

深度学习算法及其在光学的应用

周宏强, 黄玲玲, 王涌天

(北京理工大学 光电学院 光电成像技术与系统教育部重点实验室, 北京 100081)

摘要: 深度学习作为机器学习的重要分支, 自出现之初就掀起了机器学习的又一次高潮。深度学习在诸如图像识别与分类、语义分割、智能驾驶等多个领域有着优异的表现。同时, 深度学习算法以其抽象特征识别和提取特性, 极强的模型构建和泛化推广能力, 被广泛应用于光学领域, 如计算全息图产生与成像、数字全息的无参数重建和光谱共振曲线预测等方面。详细介绍了深度学习的基本原理及在图像分类、超分辨成像、计算全息和数字全息、表面等离激元共振曲线预测、超表面的结构设计等方面的典型应用研究, 并探讨了深度学习在物理光学领域未来值得研究的方向。

关键词: 深度学习; 神经网络; 光学; 成像技术

中图分类号: TN202 **文献标志码:** A **DOI:** 10.3788/IRLA201948.1226004

Deep learning algorithm and its application in optics

Zhou Hongqiang, Huang Lingling, Wang Yongtian

(Key Laboratory of Photoelectronic Imaging Technology and System, Ministry of Education, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China)

Abstract: As an important branch of machine learning, deep learning has reached another climax of machine learning since its inception. Deep learning has excellent performance in many fields such as image recognition and classification, semantic segmentation, and intelligent driving and so on. At the same time, deep learning algorithms are widely used in the field of optics such as computational hologram generation and imaging, non-parameter reconstruction of digital holography, and spectral resonance curves prediction due to their abstract feature recognition and extraction characteristics, strong model building and generalization capabilities. This article detailed the basic principles of deep learning and its typical application research in image classification, super-resolution imaging, computer generated hologram and digital holography, prediction of surface plasmonics resonance curves, and structural design of metasurfaces. And future development of deep learning in the physical optical field was worth exploring.

Key words: deep learning; neural network; optics; imaging technology

收稿日期: 2019-10-11; 修订日期: 2019-11-21

基金项目: 北京市卓越青年科学家项目(BJJWZYJH01201910007022); 国家自然科学基金面上项目(61775019);

北京市科技新星(Z171100001117047); 北京市面上项目(4172057); 教育部霍英东高校教师基金(161009)

作者简介: 周宏强(1992-), 男, 博士生, 主要从事深度学习及微纳光学方面的研究。Email: hzhou@bit.edu.cn

导师简介: 王涌天(1957-), 男, 教授, 博士生导师, 博士, 主要从事光学技术和虚拟、增强现实等方面的研究。Email: wyt@bit.edu.cn

通讯作者: 黄玲玲(1986-), 女, 教授, 博士生导师, 博士, 主要从事微纳光学、计算全息、表面等离激元及深度学习等方面的研究。

Email: huanglingling@bit.edu.cn

0 引言

机器学习是一门基于经验学习的计算机科学,它通过对输入数据的特征提取和识别来发现其内在的经验规律。目前各种机器算法结构模型在金融股市、医学图像、工业生产、科学研究等领域有着广泛的应用。最初的人工神经网络就是模拟生物的神经网络模型构建的计算结构^[1-2],具有较强的模型学习能力和鲁棒性。随后由于其自身的限制,多层感知机逐步取代人工神经网络^[3-4]。感知机模型被扩展到具有多层神经网络模型^[5],结构复杂程度进一步提高,能学习更加深层的特征,因而得到进一步的推广。由于早期的神经网络无法解决非线性问题,并且存在训练数据量大,时间长,容易过拟合等缺陷,这使得神经网络模型陷入发展的瓶颈期。

1995 年支持向量机 (Support Vector Machine, SVM) 为解决非线性分类问题被加以改进并且得到广泛关注^[6-7]。该算法基于核方法,其核心思想是将原始的数据通过某种非线性映射到多维空间中,然后再利用线性分类器进行处理,具有模型简单,解释性强的特点,更重要的是其回归分类性能比最早提出的神经网络更优。特别是 SVM 在小样本的处理上有着很好的表现。SVM 是针对二分类设计,对于多分类任务要进行专门的推广。因此在多核分类的模型数据中效率会降低。自 2008 年之后,随着大数据时代的到来,SVM 显然已经不能满足大规模数据处理和更多分类的要求。同时,神经网络处理非线性问题存在的缺陷,如非线性函数的梯度消失或梯度爆炸等问题被逐渐克服解决,以及计算机性能和显卡运算(Graphics Processing Unit, GPU)等能力的提升,使得深层的神经网络重新焕发生机。例如神经网络的规模可达到 1 207 层以上,极大地提高了大规模数据的特征提取和识别能力。这些神经网络让计算机从经验中学习,并且通过层次化的概念来理解世界。这样计算机便能够构建简单的知识来学习复杂的知识,当层次增多之后便得到一个深层的图谱,笔者称之为深度学习^[8]。目前,深度学习的主要基础模型有卷积神经网络(Convolutional Neural Network, CNN)^[9-10],循环神经网络(Recurrent Neural Network, RNN)^[11-12],生成对抗式网络

(Generative Adversarial Network, GAN)^[13],受限玻尔兹曼机 (Restricted Boltzmann Machine, RBM)^[14],深度置信网络 (Deep Belief Network, DBN)^[15]等。

