

## 基于深度学习算法的计算光学研究进展

王一同, 周宏强, 闫景逍, 合聪, 黄玲玲\*

北京理工大学光电学院北京市混合现实与新型显示工程技术中心, 北京 100081

**摘要** 随着大数据时代的来临,深度学习被广泛地应用于图像处理、语音识别、自动驾驶等领域。作为一种数据驱动的算法,深度学习应用于光学领域可以有效地提高计算效率和成像质量,接近甚至突破物理极限,并且以光子为媒介诞生的光学神经网络能够突破传统电子神经网络的限制,具有高速、低损耗的优势。通过实例详细分析了深度学习在微纳结构设计与光谱响应预测、全息成像应用、光学感知与成像技术、新型光子驱动神经网络等方向的应用,并分析了深度学习与光学结合存在的挑战,对该领域的发展方向与前景进行了展望。

**关键词** 光计算;深度学习;神经网络;超表面;计算光学

中图分类号 TN202

文献标志码 A

doi: 10.3788/CJL202148.1918004

### 1 引言

深度学习是当今发展最快的机器学习方法之一,谷歌公司(Google)开发的“AlphaGo”在人机对战中击败了职业围棋选手而被人们关注。它的优势十分明显,利用计算机强大的计算能力,结合训练好的神经网络,能够快速、准确地对输入的数据进行响应,从而避免了耗时且过量的计算。因此,凭借强大的泛化能力和良好的鲁棒性,深度学习在自动驾驶<sup>[1]</sup>、语音翻译<sup>[2]</sup>、图像处理<sup>[3]</sup>、生物医学<sup>[4]</sup>等领域均有着较为广泛的应用。

早在 1943 年,麦卡洛克和皮兹发表论文《神经活动中内在思想的逻辑演算》,提出了 McCulloch-Pitts(MP)模型<sup>[5]</sup>。MP 模型是一个模仿神经元的结构和生物特点,基于神经网络的数学模型,本质上是一种“模拟人类大脑”的神经元模型。这种模仿生物神经元工作模式的思想与方法为人工神经网络奠定了基础。1958 年,美国科学家 Rosenblatt<sup>[6]</sup>提出了一种两层神经元构成的神经网络,其被称为“感知机”。感知机是一种线性分析网络,可以对输入的训练集数据进行二分类;并且它能够在训练集中自动更新权值,是一种较为简易的人工神经网络。然而由于硬件计算能力有限,神经网络的发展进入瓶颈。

直到 1995 年支持向量机(SVM)的出现,突破性地解决非线性问题,并且可以使用核函数对高维空间进行映射,在小样本的条件下分类效果较好<sup>[7]</sup>。随着计算机技术的发展,中央处理器(CPU)和图形处理器(GPU)运算以及处理图像的能力随之变得更加强大,神经网络的性能也相应获得了提升。

人工神经网络的基本任务是模仿生物神经元和人脑的工作原理,以学习某一过程内在的原理或提取目标特征等。随着研究人员对深度学习研究的日益加深,各种复杂、混合的神经网络被提出并面向不同的应用场景。任务目标不同往往需要不同的网络架构。常见的基本神经网络包括全连接神经网络(FCNN)、卷积神经网络(CNN)、循环神经网络(RNN)、生成对抗神经网络(GAN)、深度置信网络(DBN)等。其中全连接神经网络多用于对数据按标签进行分类<sup>[8]</sup>;卷积神经网络可以提取图像的局部特征<sup>[9]</sup>;生成对抗神经网络的优势在于可以学习真实数据的特点并生成新的同类数据<sup>[10]</sup>。基于几种网络的特点和优势,全连接神经网络、卷积神经网络和生成对抗网络可以单独或多个结合使用,处理不同的任务目标<sup>[11-13]</sup>。深度学习具有从大量数据中学习特征规律的能力。基于这个优势,深度学习在光学领域也有着很好的应用前景,可被应用于纳米

收稿日期: 2021-07-02; 修回日期: 2021-08-03; 录用日期: 2021-08-16

基金项目: 国家自然科学基金(61775019, 92050117)、北京卓越青年科学家项目(BJJWZYJH01201910007022)

通信作者: \*huanglingling@bit.edu.cn

结构共振光谱预测<sup>[14]</sup>、超分辨率成像<sup>[15-16]</sup>、微纳结构参数设计<sup>[17]</sup>、图像处理<sup>[18]</sup>、全息成像<sup>[19]</sup>、新型光学神经网络<sup>[20]</sup>等方面。

本文首先介绍深度学习应用于微纳结构设计与光谱响应预测的实例,深度学习与全息相结合的典型应用;接着以实例分析了神经网络用于光学感知与成像技术的优势;最后介绍了新型的光子驱动的衍射神经网络,并对深度学习应用在光学中的发展趋势进行了展望。

## 2 光谱响应预测与微纳结构设计

### 2.1 光谱响应预测

根据已有的微纳结构预测光谱响应是微纳光学研究中的重要环节。虽然截止到目前,研究人员可以通过有限元法(FEM)、时域有限差分(FDTD)法、严格耦合波分析等复杂的模拟算法来计算微纳结构的光谱响应及其性能,但存在超大参数空间结构扫描计算耗时、占用内存大等问题<sup>[21-23]</sup>。通过深度学习辅助设计,可以大大减少计算模拟的运算量,从而提高微纳结构光谱响应预测的效率。

图 1 中的工作将任意形状的纳米结构映射到三维网格上,通过耦合偶极子近似法模拟训练数据,利用计算得到的内部场去训练神经网络模型,以此预测纳米结构内部的电极化密度,无需大量计算即可获得近场电磁响应、远场散射模式、坡印廷矢量等各种物理量<sup>[24]</sup>。结果表明,网络预测只有大约 5% 左右的误差。值得注意的是,训练好的 U-Net 神经网络<sup>[25]</sup>可以学习 3D 微纳结构中的电磁响应原理。通过足够的训练,该网络对麦克斯韦方程组也有一定的学习能力,能够准确预测微纳结构的电场响应。

从上述工作可知,数据驱动深度学习可以模拟电磁场传播和分布<sup>[26]</sup>,从而预测结构性质,并且误差可以控制在较低水平。神经网络可以将预测速度提升几个数量级。深度学习在结构响应预测中的作用可以用数值模拟等手段代替,但是对于一个需求的结构,设计过程需要大量时间和计算。这种为了得到理想的共振响应去设计微纳结构的过程被称为“逆设计问题”。为了避免逆向设计过程中大量的试错计算或是猜测所消耗的时间与计算资源,常采用伴随算法、遗传算法等逆设计方法来解决逆设计问题。随着深度学习的发展,这种数据驱动的高速算法已经开始作为解决逆问题的一种手段。

### 2.2 逆向设计微纳结构

早期神经网络的输入和输出均是数组,为了增

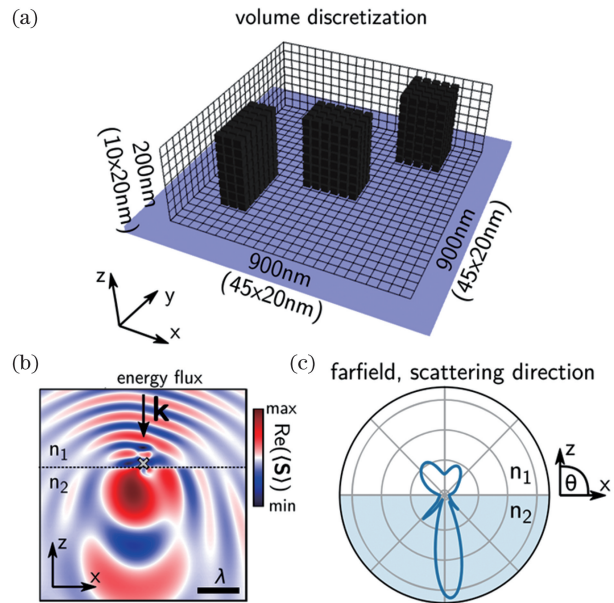


图 1 神经网络预测微纳结构响应<sup>[24]</sup>。(a) 3D 微纳结构建模示意图;(b) 神经网络预测的坡印廷矢量;(c) 神经网络预测的远场散射模式

Fig. 1 Micro-nano structure response predicted by neural network<sup>[24]</sup>. (a) Schematic of 3D micro-nano structure modeling; (b) Poynting vector acquired by neural network predicting; (c) far-field scattering model acquired by neural network predicting

加网络预测的准确度并减少计算时间,早期的网络多是针对固定的微纳结构,以部分结构参数为变量进行设计的。图 2(a)为目标结构,是 20 层 SiO<sub>2</sub> 与 Si<sub>3</sub>N<sub>4</sub> 相间的多层膜结构。微纳结构逆设计问题存在非唯一解情况,以 6 层结构为例,结构 A 与结构 B 的透射共振光谱接近重合,这将导致在训练数据过程中存在输入相同但输出标签不同的情况<sup>[27]</sup>。以各层膜厚为输入,以结构透射光谱响应为输出的全连接网络称为正向网络;以结构透射光谱响应为输入,结构设计参数为输出的全连接网络称为逆向网络。测试中,先以数值模拟计算得到的 500000 组数据作为训练集,另外 50000 组数据作为测试集,分别对正向网络和逆向网络进行训练。为了避免逆向网络不收敛,在训练逆向网络的过程中,考虑采用 500000 组数据的全训练集和只使用其中 26.1% 的一一映射的过滤训练集进行训练。将训练好的逆向网络与正向网络级联,如图 2(b)所示。在级联网络的训练中,先固定正向网络的权重,调整逆向网络的权重,以此减小输出响应与输入响应之间的误差。图 2(c)、(d)分别是对于同一个光谱响应,两个网络的输出结构仿真后的光谱响应曲线,证明了过滤训

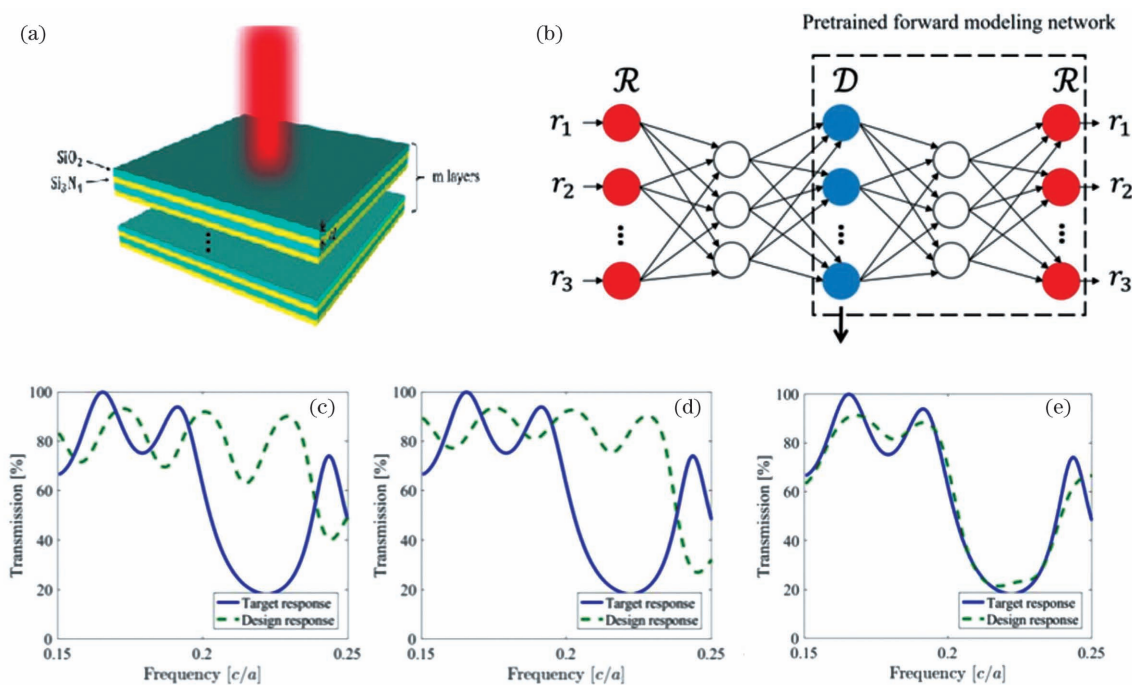


图 2 全连接神经网络应用于微纳结构的逆设计问题<sup>[27]</sup>。(a)多层膜结构参数示意图;(b)级联网络图;(c)(d)未过滤训练网络和过滤训练网络的预测结构对应的响应曲线;(e)级联网络的预测结构输出的光谱响应曲线

Fig. 2 Fully-connected neural network for inverse design of micro-nano structures<sup>[27]</sup>. (a) Schematic of parameters of structure with multi-layer film; (b) cascading network; (c) response curves corresponding to the prediction structures of unfiltered training network and filtered training network; (e) spectral response of the prediction structure of the cascade network

练数据的方法难以解决非唯一解对逆向网络训练影响的问题。图 2(e)展示了级联网络中逆向网络的学习曲线和预测效果,相较于单独训练逆向网络,级联网络收敛速度更快,预测能力也更强。该工作尝试解决了微纳结构逆设计中非唯一解问题对大训练集网络的束缚,并提供了一种新的训练逆向网络的思路。虽然准确率相较于物理仿真模型还有差距,但为之后的逆设计问题研究提供了方向。

除了可以预测多层膜结构的响应与逆设计问题,深度学习还可以预测如图 3(a)所示的多层球状纳米粒子结构<sup>[28]</sup>。该结构为 5 层 SiO<sub>2</sub> 和 TiO<sub>2</sub> 材料交替的纳米球,中心为 SiO<sub>2</sub>,完成了窄波段和宽波段范围内最大散射率结构的设计;同时还可以通过光谱响应预测周期型阵列结构,与超表面结合,实现不同功能。利用具有旋光性的光谱预测手性超表面结构的参数<sup>[29]</sup>,该手性超材料表现出圆二色性,即对左、右旋圆偏振光具有吸收差,如图 3(b)所示;利用深度学习对“H”型超表面结构的透射光谱进行预测,可以实现不同偏振、不同共振峰光谱预测及逆设计<sup>[30]</sup>,如图 3(c)所示;根据光谱响应预测多种三维阵列超表面结构,并实现超表面滤波器等功能<sup>[31]</sup>,一定程度上增加了结构多样性和增强了泛化

