation functions<sup>184</sup>, a critical component for free-space optical processors to achieve enhanced performance.

As versatile analog computing platforms with machine-learning capabilities, these free-space optical processors form powerful *visual information processing* tools that can directly access and harness the spatial, spectral, polarization and amplitude/phase information of light waves. One critical application enabled by diffractive visual processors is all-optical QPI that can convert phase-encoded information of specimen into intensity patterns in a quantitative manner, performing all-optical phase recovery<sup>93</sup>, with various applications in biological imaging. Similarly, the ability to perform imaging through unknown, random diffusers using analog visual computing via spatially engineered diffractive layers<sup>28</sup> or metasurfaces will open up new possibilities in, e.g., medical imaging, security screening and free-space communication.



**Fig. 9 | Roadmap and future outlook of diffractive surface- and metasurface-enabled computing platforms.** These free-space optical processors can have transformative impact on various technologies ranging from robotics to biomedical imaging and telecommunications, among others.

As another application, the class-specific imaging capability of diffractive visual processors further provides a novel solution to privacy protection problems that can be concerning during the deployment of data-driven imaging and surveillance technologies<sup>90</sup>. In addition to harnessing the spatial features of a scene, their inherent ability to process polarization<sup>14,112,132</sup> and spectral/temporal<sup>119,133,173,188</sup> information of light also makes these free-space optical networks promising for the development of next-generation sensing and telecommunication technologies. For example, both diffractive layers and metasurface-based designs possess the ability to manipulate structured light that carries OAM<sup>63,95,123</sup>, which is expected to be a core component of 6G communication technologies.

Finally, we have discussed in this Perspective some of the key metrics for evaluating the performance of free-space optical processors, including computation speed and scalability, power efficiency, and reconfigurability as well as the critical challenges that need to be addressed for achieving competitive scores on these metrics. In particular, the design and fabrication methods for massive-scale free-space optical processors with cascaded structures in 3D are still lacking. We envision that monolithic diffractive computing volumes<sup>169</sup>, as opposed to discrete engineered surfaces or layers, will form a powerful class of free-space computing platforms that can realize a significantly higher density of diffractive features in 3D; but such a capability needs a fast and accurate forward optical model in the design process, which will also need to carefully keep a record of the evanescent waves and their interactions with subwavelength structures within the spatially engineered volume. We expect that physics-informed deep neural networks<sup>170</sup> trained to solve Maxwell's equations rapidly would be one promising platform to be used in the forward model of such 3D computing platforms. These 3D designs will be capable of performing large-scale computing and optical information processing at low energy consumption and ideally below the diffraction limit of light. Another expected breakthrough in free-space optical computing would be the realization of dynamic tuning and on-demand

reconfiguration of diffractive<sup>18</sup> or metasurface-based<sup>27,102,121,144,147</sup> optical processors. Rapid developments in various nanofabrication techniques, including nanolithography<sup>162–165</sup> and 3D printing<sup>43,45,46,136</sup>, can accelerate the transition of dynamic diffractive and metasurface-based optical networks from GHz to visible frequencies. We also expect major advances in dynamic hybrid processors<sup>62</sup> that combine fixed/static diffractive layers or metasurfaces with subwavelength features for fine control over the optical wavefront, adaptive optical superstructures for reconfigurability<sup>27</sup>, and digital neural networks that are jointly-optimized with the free-space optical processor for the pre- and post-processing<sup>66,91,119,189</sup> of the input/output information.

## References

1. Solli, D. R. & Jalali, B. Analog optical computing. *Nat. Photonics* **9**, 704–706 (2015).
2. Abdollahramezani, S., Hemmatyar, O. & Adibi, A. Meta-optics for spatial optical analog computing. *Nanophotonics* **9**, 4075–4095 (2020).
3. Teğin, U., Yıldırım, M., Oğuz, İ., Moser, C. & Psaltis, D. Scalable optical learning operator. *Nat. Comput. Sci.* **1**, 542–549 (2021).
4. Zhou, Y., Zheng, H., Kravchenko, I. I. & Valentine, J. Flat optics for image differentiation. *Nat. Photonics* **14**, 316–323 (2020).
5. Wang, Z., Chang, L., Wang, F., Li, T. & Gu, T. Integrated photonic metasystem for image classifications at telecommunication wavelength. *Nat. Commun.* **13**, 2131 (2022).
6. Estakhri, N. M., Edwards, B. & Engheta, N. Inverse-designed metasurfaces that solve equations. *Science* **363**, 1333–1338 (2019).
7. Nikkha, V., Tzarouchis, D. C., Hoorfar, A., Engheta, N. Inverse-designed metastructures together with reconfigurable couplers to compute forward scattering. *ACS Photonics* **10**, 977–985 (2022).
8. Zhu, H. H. et al. Space-efficient optical computing with an integrated chip diffractive neural network. *Nat. Commun.* **13**, 1044 (2022).



