



ELSEVIER

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng



Research

Optical Neural Networks—Article

可见光波段的深度衍射神经网络

陈航^a, 冯佳楠^a, 江闽伟^a, 王逸群^b, 林杰^{a,c,*}, 谭久彬^a, 金鹏^{a,c,*}

^a Center of Ultra-precision Optoelectronic Instrument, Harbin Institute of Technology, Harbin 150001, China.

^b Nanofabrication Facility, Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou 215123, China.

^c Key Laboratory of Micro-Systems and Micro-Structures Manufacturing, Ministry of Education, Harbin Institute of Technology, Harbin 150001, China

ARTICLE INFO

Article history:

Received 10 February 2020

Revised 19 May 2020

Accepted 20 July 2020

Available online 13 February 2021

关键词

光计算

光学神经网络

深度学习

光学机器学习

深度衍射神经网络

摘要

基于衍射光学元件的光学深度学习在并行处理、计算速度和计算效率方面有着独特优势。深度衍射神经网络(D²NN)是其中一项具有里程碑意义的研究工作。D²NN在太赫兹波段通过3D打印进行神经网络的物理固化。鉴于太赫兹波段下存在的粒子间耦合限制和材料损耗,本文将D²NN的应用波段延展至可见光波段,并提出了包括修订公式在内的一般理论,解决了工作波长、人工神经元特征尺寸和加工制备之间的矛盾。在632.8 nm的工作波长下,本文提出了一种新颖的可见光D²NN分类器,可用于原始目标(手写数字0~9)和已更改目标(被遮盖和涂改目标)的目标识别。本文获得的实验分类精度(84%)和数值分类精度(91.57%)量化了理论设计和制造系统性能之间的匹配程度。本文所提出的一般理论模型可将D²NN应用于各种实际问题或设计全新的应用场景。

©2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. 引言

深度学习是一种通过计算机构建多层人工神经网络实现数据表示、抽象和高级任务的机器学习方法[1]。近些年来,深度学习的快速发展引起了学术界的广泛关注。在人工智能、计算机科学等相关领域,深度学习甚至已经远超人类专家的水准。目前,深度学习已经被广泛应用于计算机视觉[2]、语音或图像识别[3,5]、机器人系统[6],以及其他领域[7,9]。然而,计算机平台下的深度学习受制于von Neumann结构的限制,在处理时间和计算能耗方面难以取得突破[10–11]。在过去的几十年间,光学系统中的光学信息处理方式实现了卷积运算、相关运算和傅里叶变

换运算等诸多运算范式,光学信息处理方式在并行计算上更展现出独特的优势[12,17]。通过设计衍射光学元件,研究人员已经在光学系统中构建出传统计算机端的深度学习框架,并实现诸多智慧功能[18,21]。基于3D打印技术在太赫兹波段下实现的深度衍射神经网络(D²NN)是其中一项里程碑式的研究工作[18]。

D²NN基于深度学习算法框架和误差反向传播理论,在计算机端进行学习训练。D²NN包含多层衍射光学元件,每层衍射光学元件上的微纳结构单元充当人工神经元。在学习训练过程中,这些人工神经元的复值透射系数(振幅/相位)得以优化,进而控制层间的光学衍射传播方式,以特定的输出光场分布形式实现预期设定的智慧任

* Corresponding authors.

E-mail addresses: linjie@hit.edu.cn (J. Lin), P.jin@hit.edu.cn (P. Jin).

2095-8099/© 2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

英文原文: *Engineering* 2021, 7(10): 1483–1491

引用本文: Hang Chen, Jianan Feng, Minwei Jiang, Yiqun Wang, Jie Lin, Jiubin Tan, Peng Jin. Diffractive Deep Neural Networks at Visible Wavelengths. *Engineering*, <https://doi.org/10.1016/j.eng.2020.07.032>

务。一旦学习训练完毕,系统便可通过物理加工手段进行固化,除去初始的相干光源和用于接收输出光场的探测器需要供电外,整个系统将无电化运行。基于 D^2NN 的诸多应用,如具备非线性激活函数(引入铁电薄膜材料)并在傅里叶空间实现的 D^2NN [17],以及针对手写数字和时尚产品进行目标识别的线性 D^2NN [18,21]等,已经被广泛研究。

通过3D打印在太赫兹波段下实现的 D^2NN 已经取得巨大的成功,其有效性和良好性能已经在诸多研究中得以验证[18,20]。尽管太赫兹波段下的 D^2NN 存在着巨大优势,但其在实际应用中仍存在一些众所周知的局限性,如材料损耗和粒子间耦合[22]。因此,将 D^2NN 从太赫兹波段扩展至可见光波段或近红外波段[23]具有重要科研意义和实际应用价值;将 D^2NN 从太赫兹波段扩展至可见光波段还会赋予系统更加新颖的应用场景[24,29]。然而,将 D^2NN 从长波长扩展至短波长,必须解决一些限制和问题,如工作波长、人工神经元特征尺寸和制备加工能力间存在的矛盾。工作波长越短,人工神经元特征尺寸越小,制备加工越困难,而经典最大半锥衍射角理论无法克服这一矛盾[18,20]。

为了克服上述矛盾,本文提出了一种一般理论模型,该模型包含了 D^2NN 系统设计过程中所需考虑的完备参数。该一般理论模型是在经典最大半锥衍射角理论的基础上,通过引入修订公式得以实现的。该理论模型定量分析和明确了将 D^2NN 应用扩展至其他波段,如氦(He)-氖(Ne)激光器等可见光波段下的系统设计参数和方法。本文设计了一系列的仿真实验,对所提出的一般理论模型进行验证。具体来说,本文以 D^2NN 分类器为例来验证所提出的一般理论模型。为了减少训练时间和训练复杂度,本文仅采用手写数据集MNIST中的部分子集进行训练,实现针对手写数字0~9的识别分类功能。在此情况下,为防止过拟合,本文设计了五层纯相位调制型 D^2NN ,共计32 000个神经元。通过采用全新测试集对系统进行盲测分类,实现了对所提出的一般理论模型的验证。

基于所提出的一般理论模型,本文进一步设计实现了一种全新功能的可见光 D^2NN 分类器,并使用He-Ne激光器在工作波长632.8 nm处进行了实验验证。与已有的 D^2NN 分类器相比[18,20],本文提出的可见光 D^2NN 分类器即使在目标发生变化的情况下(如数字被遮挡或涂改),也可对目标物进行正确分类。在学习训练中,本文将MNIST数据集中的55 000个手写数字训练样本扩展到80 000个,其中新增的25 000个手写数字实物为遮挡或涂改的MNIST训练数据集。本文使用包含500万个神经元