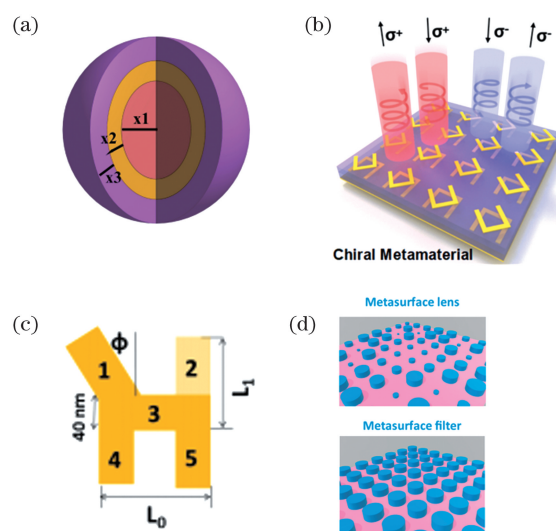


图 3 基于响应预测微纳结构的研究<sup>[28-31]</sup>。(a)多层纳米粒子结构;(b)手性材料超表面结构;(c)“H”型超表面结构;(d)电磁响应预测阵列超表面结构

Fig. 3 Research on micro-nano structure based on response prediction<sup>[28-31]</sup>. (a) Multilayer nanoparticle structure; (b) chiral metasurface structure; (c) H-type metasurface structure; (d) electromagnetic response prediction array metasurface structure

能力,但是由于设计参数有限,无法实现高自由度的按需结构设计,如图 3(d)所示。

上述工作采用单一的神经网络或是级联神经网络,以保证网络训练的稳定性,最终让网络收敛到一个可行的解。这种方法虽然一定程度上确保了预测的准确性,但是牺牲了结构的多样性,并且泛化能力不足。相比于传统神经网络和级联网络,生成对抗神经网络的优势在于可以学习样本中的特征,从而对应生成新的同类样本<sup>[32]</sup>。基于生成对抗神经网络的自生成特性,将其应用于微纳结构逆设计问题中时可以大大提高设计结构的多样性<sup>[33]</sup>。

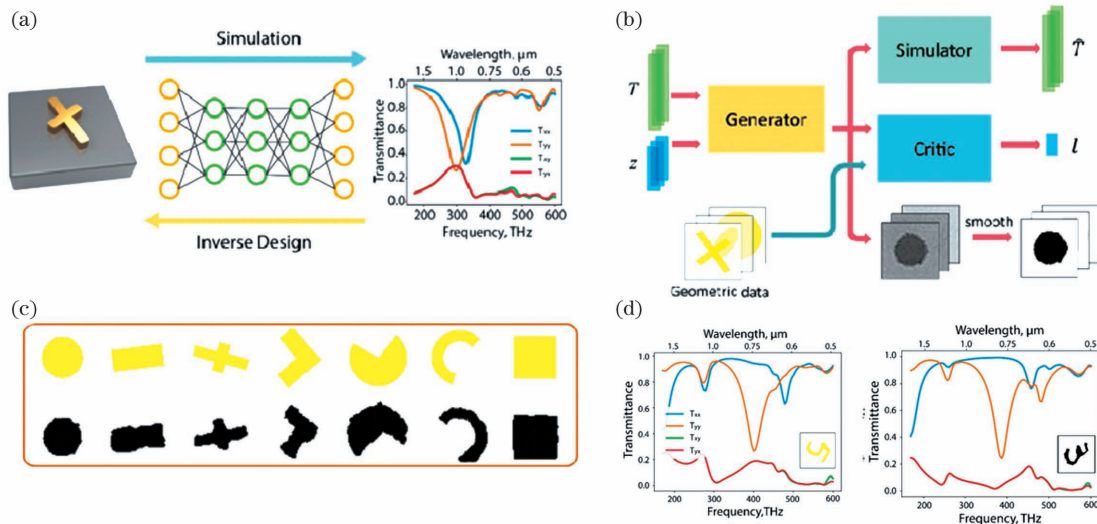


图 4 生成对抗神经网络进行不同结构的光谱曲线预测和逆向设计<sup>[34]</sup>。(a)结构预测及逆向设计示意图;(b)生成对抗神经网络原理图;(c)需求结构与网络预测结构对比;(d)需求结构与预测结构光谱响应对比

Fig. 4 Spectrum prediction and inverse design of different structures using generating adversarial neural network<sup>[34]</sup>. (a) Structure prediction and inverse design; (b) principle of generating adversarial neural network; (c) comparison of target structure and prediction structure; (d) comparison of spectrum response of target structure and prediction structure

该课题组<sup>[35]</sup>于 2021 年还提出了一种多功能超表面神经网络系统,该工作增加了网络结构的复杂度,同时也将单层的周期型结构调整为多层非周期型结构,原理如图 5 所示。系统本身是生成对抗网络,生成对抗网络的生成器是一个组合模式生成网络(CPGN),其作用是解码一些特征向量并生成相应的备选纳米结构;然后神经网络生成器计算备选纳米结构的透射光谱和反射光谱;接着 evolutionary strategy(ES)性能优化器选择其中性能符合要求的纳米结构参数。在整个训练过程的每次迭代中,不同层上不同位置的超材料单元都能得到优化,整个超表面系统的设计以循环的方式进行,直到整个系统的性能令人满意。该系统可以用作多功能光学器件,实现偏振复用双功能光束发生器、用于全光计算的二阶微分器、空间偏振波长复用全息等功能,这些

佐治亚理工大学的 Cai 课题组<sup>[34]</sup>在 2018 年使用生成对抗神经网络进行不同结构的光谱曲线预测和逆向设计,如图 4 所示。生成对抗神经网络能够随机生成图案模型,输出是结构的光谱曲线,从网络的输出端到输入端,根据光谱曲线输出结构图。值得注意的是,相比于传统神经网络,生成对抗神经网络可以不局限于固定结构,从而生成多种同类结构,增加结构预测的泛化性与目标解的多样性。但是由于不同结构可能有相似的光谱响应,该网络具有非唯一解,同时结构的复杂度对网络的优化能力也存在考验。

功能都在仿真中得到验证,用于诸如多功能光束整形、全光计算、空间波长-偏振复用全息术等应用。

该工作可以实现对微纳结构的多种功能的逆设计,对于目标光谱或电磁场空间分布,可以准确地找到相应的超表面系统组合。但是对于现阶段基于深度学习的微纳结构逆设计中的根据光谱曲线预测结构参数的非唯一解问题并没有很好解决。这种非唯一的问题也是神经网络与微纳设计相结合中需要攻克的难题。

为了提高微纳设计效率的同时保留结构设计的泛化能力,尝试降低微纳结构的参数维度,实现训练数据的简化,并达到数据与标签一一映射的效果。Ma 等<sup>[36]</sup>利用变分自编码器(VAE)编码超材料的结构及其光学响应,使得在预测复杂的结构性能关系时无需大量的数据。变分自编码器可以通过编码

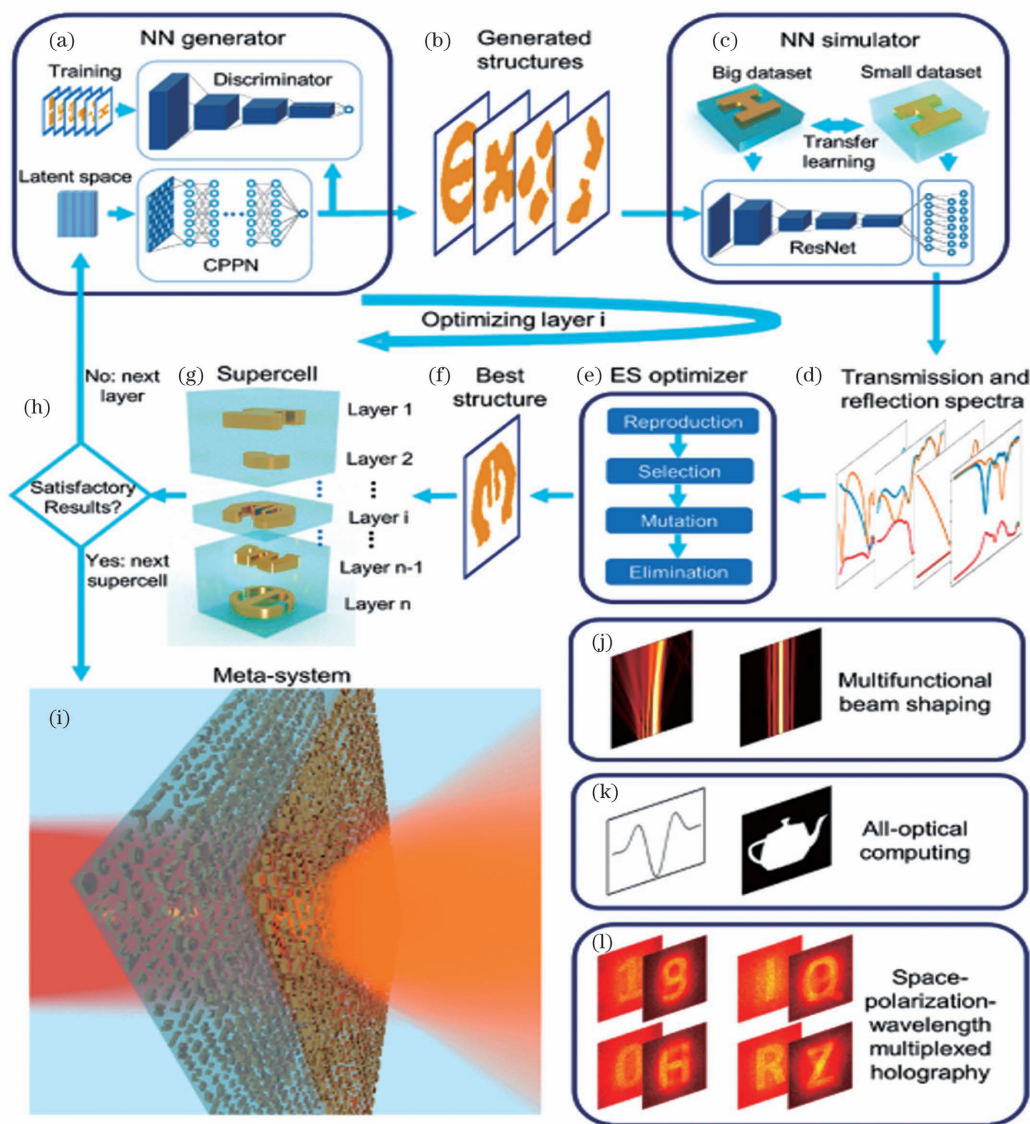


图 5 多功能超表面神经网络系统<sup>[35]</sup>。(a)~(h)神经网络系统结构示意图;(i)超表面系统;(j)~(l)多功能光束整形、全光计算和空间波长-偏振复用全息应用

Fig. 5 Multi-functional metasurface neural network system<sup>[35]</sup>. (a)–(h) Structure diagram of neural network system; (i) metasurface system; (j)–(l) application including multifunctional beam shaping, all-optical computation, and spatial wavelength-polarization multiplexing holography, respectively

器网络将真实样本转换为一个理想的数据分布,然后将这个数据分布传递给一个解码器网络,得到生成样本,当生成样本与真实样本足够接近时,自编码器模型训练完成。图 6(a)显示了超表面设计和表征的深层生成模型。这种超表面及其光学响应被编码到相同的隐空间中,这样类似的设计和光响应就自动聚集在一起。在解码过程中,通过对给定需求的解空间进行采样,生成候选设计。图 6(b)展示了通过所提生成模型进行超表面按需设计的一个例子(所需的频率为 40~100 THz 的反射光谱)和输出获得的超表面设计(模拟计算所得光谱),通过对比,可以看到设计的光谱达到了预期目标。总的来说,

前向预测可以归结为一个回归问题,为了解决逆向问题而引入了隐空间,根据先验知识,将超表面结构与光学响应一起编码。而逆向设计可以看成是一个生成过程,首先从隐空间中采样得到初始变量,然后根据采样的变量生成特定的结构。该模型提高了结构逆设计的多样性和泛化能力,相比生成对抗模型,变分自编码器的优势在于采用预训练的模拟器来保证生成样本的可行性,从而提高了模型的可解释性。

此外,应用自编码器(AE)和主成分分析对光子数据集进行分析,可以表明结构参数与其光学响应之间的内在物理联系<sup>[37]</sup>。在微纳结构设计中,不同的微纳结构可能会对应相似的光谱响应,可以将传

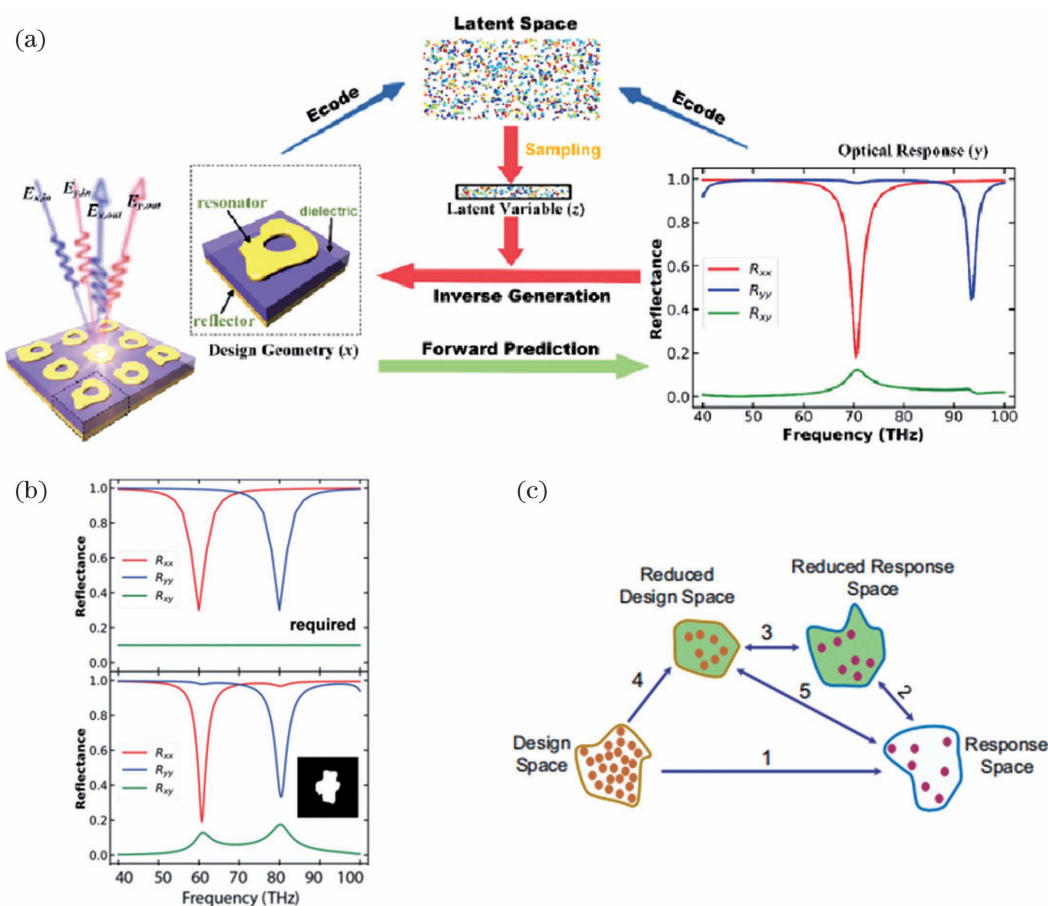


图 6 深度学习降低纳米结构数据维度<sup>[36-37]</sup>。(a)VAE 编码超材料结构,光谱响应至潜在空间并取样解码;(b)VAE 网络按需求设计的例子,频率为 40~100 THz 的反射光谱;(c)结构-响应一对一映射