9. Meng, Y. et al. Optical meta-waveguides for integrated photonics and beyond. *Light. Sci. Appl.* **10**, 235 (2021).
10. Tzarouchis, D. C., Mencagli, M. J., Edwards, B. & Engheta, N. Mathematical operations and equation solving with reconfigurable metadevices. *Light. Sci. Appl.* **11**, 263 (2022).
11. Nikkhah, V. et al. Inverse-designed low-index-contrast structures on silicon photonics platform for vector-matrix multiplication. *Nat. Photonics* (2024).
12. Imani, M. F. et al. Review of metasurface antennas for computational microwave imaging. *IEEE Trans. Antennas Propag.* **68**, 1860–1875 (2020).
13. Kulce, O., Mengu, D., Rivenson, Y. & Ozcan, A. All-optical synthesis of an arbitrary linear transformation using diffractive surfaces. *Light. Sci. Appl.* **10**, 1–21 (2021).
14. Li, J., Hung, Y.-C., Kulce, O., Mengu, D. & Ozcan, A. 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).
15. Li, Y. et al. Universal polarization transformations: spatial programming of polarization scattering matrices using a deep learning-designed diffractive polarization transformer. *Advanced Materials* **35**, 2303395 (2023).
16. Luo, Y. et al. Design of task-specific optical systems using broadband diffractive neural networks. *Light. Sci. Appl.* **8**, 112 (2019).
17. Li, J. et al. Massively parallel universal linear transformations using a wavelength-multiplexed diffractive optical network. *Adv. Photonics* **5**, 016003 (2023).
18. Veli, M. et al. Terahertz pulse shaping using diffractive surfaces. *Nat. Commun.* **12**, 37 (2021).
19. Swanson, G. J. & Veldkamp, W. B. Diffractive optical elements for use in infrared systems. *Opt. Eng.* **28**, 605–608 (1989).
20. Wyrowski, F. Diffractive optical elements: iterative calculation of quantized, blazed phase structures. *JOSA A* **7**, 961–969 (1990).
21. Lalanne, P., Astilean, S., Chavel, P., Cambriil, E. & Launois, H. Design and fabrication of blazed binary diffractive elements with sampling periods smaller than the structural cutoff. *J. Opt. Soc. Am. A* **16**, 1143–1156 (1999).
22. Poleshchuk, A. G. et al. Polar coordinate laser pattern generator for fabrication of diffractive optical elements with arbitrary structure. *Appl. Opt.* **38**, 1295–1301 (1999).
23. Bomzon, Z. E., Biener, G., Kleiner, V. & Hasman, E. Space-variant pancharatanam-berry phase optical elements with computer-generated subwavelength gratings. *Opt. Lett.* **27**, 1141–1143 (2002).
24. Wang, X.-Q., Srivastava, A. K., Chigrinov, V. G. & Kwok, H.-S. Switchable Fresnel lens based on micropatterned alignment. *Opt. Lett.* **38**, 1775–1777 (2013).
25. Chang, C. & Sakdinawat, A. Ultra-high aspect ratio high-resolution nanofabrication for hard X-Ray diffractive optics. *Nat. Commun.* **5**, 4243 (2014).
26. Lin, X. et al. All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018).
27. Zhou, T. et al. Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit. *Nat. Photonics* **15**, 367–373 (2021).
28. Luo, Y. et al. Computational imaging without a computer: seeing through random diffusers at the speed of light. *eLight* **2**, 4 (2022).
29. Smith, D. R., Vier, D., Kroll, N. & Schultz, S. Direct calculation of permeability and permittivity for a left-handed metamaterial. *Appl. Phys. Lett.* **77**, 2246–2248 (2000).
30. Kuester, E. F., Mohamed, M. A., Piket-May, M. & Holloway, C. L. Averaged transition conditions for electromagnetic fields at a metafilm. *IEEE Trans. Antennas Propag.* **51**, 2641–2651 (2003).
31. Falcone, F. et al. Cabinet principle applied to the design of metasurfaces and metamaterials. *Phys. Rev. Lett.* **93**, 197401 (2004).
32. Edwards, B., Alù, A., Silveirinha, M. G. & Engheta, N. Experimental verification of plasmonic cloaking at microwave frequencies with metamaterials. *Phys. Rev. Lett.* **103**, 153901 (2009).
33. Dicken, M. J. et al. Frequency tunable near-infrared metamaterials based on VO<sub>2</sub> phase transition. *Opt. Express* **17**, 18330–18339 (2009).
34. Yu, N. et al. Light propagation with phase discontinuities: generalized laws of reflection and refraction. *Science* **334**, 333–337 (2011).
35. Lin, D., Fan, P., Hasman, E. & Brongersma, M. L. Dielectric gradient metasurface optical elements. *Science* **345**, 298–302 (2014).
36. Arbabi, A., Horie, Y., Bagheri, M. & Faraon, A. Dielectric metasurfaces for complete control of phase and polarization with subwavelength spatial resolution and high transmission. *Nat. Nanotechnol.* **10**, 937–943 (2015).
37. Shi, Y. & Fan, S. Dynamic non-reciprocal meta-surfaces with arbitrary phase reconfigurability based on photonic transition in meta-atoms. *Appl. Phys. Lett.* **108**, 021110 (2016).
38. Khorasaninejad, M. et al. Metalenses at visible wavelengths: diffraction-limited focusing and subwavelength resolution imaging. *Science* **352**, 1190–1194 (2016).
39. Colburn, S., Zhan, A. & Majumdar, A. Metasurface optics for full-color computational imaging. *Sci. Adv.* **4**, eaar2114 (2018).
40. Wang, S. et al. A broadband achromatic metalens in the visible. *Nat. Nanotechnol.* **13**, 227–232 (2018).
41. Chen, W. T. et al. A broadband achromatic metalens for focusing and imaging in the visible. *Nat. Nanotechnol.* **13**, 220–226 (2018).
42. Luo, X. et al. Metasurface-enabled on-chip multiplexed diffractive neural networks in the visible. *Light. Sci. Appl.* **11**, 158 (2022).
43. Desponds, A. et al. Microfabrication by two-photon lithography, and characterization, of SiO<sub>2</sub>/TiO<sub>2</sub> based hybrid and ceramic microstructures. *J. Sol-Gel Sci. Technol.* **95**, 733–745 (2020).
44. Wang, H., Wang, H., Zhang, W. & Yang, J. K. W. Toward near-perfect diffractive optical elements via nanoscale 3d printing. *ACS Nano* **14**, 10452–10461 (2020).
45. Xing, J.-F., Zheng, M.-L. & Duan, X.-M. Two-photon polymerization microfabrication of hydrogels: an advanced 3d printing technology for tissue engineering and drug delivery. *Chem. Soc. Rev.* **44**, 5031–5039 (2015).
46. Geng, Q., Wang, D., Chen, P. & Chen, S.-C. Ultrafast multi-focus 3-d nano-fabrication based on two-photon polymerization. *Nat. Commun.* **10**, 2179 (2019).
47. Krueger, N. A. et al. Porous silicon gradient refractive index micro-optics. *Nano Lett.* **16**, 7402–7407 (2016).
48. Ocier, C. R. et al. Direct laser writing of volumetric gradient index lenses and waveguides. *Light. Sci. Appl.* **9**, 196 (2020).
49. Gale, M. T. Replication techniques for diffractive optical elements. *Microelectron. Eng.* **34**, 321–339 (1997).
50. Lalanne, P., Astilean, S., Chavel, P., Cambriil, E. & Launois, H. Blazed binary subwavelength gratings with efficiencies larger than those of conventional échelette gratings. *Opt. Lett.* **23**, 1081–1083 (1998).
51. Bomzon, Z. E., Kleiner, V. & Hasman, E. Formation of radially and azimuthally polarized light using space-variant subwavelength metal stripe gratings. *Appl. Phys. Lett.* **79**, 1587–1589 (2001).
52. Wu, T.-K. Frequency selective surfaces. *Encyclopedia RF Microwave Engineering* (John Wiley & Sons, 1995).
53. Kildishev, A. V., Boltasseva, A. & Shalaev, V. M. Planar photonics with metasurfaces. *Science* **339**, 1232009 (2013).
54. Kwon, H., Sounas, D., Cordaro, A., Polman, A. & Alù, A. Nonlocal metasurfaces for optical signal processing. *Phys. Rev. Lett.* **121**, 173004 (2018).
55. Murray, W. A. & Barnes, W. L. Plasmonic materials. *Adv. Mater.* **19**, 3771–3782 (2007).