的五层相位型 D^2NN ,对全新测试集内11 000幅样本图像(包含MNIST数据集中的10 000个手写数字测试样本和1000个遮挡或涂改的手写数字测试样本)进行盲测识别分类,数值分类精度为91.57%。随后本文采用多步光刻-套刻-刻蚀工艺在二氧化硅(SiO_2)基底上制备实现了上述可见光 D^2NN 。在实验过程中,本文采用50个手写数字实物作为输入,对系统进行实验测试,系统的实验分类精度为84%。可见,本文获得的实验分类精度(84%)和数值分类精度(91.57%)量化了理论设计和制造系统性能之间的匹配。

基于所提出的一般理论模型,对 D^2NN 进行仿真设计,本文实现了可见光波段下的 D^2NN 。深度神经网络有时被称为黑匣子[30],本文所提出的可见光 D^2NN 可为剖析该问题提供一定的解法思路。此外,充分了解人类大脑中生物神经元间的相互作用,能够更好地构建深度神经网络。本文所提出的可见光 D^2NN 可以为当前生物神经元之间的相互作用提供新的见解思路,并且所提出的框架理论可以将 D^2NN 应用于各种人机交互设备和人工智能交互设备下的实际问题,以促进生物学和计算科学发展。

2. 材料和方法

2.1. 深度学习和深度衍射神经网络架构

本文所提出的可见光 D^2NN 遵循光学衍射理论和深度学习框架[32,35]。与传统的深度学习系统相比, D^2NN 有以下几点不同:① D^2NN 的前向衍射传播遵循惠更斯理论(Huygens' principle)和瑞利-索末菲光学衍射理论(Rayleigh-Sommerfeld diffraction)[36];②训练用的损失函数,如交叉熵函数(SCE)基于输出平面不同位置探测器的光强分布。在此基础上,本文可以基于目标理想光强分布来计算损失函数,并由此进行误差反向传播来优化网络结构以及相应的相位值分布。

在深度学习框架中, D^2NN 每层衍射层中每个人工神经元的相位值,将依据所要实现的特定任务进行不断地学习迭代调整。系统通过输入训练数据,经光场衍射传播,最终在输出平面上得到特定的光场分布。在对 D^2NN 前向传播模型进行详细分析讨论前,本文先对 l 层进行如下相关定义(其中 l 层表示 D^2NN 中的第 l 层衍射层)。

• t^l 表示 l 层衍射层的复值传递系数。 i 表示在第 l 层空间坐标为 (x_i, y_i, z_i) 的人工神经元。对于第 i 个人工神经元,其复值传递系数由振幅和相位组成,并可进一步描述为: $t_i^l(x_i, y_i, z_i) = a_i^l(x_i, y_i, z_i) \exp(j\Phi_i^l(x_i, y_i, z_i))$,其中 $a_i^l(x_i, y_i, z_i)$ 为复值传递系数中的振幅项, $\Phi_i^l(x_i, y_i, z_i)$ 为复值传递系数中

的相位项, j 为虚数并满足 $j = \sqrt{-1}$ 。对于纯相位型 D^2NN 系统, 振幅项 $a_i^l(x_i, y_i, z_i)$ 往往取常数 1。

• W^{l+1} 表示经过 l 层衍射层传播后的次波衍射函数。对于第 l 层空间坐标为 (x_l, y_l, z_l) 的人工神经元, 其次波衍射函数为 W_i^{l+1} 。 W^{l+1} 的振幅和相位取决于输入波函数和对应的 l 层衍射层的复值传递系数 t^l 。

• Y^{l+1} 表示次波衍射函数 W^{l+1} 经过 l 和 $l+1$ 层之间的距离传输后的输出光场函数。 Y^{l+1} 也可表示为 $l+1$ 层衍射层的输入光场函数。当输入光场函数 Y^l 垂直入射到第 l 层时, W^{l+1} 为透射函数 t^l 和输入光场函数 Y^l 之间的乘积:

$$W^{l+1} = Y^l \circ t^l \quad (1)$$

式中, \circ 表示哈密顿乘积操作。

光场由第 l 层传输到第 $l+1$ 层后, 会引入一个相移因子, 如下所示:

$$\exp(jkd_l \cos \gamma) = \exp\left[jkd_l \sqrt{1 - (\cos^2 \alpha + \cos^2 \beta)}\right] \quad (2)$$

式中, d_l 表示第 l 层和第 $l+1$ 层之间的传播距离; α 、 β 和 γ 是次波衍射函数 W^{l+1} 沿 x 轴、 y 轴、 z 轴方向的各自方位角; k 为波数并满足 $k = 2\pi/\lambda$, λ 表示工作波长。

假定 $u = \frac{\cos \alpha}{\lambda}$ 、 $v = \frac{\cos \beta}{\lambda}$, 其中 u 和 v 为空间频率。

在经过由 l 层到 $l+1$ 层的距离传播后, 相应的相移因子可以表示为:

$$\exp(jkd_l \cos \gamma) = \exp\left[jkd_l \sqrt{1 - \lambda^2(u^2 + v^2)}\right] \quad (3)$$

随后, 我们可以推导得到

$$Y^{l+1} = W^{l+1} \cdot \exp\left[jkd_l \sqrt{1 - \lambda^2(u^2 + v^2)}\right] \quad (4)$$

最终, D^2NN 的前向传播模型概括如下:

$$Y^{l+1} = (Y^l \circ t^l) \cdot \exp\left[jkd_l \sqrt{1 - \lambda^2(u^2 + v^2)}\right] \quad (5)$$

2.2. 基于 TensorFlow 设计 D^2NN 和处理流程

本文中的 D^2NN 仿真设计, 是基于 Python (v3.7) 以及 TensorFlow (v1.12.0, Google Inc., USA) 框架所提出的, 并在计算机端经过了 20 个循环的学习训练。计算机的配置为 Windows 10 操作系统 (Microsoft Corporation, USA)、GeForce GTX 1080 Ti 显卡、Intel® Core™ E5-2650 @2.00 GHz CPU 以及 128 GB 内存。 D^2NN 每层训练参数通过反向传播函数和随机梯度优化算法 Adam, 在学习率为 10^{-3} 下进行学习迭代优化。五层 D^2NN 训练时间约为 20 h。

手写数字实物通过激光直写工艺在玻璃基底上进行制备, 并制作成振幅型目标样本输入到 D^2NN 中。具体工艺流程为: 首先, 使用丙酮和异丙醇清洁基底。然后, 在清洁完的基底上利用电子束蒸镀工艺镀上一层厚度约为几百