Fig. 6 Deep learning reduces the data dimension of the nano-structure<sup>[36-37]</sup>. (a) VAE encoding the metamaterial structure and spectral decoding in latent space; (b) VAE network design on demand, reflection spectra in the 40-100 THz frequency; (c) one-to-one mapping of the structure to the response

统多对一设计问题重新划分为一对一映射问题和能够用解析公式解决的多对一问题,从而在微纳结构逆设计问题中将计算复杂度降低几个数量级,如图 6(c)所示。

### 2.3 拓扑优化结构设计

拓扑优化是一种给定优化目标、性能指标、约束条件等的情况下,调整材料或结构在给定空间内分布情况的优化方法<sup>[38]</sup>。拓扑优化可以应用于微纳结构设计中,提高设计的泛化能力,常用的优化算法有梯度优化算法<sup>[39]</sup>、遗传算法<sup>[40]</sup>等。深度学习作为一种强大的优化方法,将深度学习与拓扑学结合在微纳逆设计问题中有很好的发展前景<sup>[41-43]</sup>。

拓扑优化可以在局部优化的情况下执行,其中局部梯度直接用于器件结构的梯度下降,并使用全局拓扑优化网络(GLOnets)进行全局优化,将局部梯度计算与生成神经网络训练相结合,进行基于全

局的优化<sup>[44]</sup>。图 7(a)为重新参数化的 GLOnets 流程图。初始量的分布由生成神经网络产生,每一组初始向量对应一组设计参数的宽度向量,这个宽度向量对应实际的结构尺寸。由于结构制作过程存在光刻分辨率、刻蚀长宽比等物理限制,形成了对刻蚀样品的最小特征尺寸的几何约束,因此在参数向量与物理结构映射的过程中需要添加几何约束条件。最初在一个没有约束的解空间中进行设计,然后使用解析方法将这些结构转换为具有强约束的物理空间。这些结构在物理空间中进行评估和优化,然后映射回它们的解空间。利用反向传播迭代更新网络权值,使损失函数最小化。图 7(b)展示了整个系统效率的直方图和纳米结构设计分布随迭代次数的变化趋势。在优化过程中,纳米结构的分布范围变窄并且效率更高,从而整个结构的性能得到了提升。

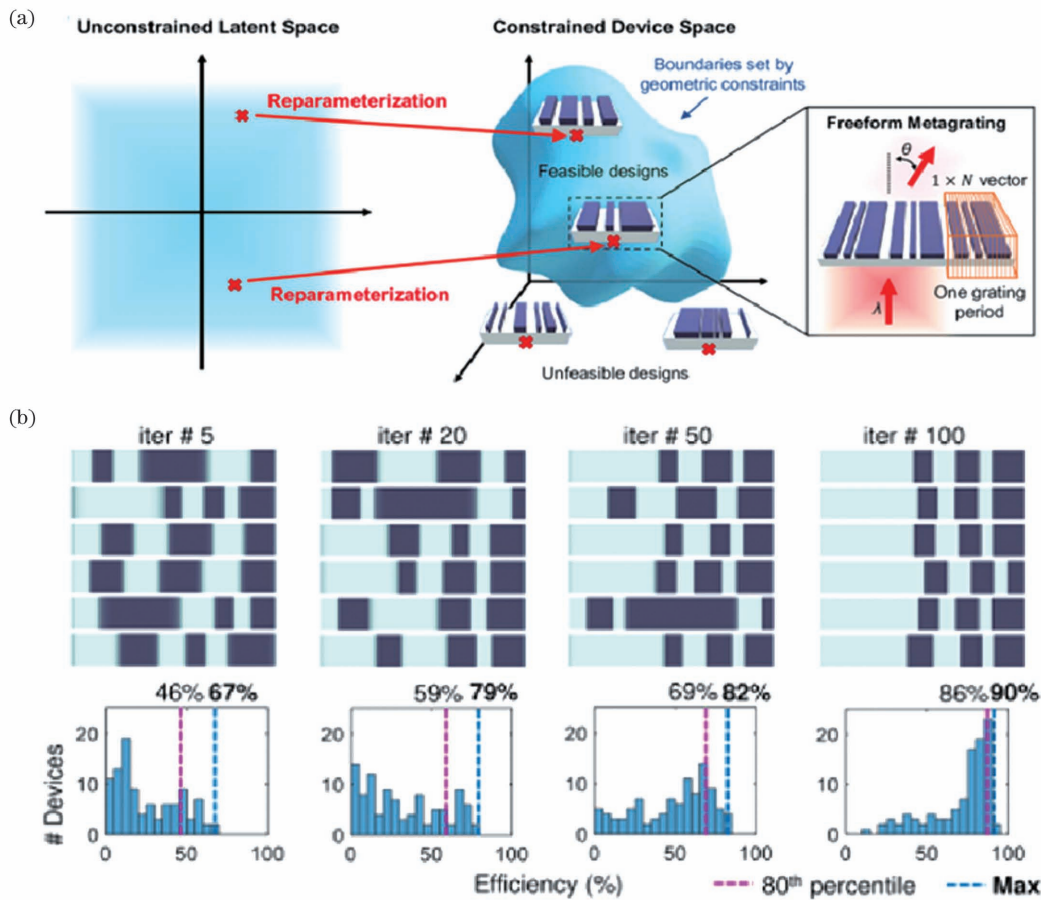


图 7 深度学习与拓扑学结合应用于逆设计问题<sup>[44]</sup>。(a)微纳结构示意图和解空间映射关系;(b)迭代次数分别为 5, 20, 50, 100 情况下,纳米结构分布和效率直方图

Fig. 7 Deep learning and topology applied to the inverse design problem<sup>[44]</sup>. (a) Schematic of micro-nano structure and spatial mapping relationship; (b) nano-structure distribution and efficiency histogram when the number of iterations is 5, 20, 50, 100, respectively

### 3 全息成像应用

深度学习能够从数据中学习规律或细节特征。利用深度学习的优势,深度学习可以替代常规的优化算法来提高图像质量和工作效率,特别是在全息成像中。全息是一种三维成像技术。根据分类方式不同,全息技术可分为传统光学全息、数字全息和计算全息 3 种<sup>[45]</sup>。随着信息时代的发展,通过计算机辅助处理,数字全息和计算全息发展迅速。利用数字全息技术获得物体的相位信息是一个非常重要的应用,特别是生物医学图像处理等领域<sup>[46]</sup>。根据记录方式的不同,全息技术又可分为同轴数字全息和离轴数字全息。离轴数字全息在记录过程中参考光和物光成一定的夹角。在数字重建过程中,频域中的衍射级次就会分离。通常基于频域滤波的方法可以去掉无用的衍射级次,以此来获得高质量重建的物像和相位信息。并且由于离轴的角度与频谱的

衍射级次位置相关,因此在实验中记录的离轴全息图通常需要进行人工的频域定位滤波<sup>[47]</sup>。在一定程度上,这些复杂的实验和后处理过程增加了时间成本,限制其进一步推广应用。另外,由于离轴的光路配置,成像系统的空间带宽积被压缩。与之相对应的同轴全息则克服了这些问题,但同轴全息有一个更加严重的缺点,就是其所有衍射级次是叠加在频谱中心的,且无法分离<sup>[48]</sup>,这就导致同轴全息重建后会伴随着孪生像的出现。传统计算方法计算同轴全息重建主要是相位恢复<sup>[49]</sup>或者强度传递方程<sup>[50]</sup>等。这些方法能够准确地得到物体的轮廓或形貌信息并且去除孪生像,但一般需要同时采集多幅强度图像并结合后处理等算法进行相位恢复。在此过程中,该实验对成像系统的稳定性和抗干扰性要求比较高,对于活体生物或动态的物体,则会变得很困难。随着深度学习技术的快速发展及普适性,其在数字全息、相位恢复等方面呈现出更多的实用价值<sup>[51]</sup>。

Rivenson 等<sup>[52]</sup>利用深度学习算法实现同轴数字全息孪生像的去除,获得了高质量的振幅和相位信息,网络结构和实验结果如图 8 所示。他们首先将拍摄的同轴全息图回传到特定距离处,得到物体的复振幅信息;由于同轴孪生像的干扰,此时得到的

复振幅信息伴随着同等程度的杂散噪声,并且由于逆向传播的非唯一解问题,采用一张全息图重建的信息图像质量较差。利用监督式的深度学习策略解决了这一问题。经过训练后的网络还能够重建类似的全息图像,提高了工作效率。

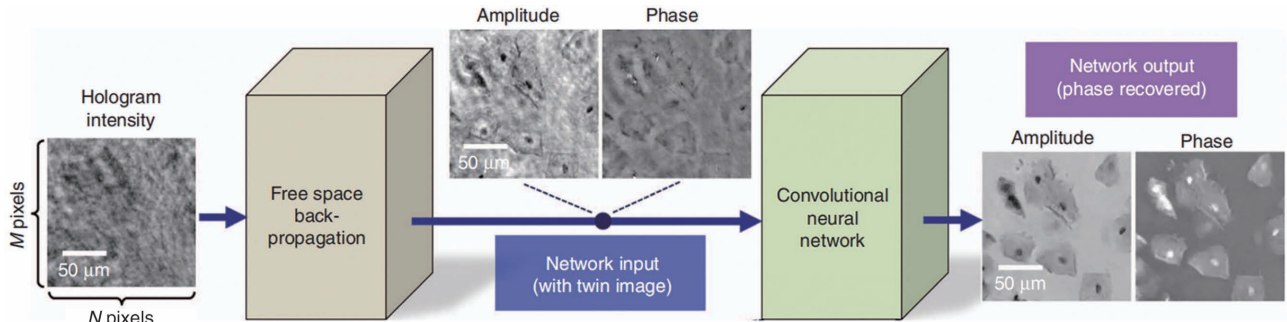


图 8 利用深度学习实现相位复原的网络结构和实验结果<sup>[52]</sup>

Fig. 8 Network structure and experimental results of phase restoration using deep learning<sup>[52]</sup>

另外一类计算机生成全息图并利用光学方法重建的技术称为计算全息<sup>[53]</sup>。计算全息在三维显示、AR/VR 等领域有着重要的应用。通常生成计算全息图的方法主要有 Gerchberg-Saxton(GS)法<sup>[54]</sup>、点源法<sup>[55]</sup>等。这类方法多是利用菲涅耳或角谱衍射的物理模型进行计算的。因此,对于每一幅全息图,都需要建模从物面传播到频谱面,并且需要多次迭代运算,以保证全息图能够有相对高的重建质量,因此计算过程的时间成本和计算成本相对较高。为了提高计算全息图生成的效率和减少计算负载,研究人员提出了一系列的改进和优化措施。特别是由于深度学习算法的普及应用,许多利用深度学习加速计算全息的方法被提出。Shi 等<sup>[56]</sup>提出一种可在手

机(如 Iphone)上进行实时三维全息运算的卷积神经网络。该方法能对手机直接拍摄的自然场景进行快速计算并生成彩色的 3D 全息图。该卷积神经网络只需要 620 kbit 的内存便可计算 1920×1080 的全息图。另外,将单一彩色图像逆向恢复成具有三维深度的全息图是比较困难的,如图 9 所示,基于卷积神经网络的方法生成全息图,对其进行三维重建后具有明显的深度信息。

相比于传统生成计算全息的方式,深度学习可以提高计算全息图的质量和计算速度。如图 10 所示,清华大学的 Wu 等<sup>[57]</sup>利用深度自编码神经网络高速生成大规模的计算全息,4000 分辨率的全息生成时间仅为 0.15 s。如此大规模的计算全息在显示

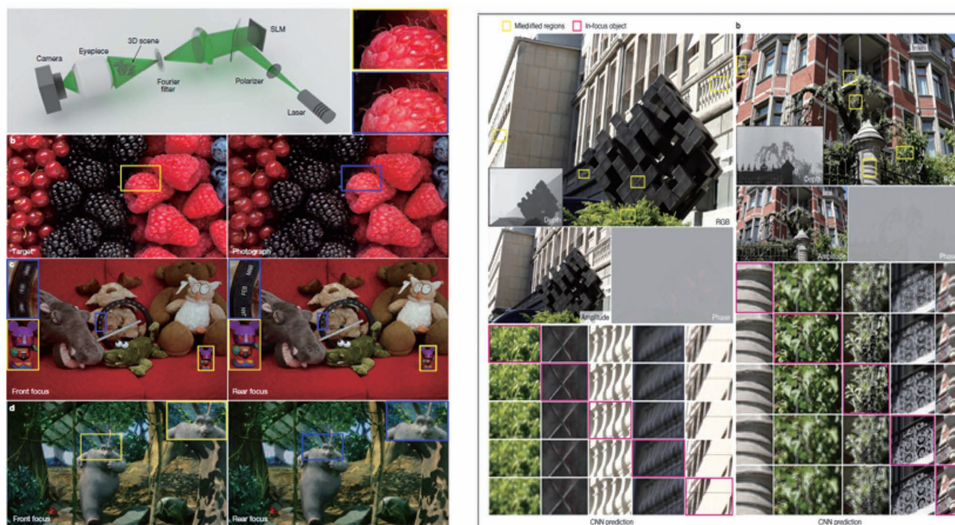


图 9 深度学习超轻量化的实时三维全息<sup>[56]</sup>

Fig. 9 Ultra-lightweight real-time 3D holography using deep learning<sup>[56]</sup>



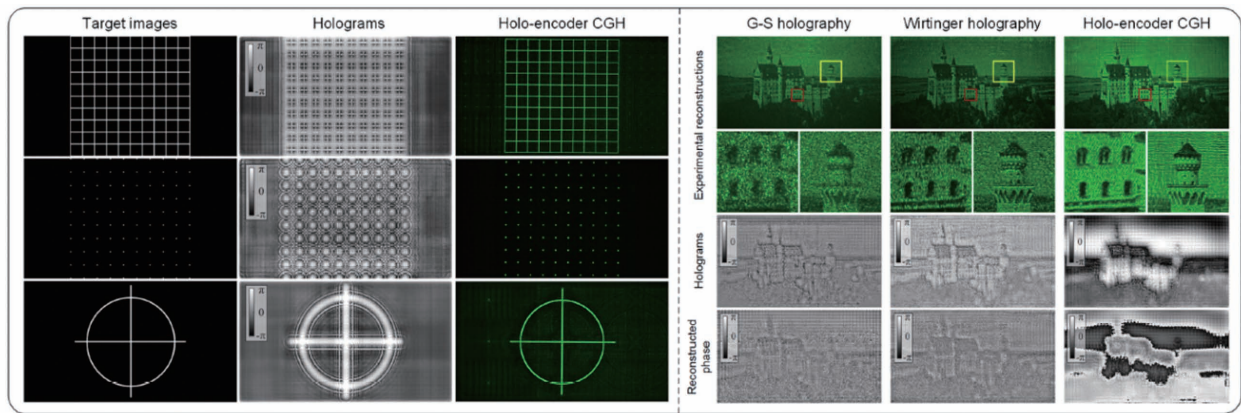


图 10 利用自编码计算全息方法高速生成全息图的实验结果<sup>[57]</sup>

Fig. 10 Experimental results of high-speed hologram generation by self-coded computer generated holography<sup>[57]</sup>