56. Knight, M. W. et al. Aluminum for plasmonics. *ACS Nano* **8**, 834–840 (2014).
57. Yang, J. et al. Photonic crystal fiber metalens. *Nanophotonics* **8**, 443–449 (2019).
58. Huang, L. et al. Dispersionless phase discontinuities for controlling light propagation. *Nano Lett.* **12**, 5750–5755 (2012).
59. Ding, X. et al. Ultrathin Pancharatnam–Berry metasurface with maximal cross-polarization efficiency. *Adv. Mater.* **27**, 1195–1200 (2015).
60. Jahani, S. & Jacob, Z. All-dielectric metamaterials. *Nat. Nanotechnol.* **11**, 23–36 (2016).
61. Zhou, Z. et al. Efficient silicon metasurfaces for visible light. *ACS Photonics* **4**, 544–551 (2017).
62. Qu, G. et al. All-dielectric metasurface empowered optical-electronic hybrid neural networks. *Laser Photonics Rev.* **12**, 2100732 (2022).
63. Devlin, R. C., Ambrosio, A., Rubin, N. A., Mueller, J. P. B. & Capasso, F. Arbitrary spin-to-orbital angular momentum conversion of light. *Science* **358**, 896–901 (2017).
64. Arbabi, A. et al. Miniature optical planar camera based on a wide-angle metasurface doublet corrected for monochromatic aberrations. *Nat. Commun.* **7**, 1–9 (2016).
65. Devlin, R. C., Khorasaninejad, M., Chen, W. T., Oh, J. & Capasso, F. Broadband high-efficiency dielectric metasurfaces for the visible spectrum. *Proc. Natl Acad. Sci.* **113**, 10473–10478 (2016).
66. Mengu, D., Luo, Y., Rivenson, Y. & Ozcan, A. Analysis of diffractive optical neural networks and their integration with electronic neural networks. *IEEE J. Sel. Top. Quantum Electron.* **26**, 1–14 (2019).
67. Liao, D., Chan, K. F., Chan, C. H., Zhang, Q. & Wang, H. All-optical diffractive neural networked terahertz hologram. *Opt. Lett.* **45**, 2906–2909 (2020).
68. Li, J., Mengu, D., Luo, Y., Rivenson, Y. & Ozcan, A. Class-specific differential detection in diffractive optical neural networks improves inference accuracy. *Adv. Photonics* **1**, 046001 (2019).
69. Rahman, M. S. S., Li, J., Mengu, D., Rivenson, Y. & Ozcan, A. Ensemble learning of diffractive optical networks. *Light. Sci. Appl.* **10**, 1–13 (2021).
70. Sagi, O. & Rokach, L. Ensemble learning: a survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **8**, e1249 (2018).
71. Luo, Y., Mengu, D. & Ozcan, A. Cascadable all-optical nand gates using diffractive networks. *Sci. Rep.* **12**, 7121 (2022).
72. Qian, C. et al. Performing optical logic operations by a diffractive neural network. *Light. Sci. Appl.* **9**, 59 (2020).
73. Lin, X. et al. Supplementary Information: all-optical machine learning using diffractive deep neural networks. <https://www.science.org/action/downloadSupplement?doi=10.1126%2Fscience.aat8084&file=aat8084-lin-sm-rev-3.pdf> (accessed 6 September 2018).
74. Feldmann, J. et al. Parallel convolutional processing using an integrated photonic tensor core. *Nature* **589**, 52–58 (2021).
75. Kulce, O., Mengu, D., Rivenson, Y. & Ozcan, A. All-optical information-processing capacity of diffractive surfaces. *Light. Sci. Appl.* **10**, 1–17 (2021).
76. Miller, D. A. Self-configuring universal linear optical component. *Photon. Res.* **1**, 01000001 (2013).
77. Xie, Z., Short, R. T. & Rushforth, C. K. A family of suboptimum detectors for coherent multiuser communications. *IEEE J. Sel. Areas Commun.* **8**, 683–690 (1990).
78. Vojcic, B. R. & Jang, W. M. Transmitter precoding in synchronous multiuser communications. *IEEE Trans. Commun.* **46**, 1346–1355 (1998).
79. Soltanian-Zadeh, H., Windham, J. P. & Peck, D. J. Optimal linear transformation for MRI feature extraction. *IEEE Trans. Med. Imaging* **15**, 749–767 (1996).
80. Wang, W., Yuan, X., Wu, X. & Liu, Y. Fast image dehazing method based on linear transformation. *IEEE Trans. Multimed.* **19**, 1142–1155 (2017).
81. Deisenroth, M. P., Faisal, A. A. & Ong, C. S. *Mathematics for Machine Learning* (Cambridge University Press, 2020).
82. Mengu, D., Zhao, Y., Tabassum, A., Jarrahi, M. & Ozcan, A. Diffractive interconnects: all-optical permutation operation using diffractive networks. *Nanophotonics* **12**, 905–923 (2022).
83. Rahman, M. S. S., Yang, X., Li, J., Bai, B. & Ozcan, A. Universal linear intensity transformations using spatially incoherent diffractive processors. *Light. Sci. Appl.* **12**, 195 (2023).
84. Bai, B. et al. Data class-specific all-optical transformations and encryption. *Adv. Mater.* **35**, 2212091 (2023).
85. Cordaro, A. et al. Solving integral equations in free space with inverse-designed ultrathin optical metagratings. *Nat. Nanotechnol.* **18**, 365–372 (2023).
86. Ji, N., Milkie, D. E. & Betzig, E. Adaptive optics via pupil segmentation for high-resolution imaging in biological tissues. *Nat. Methods* **7**, 141–147 (2010).
87. Johnson-Roberson, M. et al. High-resolution underwater robotic vision-based mapping and three-dimensional reconstruction for archaeology. *J. Field Robot.* **34**, 625–643 (2017).
88. Popoff, S. M. et al. Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media. *Phys. Rev. Lett.* **104**, 100601 (2010).
89. Goodman, J. W. *Speckle Phenomena in Optics: Theory and Applications* (Roberts & Company, 2007).
90. Bai, B. et al. To image, or not to image: class-specific diffractive cameras with all-optical erasure of undesired objects. *eLight* **2**, 14 (2022).
91. Işıl, Ç. et al. Super-resolution image display using diffractive decoders. *Sci. Adv.* **8**, eadd3433 (2022).
92. Sakib Rahman, M. S. & Ozcan, A. Computer-free, all-optical reconstruction of holograms using diffractive networks. *ACS Photonics* **8**, 3375–3384 (2021).
93. Mengu, D. & Ozcan, A. All-optical phase recovery: diffractive computing for quantitative phase imaging. *Adv. Opt. Mater.* **10**, 2200281 (2022).
94. Zhang, Z. et al. 6G wireless networks: vision, requirements, architecture, and key technologies. *IEEE Veh. Technol. Mag.* **14**, 28–41 (2019).
95. Huang, Z. et al. Orbital angular momentum deep multiplexing holography via an optical diffractive neural network. *Opt. Express* **30**, 5569–5584 (2022).
96. Chang, T. et al. Universal metasurfaces for complete linear control of coherent light transmission. *Adv. Mater.* **34**, 2204085 (2022).
97. Ni, X., Kildishev, A. V. & Shalaev, V. M. Metasurface holograms for visible light. *Nat. Commun.* **4**, 2807 (2013).
98. Zheng, G. et al. Metasurface holograms reaching 80% efficiency. *Nat. Nanotechnol.* **10**, 308–312 (2015).
99. Ren, H. et al. Metasurface orbital angular momentum holography. *Nat. Commun.* **10**, 2986 (2019).
100. Li, S.-Q. et al. Phase-only transmissive spatial light modulator based on tunable dielectric metasurface. *Science* **364**, 1087–1090 (2019).
101. She, A., Zhang, S., Shian, S., Clarke, D. R. & Capasso, F. Adaptive metalenses with simultaneous electrical control of focal length, astigmatism, and shift. *Sci. Adv.* **4**, eaap9957 (2018).
102. Arbabi, E. et al. Mems-tunable dielectric metasurface lens. *Nat. Commun.* **9**, 812 (2018).
103. Ee, H.-S. & Agarwal, R. Tunable metasurface and flat optical zoom lens on a stretchable substrate. *Nano Lett.* **16**, 2818–2823 (2016).
104. Li, J. et al. Addressable metasurfaces for dynamic holography and optical information encryption. *Sci. Adv.* **4**, eaar6768 (2018).