纳米的铬。随后旋涂上一层正向光刻胶进行预烘焙后, 使用激光直写工艺对数字图案进行曝光。最后, 使用显影剂剥离曝光后的光刻胶, 并使用铬媒染剂剥离未覆盖的铬。通过使用丙酮和异丙醇再次清洗基底, 完成整个加工流程。

为了更好地制备可见光波段的 D^2NN , 本文使用 sigmoid 函数将每个人工神经元的相位值约束到 $0 \sim \pi$, 经过经典的多步光刻-套刻-刻蚀工艺实现制备。在制备之前, 还需要将人工神经元的相位值 Φ 进一步转化为对应的台阶高度 Δh ($\Delta h = \lambda \Phi / 2\pi \Delta n$, 其中 Δn 是基底和空气之间的相对折射率差)。 D^2NN 在制备前需要通过上述方法在 SiO_2 基底上进行清洗。清洗过后, 先使用六甲基二硅氮烷预处理来改变 SiO_2 基底表面的活性, 以增强光刻胶和基底的附着力。随后再旋涂光刻胶并进行曝光, 并使用显影剂将曝光的光刻胶进行剥离。紧接着, 基底经过氧等离子体处理后, 采用磁中性环放电刻蚀。重复上述操作, 最终可以制备得到设计的 D^2NN 。更多的设计处理过程可以参考附录 A 中的图 S1 和图 S2。

3. 结果

本文通过在 SiO_2 基底上采用多步光刻-套刻-刻蚀工艺, 将设计的 D^2NN 制成如图 1 所示的五层光学衍射层结构。对于被遮挡或涂改的目标, 现有的深度学习系统会出现误分类 [18,20]。即使在足够高的目标分类识别精度下 [20], 现有系统仍然会出现错误分类。针对这种情况, 如图 1 (a) 所示, 本文设计的手写数字 D^2NN 系统可以对出现的所有情况进行分类识别: 既可以对原始数据集中的数字 (0~9) 进行正确分类, 也可以对遮挡或涂改的数字样本进行分类识别。为了完成上述分类识别任务, 本文设计了一种纯相位五层 D^2NN 系统。系统将对 80 000 张手写数字实物进行学习训练, 其组成包含 MNIST 数据集中全部 55 000 张原始的手写数字训练样本, 以及 25 000 张对 MNIST 数据集进行遮挡和涂改变形后的手写数字实物。在仿真阶段, 手写数字实物通过振幅编码的形式输入 D^2NN 系统, 经过五层衍射层的传播, 在输出平面 11 个不同的探测区域实现映射。学习训练过程如图 1 (a) 所示: 我们将原始的手写数字 (0~9) 输出映射到编号为 0 至编号为 9 所对应的探测区域, 对于遮挡或涂改的手写数字实物统一映射到编号为 X 所对应的探测区域。探测区域的空间位置分布如图 1 (a) 所示。系统实现正确分类的判断依据为对应探测区域采集到最大的光强能量分布。

训练完成后, 使用 11 000 张全新的测试样本对系统的

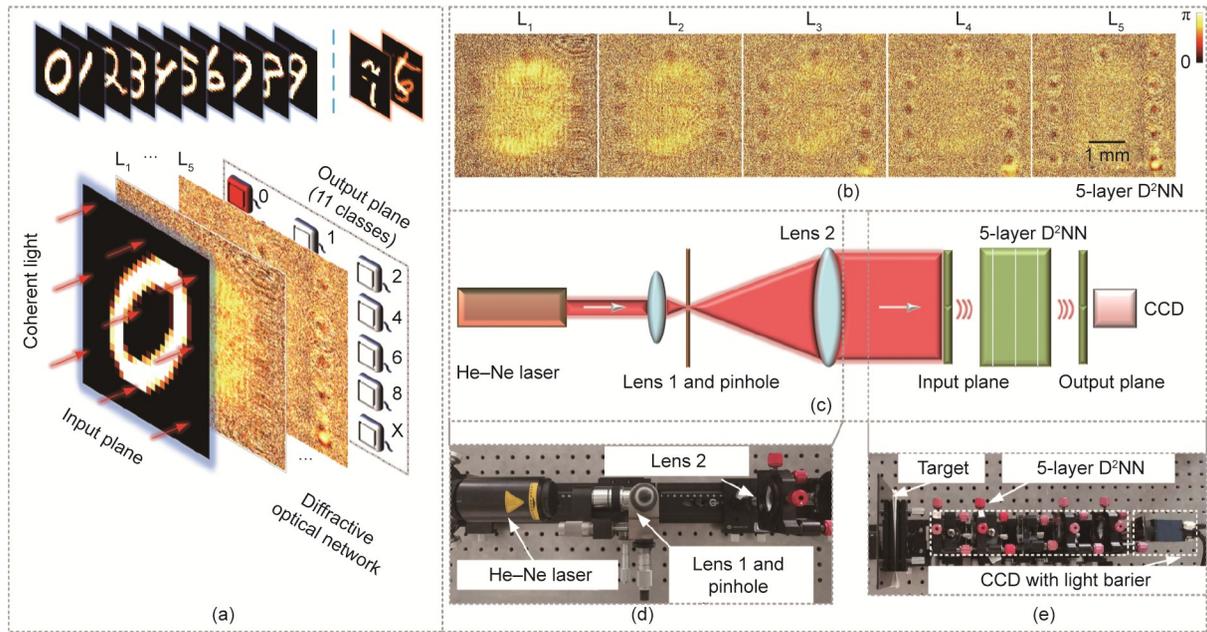


图1. 可见光D²NN实验原理图及实验装置示意图。(a)对于原始手写数字0~9以及遮挡涂改的数字(如遮挡的手写数字7和涂改的手写数字5)的分类器原理图;(b)探测器的空间位置分布图;(c)五层衍射层 L_1 、 L_2 、 L_3 、 L_4 以及 L_5 仿真相位值分布图;(d)、(e)实验装置示意图。CCD:电荷耦合器件。

数值分类精度进行测试,这些测试样本包括从MNIST测试数据集获得的全部10 000张原始手写数字测试样本以及1000张经过遮挡和涂改的手写数字实物,这些遮挡和涂改的数字样本是在MNIST测试数据集的基础上变形的。对于上述测试集,系统达到了91.57%的数值分类精度。