领域有着极其重要的作用,并且在成像表现上,相比于传统的 GS 全息和 Wirtinger 全息,具有更好的图像质量。由于在损失函数中加入了正则化项,因此自编码的全息在相干重建中具有更少的散斑噪声的干扰。该方法可以应用在实时的 3D 显示、AR/VR 或衍射光学元件的设计中。

#### 4 基于深度学习的光学感知与成像

单像素成像作为一种新型的计算成像技术,通过一系列掩模图案对目标物体的场景进行调制,使用一个没有空间分辨率的单像素探测器获取光场信息,然后将掩模图案的信息和单像素探测器所测得的光强做关联计算,来重建物体的图像。通常,单像素成像可以通过关联算法、傅里叶变换算法、压缩感知算法等相关算法实现对目标物体的图像重建<sup>[58-60]</sup>。然而,基于传统算法的单像素成像往往需要扫描采集和计算,消耗大量的计算时间和计算力,并且在低采样率的情况下,成像质量也不能够保证。结合深度学习数据驱动的优点,神经网络为单像素成像技术添加了新的活力。

将深度学习与单像素成像技术结合,最早提出的方法先用单像素成像系统对目标物体的光场信息进行获取,利用关联算法等完成图像重建;之后将在低采样率下重建的模糊图像作为输入,网络输出的图像被恢复成更高分辨率、更高质量的图像。这种基于深度学习的单像素成像方法可以使用 DNN<sup>[61]</sup>、CNN<sup>[62]</sup> 等网络结构实现,在低采样率的前提下可以明显地提高图像的重建质量。为了进一步提高图像的质量,同时实现图像去噪和图像细节保留,Rizvi 等<sup>[63]</sup> 在卷积神经网络的基础上提出了一种深度卷积自动编码网络(DCAN)模型,如

图 11(a)所示。该网络由两个阶段组成,下层阶段(E1-E2-D2-D1)是用于抑制孪生像和图像去噪的自动编码器,为了恢复细节,上层阶段(F1-F2-F3)很大程度上能够保留更多的细节。各个阶段的特征图如图 11(b)所示,高频信息由特征图的上层捕获,在下层中,信息首先被编码(E1 到中间),然后被解码(中间到 D1)以消除噪声。该网络同时恢复上层阶段图像细节和抑制下层阶段图像噪声,最终经过整合,可以实现图像的清晰化。

在单像素成像系统中,深度学习能够用于产生掩模图案,从而获取更有效的采样场景,这样就可以用更少的测量次数来恢复高质量图像。Higham 等<sup>[64]</sup> 利用深度卷积自动编码网络证明了这种方法的可行性,网络结构如图 12 所示。该网络第一个编码层由  $M$  个可变化的二值图案组成,用来模拟传统单像素成像系统中数字微镜器件(DMD)的作用,产生光强信号获取场景信息,随后从压缩信号中重建图像。通过在一个大型图像数据库上训练网络,可以优化掩模图案并实现对目标物体的图像重建。深度学习产生的掩模图案和神经网络可以在 4% 采样率的情况下恢复出视频传输速率的图像。并且这种优化方法的网络结构还可以用于产生一个场景的三维图像,将深度学习得到的掩模图像投影到目标物体上,恢复出场景的深度信息<sup>[65]</sup>。

在单像素成像系统中,将掩模图案和单像素探测器获取的强度信号相关联,以重建目标图像。深度学习还可以直接从一维的强度信号中直接恢复出二维的目标图像,从而实现在一些复杂环境下的应用。Wang 等<sup>[66]</sup> 研究了一种端到端卷积神经网络模型,通过模拟训练数据。训练后的网络可以直接从实验获得的一维强度信号中重建二维图像,而不需

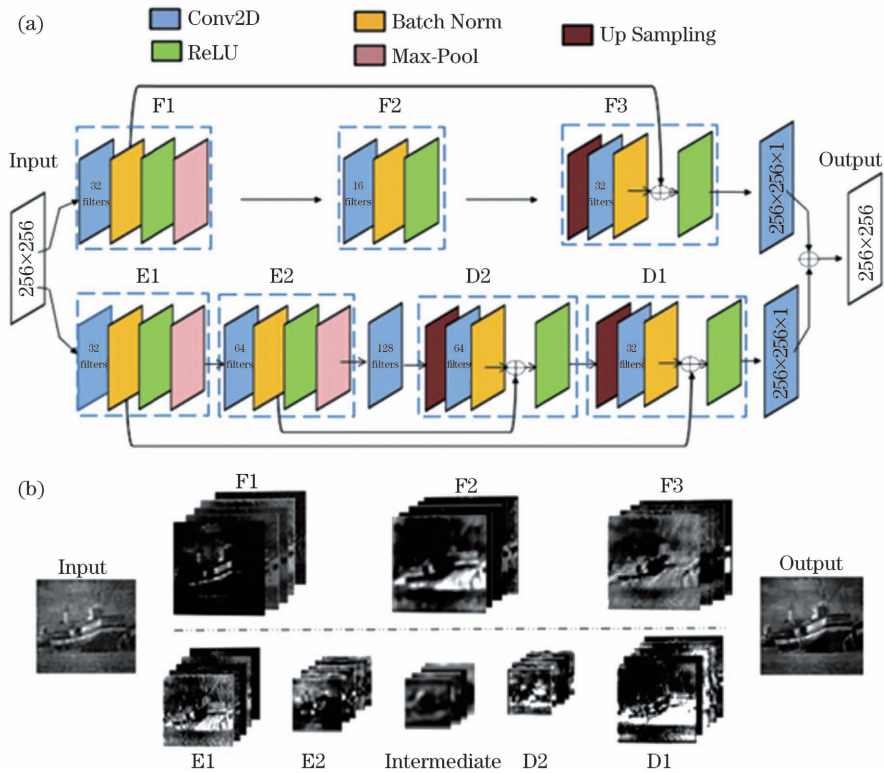


图 11 DCAN 模型和特征图<sup>[63]</sup>。(a)DCAN 模型结构图;(b)DCAN 各阶段的“船”测试图像特征图  
Fig. 11 DCAN model and feature maps<sup>[63]</sup>. (a) Schematic of DCAN model; (b) feature maps of ‘boats’ test image in DCAN each-stage

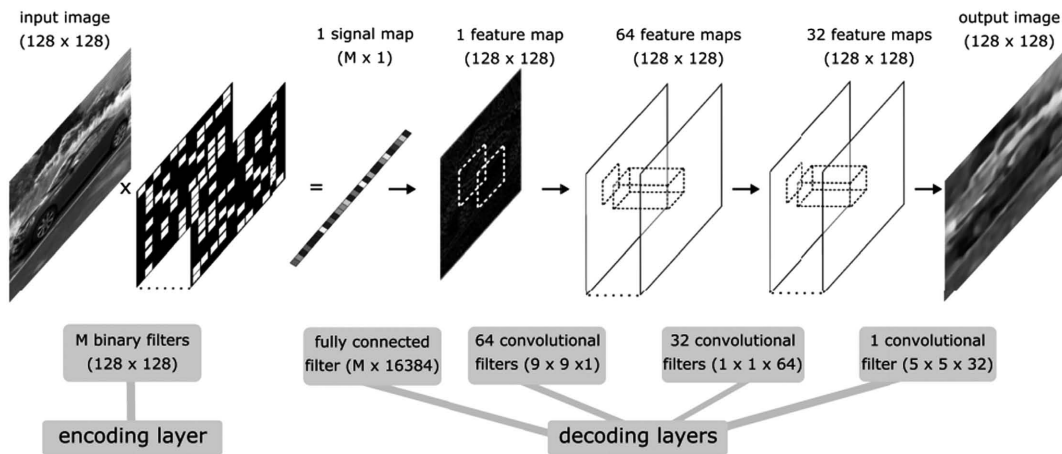


图 12 用于优化掩模图案的深度卷积自动编码网络模型<sup>[64]</sup>

Fig. 12 Deep convolutional auto-encoder network model for optimizing mask pattern<sup>[64]</sup>

要照明模式序列,网络结构如图 13 所示。该网络将获取的强度信号归一化后送入到网络,利用 2 个全连接层来检测输入强度信号的关联,并将之重塑为二维信号,通过卷积神经网络重建目标图像;使用 4 个独立的路径,利用最大池化方法对传入的图像进行下采样,获取不同尺度下的图像特征;之后进行上采样以恢复图像的大小,通过 1 个连接层对不同通道中提取的图像特征进行整合;最终由 4 个卷积层

和 1 个最大池化层生成重建的图像。利用这种端到端的神经网络,可以实现采样率为 6.25% 的单像素成像系统的图像高质量重建,并且不需要掩模图案的信息,直接从一维强度信息中恢复出二维目标图像。另外, Wu 等<sup>[67]</sup>提出了一种深度神经网络 (DAttNet),同样将单像素探测器获取的一维强度信号作为输入,可以在采样率仅为 5.45% 时重建出高质量的目标图像。

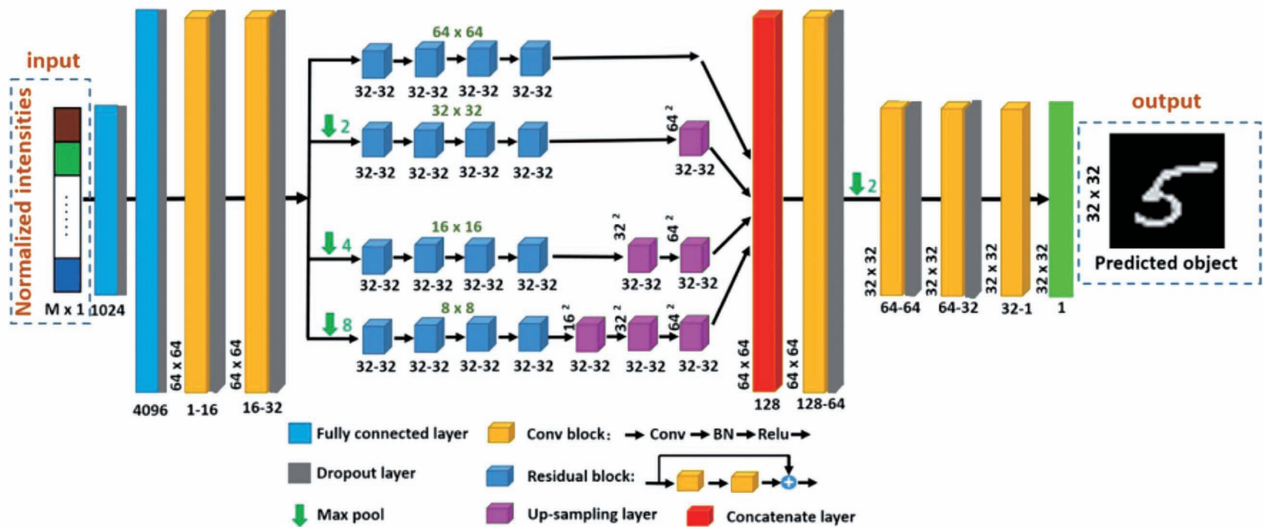


图 13 一种端到端的神经网络模型结构图<sup>[66]</sup>

Fig. 13 Schematic of an end-to-end neural network model<sup>[66]</sup>

深度学习除了可以提高单像素成像技术的成像质量,还广泛应用于突破衍射极限的超分辨率成像<sup>[68]</sup>、多源图像融合<sup>[69-71]</sup>等计算成像领域,并且取得了一些成果。相比于传统的器件成像与计算成像方法,神经网络可以更好地提取图像特征,在训练数据充足的情况下,深度学习方法拥有更快的计算速度和更好的成像质量。获取训练数据和搭建网络,以避免过量计算的同时提取图像细节特征,从而达到由低质量成像到高质量成像转换的目的,是如今深度学习在计算成像技术领域的发展方向。

区别于计算成像,光学感知成像技术采用激光主动照明,通过光学传感器或接收器采集光信号,并通过对目标信号进行处理复现物体的信息。非视域成像技术作为光学成像的一个重要分支,主要技术包括超快激光瞬时成像<sup>[72]</sup>、传输选通成像<sup>[73]</sup>、单光子信号计数成像<sup>[74]</sup>等,存在散射光难以克服<sup>[75]</sup>、视

场角受限<sup>[76]</sup>、波前整形速度慢<sup>[77]</sup>等问题。将深度学习引入到非视域成像,用数据驱动的方法代替单一的物理模型,往往具有较高的速度和较好的适应性<sup>[78]</sup>。

在光学领域,存在着厚散射介质成像的问题,光子通过光学厚度超过散射平均自由路径 10 倍的散射介质后变得扩散,只能得到类似噪声的图像。在光学或计算成像方面有着很多优秀的方法解决厚散射介质问题,例如用全息术产生散射光子的相位共轭光束<sup>[79]</sup>,当光束再次传播于散射介质时失真就可以消除,但没有记录任何关于内部散射过程的信息;根据记忆效应,能实现通过不透明散射层的非侵入性成像等应用<sup>[80]</sup>,但通过散射板的光场记忆效应范围随着散射板厚度增加迅速下降到波长范围<sup>[81]</sup>。Lyu 等<sup>[82]</sup>提出了一种混合神经网络,如图 14 所示,网络主要包含 5 个二维卷积层和 3 个全连接层,