为方便和加快依托和借助深度学习优势的产业和各个领域的推广应用,诸如 Google、Facebook、Microsoft 等公司都纷纷研发并开源了深度学习的框架。Google 的 TensorFlow 框架、Facebook 的 Caffe 框架、百度的飞桨 PaddlePaddle 等都是目前主流的深度学习框架。TensorFlow 编程接口都是基于图形界面,可在 Python 平台方便地运行,算法结构简单清晰,再加上 Python 可以调用 GPU 进行并行运算,结合 Nvidia 公司提供的深度学习的高性能库单元—cuDNN 等大大提高了深度学习训练速度和模型性能。还有一种是基于常用的机器学习算法的 Scikit-Learn^[16]。这些学习框架基本上能满足各类深度学习开发者的需求,提高工作效率。而在物理科学研究方面,深度学习有着广泛的应用。如计算全息成像,超分辨率成像,表面等离激元共振图谱,逆向设计等都有极好的表现。文中首先介绍了在光学成像分类和微纳光学中的逆向设计中常用的卷积神经网络和对抗神经网络算法的原理;接着从传统的物理模型与神经网络结合的案例详细介绍了深度学习在物理光学中的典型应用研究;最后介绍了深度学习在物理光学科学研究的应用前景。

1 深度学习算法基本原理

1.1 卷积神经网络

卷积神经网络(CNN)是一种基于卷积运算的神经网络模型^[9-10, 17]。卷积网络在车牌文字识别、人脸识别、自动驾驶等多个领域都有优异的表现。经典的 LetNet-5 是由卷积网络之父的 Yann LeCun 于 1995 年提出的,奠定了卷积神经网络的基础^[18]。卷积神经网络最初是用于手写字符的识别与分类,其分类的精度能到达 98% 以上,在当时机器学习领域轰动一时。其实早在 20 世纪 90 年代,CNN 网络模型被用于简单的文字识别,语音识别,人脸识别等场景的应用^[19-24]。到 21 世纪之后,随着计算能力的提升,10~20 层的卷积神经网络就能达到上百万的神经元相连接,而训练时间缩短到几个小时。

1.1.1 卷积运算

如图 1 所示,经典的 LetNet-5 神经网络主要由输入层 (Input), 亚采样层 (Subsampling), 卷积层 (Convolutions) 和全连接层 (Full connection) 及输出层 (Output) 构成。卷积是 CNN 计算的核心所在。卷

积神经网络的特征提取是基于卷积核对图像的卷积运算操作完成的,提取的图像特征也被储存在核函数集中。在机器学习的应用中,输入的数据通常是高维度的,而核函数也是高维度的数据组。一般把这种高维数据组称为张量。这样便可以对核函数进行拆

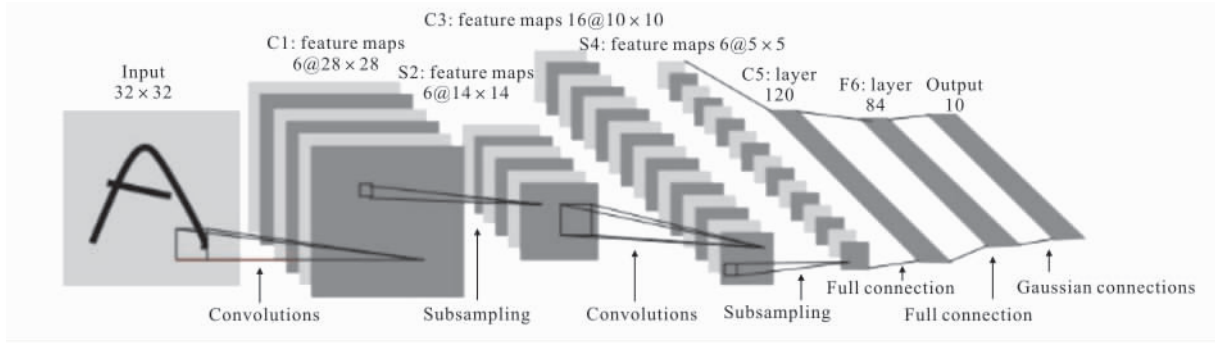


图 1 LetNet-5 神经网络结构^[18]
Fig.1 LetNet-5 neural network structure^[18]

分处理,提高运算效率。

1.1.2 卷积神经网络特征

卷积神经网络有三个主要特征:稀疏交互、参数共享、线性平移^[25]。传统的神经网络主要是通过多个矩阵连续相乘来建立输入和输出的关系。结构特征通过矩阵存储的数据来表示。

稀疏交互式,也叫稀疏连接,是指网络模型中前后相连接的矩阵大小不同,且矩阵将会变得越来越小,从而大大减小了运算量。一般而言,每一个卷积层之后都会有一个池化层。池化层的作用就是对矩阵降维运算,进一步减轻数据冗余,减小计算量,提高计算速度。

参数共享是指在 CNN 模型运算过程中,多个函数使用相同的参数,如权值、偏置等。卷积运算中的参数共享可以使得每一层的卷积运算都针对一个参数集合进行学习,从而能够显著提高 CNN 模型的计算效率,节省内存空间。

平移特性就是利用卷积和相关的平移特性就使得图像在平移、缩放、旋转等过程中不影响模型的性能,从而符合实际的应用需求。

1.1.3 卷积神经网络的基本结构

现在沿用的 CNN 一般包含五层结构:输入层、卷积层、激活层、池化层、全连接层。根据不同任务目标的需要,CNN 的结构会进行优化和改进,其中卷

积层、激活层和池化层都可以循环叠加,但总体基本沿用这个结构^[26-30]。

激活层中常用的激活函数有:

(1) Sigmoid 函数^[31]

$$f(z) = \frac{1}{1 + e^{-z}} \quad (1)$$

如图 2(a)所示,Sigmoid 函数能将输入的任意实值函数转化为 0~1 之间的输出,以便于统计概率的表示。缺点是在神经网络反向梯度计算的过程中极易出现梯度消失和梯度爆炸等问题。另外,Sigmoid 函数的输出不是零对称的正态分布。这样就使得在梯度计算后期收敛速度相当慢,影响训练效率。另一个缺点就是其解析函数含有幂函数,不利于计算机的梯度求导运算。