仿真测试后,本文将图1(b)中五层D²NN的相位分布进行物理固化。在此之前,需要将人工神经元的相位分布转化为对应位置的台阶高度,经由多步光刻-套刻-刻蚀工艺在SiO₂基底上完成加工制备。制备完成后的系统,依据图1(c)的实验原理图进行测试平台的搭建和表征。在实验测试中,本文采用一台工作波长为632.8 nm的He-Ne激光器(25-STP-912-230, Melles Griot, USA)作为入射光源。He-Ne激光器的输出经透镜1和透镜2进行准直扩束,并通过针孔进行空间滤波。经准直扩束和空间滤波后的激光入射至输入平面的目标图像,经过五层D²NN的层层衍射调制,在电荷耦合器件(CCD,大恒光电)探测平面进行接收。考虑到训练样本集的样本数量,本文设计的D²NN的每层结构包含100万个人工神经元(1000×1000),五层共计500万个人工神经元。激光器的输出功率为5 mW。每个人工神经元的特征尺寸为4 μm×4 μm,单层D²NN的结构尺寸为4 mm×4 mm。D²NN的层间距离为5 cm。对于五层D²NN的对准过程细节详见附录A中的图S3。

对于11 000张全新测试样本,本文设计的五层D²NN

数值分类精度为91.57%。本文采用图2(a)中混淆矩阵的方式对11 000张测试样本的数值分类精度进行直观展示。通过激光直写技术制备的50个手写数字实物,实验分类精度达到84%。实验分类精度相较于数值分类精度仅有较小幅度的降低。本文设计了一种特殊的遮挡器件用于获得输出光场,其中,编号为0至编号为X的探测区域可透光,并由CCD进行探测,其余部分光场信息由遮挡器件遮挡,无法透光。在测试过程中,本文首先测试原始手写数字实物和遮挡涂改后的手写数字实物。本文将手写数字3、涂改的手写数字3、手写数字4、遮挡的手写数字4作为输入目标输入系统,如图3(a)所示。图3(b)所展示的仿真结果和实验测试结果表明,可见光D²NN可以很容易对原始目标和遮挡涂改目标进行分类。如图3(c)所示,光强能量分布表明系统能够正确识别输入目标。随后本文使用四种不同形式的手写数字6进行测试,如图4所示。图4(c)的仿真结果和实验结果表明,所制备的可见光D²NN实现了正确分类。可见光D²NN可以将输出平面的光强集中到对应数字的探测区域。

综上所述,本文提出的可见光D²NN在工作波长632.8 nm处能够成功识别原始目标(手写数字0~9)和被遮挡涂改的目标。此外,本文提出的可见光D²NN系统还具有迁移学习的能力,如图5所示。当激光不经过任何的手写数字实物直接照射到D²NN系统时,之前的系统仍会将其衍射到某一探测区域[18,20],这表明此时的入射光被

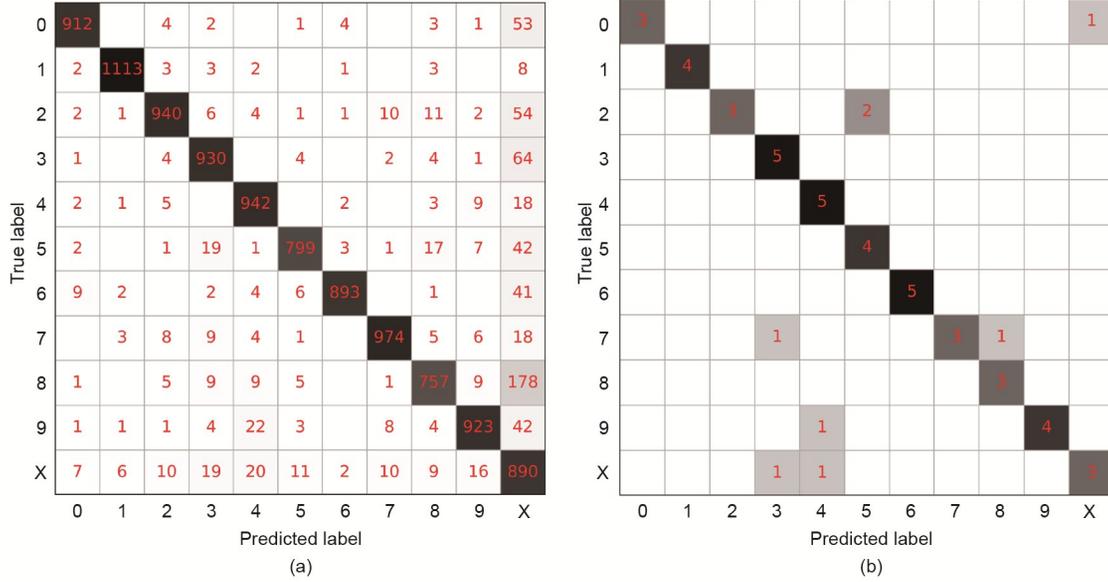


图2. 仿真结果和实验测试结果的混淆矩阵。(a) 仿真结果的混淆矩阵。对于11 000张全新测试样本，五层D²NN的数值分类精度为91.57%。(b) 对于50个手写数字实物，系统的实验分类精度为84%。

误判为某一类手写数字。当激光直接入射到本文提出的D²NN系统时，系统将能量聚焦到编号为X对应的探测区域。图5(b)所示的实验测试结果和仿真结果相符合。

本文提出的D²NN一般理论可以解决系统从长波长延展至短波长应用时出现的矛盾。本文所进行的定量分析也证明了依据该理论构建可见光D²NN系统的正确性，该理论解决了入射波长、人工神经元特征尺寸和加工难度之间的矛盾。

层间人工神经元的连通性将直接影响信息的衍射传播，进而影响D²NN系统的信息传递和推理性能。全连通D²NN可以实现人工神经元间充分的信息传递。全连通D²NN意味着任一人工神经元的衍射角足够大并可覆盖下一层的全部人工神经元。全连通系统可通过最大半锥衍射角(φ_{\max})定性描述，其受入射波长和人工神经元特征尺寸调控，具体如下所示：

$$\varphi_{\max} = \sin^{-1}(\lambda/2d_f) \quad (6)$$

式中， d_f 表示人工神经元特征尺寸。

为了获得较大的衍射角，需要研究人员设计较小的人工神经元尺寸和考虑较大的工作波长。在之前的研究工作[18]中，研究人员提出的D²NN系统使用0.75 mm的太赫兹波长，人工神经元特征尺寸为0.4 mm × 0.4 mm，最大半锥衍射角约为70°。对于可见光，如本文使用的He-Nc激光器，其工作波长为632.8 nm，约为太赫兹波长的1/1200。为了能够获得70°的最大半锥衍射角，需要将人工神经元的特征尺寸设计为小于330 nm，这也约为太赫兹波段下人工神经元特征尺寸的1/1200。如果采用330 nm作为人

工神经元的特征尺寸来设计制造系统，将对设计制作工艺提出巨大挑战，即出现了工作波长、人工神经元特征尺寸和加工难度之间的矛盾。因此，在设计可见光D²NN时，需要提出一种更为通用的设计思路。