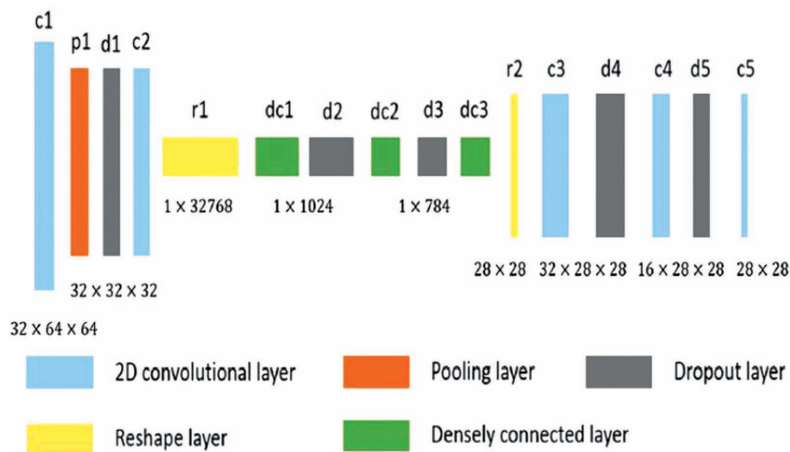


图 14 混合神经网络结构<sup>[82]</sup>

Fig. 14 Structure of hybrid neural network<sup>[82]</sup>

3 个全连接层置于 5 个卷积层之间,实现了经厚散射介质后的计算成像。

实验演示结果如图 15 所示。图 15(a)为散射介质成像的实验装置,空间光调制器(SLM)为振幅型,P1 和 P2 是线性偏振器,散射介质板是一个

3 mm 厚的白色聚苯乙烯。图 15(b)为图像经白色聚苯乙烯板后得到的散斑图像。通过对比图 15(c)混合神经网络重建出的图像和图 15(d)真实目标图像,图像中所有可见特征特别是目标的边缘,都得到了成功的恢复。

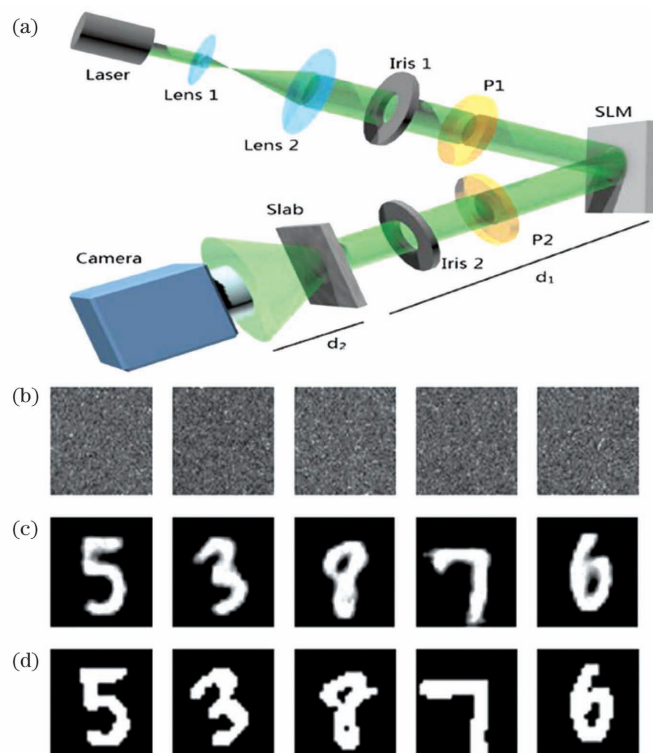


图 15 混合神经网络的厚散射介质成像<sup>[82]</sup>。(a)散射介质成像实验装置;(b)透过散射介质后的散斑图像;(c)混合神经网络重构图像;(d)真实目标图像

Fig. 15 Thick scattering media imaging using hybrid neural networks<sup>[82]</sup>. (a) Scattering medium imaging experimental device; (b) speckle images after scattering medium; (c) reconstructed images of mixed neural network; (d) real target images

Zhou 等<sup>[83]</sup>提出了一种基于深度神经网络的端到端可压缩相位成像方法,该方法通过弱散射,只使用一个低位或二进制模式来重建复杂的对象。如图 16(a)所示,作为预处理器的薄扩散器被放置在图像传感器的前面,在参考模式(移除物体,强度分布完全来自扩散器)和位移模式(加入物体)的位移不变范围情况下,能够在视觉上观察到由物体波前引入的斑点的局部轻微位移,可以建立该模式中局部特征的空间变化与相应波前的相位分布之间的隐含关系,就可以将输入的波前信息编码为生成的斑点的失真和局部变化。如图 16(b)所示,压缩的(低比特的)模式或捕获的原始散斑模式首先被放到一个 DNN 中进行训练,该网络具有一个典型的 U-Net 结构,由一个编码器和一个解码器组成,可以从低比特深度模式下分布的离散颗粒中提取出物体的相位轮廓。

深度神经网络还能应用于光学显微成像领域,显著改善其成像效果。Rivenson 等<sup>[84]</sup>使用了传统的卷积神经网络进行训练后,将普通光学显微镜获得的图像作为输入,在不改变任何设计情况下图像的分辨率得到了提升。图 17(a)为显微成像神经网络训练流程,将一组由低分辨率图像组成的数据作为输入,训练标签为其对应的高分辨率图像,通过优化各种参数对神经网络进行训练,使网络输出与对应的高分辨率训练标签之间的损失函数最小化。将普通光学显微镜图像放入到显微成像神经网络中,网络迅速输出图像,有更好的分辨率,匹配性的高数值孔径镜头也有超过有限的景深和较大的视野范围,如图 17(b)所示。光学显微镜被认为是最广泛使用的技术之一,在神经网络的协助下使用光学显微镜的相关学科,例如生命科学<sup>[85]</sup>,将会有更加迅速的发展。

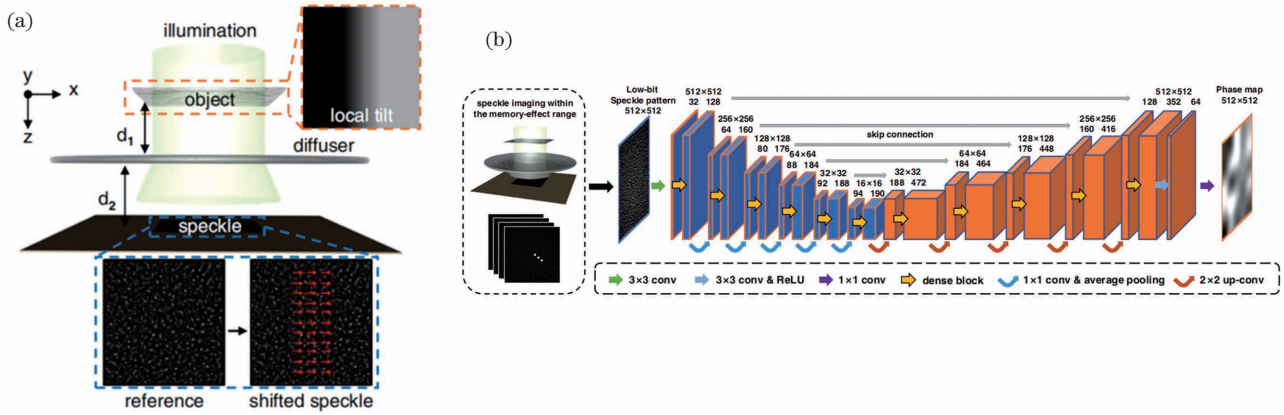


图 16 深度神经网络通过散射实现相位成像<sup>[83]</sup>。(a)通过散射重建相位图像的原理；(b)深度神经网络模型结构图及其训练过程

Fig. 16 Using depth neural network to achieve phase imaging by scattering<sup>[83]</sup>. (a) Schematic of phase image reconstructed by scattering; (b) structure of the deep neural network model and its training process

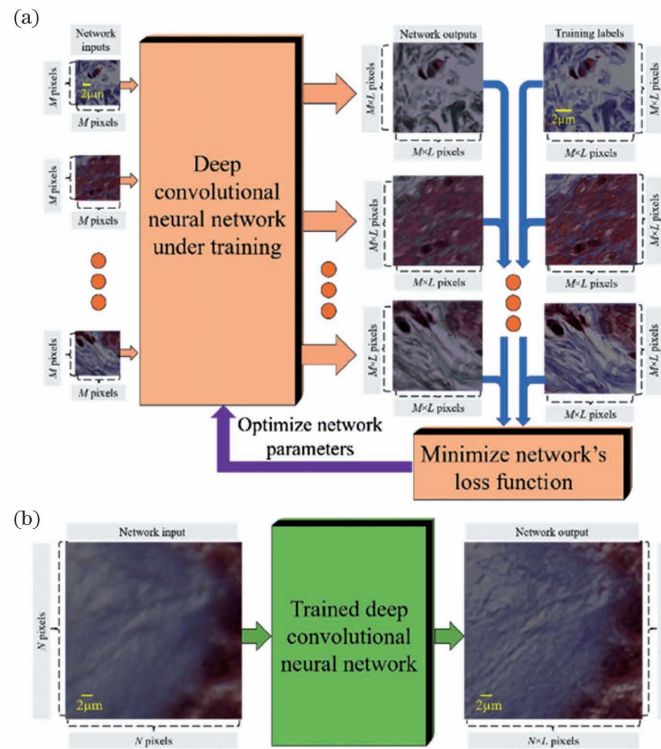


图 17 显微成像神经网络训练流程<sup>[84]</sup>。(a)显微成像神经网络训练示意；(b)显微成像神经网络结果展示  
Fig. 17 Training flow of microscopic imaging neural network<sup>[84]</sup>. (a) Training of microscopic imaging neural network; (b) results of microscopic imaging neural network

除此之外,其他成像技术如压缩感知技术与深度学习可以很好结合<sup>[86]</sup>,应用于三维成像<sup>[87]</sup>、全息像复现<sup>[88]</sup>、噪声抑制<sup>[89]</sup>等方面。

### 5 新型光子驱动神经网络

目前,各种基于深度学习算法的神经网络都是在 CPU、GPU、可编程门电路以及一些特殊电子芯片上运行的,由于冯·诺依曼架构中数据与程序的

存储分离<sup>[90]</sup>,在运算过程中需要进行大量的数据读写,数据的处理速度与能量损失成正比,并且电信号容易被干扰<sup>[91]</sup>。随着深度学习的不断发展,所需的计算能力不断上升,能量的消耗也不断扩大。目前基于硅的电子元件仍然作为主要元件,但其发展的瓶颈已逐渐展现,需要寻找替代方法来解决电子元件发展的瓶颈问题。光是一种电磁波,在传播的过程中与电信号不同,不受外界电磁场干扰<sup>[92]</sup>,同时

作为信息传输的载体时有着高带宽、高传输速度、低损耗等特点,研究人员开始考虑使用光来代替电子,希望在光学领域建立神经网络。

2018 年, Lin 等<sup>[93]</sup>提出了一种全光衍射深度神经网络(D2NN),此网络能够实现手写数字和 fashion 图像分类、太赫兹波段成像透镜的功能,网络结构如图 18 所示。类比神经网络每层之间神经元相互连接的想法,此光学神经网络每一层上的每

一点代表一个神经元,能够透射或反射入射波,依据惠更斯-菲涅耳原理,每一点都作为一个次波源,实现与下一层各神经元之间的相互连接。在数值计算中,D2NN 对手写数字和 fashion 图像的分类准确率分别达 91.75% 和 81.13%,与传统神经网络实现的精确度已较为接近。接着他们利用 3D 打印的方法控制不同位置像素块的厚度,调节光通过时产生的相位差,制作出了 5 层无源衍射元件进行实验验证。

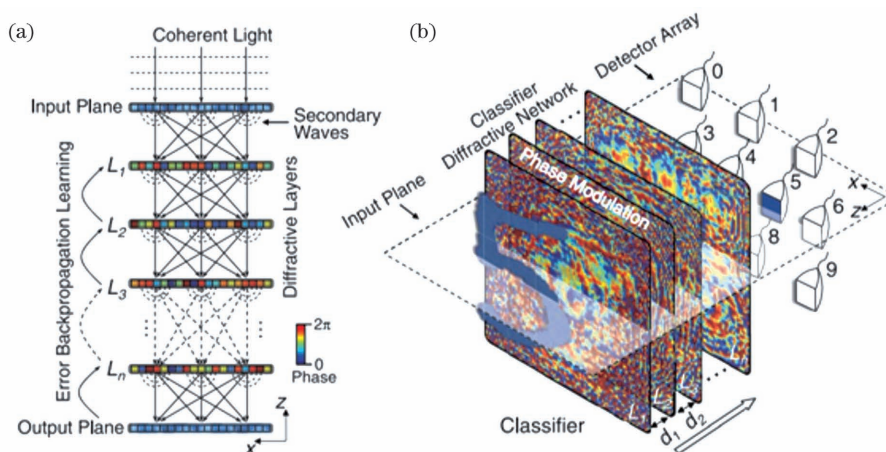


图 18 衍射深度神经网络<sup>[93]</sup>。(a)衍射神经网络入射波传播示意图;(b)分类衍射神经网络系统结构

Fig. 18 Diffraction deep neural network<sup>[93]</sup>. (a) Incident wave propagation of diffraction neural network; (b) system structure of classified diffraction neural network

随后 D2NN 也被用于其他的任务, Veli 等<sup>[94]</sup>利用 D2NN 将任意宽带脉冲整形为理想的光波形状,根据衍射神经网络无源的特性,实现一个微型的无源光脉冲整形系统。脉冲整形衍射网络由 4 个衍射层构成,如图 19(a)所示,可看出结构十分紧凑,太赫兹波段的脉冲入射直径为 0.8 cm 的矩孔,经过衍射层在距输出面 0.2 cm 的距孔后方成像。网络中的衍射层相位分布作为可训练参数,而输出的时域

波形的载频是预先设定的参数,为 0.35 THz,网络经过训练后,能够在不使用其他传统光学元件的情况下将输入脉冲合成为脉冲宽度为 16.59 ps 的方形时域脉冲,如图 19(b)所示。3D 打印的衍射层在实验中产生了脉冲宽度为 15.52 ps 的时域波形,与数值模拟的结果十分接近。

2020 年浙江大学 Qian 等<sup>[95]</sup>利用 D2NN,巧妙地将光学逻辑运算转换为一个分类的任务,实现了

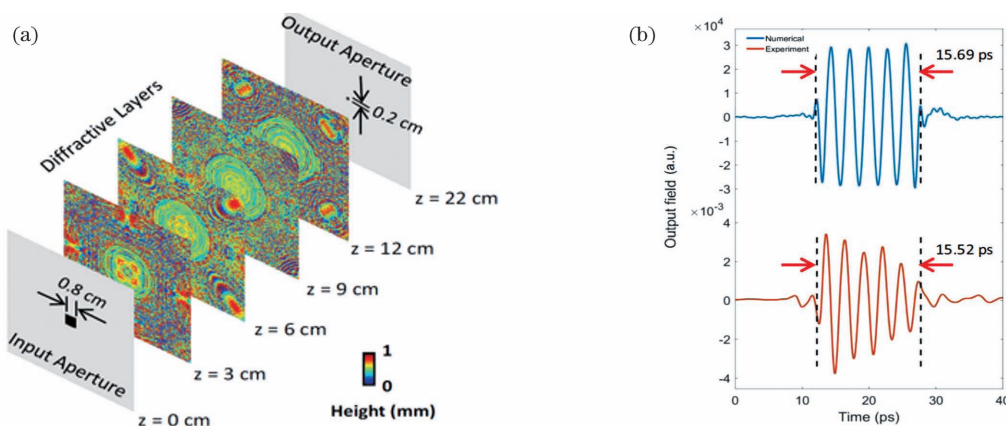


图 19 D2NN 实现光脉冲整形系统<sup>[94]</sup>。(a)脉冲整形衍射网络系统结构;(b)脉冲整形数值模拟与实验结果

Fig. 19 Optical pulse shaping system with D2NN<sup>[94]</sup>. (a) Structure of pulse shaping diffraction network; (b) numerical simulation and experimental results of pulse shaping