(2) Tanh 函数^[32]

$$\text{Tanh}(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (2)$$

它解决了 Sigmoid 函数的非零值正态分布问题,但是仍然无法梯度消失和幂函数的问题。

(3) ReLu 函数^[33]:

$$\text{ReLu}(z) = \max(0, z) \quad (3)$$

如图 2(c)所示,ReLu 函数在向右方向已经解决了梯度消失的问题。实际上,从表达式可以看出,ReLu 函数事实上是一个取最大值函数,输出值是非

负值。其解析式很简单,解决了梯度消失的问题,并且具有计算速度快和模型收敛速度快等优点。

降低运算效率等问题。

1.1.4 损失函数

损失函数对 CNN 网络的训练和优化收敛有着至关重要的影响。常用的损失函数如下:

(1) 均方差损失函数(Mean Square Error, MSE)

均方差是比较常用的损失函数,用以评价测试数据与目标数据的差异。在神经网络中的损失函数的定义如下:

$$L(w, b) = \frac{1}{2N} \sum_1^N ||y-a||^2 \quad (4)$$

式中: N 为神经元数量; y 为目标数据; a 为测试数据。测试数据一般来自于网络的预测输出 $a=f(w \cdot x+b)$ 。 x 为训练数据, w 为权值, b 为偏置, f 为激活函数。MSE 在线性回归中表现较好,可以有效地计算反向的梯度传播。当激活函数为 Sigmoid 函数时,容易造成梯度的损失,从而导致层数较浅的权值没有被更新,即存在梯度消失的问题。因此在逻辑回归中选择 MSE 需要考虑梯度的损失情况。

(2) 交叉熵函数(Cross-Entropy)

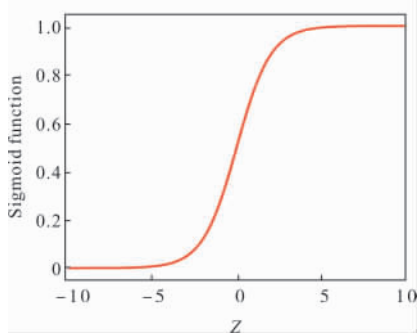
交叉熵函数作为损失函数是基于 Softmax 计算的,Softmax 将网络的输出转换成概率的形式作为预测输出。Softmax 的函数表示为: $p_i = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}}$ 。其中,

e^{z_i} 是某一计算的分类输出的数值,分母则是所有计算分类的总和。这样就得到了类别的输出概率 p_i 。那么交叉熵的损失函数定义为:

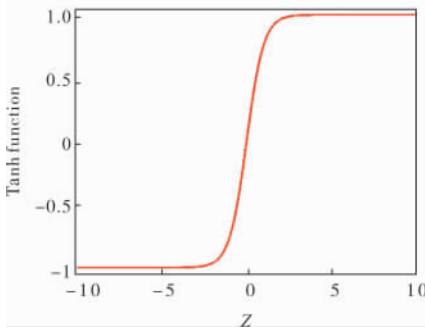
$$J(o, y) = -\frac{1}{N} \sum_1^N \sum_{j=1}^k (o_j \cdot \log y_j + (1-o_j) \log(1-y_j)) \quad (5)$$

式中: o_i 为目标值; y_i 为预测值; N 为数据总量; k 为类别数。在计算中,仍然以 Sigmoid 函数作为激活函数,将公式(1)代入公式(5),计算最后的反向传播梯度可知隐层的梯度仍然存在 Sigmoid 函数,这就解决了输出层神经元学习率缓慢的问题。

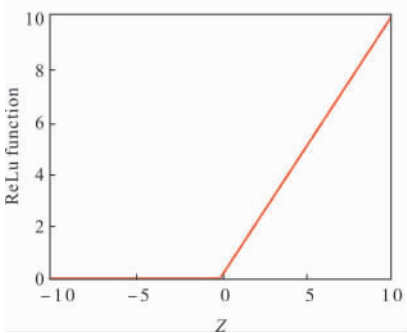
上面也提到有多种激活函数可供选择,其中 ReLu 函数形式相对简单,在梯度计算过程中的内存消耗较少,并且 ReLu 函数将一部分神经元的输出变成 0,使得网络变得稀疏了,在一定程度上减缓了过拟合问题。在实际的训练过程中,均方差的误差函数一般与 Sigmoid 函数结合适用于线性回归问题,



(a) Sigmoid 函数
(a) Sigmoid function



(b) Tanh 函数
(b) Tanh function



(c) ReLu 函数
(c) ReLu function

图 2 三种主要激活函数

Fig.2 Three main activation functions

这些激活函数在大多数开源的深度学习框架中都被集成,调用运算极其方便。全连接层的作用是实现分类,对手写数字识别、动物分类、图像空间语义分割等任务至关重要。全连接层一般为二维长矩阵,与激活函数组合能够增加模型的非线性表达分类能力,但是存在容易造成模型过拟合,增加计算负担,

对于真实结果的差别越大,则区别就会越大,回归精度越高。MSE 在分类问题中则容易出现梯度消失的情况,而交叉熵函数可以解决分类模型中的梯度消失的问题。交叉熵函数只关心分类的结果是否正确,而 MSE 函数则注重每个类别的大小,这在实际的分类问题中是没有必要的。因此交叉熵函数更适合逻辑回归问题。

1.1.5 优化算法

(1) 随机梯度下降(SGD)

这种优化算法的优点在于可以分散训练的数据量,减少计算机的负载,提高计算效率。特别是当训练数据出现重复时,SGD 在训练中也不会因为分块训练模式而降低效率。学习率一般通过经验和误差来选取。过大的学习率则会造成目标曲线的剧烈震荡,造成神经网络无法正常的更新参数,使得训练模型不能够正常收敛。而过小的学习率则容易使系统局限于局部最小值无法跳出。而这个最小值往往使得系统的损失值较大,并不能完成对神经网络的优化。

(2) 动量(Momentum)

根据牛顿运动定律,动量表示物质的质量与速度的乘积。假设质量为单位 1,那么速度在数值上等于物质运动的动量。引入动量到 SGD 中可以一定程度地解决其随机更新的不稳定性,因为它在更新时在一定程度上保留之前梯度的方向,同时利用当前分块来微调更新方向。