对于相邻两个衍射层层间的传播距离 D ，每个人工神经元的衍射半径 R 可以表示为：

$$R = D \times \tan \varphi_{\max} \quad (7)$$

如果 d_f 为人工神经元的特征尺寸， N 为每层人工神经元的总数，则每个衍射层的尺寸为：

$$w = \sqrt{N \times d_f^2} \quad (8)$$

式中，通常假定衍射层为方形[18,20]。为了能够得到全连通的D²NN，每个人工神经元的衍射半径应满足 $R \geq w$ 。因此在传统最大半锥衍射角理论的基础上，本文提出一个全新的公式，如下所示：

$$\begin{cases} \varphi_{\max} = \sin^{-1}(\lambda/2d_f) \\ D \geq \sqrt{N} \cdot d_f \cdot \sqrt{4d_f^2/\lambda^2 - 1} \end{cases} \quad (9)$$

改进后得出的一般公式可以用于定量分析D²NN的连通性。当参数满足公式(9)时，全连通的D²NN得以实现。该公式表明，D²NN的连通性受工作波长、人工神经元特征尺寸、人工神经元数量、层间距离等的共同影响。可以通过调节人工神经元数量、层间距离等解决在短波长，如可见光波段下，人工神经元大小和加工制备能力之间的矛盾。该公式可以将D²NN应用于任何工作波长下的场景。

图2至图5中的实验参数皆是依据公式(9)进行选

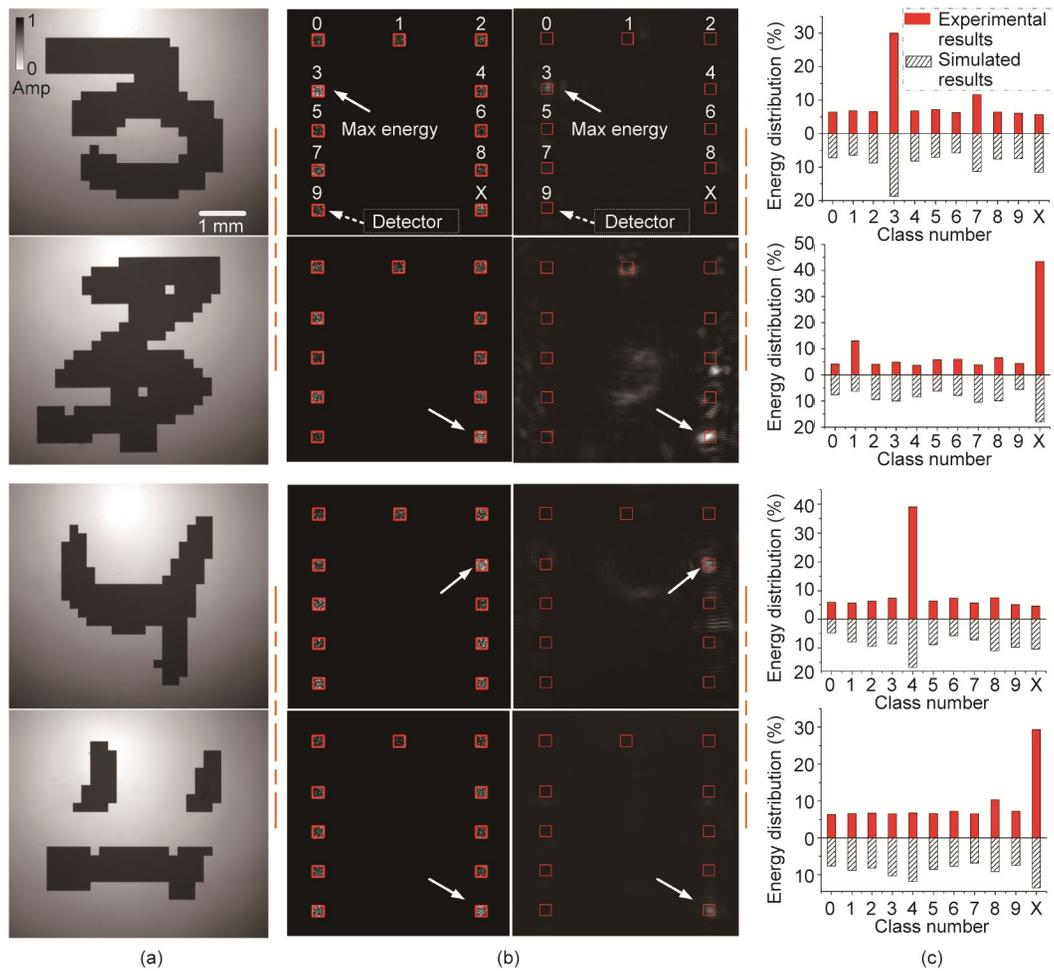


图3. 用于可见光 D^2NN 手写数字的分类结果。(a) 在光学显微镜下的样本图像, 包括手写数字3、涂改的手写数字3、手写数字4、遮挡的手写数字4。Amp: 振幅。(b) 仿真结果和实验测试结果表明了 D^2NN 成功将样本分类到对应的探测区域, 其中每个探测区域对应一种手写数字类型。例如, 手写数字3和手写数字4输出能量集中到编号为3和编号为4对应的探测区域(如白色箭头所示区域)。涂改的手写数字3和遮挡的手写数字4的输出能量均集中到编号为X的探测区域。Max: 最大值。(c) 实验测试和仿真测试结果的能量分布直方图, 表明了可见光 D^2NN 的有效性以及推理能力。

取, 所产生的实验结果验证了公式(9)的正确性。为了进一步检验公式(9)的正确性, 本文进行了一系列的仿真分析。为了减少训练时长和复杂度, 本文使用MNIST数据集中的一个小子集, 设计训练了一个五层纯相位型 D^2NN , 用于对手写数字0~9进行分类识别。如图6所示, 训练集包含了10 000张手写数字0~9; 每一类数字大约为1000张随机抽取的样本图片。将这10 000张手写数字实物以振幅编码的形式输入 D^2NN 中, 进行传播训练, 输出光场映射到探测平面10个不同区域。训练完成后, 本文使用500张样本对 D^2NN 进行分类识别性能测试。其中用于测试的500张样本均为MNIST数据集中随机抽取的全新样本。依据系统对测试集的盲测分类结果来验证所提出的理论的正确性。