105. Malek, S. C., Ee, H.-S. & Agarwal, R. Strain multiplexed metasurface holograms on a stretchable substrate. *Nano Lett.* **17**, 3641–3645 (2017).
106. Mínguez-Vega, G., Mendoza-Yero, O., Lancis, J., Gisbert, R. & Andrés, P. Diffractive optics for quasi-direct space-to-time pulse shaping. *Opt. Express* **16**, 16993–16998 (2008).
107. Weiner, A. M. Femtosecond pulse shaping using spatial light modulators. *Rev. Sci. Instrum.* **71**, 1929–1960 (2000).
108. Divitt, S., Zhu, W., Zhang, C., Lezec, H. J. & Agrawal, A. Ultrafast optical pulse shaping using dielectric metasurfaces. *Science* **364**, 890–894 (2019).
109. Xiao, S., Weiner, A. M. & Lin, C. A dispersion law for virtually imaged phased-array spectral dispersers based on paraxial wave theory. *IEEE J. Quantum Electron.* **40**, 420–426 (2004).
110. Zhao, Y. & Alù, A. Manipulating light polarization with ultrathin plasmonic metasurfaces. *Phys. Rev. B* **84**, 205428 (2011).
111. Yu, N. et al. A broadband, background-free quarter-wave plate based on plasmonic metasurfaces. *Nano Lett.* **12**, 6328–6333 (2012).
112. Wu, P. C. et al. Versatile polarization generation with an aluminum plasmonic metasurface. *Nano Lett.* **17**, 445–452 (2017).
113. Balthasar Mueller, J. P., Rubin, N. A., Devlin, R. C., Groever, B. & Capasso, F. Metasurface polarization optics: independent phase control of arbitrary orthogonal states of polarization. *Phys. Rev. Lett.* **118**, 113901 (2017).
114. Zhao, Y., Belkin, M. A. & Alù, A. Twisted optical metamaterials for planarized ultrathin broadband circular polarizers. *Nat. Commun.* **3**, 870 (2012).
115. Hu, G., Wang, M., Mazor, Y., Qiu, C.-W. & Alù, A. Tailoring light with layered and moiré metasurfaces. *Trends Chem.* **3**, 342–358 (2021).
116. Chang, J., Sitzmann, V., Dun, X., Heidrich, W. & Wetzstein, G. Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification. *Sci. Rep.* **8**, 12324 (2018).
117. Lee, Y.-J., On, M. B., Xiao, X., Proietti, R. & Yoo, S. J. B. Photonic spiking neural networks with event-driven femtojoule optoelectronic neurons based on izhikevich-inspired model. *Opt. Express* **30**, 19360–19389 (2022).
118. Wu, C. et al. Harnessing optoelectronic noises in a photonic generative network. *Sci. Adv.* **8**, eabm2956 (2022).
119. Li, J. et al. Spectrally encoded single-pixel machine vision using diffractive networks. *Sci. Adv.* **7**, eabd7690 (2021).
120. Shan, T., Pan, X., Li, M., Xu, S. & Yang, F. Coding programmable metasurfaces based on deep learning techniques. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **10**, 114–125 (2020).
121. Liu, C. et al. A programmable diffractive deep neural network based on a digital-coding metasurface array. *Nat. Electron.* **5**, 113–122 (2022).
122. Mengu, D., Veli, M., Rivenson, Y. & Ozcan, A. Classification and reconstruction of spatially overlapping phase images using diffractive optical networks. *Sci. Rep.* **12**, 8446 (2022).
123. Wang, H. et al. Intelligent optoelectronic processor for orbital angular momentum spectrum measurement. *Photonix* **4**, 9 (2023).
124. Wu, J., Zhang, Z., Ren, X., Huang, Z. & Wu, X. A broadband electronically mode-reconfigurable orbital angular momentum metasurface antenna. *IEEE Antennas Wirel. Propag. Lett.* **18**, 1482–1486 (2019).
125. He, Q., Sun, S., Xiao, S. & Zhou, L. High-efficiency metasurfaces: principles, realizations, and applications. *Adv. Opt. Mater.* **6**, 1800415 (2018).
126. West, P. R. et al. Searching for better plasmonic materials. *Laser Photonics Rev.* **4**, 795–808 (2010).
127. Boltasseva, A. & Atwater, H. A. Low-loss plasmonic metamaterials. *Science* **331**, 290–291 (2011).