(3) 自适应矩估计(Adam)

这种优化方法主要是对学习率的调整,因为 SGD 中的学习率是固定或者依据训练者经验调节的,训练结果千差万别。Adam 方法借鉴了自适应梯度(Adagrad)、Adadelta 和 RMSprop 等算法对学习率的自更新方式。

这种优化方法的好处是可以动态地调控学习率,使其朝着稳定的方向变化,在优化网络参数的过程中能够根据输入数据高效地寻找到全局最优解。而在实际的训练和应用中,没有一种函数或者算法可以完美地解决所有问题,因而,根据实际的应用需求在了解算法原理的基础上选择合适的函数和参数是重要的。

卷积神经网络在深度学习的发展中发挥了重要

作用。CNN 具有极好的模型泛化能力,其收敛速度快,模型逼近能力强。然而,CNN 是一种监督学习算法,训练数据要求有标签,也就是参考标准。事实上,大量的训练数据产生都不可能自动附带标签,这就需要人为进行标记,需要耗费大量的人力和时间成本。接下来介绍一种高效的半监督的神经网络。

1.2 生成对抗式网络

1.2.1 生成对抗网络基本结构

生成对抗式网络(GAN)是基于博弈对抗思想的一种生成式模型^[13, 34-36]。它有两个子网络,即生成模型和判别模型。如图 3 是 GAN 的博弈过程示意图。

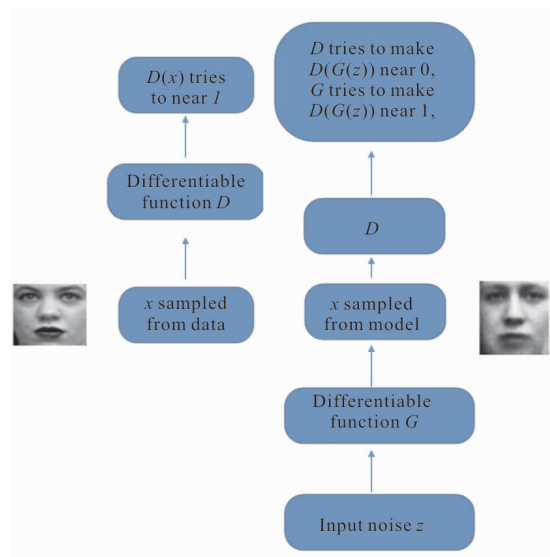


图 3 GAN 博弈示意图^[13]

Fig.3 Schematic of GAN game^[13]

生成模型首先对真实的数据进行建模,随后在网络中不断产生全新的数据,而判别模型则根据部分已知条件来识别生成模型的真假,即是否满足模型的要求。经过不断的判断结果反馈,生成模型能够生成精度越来越高的数据,从而判别模型也无法识别生产数据的真假。这实际上是生成模型与判别模型的“极大极小博弈”。从 GAN 模型的判别目标函数能直观分析此博弈过程^[13],如图 4 所示。

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}} [\log D(x) + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (6)$$

GAN 中的生成模型 G 和判别模型 D 均采用了多层感知机模型。式中 p_z 是噪声先验条件,它的作用

主要是学习生成模型在训练数据集上的概率分布。 $G(z)$ 表示生成模型,产生新的数据。 $D(x)$ 则是指数据 x 来自真实训练数据的分布概率 p_{data} 。该式左侧表示更新 D 时需要最大化公式右侧的表达,更新 G 时需要最小化右侧的表达。即在对判别模型 D 进行更新的时候,对于来自真实数据 p_{data} 而言,模型的趋势是 $\log D(x)$ 越大越好,即 $D(x)$ 趋向于 1。生成模型产生的数据 $G(z)$ 则希望 $D(G(z))$ 越来越接近于 0,即 $\log D(1-G(z))$ 越大越好,从而 D 模型整体趋向于越来越大, D 模型的判别能力就越强。在对生产模型 G 进行更新时,模型趋向于 G 产生的数据与真实数据相同。 $D(G(z))$ 趋向于 1,那么 $\log(1-D(G(z)))$ 趋向于越来越小。此外值得注意的是 GAN 中采用的损失函数和优化算法多与 CNN 中的方法近似。因此在训练网络和参数优化的过程可以参考上一小结。

Algorithm 1 Minibatch stochastic gradient descent training of generative adversarial nets. The number of steps to apply to the discriminator, k , is a hyperparameter. We used $k=1$, the least expensive option, in our experiments.

```

For number of training iterations do
  for  $k$  steps do
    · Sample minibatch of  $m$  noise samples  $\{z^{(1)}, \dots, z^{(m)}\}$  from noise prior  $p_g(z)$ .
    · Sample minibatch of  $m$  examples  $\{x^{(1)}, \dots, x^{(m)}\}$  from data generating distribution  $p_{data}(x)$ .
    · Update the discriminator by ascending its stochastic gradient:
      
$$\Delta_{\theta_d} \frac{1}{m} \sum_{i=1}^m [\log D(x^{(i)}) + \log(1 - D(G(z^{(i)})))]$$

    end for
    · Sample minibatch of  $m$  noise samples  $\{z^{(1)}, \dots, z^{(m)}\}$  from noise prior  $p_g(z)$ .
    · Update the generator by descending its stochastic gradient:
      
$$\Delta_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log(1 - D(G(z^{(i)})))$$

    end for
    The gradient-based updates can use any standard gradient-based learning rule. We used momentum in our experiments.
  