光学逻辑运算。图 20(a)为实现光学逻辑运算的一个网络结构展示,输入层的每个区域都被设置为一个特定的逻辑状态或逻辑运算符,通过控制输入平面设定区域的透射与不透射两种状态来调整输入信息,换句话说,通过调整输入面的传输状态来对输入的平面波进行空间编码,实现特定的光学逻辑操作。而输出平面被设置为两个区域,

分别代表“0”输入和“1”输入,通过中间的隐藏层也就是衍射光学元件对编码后的输入光进行解码,将入射光散射到输出层的指定区域。以“与”、“或”和“非”三种逻辑运算为例,3种运算符和4个逻辑状态被设置到特定的区域,通过激活不同的输入层区域,可以实现不同的逻辑计算功能,如图 20(b)所示。

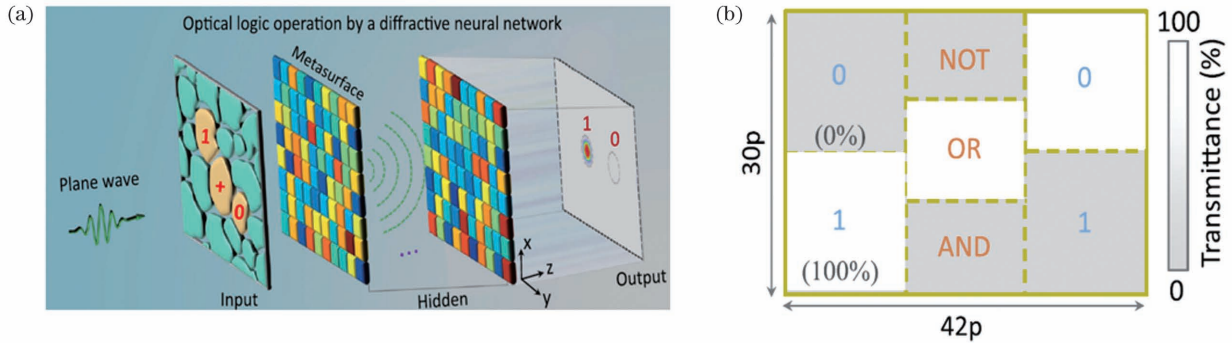


图 20 D2NN 实现光学逻辑运算<sup>[95]</sup>。(a)光学逻辑神经网络系统结构;(b)入射平面区域功能划分  
Fig. 20 Optical logic operation with D2NN<sup>[95]</sup>. (a) Optical logic neural network system structure; (b) regional division of incident plane

2021年,Gu 课题组<sup>[96]</sup>利用具有单层纳米尺度的全息感知器,也就是单层无源衍射元件来进行光学推理,他们使用双光子纳米光刻的方式将单层衍射元件打印至 CMOS 芯片上,实现了一种紧凑的光学解密器,利用单层的无源器件实现对单个或整类图像进行关键性解密,为器件的紧凑和高度集成性提供了一种解决方式。图 21 为解密器的工作展示,本质上可以将其解释为二分类和多分类。

为单个图像解密的示意图,当输入的图像为指定的图像时,解密器才会给出正确的反馈,输入其他任何图像时都只会反馈错误的答案。图 21(b)则为整类图像的解密,以 A 字母为例,只要输入的图像为 A,无论其书写的样式如何,都将入射光散射到对应预先设定的区域,依据这样的特性,未来可以考虑将整类解密的衍射元件用于人脸识别中,相较于单个图像解密,可以用不同角度人脸识别来成功解锁设备。

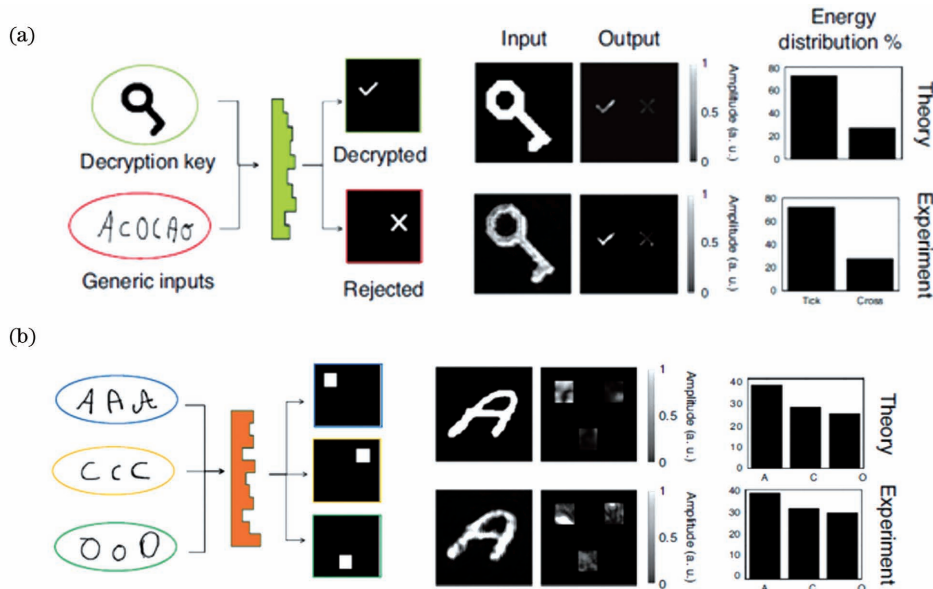


图 21 单层衍射元件实现光学解密器<sup>[96]</sup>。(a)单个图像解密;(b)同类图像解密  
Fig. 21 Optical decoder using single layer diffraction element<sup>[96]</sup>. (a) Decryption of a single image; (b) decryption of a class of images

D2NN 能够实现传统神经网络中的各种复杂函数,并且能够实现光速并行运算,这为基于深度学习的无源组件快速分析数据、分类图像和目标创造了新的机遇。而上述网络仍存在一个问题,只能应用于单任务,并且训练完成并打印好之后就无法进行修改。

在此方面,清华大学 Dai 课题组<sup>[97]</sup>提出了一种光电可重构的计算模型,通过构造一个衍射处理器,能够支持不同类型神经网络的构成,同时也可以应

用于不同的任务之中。如图 22 所示,文献[98]提出了一种衍射处理器 DPU,选择 DMD 和 SLM 作为光调制器输入节点,采用 CMOS 传感器作为光电探测器输入光电神经元,实现了具有视频速率推理能力的 DPU。每个 DPU 可以作为衍射神经网络的一层,通过多个 DPU 的不同组合来实现各种神经网络的组成,最终手写数字识别和视频人物动作识别的计算精度分别达到了 97.6% 和 96.3%,与现有的电子神经网络识别准确率相当。

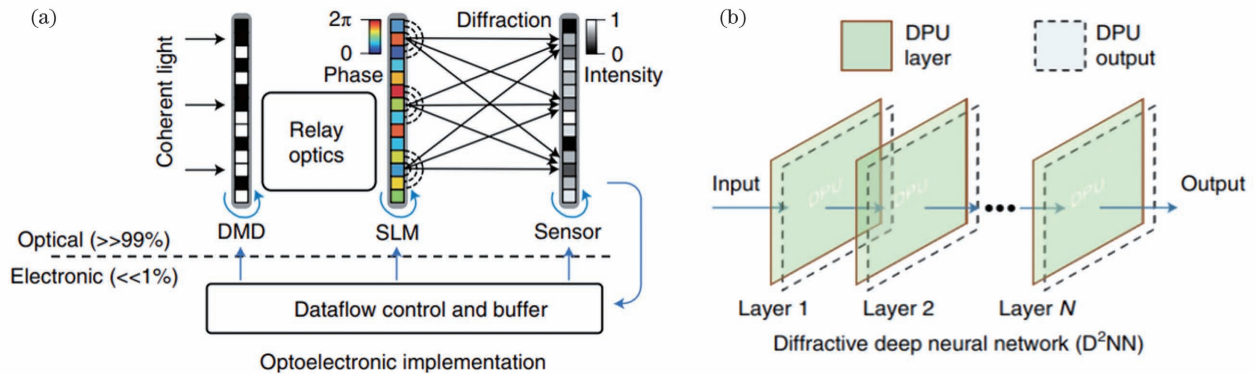


图 22 可重构衍射处理器<sup>[97]</sup>。(a)DMD、相位 SLM 和 CMOS 传感器实现 DPU 内部结构;(b)DPU 构建为 DNN 结构

Fig. 22 Reconfigurable diffraction processor<sup>[97]</sup>. (a) DMD, phase-type SLM, and CMOS sensors constitute the internal structure of DPU; (b) DPU is constructed as DNN structure

此外,Ozcan 课题组<sup>[98]</sup>将单像素与衍射神经网络相结合,通过将物体的空间信息编码到衍射光的功率谱中,对在输出平面用单像素光谱探测器得到的光谱进行物体分类,并在太赫兹光谱下验证了与单像素结合的衍射神经网络框架。通过检测 10 个不同波长的衍射光的光谱功率对手写数字图像进行光学分类,其中每个波长代表一个数字。光谱类分

数  $s$  定义在一组离散的波长上,最大值决定物体所属类别,用宽带脉冲光照射手写数字 8,由单个探测器收集衍射光的功率谱,得到  $\lambda$  对应的  $S_{\lambda}$  最大,则认为识别出数字 8,如图 23 所示。

除了以光和被动式无源衍射器件为载体的全光神经网络外,早在 2017 年就已提出了一种全新的全光神经网络架构,将神经网络集成于片上光子芯

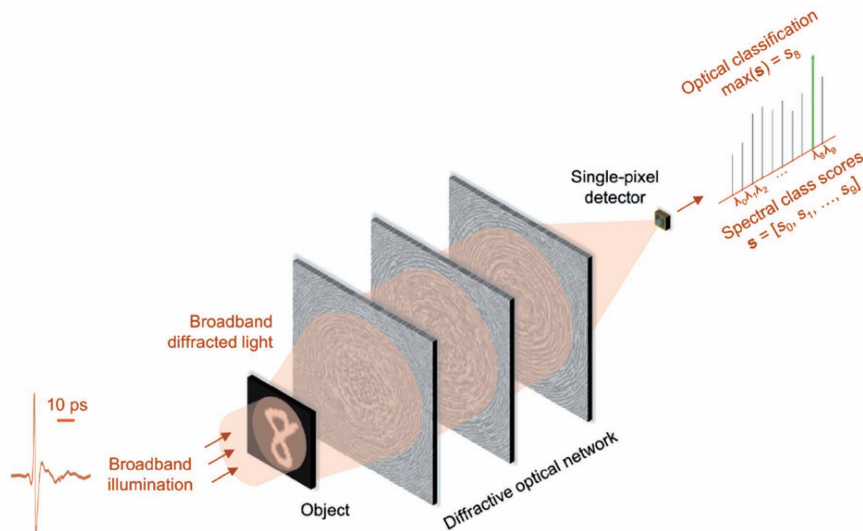


图 23 光谱编码分类的单像素衍射神经网络结构<sup>[98]</sup>

Fig. 23 Structure of single-pixel diffraction neural network for spectral coding classification<sup>[98]</sup>



片<sup>[99]</sup>, 利用光学的独特优势, 有望在传统学习任务的计算速度上至少提高 2 个数量级, 在功率效率上提高 3 个数量级, 如图 24(a) 所示。用一个可编程纳米光子处理器, 在硅光子集成电路中串联 56 个可编程马赫-曾德尔干涉仪, 处理器之间通过光波导来传输光。图 24(b) 展示了物理波(声波)模拟递归神

经网络<sup>[100]</sup>, 其原理是由于物理波传播方程与深度学习中递归神经网络的原理公式类似, 采用声波传播的过程模拟递归神经网络的运算过程, 实现对元音字母发音进行分类的功能。训练集为现实的元音 iy ei ae 的发音, 通过网络训练的结果生成中间的声波吸收模块。

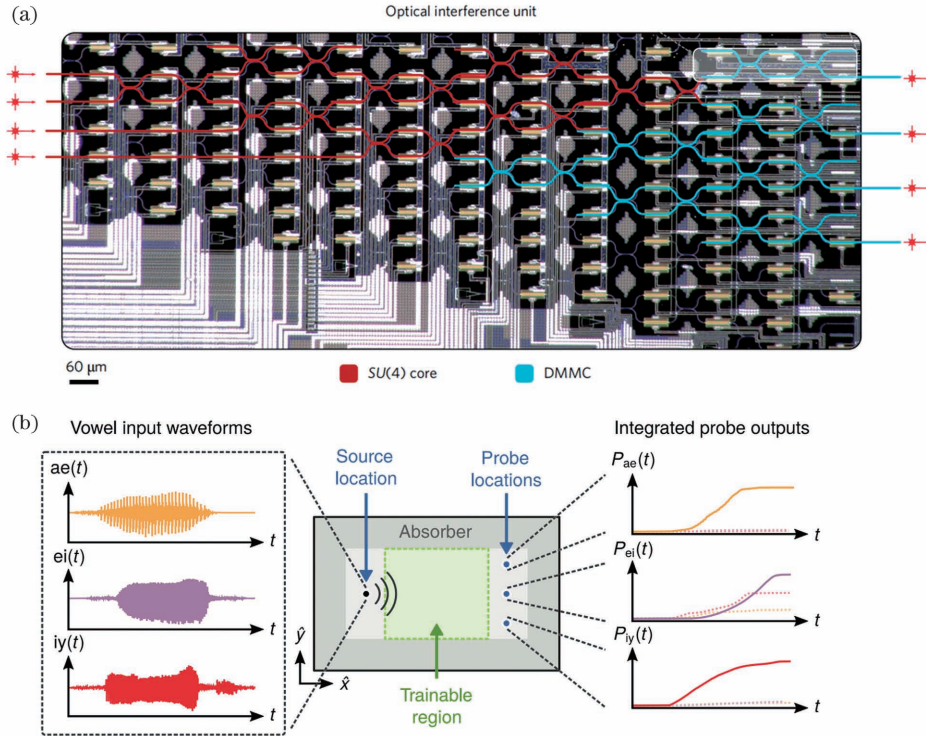


图 24 光子神经网络与物理波神经网络系统实现元音语音分类<sup>[99-100]</sup>。(a)集成片上光子神经网络单元;(b)声波神经网络系统

Fig. 24 Classification of vowels using photon neural network and physical wave neural network<sup>[99-100]</sup>. (a) Photon neural network units on-chip; (b) acoustic neural network system