128. Naik, G. V., Shalae, V. M. & Boltasseva, A. Alternative plasmonic materials: beyond gold and silver. *Adv. Mater.* **25**, 3264–3294 (2013).
129. Sun, S. et al. High-efficiency broadband anomalous reflection by gradient meta-surfaces. *Nano Lett.* **12**, 6223–6229 (2012).
130. Miller, D. A. B. Waves, modes, communications, and optics: a tutorial. *Adv. Opt. Photon.* **11**, 679–825 (2019).
131. Khorasaninejad, M. et al. Polarization-insensitive metalenses at visible wavelengths. *Nano Lett.* **16**, 7229–7234 (2016).
132. Chen, M.-H. et al. Polarization-insensitive gan metalenses at visible wavelengths. *Sci. Rep.* **11**, 14541 (2021).
133. Chen, W. T. et al. Dispersion-engineered metasurfaces reaching broadband 90% relative diffraction efficiency. *Nat. Commun.* **14**, 2544 (2023).
134. Mengu, D. et al. At the intersection of optics and deep learning: statistical inference, computing, and inverse design. *Adv. Opt. Photon.* **14**, 209–290 (2022).
135. Ancombe, N. Direct laser writing. *Nat. Photonics* **4**, 22–23 (2010).
136. Fullager, D. B., Boreman, G. D. & Hofmann, T. Infrared dielectric response of nanoscribe Ip-Dip and Ip-L monomers after polymerization from 250 cm<sup>-1</sup> to 6000 cm<sup>-1</sup>. *Opt. Mater. Express* **7**, 888–894 (2017).
137. Oran, D. et al. 3d Nanofabrication by volumetric deposition and controlled shrinkage of patterned scaffolds. *Science* **362**, 1281–1285 (2018).
138. Shaltout, A. M., Shalae, V. M. & Brongersma, M. L. Spatiotemporal light control with active metasurfaces. *Science* **364**, eaat3100 (2019).
139. Lian, C. et al. Photonic (computational) memories: tunable nanophotonics for data storage and computing. *Nanophotonics* **11**, 3823–3854 (2022).
140. Ren, Z. et al. Leveraging of mems technologies for optical metamaterials applications. *Adv. Opt. Mater.* **8**, 1900653 (2020).
141. Guo, Q. et al. Femtojoule femtosecond all-optical switching in lithium niobate nanophotonics. *Nat. Photonics* **16**, 625–631 (2022).
142. Koch, U., Hoessbacher, C., Emboras, A. & Leuthold, J. Optical memristive switches. *J. Electroceram.* **39**, 239–250 (2017).
143. Berggren, K. et al. Roadmap on emerging hardware and technology for machine learning. *Nanotechnology* **32**, 012002 (2021).
144. Wang, Q. et al. Optically reconfigurable metasurfaces and photonic devices based on phase change materials. *Nat. Photonics* **10**, 60–65 (2016).
145. Yang, A. et al. Programmable and reversible plasmon mode engineering. *Proc. Natl Acad. Sci.* **113**, 14201–14206 (2016).
146. Zhang, C. et al. Stretchable all-dielectric metasurfaces with polarization-insensitive and full-spectrum response. *ACS Nano* **14**, 1418–1426 (2020).
147. Sherrott, M. C. et al. Experimental demonstration of >230° phase modulation in gate-tunable graphene–gold reconfigurable mid-infrared metasurfaces. *Nano Lett.* **17**, 3027–3034 (2017).
148. Fan, Y. et al. Graphene plasmonics: a platform for 2D optics. *Adv. Opt. Mater.* **7**, 1800537 (2019).
149. Lee, S., Peng, R., Wu, C. & Li, M. Programmable black phosphorus image sensor for broadband optoelectronic edge computing. *Nat. Commun.* **13**, 1485 (2022).
150. Barton III, D., Lawrence, M. & Dionne, J. Wavefront shaping and modulation with resonant electro-optic phase gradient metasurfaces. *Appl. Phys. Lett.* **118**, 071104 (2021).
151. Tzarouchis, D. C., Edwards, B. & Engheta, N. Programmable Wave-based analog computing machine: a metastructure that designs metastructures. *arXiv:2301.02850* (2022).
152. Tsilipakos, O. et al. Toward intelligent metasurfaces: the progress from globally tunable metasurfaces to software-defined metasurfaces with an embedded network of controllers. *Adv. Opt. Mater.* **8**, 2000783 (2020).