为了对公式(9)进行定量分析, 本文将其拟合成分

线, 如图6(b)所示。为了防止过拟合, 根据设计经验, 本文设定每层人工神经元数量为6400(80×80)。根据图6(b)的拟合曲线所示, 当工作波长、衍射层层间距离和人工神经元特征尺寸等参数选择位于拟合曲线本身或拟合曲线上区域(如绿色箭头所指)时, D^2NN 将实现层间全连通; 当上述参数落在拟合曲线下方区域(如红色箭头所指)时, D^2NN 将无法实现全连通。在图6(b)中, 示例1和示例2所选取的参数位于拟合曲线本身, 示例3选取的参数位于拟合曲线上方, 示例4选取的参数位于拟合曲线下方。示例1至示例3所示系统, 对样本的数值分类精度均在90%以上, 但是对于示例4系统, 其数值分类精度仅约为0.1%。示例1至示例4的混淆矩阵对应图6(d)~(g)。这些结果证明了所提出的理论的正确性, 即改进后的理论是一般理论模型, 为构建 D^2NN 提供了定量分析

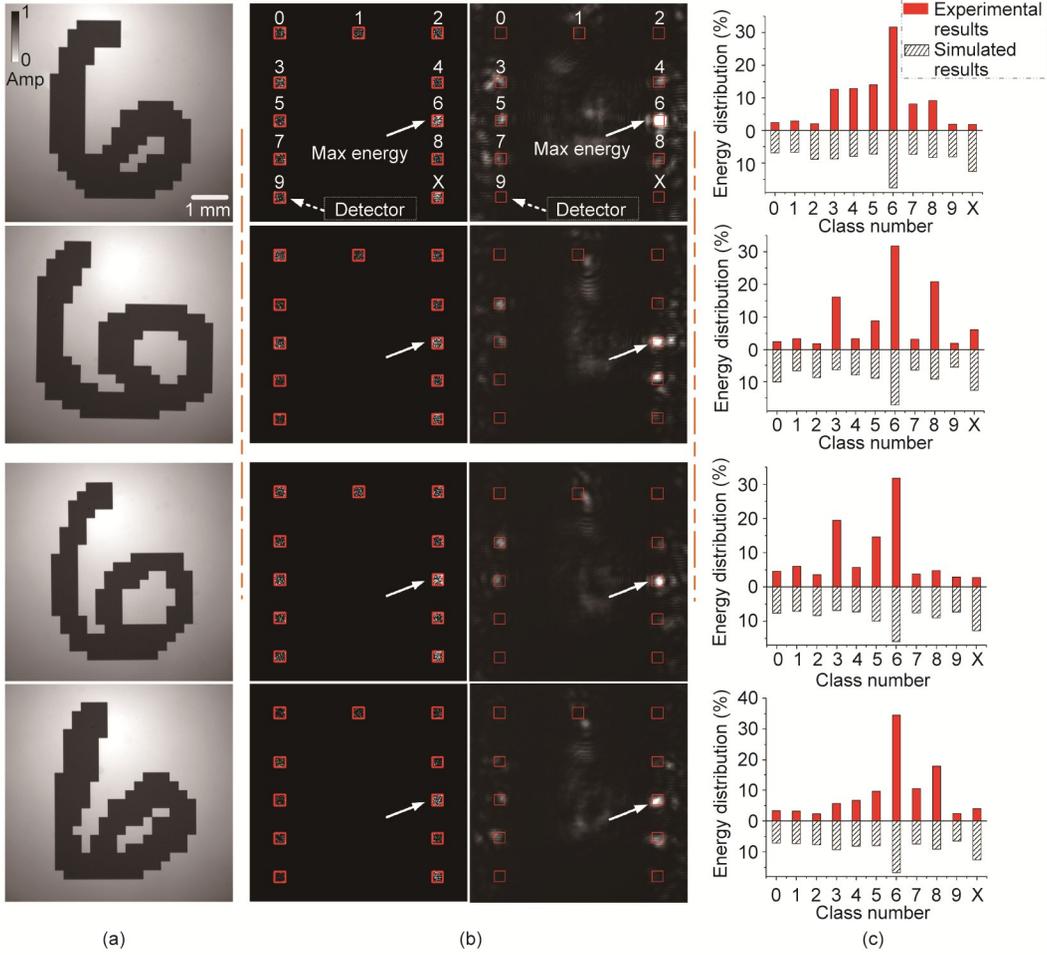


图4. 可见光D²NN手写数字分类结果。(a) 在光学显微镜下, 不同形式的手写数字6。(b) 仿真结果和实验测试结果证明D²NN对于不同类型的手写数字6能够正确分类识别。四种不同形式的手写数字6, 其输出能量均聚焦到编号为6的探测区域, 如白色箭头所示。(c) 实验测试和仿真测试结果的能量分布直方图, 表明了可见光D²NN的有效性以及推理能力。

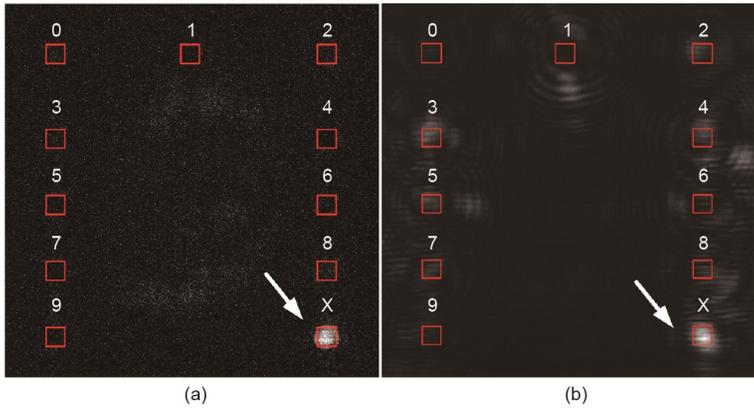
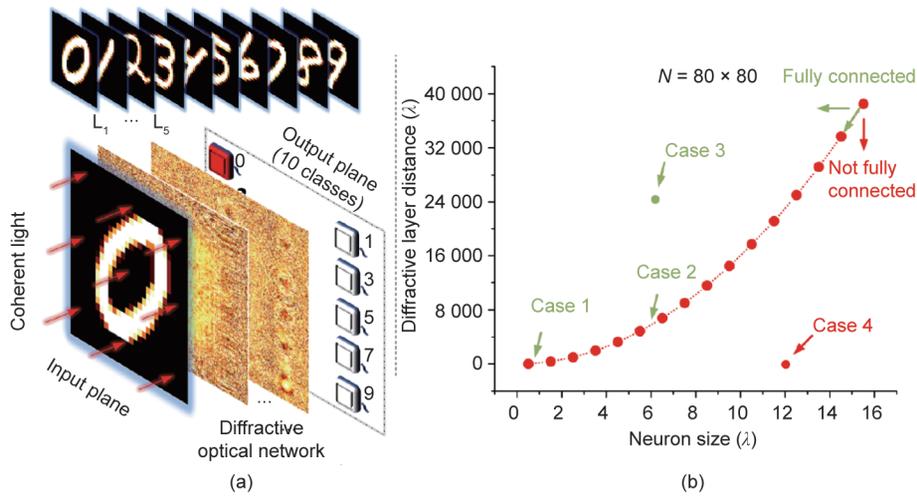