光学神经网络结合了光学和深度学习, 相较于传统网络结构, 能实现高带宽、高传输速度以及低损耗的信号传输, 但在集成化、规模化以及可训练性上仍与电子神经网络有着一定的差距, 设计时也需要依赖电子神经网络。并且现有的光学神经网络大都是通过线性计算来实现具体任务的。如何从根源上突破光神经网络的瓶颈, 还需要大量的研究和思考, 相信在将来, 光神经网络能够更好地发挥光学与深度学习结合带来的优势, 实现真正的全光计算系统。深度学习也被应用于其他光学领域, 如光学隐身衣<sup>[101]</sup>、光学防伪<sup>[102]</sup>、光纤信道建模<sup>[103]</sup>等。

## 6 总结与展望

深度学习作为近几年十分热门的话题, 在诸多领域均已展现了优势, 并在光学领域越来越被重视。深度学习与光学结合还可用于轨道角动量通信、校

正像差、微纳结构设计、全息图产生与再现、图像处理、超分辨率成像、计算成像、虚拟现实、增强现实、光学传感等领域<sup>[104-107]</sup>。此外, 本课题组将深度学习与光学加密结合, 生成了轨道角动量加密的计算全息轨道角动量通信网络, 将拓扑电荷数编码成序列信息作为另一个扩展的数字信道, 通过训练两个 CNN 重建加密全息图, 并解码拓扑信息, 提高了图像信息存储的安全性。

深度学习与光衍射和计算相结合解决一些实际的应用。然而, 光学与深度学习算法相结合的过程中仍然有一些问题。解决这些问题有利于提高深度学习算法在实际光学应用场景中的泛化能力和工作效率<sup>[108]</sup>。1)深度学习框架与物理模型的融合。目前深度学习功能强大, 但缺乏可解释性。在深度学习计算过程中融入物理模型, 从而增加模型的可解释性。推广利用神经网络学习物理规律, 尝试解释

复杂的物理现象,以此降低网络对数据的依赖程度。2)实验数据和仿真数据的融合。常规的深度学习网络能够实现精准的任务时往往需要大量的同类数据的计算和训练的优化。而随着深度学习的发展和光学场合应用,神经网络的规模和复杂程度提升的同时,往往伴随着数据量需求的提升。在实验中获取大量的训练数据往往是相当耗时或不容易实现的。因此,研究人员尝试了物理解析方法和数值模拟仿真获取训练数据。仿真数据与实验中采集的数据误差较小。采用基于真实数据采用多种算法生成新数据等方式来扩大训练集。3)深度学习与其他优化方法结合。深度学习优化算法与传统优化算法,如压缩感知等可以并行优化。深度学习加入传统优化算法可以在一定程度上进行优势互补,提高网络模型的性能并大大缩短计算时间,并且也有利于增强网络的泛化性和鲁棒性。将深度学习与传统算法更好结合,在图像处理、显微成像、计算成像等领域均有很好的发展前景。4)光衍射神经网络突破传统电子神经网络限制。光衍射神经网络作为深度学习与光学结合的产物,有着高速、低功耗的特点,并能突破传统冯·诺依曼架构的约束。但到目前为止,光驱动神经网络在可重构性、集成性、稳定性、实用性上有很大的提升空间。解决全光学神经网络只针对单一任务、可重构性差,片上光子神经网络稳定性低等问题,还需要更多的尝试与研究,距离稳定的实际应用还有一段路要走,因此还有着广泛的提升空间。

### 参 考 文 献

- [1] Al-Qizwini M, Barjasteh I, Al-Qassab H, et al. Deep learning algorithm for autonomous driving using GoogLeNet [C] // 2017 IEEE Intelligent Vehicles Symposium (IV), June 11-14, 2017, Los Angeles, CA, USA. New York: IEEE Press, 2017: 89-96.
- [2] Kotz S A, Cappa S F, von Cramon D Y, et al. Modulation of the lexical-semantic network by auditory semantic priming: an event-related functional MRI study [J]. *NeuroImage*, 2002, 17 (4): 1761-1772.
- [3] Bui H M, Lech M, Cheng E, et al. Object recognition using deep convolutional features transformed by a recursive network structure [J]. *IEEE Access*, 2016, 4: 10059-10066.
- [4] Leibig C, Allken V, Ayhan M S, et al. Leveraging uncertainty information from deep neural networks for disease detection [J]. *Scientific Reports*, 2017, 7 (1): 17816.
- [5] Rashevsky N. Some remarks on the Boolean algebra of nervous nets in mathematical biophysics [J]. *The Bulletin of Mathematical Biophysics*, 1945, 7 (4): 203-211.
- [6] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain [J]. *Psychological Review*, 1958, 65(6): 386-408.
- [7] Cortes C, Vapnik V. Support-vector networks [J]. *Machine Learning*, 1995, 20(3): 273-297.
- [8] Smirnov E A, Timoshenko D M, Andrianov S N. Comparison of regularization methods for ImageNet classification with deep convolutional neural networks [J]. *AASRI Procedia*, 2014, 6: 89-94.
- [9] Tran D, Bourdev L, Fergus R, et al. Learning spatiotemporal features with 3D convolutional networks [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 4489-4497.
- [10] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 105-114.
- [11] Genty G, Salmela L, Dudley J M, et al. Machine learning and applications in ultrafast photonics [J]. *Nature Photonics*, 2021, 15(2): 91-101.
- [12] Wiecha P R, Arbouet A, Girard C, et al. Deep learning in nano-photonics: inverse design and beyond [J]. *Photonics Research*, 2021, 9(5): B182-B200.
- [13] Liu Z, Zhu D, Raju L, et al. Tackling photonic inverse design with machine learning [J]. *Advanced Science*, 2021, 8(5): 2002923.
- [14] Sajedian I, Kim J, Rho J. Predicting resonant properties of plasmonic structures by deep learning [EB/OL]. (2018-05-19) [2021-06-25]. <https://arxiv.org/abs/1805.00312>.
- [15] Dardikman-Yoffe G, Eldar Y C. Learned SPARCOM: unfolded deep super-resolution microscopy [J]. *Optics Express*, 2020, 28(19):27736-27763.
- [16] Dong C, Loy C C, He K M, et al. Image super-resolution using deep convolutional networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295-307.
- [17] Dong Xu, Yu Luo, Jun Luo, et al. Efficient design of a dielectric metasurface with transfer learning and genetic algorithm [J]. *Optical Materials Express*, 2021, 11(7):1852-1862.
- [18] Jiao L C, Zhao J. A survey on the new generation of deep learning in image processing [J]. *IEEE Access*,

- 2019, 7: 172231-172263.
- [19] Horisaki R, Takagi R, Tanida J. Deep-learning-generated holography[J]. *Applied Optics*, 2018, 57(14): 3859-3863.
- [20] Qu Y R, Zhu H Z, Shen Y C, et al. Inverse design of an integrated-nanophotonics optical neural network[J]. *Science Bulletin*, 2020, 65(14): 1177-1183.
- [21] Oskooi A F, Roundy D, Ibanescu M, et al. Meep: a flexible free-software package for electromagnetic simulations by the FDTD method[J]. *Computer Physics Communications*, 2010, 181(3): 687-702.
- [22] Deslandes D, Wu K. Integrated microstrip and rectangular waveguide in planar form[J]. *IEEE Microwave and Wireless Components Letters*, 2001, 11(2): 68-70.
- [23] Kogelnik H. Coupled wave theory for thick hologram gratings[J]. *The Bell System Technical Journal*, 1969, 48(9): 2909-2947.
- [24] Wiecha P R, Muskens O L. Deep learning meets nanophotonics: a generalized accurate predictor for near fields and far fields of arbitrary 3D nanostructures[J]. *Nano Letters*, 2020, 20(1): 329-338.
- [25] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. *Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science*. Cham: Springer, 2015, 9351: 234-241.
- [26] Rudy S H, Brunton S L, Proctor J L, et al. Data-driven discovery of partial differential equations[J]. *Science Advances*, 2017, 3(4): e1602614.
- [27] Liu D J, Tan Y X, Khoram E, et al. Training deep neural networks for the inverse design of nanophotonic structures[C]//2019 Conference on Lasers and Electro-Optics (CLEO), May 5-10, 2019, San Jose, California, United States. Washington, D.C.: OSA, 2019: 7b01377.
- [28] Peurifoy J, Shen Y, Jing L, et al. Nanophotonic particle simulation and inverse design using artificial neural networks[J]. *Science Advances*, 2018, 4(6): eaar4206.
- [29] Ma W, Cheng F, Liu Y. Deep-learning-enabled on-demand design of chiral metamaterials[J]. *ACS Nano*, 2018, 12(6): 6326-6334.
- [30] Malkiel I, Mrejen M, Nagler A, et al. Plasmonic nano-structure design and characterization via Deep Learning[J]. *Light, Science & Applications*, 2018, 7: 60.
- [31] An S S, Fowler C, Zheng B W, et al. A deep learning approach for objective-driven all-dielectric metasurface design[J]. *ACS Photonics*, 2019, 6(12): 3196-3207.
- [32] Wang H, Jia W, Wang J, et al. Learning graph representation with generative adversarial nets[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021, 33(8):3090-3103.
- [33] Liu C, Yu W M, Ma Q, et al. Intelligent coding metasurface holograms by physics-assisted unsupervised generative adversarial network[J]. *Photonics Research*, 2021, 9(4): B159-B167.
- [34] Liu Z, Zhu D, Rodrigues S P, et al. Generative model for the inverse design of metasurfaces[J]. *Nano Letters*, 2018, 18(10): 6570-6576.
- [35] Zhu D, Liu Z, Raju L, et al. Building multifunctional metasystems via algorithmic construction[J]. *ACS Nano*, 2021, 15(2): 2318-2326.
- [36] Ma W, Cheng F, Xu Y, et al. Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy[J]. *Advanced Materials*, 2019, 31(35): e1901111.
- [37] Kiarashinejad Y, Abdollahramezani S, Adibi A. Deep learning approach based on dimensionality reduction for designing electromagnetic nanostructures[J]. *Npj Computational Materials*, 2020, 6: 12.
- [38] Jorge L P, Meeks W H. The topology of complete minimal surfaces of finite total Gaussian curvature[J]. *Topology*, 1983, 22(2): 203-221.
- [39] Camayd-Muñoz P, Ballew C, Roberts G, et al. Multifunctional volumetric meta-optics for color and polarization image sensors[J]. *Optica*, 2020, 7(4): 280-283.
- [40] Tamura K, Peterson D, Peterson N, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods[J]. *Molecular Biology and Evolution*, 2011, 28(10): 2731-2739.
- [41] Augenstein Y, Rockstuhl C. Inverse design of nanophotonic devices with structural integrity[J]. *ACS Photonics*, 2020, 7(8): 2190-2196.
- [42] Zhan A, Gibson R, Whitehead J, et al. Controlling three-dimensional optical fields via inverse Mie scattering[J]. *Science Advances*, 2019, 5(10): eaax4769.
- [43] Mansouree M, McClung A, Samudrala S, et al. Large-scale parametrized metasurface design using adjoint optimization[J]. *ACS Photonics*, 2021, 8(2): 455-463.
- [44] Chen M, Jiang J, Fan J. Design space

- reparameterization enforces hard geometric constraints in inverse-designed nanophotonic devices [J]. *ACS Photonics*, 2020, 7 (11): 3141-3151.
- [45] Kubota T. 48 years with holography [J]. *Optical Review*, 2014, 21(6): 883-892.
- [46] Wan M, Healy J J, Sheridan J T. Terahertz phase imaging and biomedical applications [J]. *Optics & Laser Technology*, 2020, 122: 105859.
- [47] Cuche E, Marquet P, Depeursinge C. Spatial filtering for zero-order and twin-image elimination in digital off-axis holography [J]. *Applied Optics*, 2000, 39 (23): 4070-4075.
- [48] Tanaka K, Hara M, Tokuyama K, et al. Improved performance in coaxial holographic data recording [J]. *Optics Express*, 2007, 15(24): 16196-16209.
- [49] Tanaka Y, Tani S, Murata S. Phase retrieval method for digital holography with two cameras in particle measurement [J]. *Optics Express*, 2016, 24 (22): 25233-25241.
- [50] Zhou W J, Guan X F, Liu F F, et al. Phase retrieval based on transport of intensity and digital holography [J]. *Applied Optics*, 2017, 57(1): A229-A234.
- [51] Khan A, Zhang Z J, Yu Y J, et al. GAN-holo: generative adversarial networks-based generated holography using deep learning [J]. *Complexity*, 2021, 2021: 1-7.
- [52] Rivenson Y, Zhang Y, Günaydın H, et al. Phase recovery and holographic image reconstruction using deep learning in neural networks [J]. *Light, Science & Applications*, 2018, 7: 17141.
- [53] Schnars U, Jptner W P O. Digital recording and numerical reconstruction of holograms [J]. *Measurement Science and Technology*, 2002, 13(9): R85-R101.
- [54] Chang C, Xia J, Yang L, et al. Speckle-suppressed phase-only holographic three-dimensional display based on double-constraint Gerchberg-Saxton algorithm [J]. *Applied Optics*, 2015, 54(23): 6994-7001.
- [55] Jin X Y, Gui J B, Liu C, et al. Progress of fast generation algorithm of computer-generated hologram based on point source model [J]. *Laser & Optoelectronics Progress*, 2018, 55(10): 100005.  
金晓宇, 桂进斌, 刘超, 等. 基于点源模型计算全息图快速生成算法的研究进展 [J]. *激光与光电子学进展*, 2018, 55(10): 100005.
- [56] Shi L, Li B, Kim C, et al. Towards real-time photorealistic 3D holography with deep neural networks [J]. *Nature*, 2021, 591(7849): 234-239.
- [57] Wu J, Liu K, Sui X, et al. High-speed computer-generated holography using an autoencoder-based deep neural network [J]. *Optics Letters*, 2021, 46 (12): 2908-2911.
- [58] Edgar M P, Gibson G M, Padgett M J. Principles and prospects for single-pixel imaging [J]. *Nature Photonics*, 2019, 13(1): 13-20.
- [59] Duarte M F, Davenport M A, Takhar D, et al. Single-pixel imaging via compressive sampling [J]. *IEEE Signal Processing Magazine*, 2008, 25(2): 83-91.
- [60] Zhang Z, Ma X, Zhong J. Single-pixel imaging by means of Fourier spectrum acquisition [J]. *Nature Communications*, 2015, 6: 6225.
- [61] Lyu M, Wang W, Wang H, et al. Deep-learning-based ghost imaging [J]. *Scientific Reports*, 2017, 7 (1): 17865.
- [62] He Y, Wang G, Dong G, et al. Ghost imaging based on deep learning [J]. *Scientific Reports*, 2018, 8(1): 6469.
- [63] Rizvi S, Cao J, Zhang K Y, et al. Deringing and denoising in extremely under-sampled Fourier single pixel imaging [J]. *Optics Express*, 2020, 28 (5): 7360-7374.
- [64] Higham C F, Murray-Smith R, Padgett M J, et al. Deep learning for real-time single-pixel video [J]. *Scientific Reports*, 2018, 8(1): 2369.
- [65] Radwell N, Johnson S D, Edgar M P, et al. Deep learning optimized single-pixel LiDAR [J]. *Applied Physics Letters*, 2019, 115(23): 231101.
- [66] Wang F, Wang H, Wang H C, et al. Learning from simulation: an end-to-end deep-learning approach for computational ghost imaging [J]. *Optics Express*, 2019, 27(18): 25560-25572.
- [67] Wu H, Wang R Z, Zhao G P, et al. Sub-Nyquist computational ghost imaging with deep learning [J]. *Optics Express*, 2020, 28(3): 3846-3853.
- [68] Luo Z, Yurt A, Stahl R, et al. Pixel super-resolution for lens-free holographic microscopy using deep learning neural networks [J]. *Optics Express*, 2019, 27(10): 13581.
- [69] Liu X Y, Wang Y H, Liu Q J. Remote sensing image fusion based on two-stream fusion network [M] // Schoeffmann K, Chalidabhongse T H, Ngo C W, et al. *MultiMedia modeling. Lecture notes in computer science*. Cham: Springer, 2018, 10704: 428-439.
- [70] Liu Y, Chen X, Peng H, et al. Multi-focus image fusion with a deep convolutional neural network [J]. *Information Fusion*, 2017, 36: 191-207.
- [71] Ma J Y, Yu W, Liang P W, et al. FusionGAN: a generative adversarial network for infrared and visible image fusion [J]. *Information Fusion*, 2019, 48: 11-