153. Venkatesh, S., Lu, X., Saeidi, H. & Sengupta, K. A programmable terahertz metasurface with circuit-coupled meta-elements in silicon chips: creating low-cost, large-scale, reconfigurable terahertz metasurfaces. *IEEE Antennas Propag. Mag.* **64**, 110–122 (2022).
154. Wang, Z. et al. Resistive switching materials for information processing. *Nat. Rev. Mater.* **5**, 173–195 (2020).
155. Liu, Y. et al. Structural color three-dimensional printing by shrinking photonic crystals. *Nat. Commun.* **10**, 4340 (2019).
156. Avayu, O., Almeida, E., Prior, Y. & Ellenbogen, T. Composite functional metasurfaces for multispectral achromatic optics. *Nat. Commun.* **8**, 14992 (2017).
157. Zhou, Y. et al. Multilayer noninteracting dielectric metasurfaces for multiwavelength metaoptics. *Nano Lett.* **18**, 7529–7537 (2018).
158. Groever, B., Chen, W. T. & Capasso, F. Meta-lens doublet in the visible region. *Nano Lett.* **17**, 4902–4907 (2017).
159. Johnson, R. W., Hultqvist, A., Bent, S. F. & Brief, A. Review of atomic layer deposition: from fundamentals to applications. *Mater. Today* **17**, 236–246 (2014).
160. Zhang, X., Chen, K. S., Ghodssi, R., Ayón, A. A. & Spearing, S. M. Residual stress and fracture in thick tetraethylorthosilicate (Teos) and silane-based pecvd oxide films. *Sens. Actuators A: Phys.* **91**, 373–380 (2001).
161. Xie, H., Fedder, G. K. & Sulouff, R. E. 2.05 - Accelerometers. In *Comprehensive Microsystems* (Eds. Gianchandani, Y. B.; Tabata, O.; Zappe, H.) 135–180 (Elsevier: Oxford, 2008).
162. She, A., Zhang, S., Shian, S., Clarke, D. R. & Capasso, F. Large area metalenses: design, characterization, and mass manufacturing. *Opt. Express* **26**, 1573–1585 (2018).
163. Totzeck, M., Ulrich, W., Göhnermeier, A. & Kaiser, W. Pushing deep ultraviolet lithography to its limits. *Nat. Photonics* **1**, 629–631 (2007).
164. Cecil, T., Peng, D., Abrams, D., Osher, S. J. & Yablonovitch, E. Advances in inverse lithography. *ACS Photonics* **10**, 910–918 (2022).
165. Einck, V. J. et al. Scalable nanoimprint lithography process for manufacturing visible metasurfaces composed of high aspect ratio TiO<sub>2</sub> meta-atoms. *ACS Photonics* **8**, 2400–2409 (2021).
166. Lee, M. H., Huntington, M. D., Zhou, W., Yang, J.-C. & Odom, T. W. Programmable soft lithography: solvent-assisted nanoscale embossing. *Nano Lett.* **11**, 311–315 (2011).
167. Mengu, D., Rivenson, Y. & Ozcan, A. Scale-, shift-, and rotation-invariant diffractive optical networks. *ACS Photonics* **8**, 324–334 (2021).
168. Mengu, D. et al. Misalignment resilient diffractive optical networks. *Nanophotonics* **9**, 4207–4219 (2020).
169. Khoram, E. et al. Nanophotonic media for artificial neural inference. *Photon. Res.* **7**, 823–827 (2019).
170. Lim, J. & Psaltis, D. Maxwellnet: physics-driven deep neural network training based on Maxwell's equations. *APL Photonics* **7**, 011301 (2022).
171. Jin, J.-M. *The Finite Element Method in Electromagnetics* (John Wiley & Sons, 2015).
172. Kunz, K. S., & Luebbers, R. J. *The Finite Difference Time Domain Method for Electromagnetics* (CRC press, 1993).
173. Chen, W. T., Zhu, A. Y. & Capasso, F. Flat optics with dispersion-engineered metasurfaces. *Nat. Rev. Mater.* **5**, 604–620 (2020).
174. Moughames, J. et al. Three-dimensional waveguide interconnects for scalable integration of photonic neural networks. *Optica* **7**, 640–646 (2020).
175. Bubeck, S. et al. Sparks of Artificial General Intelligence: Early Experiments with Gpt-4. (2023).
176. MacLennan, B. J. A Review of Analog Computing. *Technical Report UT-CS-07-601* (Department of Electrical Engineering & Computer Science, University of Tennessee, 2007).