图5. 本文提出的D²NN具备迁移学习能力。(a) 将入射光直接输入系统时输出平面光场能量分布的仿真结果: 在输出平面上大部分能量集中到编号为X对应的探测区域, 表示系统将此种情况判断为错误分类。(b) 输出平面光场分布的实验测试结果, 与仿真结果相符合。

手段, 并进一步证明了全连通D²NN所带来的性能优势。图6(c)仿真结果与前人的研究[18,20]结果一致。通过对比示例1和示例2可以看出, 本文所提出的一般理论可以

通过调整衍射层层间距离, 解决人工神经元特征尺寸和加工制备难度之间的矛盾。大的层间距离 ($5.7 \times 10^3 \lambda$) 和较大的人工神经元特征尺寸 (6λ) 与小的层间距离 (15λ)



Case	Neuron size (λ)	Diffractive layer distance (λ)	Blind testing accuracy (%)
Case 1	0.53	15	90.0
Case 2	6.00	$\sim 5.7 \times 10^3$	91.0
Case 3	6.00	$\sim 2.4 \times 10^3$	91.2
Case 4	12.00	15	~ 0.1

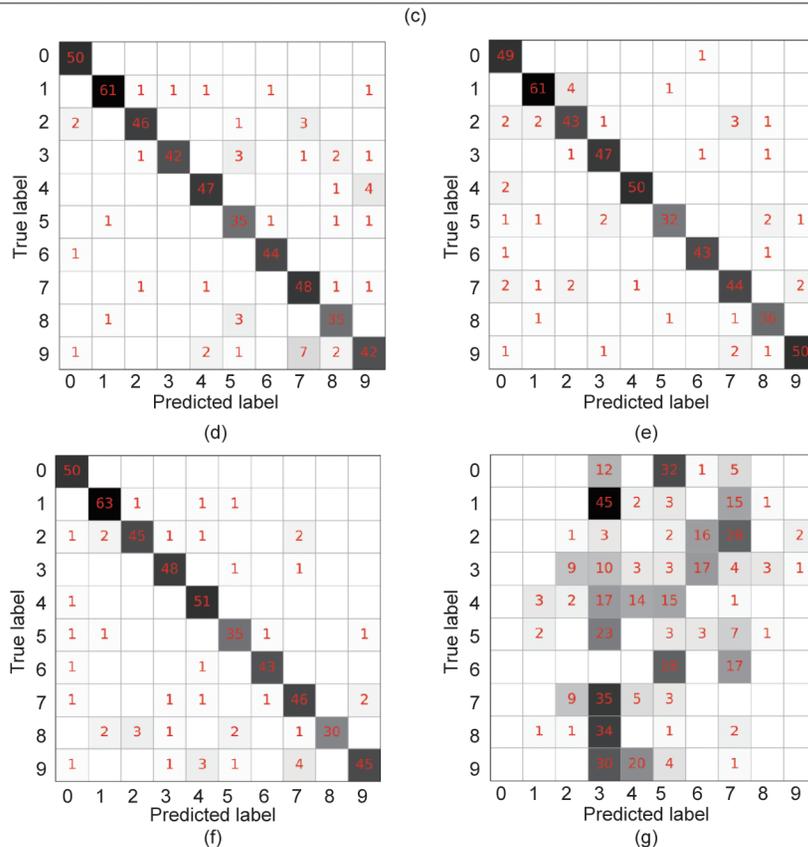


图6. 依据公式(9)设计的D²NN系统。(a)对于手写数字0~9实现分类的D²NN原理图,探测区域分布如图所示; (b)依据公式(9)设计的拟合曲线; (c)四种不同参数的测试情况。对于(c)中四种不同参数的数值分类精度进行分析研究,并给出示例1(d)、示例2(e)、示例3(f)和示例4(g)的混淆矩阵,证明所提理论的正确性。

和小的神经元特征尺寸 (0.53 λ) 性能保持一致。通过调整神经元特征尺寸和衍射层层间距离可以解决加工制备难题。本文还分析了衍射层层间对准误差和衍射层相位加工误差带来的影响, 详细的分析可以查看附录 A 中的图 S4 和图 S5。

4. 分析和结论

本文提出了构建 D²NN 的一般设计理论, 并将 D²NN 扩展到可见光波段。可见光波段的 D²NN 可以克服太赫兹波段固有的一些问题并存在诸多潜在应用价值[25,29]。然而, 直接将系统从太赫兹波段扩展到可见光波段存在一些困难, 这体现在工作波长、神经元特征尺寸和加工制备能力间的矛盾。更短的波长需要更小的人工神经元特征尺寸, 这会使加工变得愈发困难。对此, 本文提出了系统设计的一般性理论用于解决上述矛盾。通过一系列的仿真分析, 本文验证了所提理论的正确性。基于所提理论, 本文设计了一种全新的可见光波段的 D²NN 识别分类器, 其可以在工作波长为 632.8 nm 处对原始手写数字 (0~9) 和变化目标 (遮挡或涂改的目标) 进行分类识别。数值分类精度为 91.57%, 实验分类精度为 84%。数值分类精度和实验分类精度进一步验证了理论分析和系统设计制备的准确性。

尽管目前已经在光学平台上实现了 D²NN [18,22,24], 但是距离真正的全光学设计处理还有一定距离。例如, 在学习训练阶段, 系统仍需要借助计算机进行仿真。光学信息处理所带来的低能耗和高速处理优势未完全体现。此外, 光学深度学习技术的应用仍处在新兴阶段, 许多早期尝试[18,20]仍然使用的是标准机器学习模型, 这可能不是光学深度学习的最佳模型。可以尝试将其他模型, 如无监督学习[37]、生成对抗网络[38]、强化学习[39-40]等整合到光学神经网络中。预计未来, 研究人员会提出更快、更准确的光学深度学习框架, 直至超越人类现有的知识能力。

致谢

本研究得到国家自然科学基金项目 (61675056、61875048) 支持。

Compliance with ethics guidelines

Hang Chen, Jianan Feng, Minwei Jiang, Yiqun Wang,

Jie Lin, Jiubin Tan, and Peng Jin declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.eng.2020.07.032>.