- 26.
- [72] Naik N, Zhao S, Velten A, et al. Single view reflectance capture using multiplexed scattering and time-of-flight imaging [C] // Proceedings of the 2011 SIGGRAPH Asia Conference on - SA '11, December 12-15, 2011, Hong Kong, China. New York: ACM Press, 2011: 1-10.
- [73] Repasi E, Lutzmann P, Steinvall O, et al. Advanced short-wavelength infrared range-gated imaging for ground applications in monostatic and bistatic configurations [J]. *Applied Optics*, 2009, 48(31): 5956-5969.
- [74] Garipey G, Tonolini F, Henderson R, et al. Detection and tracking of moving objects hidden from view [J]. *Nature Photonics*, 2016, 10(1): 23-26.
- [75] Vellekoop I M, Lagendijk A, Mosk A P. Exploiting disorder for perfect focusing [J]. *Nature Photonics*, 2010, 4(5): 320-322.
- [76] Judkewitz B, Horstmeyer R, Vellekoop I M, et al. Translation correlations in anisotropically scattering media [J]. *Nature Physics*, 2015, 11(8): 684-689.
- [77] Li L, Li Q, Sun S, et al. Imaging through scattering layers exceeding memory effect range with spatial-correlation-achieved point-spread-function [J]. *Optics Letters*, 2018, 43(8): 1670-1673.
- [78] Li Y Z, Xue Y J, Tian L. Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media [J]. *Optica*, 2018, 5(10): 1181-1190.
- [79] Yaqoob Z, Psaltis D, Feld M S, et al. Optical phase conjugation for turbidity suppression in biological samples [J]. *Nature Photonics*, 2008, 2(2): 110-115.
- [80] Bertolotti J, van Putten E G, Blum C, et al. Non-invasive imaging through opaque scattering layers [J]. *Nature*, 2012, 491(7423): 232-234.
- [81] Schott S, Bertolotti J, Léger J F, et al. Characterization of the angular memory effect of scattered light in biological tissues [J]. *Optics Express*, 2015, 23(10): 13505-13516.
- [82] Lyu M, Wang H, Li G W, et al. Learning-based lensless imaging through optically thick scattering media [J]. *Advanced Photonics*, 2019, 1(3): 036002.
- [83] Zhou Z Y, Xia J, Wu J, et al. Learning-based phase imaging using a low-bit-depth pattern [J]. *Photonics Research*, 2020, 8(10): 1624-1633.
- [84] Rivenson Y, Göröcs Z, Günaydin H, et al. Deep learning microscopy [J]. *Optica*, 2017, 4(11): 1437-1443.
- [85] Chen X, Li Y, Wyman N, et al. Deep learning provides high accuracy in automated chondrocyte viability assessment in articular cartilage using nonlinear optical microscopy [J]. *Biomedical Optics Express*, 2021, 12(5): 2759-2722.
- [86] Zhang H, Cao L C, Jin G F, et al. Progress on lensless digital holography imaging based on compressive holographic algorithm [J]. *Laser & Optoelectronics Progress*, 2020, 57(8): 080001. 张华, 曹良才, 金国藩, 等. 基于压缩感知算法的无透镜数字全息成像研究 [J]. *激光与光电子学进展*, 2020, 57(8): 080001.
- [87] Wu X Y, Yu Y J, Zhou W J, et al. 4f amplified in-line compressive holography [J]. *Optics Express*, 2014, 22(17): 19860-19872.
- [88] Moon I, Jaferzadeh K, Kim Y, et al. Noise-free quantitative phase imaging in Gabor holography with conditional generative adversarial network [J]. *Optics Express*, 2020, 28(18): 26284-26301.
- [89] Zhang H, Liu S W, Cao L C, et al. Noise suppression for ballistic-photons based on compressive in-line holographic imaging through an inhomogeneous medium [J]. *Optics Express*, 2020, 28(7): 10337-10349.
- [90] Zaharia M, Chowdhury M, Das T, et al. Resilient distributed datasets: a fault-tolerant abstraction for in-memory cluster computing [C] // Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation, April 25-27, 2012, San Jose, CA, USA. Berkeley: USENIX Association, 2012.
- [91] Schwabe R J, Zelinger S, Key T S, et al. Electronic lighting interference [J]. *IEEE Industry Applications Magazine*, 1998, 4(4): 43-48.
- [92] Hu W, Li X, Yang J, et al. Crosstalk analysis of aligned and misaligned free-space optical interconnect systems [J]. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 2010, 27(2): 200-205.
- [93] Lin X, Rivenson Y, Yardimci N T, et al. All-optical machine learning using diffractive deep neural networks [J]. *Science*, 2018, 361(6406): 1004-1008.
- [94] Veli M, Mengu D, Yardimci N T, et al. Terahertz pulse shaping using diffractive surfaces [J]. *Nature Communications*, 2021, 12(1): 37.
- [95] Qian C, Lin X, Lin X, et al. Performing optical logic operations by a diffractive neural network [J]. *Light, Science & Applications*, 2020, 9: 59.
- [96] Goi E, Chen X, Zhang Q, et al. Nanoprinted high-neuron-density optical linear perceptrons performing near-infrared inference on a CMOS chip [J]. *Light, Science & Applications*, 2021, 10(1): 40.

- [97] Zhou T K, Lin X, Wu J M, et al. Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit[J]. *Nature Photonics*, 2021, 15(5): 367-373.
- [98] Li J X, Mengu D, Yardimci N T, et al. Spectrally encoded single-pixel machine vision using diffractive networks [J]. *Science Advances*, 2021, 7 (13): eabd7690.
- [99] Shen Y C, Harris N C, Skirlo S, et al. Deep learning with coherent nanophotonic circuits [C] // 2017 IEEE Photonics Society Summer Topical Meeting Series (SUM), July 10-12, 2017, San Juan, PR, USA. New York: IEEE Press, 2017: 189-190.
- [100] Hughes T W, Williamson I A D, Minkov M, et al. Wave physics as an analog recurrent neural network [J]. *Science Advances*, 2019, 5(12): eaay6946.
- [101] Qian C, Zheng B, Shen Y C, et al. Deep-learning-enabled self-adaptive microwave cloak without human intervention[J]. *Nature Photonics*, 2020, 14 (6): 383-390.
- [102] Liu H L, Xie D, Shen H Y, et al. Functional micro-nano structure with variable colour: applications for anti-counterfeiting [J]. *Advances in Polymer Technology*, 2019, 2019: 1-26.
- [103] Yang H, Niu Z K, Xiao S L, et al. Fast and accurate optical fiber channel modeling using generative adversarial network [J]. *Journal of Lightwave Technology*, 2021, 39(5): 1322-1333.
- [104] Ibáñez M B, Di Serio Á, Villarán D, et al. Experimenting with electromagnetism using augmented reality: impact on flow student experience and educational effectiveness [J]. *Computers & Education*, 2014, 71: 1-13.
- [105] Zhang L P, Zhang L F, Du B. Deep learning for remote sensing data: a technical tutorial on the state of the art[J]. *IEEE Geoscience and Remote Sensing Magazine*, 2016, 4(2): 22-40.
- [106] Zhou Y, Tuzel O. VoxelNet: end-to-end learning for point cloud based 3D object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 4490-4499.
- [107] Liu J M, Wang P P, Zhang X K, et al. Deep learning based atmospheric turbulence compensation for orbital angular momentum beam distortion and communication[J]. *Optics Express*, 2019, 27(12): 16671-16688.
- [108] Zhou H Q, Huang L L, Wang Y T. Deep learning algorithm and its application in optics [J]. *Infrared and Laser Engineering*, 2019, 48(12): 1226004. 周宏强, 黄玲玲, 王涌天. 深度学习算法及其在光学的应用 [J]. *红外与激光工程*, 2019, 48(12): 1226004.

## Advances in Computational Optics Based on Deep Learning

Wang Yitong, Zhou Hongqiang, Yan Jingxiao, He Cong, Huang Lingling\*

*Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China*

### Abstract

**Significance** With the advent of the period of big data, deep learning is playing a very important role in daily life and scientific research, and it has been widely applied in image processing, speech recognition, autonomous driving, and other fields. Deep learning in the optical field, as a data-driven algorithm, can effectively improve computational efficiency and imaging quality, approaching or even breaking through physical limits. An artificial neural network is a basic form of imitating biological neurons and the working principle of the human brain to complete learning process of internal principle or to extract target features. Different kinds of complicated neural networks are put forward and geared to the demands of different application scenarios for the target that often requiring different network architectures. Furthermore, an optical neural network using photons as the medium can break through the limitations of a traditional electronic neural network and provide high speed and low loss advantages.

**Progress** In this paper, we analyze the applications of deep learning in micro-nano structure design and spectral response prediction, holographic imaging application, optical sensing and imaging technology, new photon-driven neural network, and other directions in detail through examples. We also list the challenges existing in the combination of deep learning and optics, and the development directions of this field have prospected. First,

predicting spectral response based on existing micro-nano structures is an important step in micro-nano optics. The neural network can predict the near-field electromagnetic response, far-field scattering mode, Poynting vector, and other physical quantities without time-consuming calculations, which indicates that data-driven deep learning can simulate the propagation and distribution of electromagnetic fields (Fig. 1). The use of neural networks can significantly improve prediction speed by several orders of magnitude. The process of designing micro-nano structures to get the ideal resonance response is called inverse design. Most early applications are designed for fixed micro-nano structures with some structural parameters as variables (Fig. 2). There are non-unique solutions in the inverse design of micro-nano structures, which cause that the neural network is hard to converge. To solve the problem, a cascaded neural network is adopted, and deep learning is used in other complex response predictions for micro-nano structure (Fig. 3). In addition, a generative adversarial network (GAN) can be used to improve the generalization and diversity of inverse design (Fig. 4), but its shortcoming is poor interpretability. GAN can be upgraded to a multi-layer GAN system in the following work, which can predict more structural parameters and have various functions (Fig. 5). Optimization algorithms commonly used in the inverse design include gradient optimization algorithms, genetic algorithms, etc. As a powerful optimization method, deep learning combined with topology is a good development prospect in micro-nano inverse design (Fig. 7).

Second, deep learning can learn rules or detailed features from data. The advantages of deep learning can replace conventional optimization algorithms to improve image quality and work efficiently, especially in holographic imaging. Deep learning is widely used in phase recovery and image reconstruction of holograms (Fig. 8). Compared with the conventional way of generating computer-generated hologram (CGH), deep learning can improve the quality and computing speed of CGH, and ultra-lightweight real-time 3D holography can be implemented (Fig. 9).

Moreover, optical sensing and computational imaging techniques based on deep learning are a hot topic. Computational imaging technique such as single-pixel imaging based on traditional algorithms often requires scanning acquisition and calculation, which consumes a lot of computing time and power. In the case of a low sampling rate, its imaging quality cannot be guaranteed. With the advantages of data-driven deep learning, the neural network adds new vitality to single-pixel imaging technology (Figs. 11–13). Unlike computational imaging, optical sensing imaging technology employs laser-active lighting to collect optical signals via optical sensors or receivers and reproduce object information via target signal processing. Combining deep learning with non-visual imaging and replacing a single optical model with a data-driven method frequently results in high speed and adaptability, as illustrated (Figs. 15–16). Deep learning also has good applications in other optical sensing imaging fields, such as microscopic imaging, three-dimensional imaging, fusion imaging, etc.

As a kind of electromagnetic wave, light is different from an electrical signal in the propagation process and is not interfered by an external electromagnetic field. Meanwhile, as a carrier of information transmission, light has high bandwidth, high transmission speed, and low loss. Researchers began to consider using light to replace electrons to establish a neural network in the optical field. A diffractive all-optical neural network can be used in image classification, optical logic operation, spectral aiming to classification, and other tasks (Figs. 18–23). Deep learning has also been applied to other optical fields, such as optical cloaking, optical anti-counterfeiting, multichannel modeling, etc.

**Conclusions and Prospects** In conclusion, there are still some challenges in the combination of optics and deep learning algorithms. Here we put forward several prospects. First, physical models are incorporated into the deep learning computing process to increase the interpretability of the models to reduce the network's dependence on data. Second, the combination of experimental and simulation data is necessary. Expanding the training set by using multiple algorithms to generate new data from real data can reduce the difficulty of training. Third, the combination of deep learning and traditional optimization algorithms can complement each other to a certain extent, improving the performance of the network model and greatly shorten the computing time. It is also beneficial to enhance the generalization and robustness of the network. Finally, an optical diffractive neural network breaks through the limitations of traditional electronic neural networks. But currently, most of the all-optical neural network is only for a single task, and it has low stability on-chip photonic neural network and other problems. More attempts and studies are needed, and there is still plenty of room for improvement.

**Key words** optics in computing; deep learning; neural network; metasurface; computational optics

**OCIS codes** 200.4260; 230.4000; 090.1760