177. MacLennan, B. J. The promise of analog computation. *Int. J. Gen. Syst.* **43**, 682–696 (2014).
178. Zangeneh-Nejad, F., Sounas, D. L., Alù, A. & Fleury, R. Analogue computing with metamaterials. *Nat. Rev. Mater.* **6**, 207–225 (2021).
179. Wetzstein, G. et al. Inference in artificial intelligence with deep optics and photonics. *Nature* **588**, 39–47 (2020).
180. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.* **2**, 303–314 (1989).
181. Leshno, M., Lin, V. Y., Pinkus, A. & Schocken, S. Multilayer feed-forward networks with a nonpolynomial activation function can approximate any function. *Neural Netw.* **6**, 861–867 (1993).
182. Voigtlaender, F. The universal approximation theorem for complex-valued neural networks. *Appl. Comput. Harmon. Anal.* **64**, 33–61 (2023).
183. Yan, T. et al. Fourier-space diffractive deep neural network. *Phys. Rev. Lett.* **123**, 023901 (2019).
184. Zuo, Y. et al. All-optical neural network with nonlinear activation functions. *Optica* **6**, 1132–1137 (2019).
185. Zhang, D. et al. Broadband nonlinear modulation of incoherent light using a transparent optoelectronic neuron array. *arXiv:2304.13298* (2023).
186. Bishara, W., Zhu, H. & Ozcan, A. Holographic opto-fluidic microscopy. *Opt. Express* **18**, 27499–27510 (2010).
187. Rahman, M. S. S. & Ozcan, A. Time-lapse image classification using a diffractive neural network. *Adv. Intell. Syst.* **5**, 2200387 (2023).
188. Khorasaninejad, M. et al. Multispectral chiral imaging with a metalens. *Nano Lett.* **16**, 4595–4600 (2016).
189. Rahman, M. S. et al. Learning diffractive optical communication around arbitrary opaque occlusions. *Nat. Commun.* **14**, 6830 (2023).
190. Hernandez-Serrano, A. I., Weidenbach, M., Busch, S. F., Koch, M., & Castro-Camus, E. Terahertz Grin Lenses Fabricated by 3D-printing. In *Proc. 41st International Conference on Infrared, Millimeter, and Terahertz waves (IRMMW-THz)*, 25–30 Sept 2016, 1–2 (IEEE, 2016).
191. Han, J., Lakhtakia, A. & Qiu, C.-W. Terahertz metamaterials with semiconductor split-ring resonators for magnetostatic tunability. *Opt. Express* **16**, 14390–14396 (2008).
192. Zeng, Y. et al. Laser-printed terahertz plasmonic phase-change metasurfaces. *Adv. Opt. Mater.* **11**, 2202651 (2023).
193. Yang, Q. et al. Broadband terahertz rotator with an all-dielectric metasurface. *Photon. Res.* **6**, 1056–1061 (2018).
194. Leitis, A., Tseng, M. L., John-Herpin, A., Kivshar, Y. S. & Altug, H. Wafer-scale functional metasurfaces for mid-infrared photonics and biosensing. *Adv. Mater.* **33**, 2102232 (2021).
195. Balli, F., Sultan, M., Lami, S. K. & Hastings, J. T. A hybrid achromatic metalens. *Nat. Commun.* **11**, 3892 (2020).
196. Hernandez-Serrano, A. I., Weidenbach, M., Busch, S. F., Koch, M. & Castro-Camus, E. Fabrication of gradient-refractive-index lenses for terahertz applications by three-dimensional printing. *J. Opt. Soc. Am. B* **33**, 928–931 (2016).
197. Li, T., Wang, H., Ling, F., Zhong, Z. & Zhang, B. High-efficiency terahertz metasurface with independently controlled and switchable function in transmission and reflection modes. *Superlattices Microstruct.* **146**, 106653 (2020).
198. Dylla-Spears, R. et al. 3D printed gradient index glass optics. *Sci. Adv.* **6**, eabc7429 (2020).
199. Garcia, N. C. & Chisum, J. D. High-efficiency, wideband grin lenses with intrinsically matched unit cells. *IEEE Trans. Antennas Propag.* **68**, 5965–5977 (2020).
200. Ji, Y. et al. Active terahertz beam deflection based on a phase gradient metasurface with liquid crystal-enhanced cavity mode conversion. *Opt. Express* **31**, 1269–1281 (2023).