References

- [1] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521(7553):436–44.
- [2] Bohannon J. Helping robots see the big picture. *Science* 2014;346(6206):186–7.
- [3] Hinton G, Deng L, Yu D, Dahl GE, Mohamed A, Jaitly N, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Process Mag* 2012;29(6):82–97.
- [4] Schofield D, Nagrani A, Zisserman A, Hayashi M, Matsuzawa T, Biro D, et al. Chimpanzee face recognition from videos in the wild using deep learning. *Sci Adv* 2019;5(9):eaaw0736.
- [5] Barbastathis G, Ozcan A, Situ G. On the use of deep learning for computational imaging. *Optica* 2019;6(8):921–43.
- [6] Giusti A, Guzzi J, Cireşan DC, He FL, Rodriguez JP, Fontana F, et al. A machine learning approach to visual perception of forest trails for mobile robots. *IEEE Robot Autom Lett* 2016;1(2):661–7.
- [7] Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, et al. Mastering the game of Go with deep neural networks and tree search. *Nature* 2016;529(7587):484–9.
- [8] Topol EJ. A decade of digital medicine innovation. *Sci Transl Med* 2019;11(498):eaaw7610.
- [9] Akhtar N, Mian A. Threat of adversarial attacks on deep learning in computer vision: a survey. *IEEE Access* 2018;6:14410–30.
- [10] Markram H. The blue brain project. *Nat Rev Neurosci* 2006;7(2):153–60.
- [11] Hines ML, Carnevale NT. The NEURON simulation environment. *Neural Comput* 1997;9(6):1179–209.
- [12] Deng R, Liu X. Optical multiplexing: tunable lifetime nanocrystals. *Nat Photonics* 2014;8:10–2.
- [13] Zijlstra P, Chon JWM, Gu M. Five-dimensional optical recording mediated by surface plasmons in gold nanorods. *Nature* 2009;459(7245):410–3.
- [14] Li X, Ren H, Chen X, Liu J, Li Q, Li C, et al. Athermally photoreduced graphene oxides for three-dimensional holographic images. *Nat Commun* 2015;6:6984.
- [15] Ren H, Li X, Zhang Q, Gu M. On-chip noninterference angular momentum multiplexing of broadband light. *Science* 2016;352(6287):805–9.
- [16] Deng R, Qin F, Chen R, Huang W, Hong M, Liu X. Temporal full-colour tuning through non-steady-state upconversion. *Nat Nanotechnol* 2015;10(3):237–42.
- [17] Yan T, Wu J, Zhou T, Xie H, Xu F, Fan J, et al. Fourier-space diffractive deep neural network. *Phys Rev Lett* 2019;123(2):023901.
- [18] Lin X, Rivenson Y, Yardimci NT, Veli M, Luo Y, Jarrahi M, et al. All-optical machine learning using diffractive deep neural networks. *Science* 2018; 361(6406):1004–8.
- [19] Mengü 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* 2019;26(1):3700114.
- [20] Li J, Mengü D, Luo Y, Rivenson Y, Ozcan A. Class-specific differential detection in diffractive optical neural networks improves inference accuracy. *Adv Photonics* 2019;1(4):46001.
- [21] Khoram E, Chen A, Liu D, Ying L, Wang Q, Yuan M, et al. Nanophotonic media for artificial neural inference. *Photonics Res* 2019;7(8):823–7.
- [22] Yao K, Unni R, Zheng Y. Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale. *Nanophotonics* 2019;8(3):339–66.
- [23] Goi E, Gu M. Laser printing of a nano-imager to perform full optical machine learning [presentation]. In: Conference on Lasers and Electro-Optics/Europe; 2019 Jun 23–27; Munich, Germany; 2019.
- [24] Gu M, Fang X, Ren H, Goi E. Optically digitalized holography: a perspective for all-optical machine learning. *Engineering* 2019;5(3):363–5.
- [25] Yamazato T, Takai I, Okada H, Fujii T, Yendo T, Arai S, et al. Image-sensor-

- based visible light communication for automotive applications. *IEEE Commun Mag* 2014;52(7):88–97.
- [26] Khan LU. Visible light communication: applications, architecture, standardization and research challenges. *Digit Commun Netw* 2017;3(2):78–88.
- [27] Uysal M, Ghassemlooy Z, Bekkali A, Kadri A, Menouar H. Visible light communication for vehicular networking: performance study of a V2V system using a measured headlamp beam pattern model. *IEEE Veh Technol Mag* 2015; 10(4):45–53.
- [28] Arsalan M, Hong HG, Naqvi RA, Lee MB, Kim MC, Kim DS, et al. Deep learning-based iris segmentation for iris recognition in visible light environment. *Symmetry* 2017;9(11):263.
- [29] Komine T, Nakagawa M. Fundamental analysis for visible-light communication system using LED lights. *IEEE Trans Consum Electron* 2004;50(1):100–7.
- [30] Benítez JM, Castro JL, Requena I. Are artificial neural networks black boxes? *IEEE Trans Neural Netw* 1997;8(5):1156–64.
- [31] Zhang Q, Yu H, Barbiero M, Wang B, Gu M. Artificial neural networks enabled by nanophotonics. *Light Sci Appl* 2019;8:42.
- [32] Rivenson Y, Göröcs Z, Günaydin H, Zhang Y, Wang H, Ozcan A. Deep learning microscopy. *Optica* 2017;4(11):1437–43.
- [33] Li Y, Xue Y, Tian L. Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media. *Optica* 2018;5(10):1181–90.
- [34] Lyu M, Wang W, Wang H, Wang H, Li G, Chen N, et al. Deep-learning-based ghost imaging. *Sci Rep* 2017;7:17865.
- [35] Nehme E, Weiss LE, Michaeli T, Shechtman Y. Deep-STORM: super-resolution single-molecule microscopy by deep learning. *Optica* 2018;5(4):458–64.
- [36] Goodman JW. *Introduction to Fourier optics*. Colorado: Roberts & Company Publishers, 2005.
- [37] Barlow HB. Unsupervised learning. *Neural Comput* 1989;1(3):295–311.
- [38] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: *Proceedings of the 27th International Conference on Neural Information Processing Systems*; December 8–13 2014; Montreal, Quebec, Canada 2014. p. 2672–2680.
- [39] Bueno J, Maktoobi S, Froehly L, Fischer I, Jacquot M, Larger L, et al. Reinforcement learning in a large-scale photonic recurrent neural network. *Optica* 2018;5(6):756–60.
- [40] Naruse M, Terashima Y, Uchida A, Kim SJ. Ultrafast photonic reinforcement learning based on laser chaos. *Sci Rep* 2017;7:8772.