```

图 4 GAN 算法流程图^[13]

Fig.4 GAN algorithm flow chart^[13]

1.2.2 生成对抗网络的应用

GAN 的提出引起了许多开发者和学者关注。经过诸多研究人员的发展,GAN 已经有了各方面的拓展,如三维物体的产生与重建^[37-38]、文本语义分析^[39]、图-图翻译转换^[40]、图像超分辨率^[41]、视频动态分析^[42]、人脸处理^[43-48]等。特别地,在参考文献^[35]中提出的一种深度卷积生成对抗式网络(Deep

Convolutional Generative Adversarial Networks, DCGAN) 是将卷积神经网络与 GAN 相结合的网络模型。其模型结构如图 5 所示。

该模型的基本结构是用多层反卷积层,此反卷积只能还原出原始信号的大小,并不能得到真值,因此不是真正的反卷积。DCGAN 的另一个特点是对池化层的操作。在反卷积操作的过程中,使得池化核函数的步长增大,从而使得输出矩阵增大。

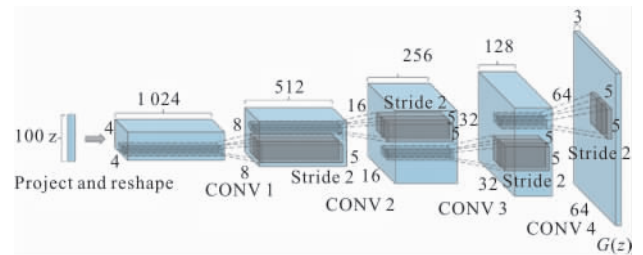


图 5 DCGAN 模型结构图^[35]

Fig.5 Schematic of DCGAN model^[35]

如图 6 所示,借助 DCGAN 模型能够基本实现图像的匹配过滤等功能。如将戴眼镜的男人与不戴眼镜的男人相减,提取眼镜信息;然后与不戴眼镜的女人相叠加,输出女人像戴着类似眼镜。但是该文章

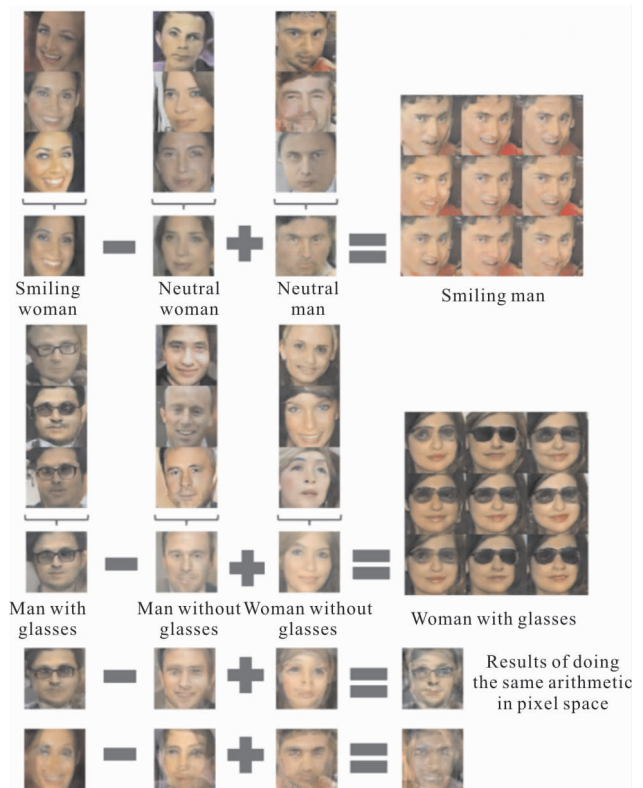


图 6 GAN 人脸识别与编辑输出结果^[35]

Fig.6 Result of GAN human face recognition and edition^[35]

指出模型仍然存在人脸无法准确对准和缩放, 输出像模糊等问题^[35]。

GAN 能够根据先验知识自动生成模型数据的巨大优势将为深度学习带来更多可能。由于其训练过程中不需要进行复杂的近似, 训练数据来源于生成模型的随机采样, 大大提高了模型训练的效率。GAN 模型也并不是完美的。GAN 采用对抗学习的原则, 在模型训练过程中生成网络和判别网络同时进行训练, 而在实际的过程中, 很难把握其训练的同步性, 模型稳定性差。另外, GAN 模型的可解释性差, 容易出现模型坍塌的情况。在 GAN 随后的发展过程中, 模型得到不断的优化改进, WGAN 的出现基本解决了上述问题^[49]。但是关于 GAN 的收敛性和博弈均衡仍是值得探讨的问题。没有任何一种网络模型能够满足设计者的所有要求。不可否认的是, GAN 模型的思想是先进的, 将 GAN 与其他模型思想相融合在实际应用中解决实际问题才是模型存在的意义。

2 深度学习算法在光学中的应用

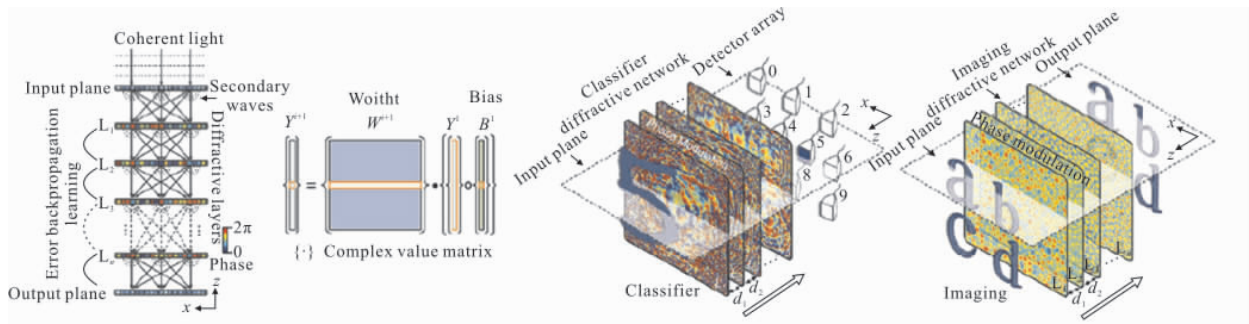
近些年来, 机器学习, 特别是由于深度学习的发展, 使得其在光学成像探测和光计算中也获得广泛的应用^[50-54]。特别的, 通过借鉴深度学习的方法, 设定无参数衍射聚焦、提高分辨率等预期目标, 从而在光学迭代相位成像、超分辨率成像、鬼成像等方面得到有力的拓展^[55-58]。文中笔者将重点介绍基于深度学习在光学计算中的应用, 如全光学网络中的图像分类和成像, 生物荧光组织的超分辨率成像, 数字全息和计算全息的再现, 微纳结构的光谱曲线预测与