201. Li, J. et al. Unidirectional imaging using deep learning–designed materials. *Sci. Adv.* **9**, eadg1505 (2023).
202. Cheng, Q. et al. Broadband achromatic metalens in terahertz regime. *Sci. Bull.* **64**, 1525–1531 (2019).
203. Lin, M.-Y., Chuang, C.-H., Chou, T.-A. & Chen, C.-Y. A theoretical framework for general design of two-materials composed diffractive Fresnel lens. *Sci. Rep.* **11**, 15466 (2021).
204. Matz, G., Messerschmidt, B. & Gross, H. Design and evaluation of new color-corrected rigid endomicroscopic high Na grin-objectives with a sub-micron resolution and large field of view. *Opt. Express* **24**, 10987–11001 (2016).
205. Zhan, T. et al. Practical chromatic aberration correction in virtual reality displays enabled by cost-effective ultra-broadband liquid crystal polymer lenses. *Adv. Opt. Mater.* **8**, 1901360 (2020).
206. Wang, S. et al. Broadband achromatic optical metasurface devices. *Nat. Commun.* **8**, 187 (2017).
207. Lan, F. et al. Real-time programmable metasurface for terahertz multifunctional wave front engineering. *Light.: Sci. Appl.* **12**, 191 (2023).
208. Guo, J. et al. Reconfigurable terahertz metasurface pure phase holograms. *Adv. Opt. Mater.* **7**, 1801696 (2019).
209. Sypek, M. et al. Highly efficient broadband double-sided Fresnel lens for THz range. *Opt. Lett.* **37**, 2214–2216 (2012).
210. Zhou, F. et al. Additive manufacturing of a 3D terahertz gradient-refractive index lens. *Adv. Optical Mater.* **4**, 1034–1040 (2016).
211. Yang, C. et al. Terahertz planar lenses based on plasmonic metasurfaces. *Phys. Lett. A* **383**, 789–792 (2019).
212. Liu, L. et al. Broadband metasurfaces with simultaneous control of phase and amplitude. *Adv. Mater.* **26**, 5031–5036 (2014).
213. Visconti, A. & Bentley, J. Design of radial gradient-index lenses for dual-band visible to short-wave infrared imaging systems. *Opt. Eng.* **52**, 112106 (2013).
214. Sun, P. et al. Ultra-broadband holography in visible and infrared regions with full-polarization nondispersive response. *Opt. Lett.* **48**, 3083–3086 (2023).
215. Khorasaninejad, M., Ambrosio, A., Kanhaiya, P. & Capasso, F. Broadband and chiral binary dielectric meta-holograms. *Sci. Adv.* **2**, e1501258 (2016).
216. Oliva, M., Michaelis, D., Fuchs, F., Tünnermann, A. & Zeitner, U. D. Highly efficient broadband blazed grating in resonance domain. *Appl. Phys. Lett.* **102**, 203114 (2013).
217. Li, Z., Palacios, E., Butun, S. & Aydin, K. Visible-frequency metasurfaces for broadband anomalous reflection and high-efficiency spectrum splitting. *Nano Lett.* **15**, 1615–1621 (2015).
218. Sun, Y. et al. Broadband anomalous reflective metasurface for complementary conversion of arbitrary incident polarization angles. *Opt. Express* **29**, 38404–38414 (2021).
219. Mo, W., Wei, X., Wang, K., Li, Y. & Liu, J. Ultrathin flexible terahertz polarization converter based on metasurfaces. *Opt. Express* **24**, 13621–13627 (2016).
220. Zang, X. et al. Polarization encoded color image embedded in a dielectric metasurface. *Adv. Mater.* **30**, 1707499 (2018).

## Acknowledgements

The Ozcan Research Group at UCLA acknowledges the support of the U.S. Department of Energy (DOE), Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Award # DE-SC0023088. The Engheta Group at the University of Pennsylvania acknowledges partial support from the Air Force Office of Scientific Research (AFOSR) Multidisciplinary University Research Initiative (MURI) grant # FA9550-21-1-0312, and the AFOSR grant # FA9550-23-1-0307.

## Author contributions

N.E. and A.O. supervised the research. J.H. and A.O. conceived the project, coordinated the writing process, wrote parts of the paper, and edited all sections. All the authors contributed to the preparation of the manuscript; D.M., D.C.T., B.E., and N.E. wrote parts of the manuscript and edited all sections.

## Competing interests

The authors have pending and issued patent applications on analog optical computing systems, including U.S. Patent No. 11,392,830, and US Patent No. 11,494,461 B2.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-45982-w>.

**Correspondence** and requests for materials should be addressed to Aydogan Ozcan.

**Peer review information** *Nature Communications* thanks Jianji Dong and the other anonymous reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024