逆向结构设计。

2.1 深度学习实现图像分类和成像

图像分类是机器学习的基本任务。将算法应用到光路中, 利用光的衍射传播和机器学习算法计算分类图像是一个新颖的课题^[59]。

图 7^[59]所示为利用深度学习算法训练神经网络, 将训练好的神经网络每层作为相位板, 用 3D 打印技术实体化, 从而可利用“光计算”实现图像分类和单透镜成像功能。图 7 中利用手写数字数据集 MNIST 和 Fashion-MNIST 数据集训练神经网络分类模型, 并在测试模型中得到极好的表现效果^[60-61], 实验能够达到 90% 的预测精度。图 7(c)、(e) 通过设计单透镜模型, 物体通过模型能够在特定的距离处成清晰的图像, 图 7(e) 中给出了光在传播方向上振幅和相位的分布。该工作集成了神经网络训练、设计、相位板加工等方法, 为相关领域采用神经网络并进行实体化应用提供依据和参考。该课题组同时还提出一种类依赖的微分探测来提高光学衍射神经网络的预测精度^[62]。这种方法在一定程度上缓解了分类探测器只能够探测强度的局限性。每组分类分成一正一负两个探测器。训练时, 将反馈的正负探测结果进行比对, 进一步提高预测精度。文章中盲测了常用的 MNIST, Fashion-MNIST, 和 grayscale CIFAR-10 数据集能够分别达到 98.52%, 91.48% 和 50.82% 的准确率。这非常接近利用全计算机分类预测精度 98.77%, 90.27%, 和 55.21%。将机器学习算法, 特别是深度学习算法用于光学网络实现图像分类和成像将有助于提高光学系统的灵活性。这在许多光学计算中是非常值得借鉴和推广的。



(a) 光学衍射神经网络示意图

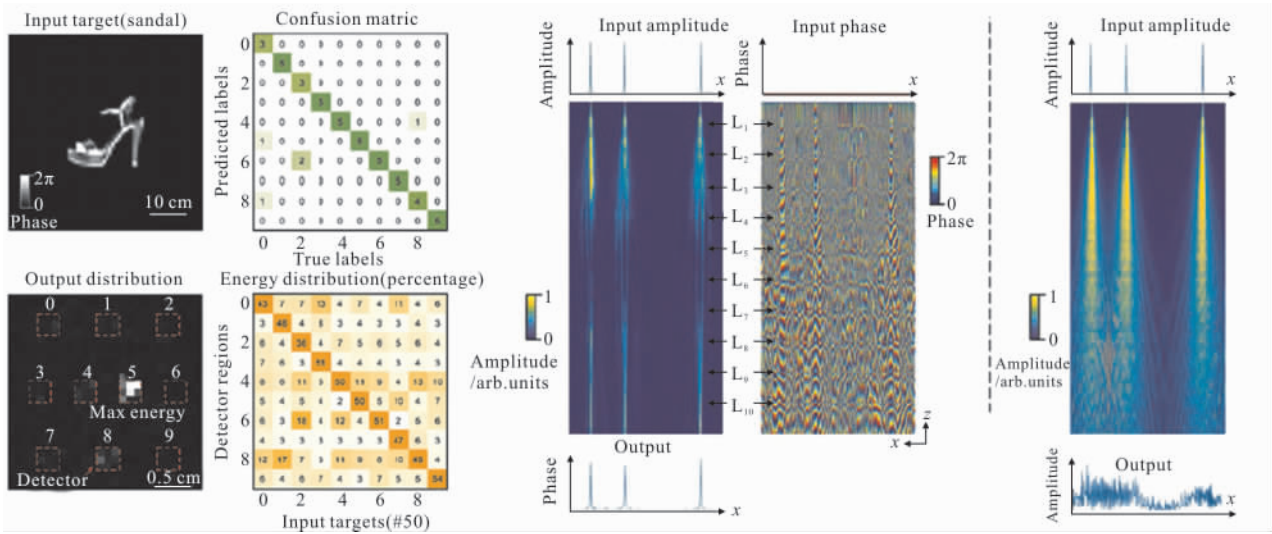
(a) Schematic diagram of all optical diffraction neural network

(b) 图像分类示意图

(b) Schematic diagram of image classification

(c) 神经网络透镜成像示意图

(c) Schematic diagram of neural network lens imaging



(d) 图像分类测试结果

(e) 神经网络透镜轴向衍射强度图

(d) Image classification test results

(e) Neural network lens axial diffraction intensity map

图 7 全光学神经网络结构及应用^[59]

Fig.7 All optical neural network structure and applications^[59]

2.2 超分辨率成像

超分辨成像一直以来都是光学成像中的热点话题。根据瑞利衍射的规则,普通光学系统的成像分辨率有一个上限,称为衍射极限。诸多学者都做了大量的工作力求突破衍射极限成像,获得更高的分辨率,如受激发射损耗 (Stimulated Emission Depletion Microscopy, STED), 随机光学重建显微(Stochastic Optical Reconstruction Microscopy, STORM) 等实现了衍射极限的突破^[63-66]。但是这类光学系统相对复杂,操作难度大,设备昂贵等问题让许多研究组望而却步。深度学习的出现给超分辨率成像和全息重建提供了一条新的思路^[67]。

将深度学习与生物显微成像相融合,提高成像深度和分辨率是一个极好的方案^[68]。如图 8 所示,该网络模型基于卷积神经网络框架,由于卷积运算的效率高和平移不变等特性,能够对输入图中的细节和特征信息进行快速的学习。经过反复多样化训练的网络还能重建出高分辨率的图像。因此 CNN 在图像处理领域有着极好的适应性。

将 40×/0.95 NA 显微镜采集到的图像,如图 9 (b)、(g)、(l)作为输入进入网络模型,经过网络输出的图像的分辨率有了明显提升,如图 9(c)、(h)、(m)。将第一次网络输出图像作为输入再一次进入网络模型后,得到的输出图像能够比输入像的分辨率

高,如图 9(d)、(i)、(n)。可见,经过优化训练后的超分辨率深度学习网络能够有效地提高同类生物显微成像的分辨率。为进一步展现网络模型提高分辨率的能力,文中^[68]获取了 100×/1.4 NA 显微物镜下的图像作为数据训练并输出,如图 9(f)、(k)、(p)所示。输出像的分辨率的对比度有了明显的提升。文中^[68]还给出了神经网络的极限分辨率,在 100×/1.4 NA 的油浸物镜下采集的图像分辨率能达到分辨率板的第十组第六线对,而神经网络的输出分辨率则可达第十一组第四线对(线宽 0.345 μm)的超分辨率成像,如图 10 所示。相比而言,显微系统的衍射极限大约为 0.355 μm。因此神经网络的输出已经明显地突破了成像系统的衍射极限,实现了超分辨成像。

生物组织和细胞的全息成像由于组织透明,像差复杂等问题,一直以来其准确成像是大难题^[69-70]。文中^[70]采用深度学习算法对单一强度全息图相位复原和振幅快速重建,并且取得了极好的表现效果。该文章采用的基本框架是卷积神经网络。通过采集大量的显微实验数据进行训练优化网络,数据包括不同生物样品,不同距离,不同放大倍率等条件下的振幅和相位像。最终经过大量数据的训练过程,使得该网络具有了极强的泛化能力,能够在多种不同的输入全息图的情况下准确重建出无孪生的再现象。如图 11 是神经网络训练和重建的示意图。

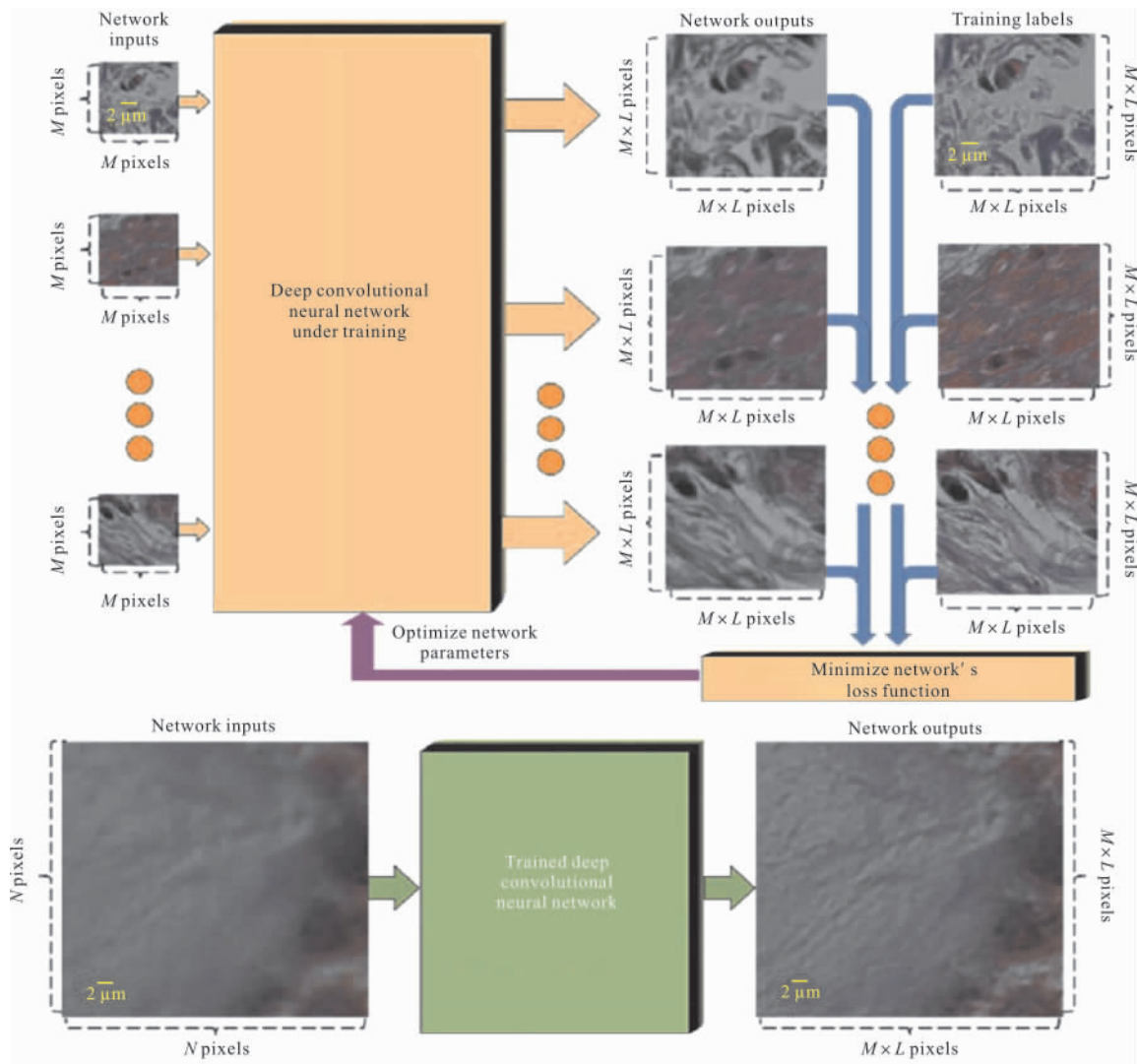


图 8 深度卷积神经网络训练流程及超分辨率成像^[68]

Fig.8 Deep convolutional neural network training process and super-resolution imaging^[68]

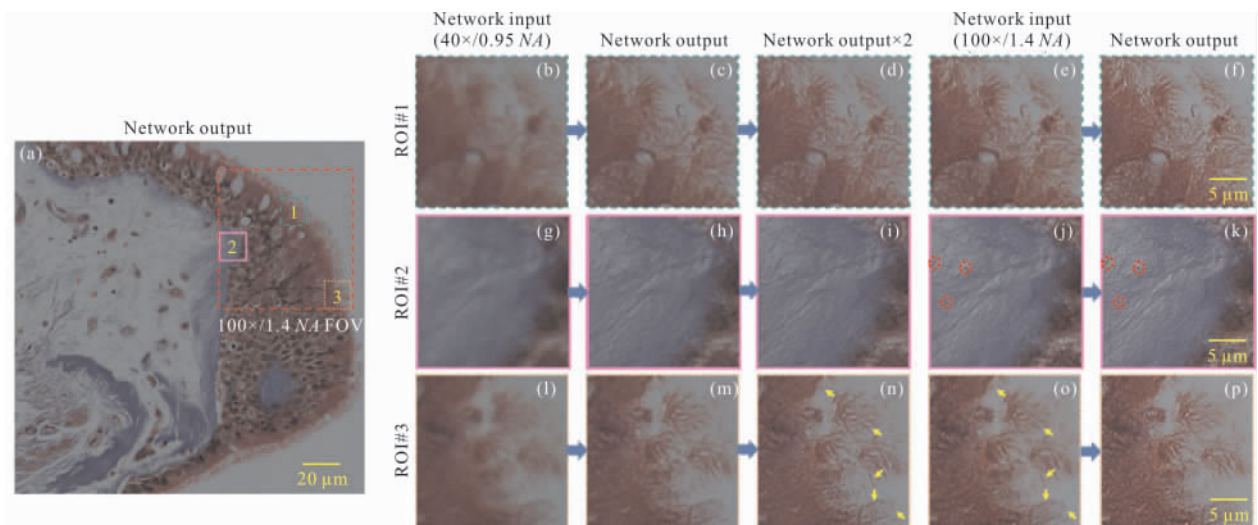


图 9 神经网络超分辨率成像结果^[68]

Fig.9 Neural network super-resolution imaging results^[68]

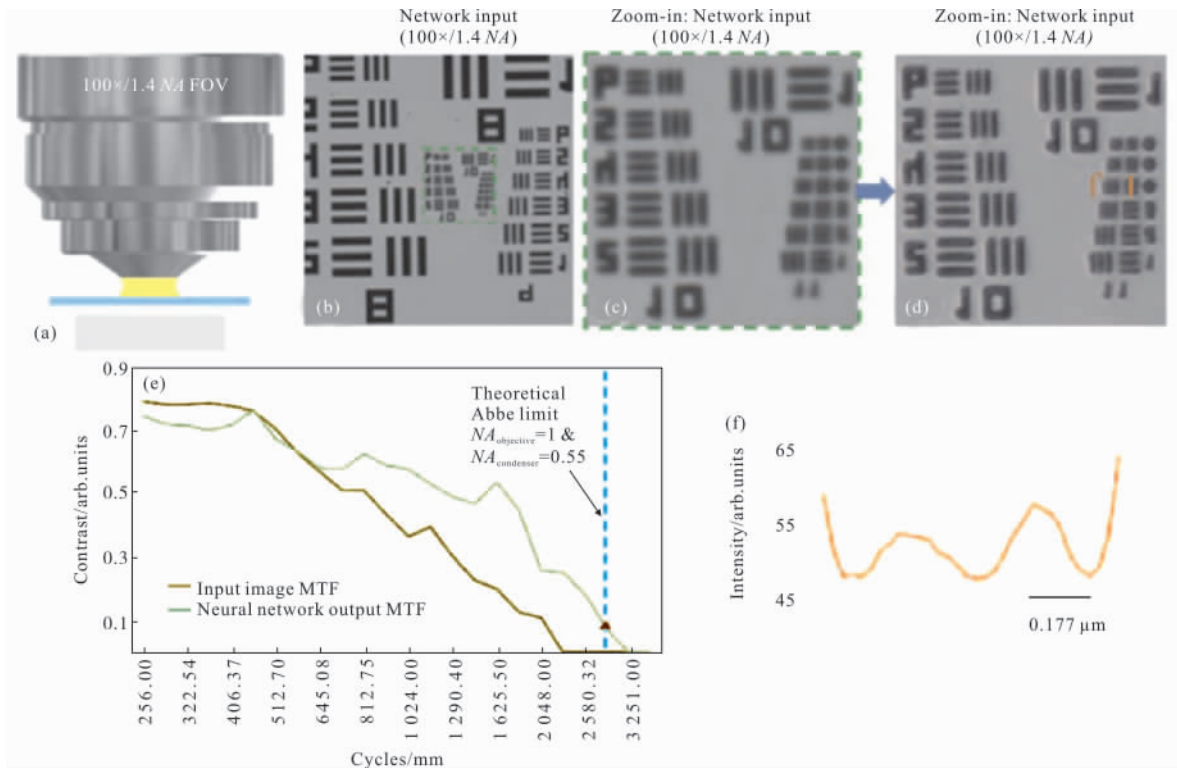


图 10 神经网络系统分辨率分析:(a) 100x/1.4 NA 显微镜; (b)~(d) 明场显微图像;
(e) 输入和输出图像的调制传递函数曲线;(f)第十一组第四线对刻线截面强度^[65]

Fig.10 Analysis of neural network resolution:(a) 100x/1.4 NA objective lens; (b)~(d) Bright field microscopy image;(e) Input and output image MTF curve;(f) Eleventh group of fourth line pairs of line cross-section intensity^[65]

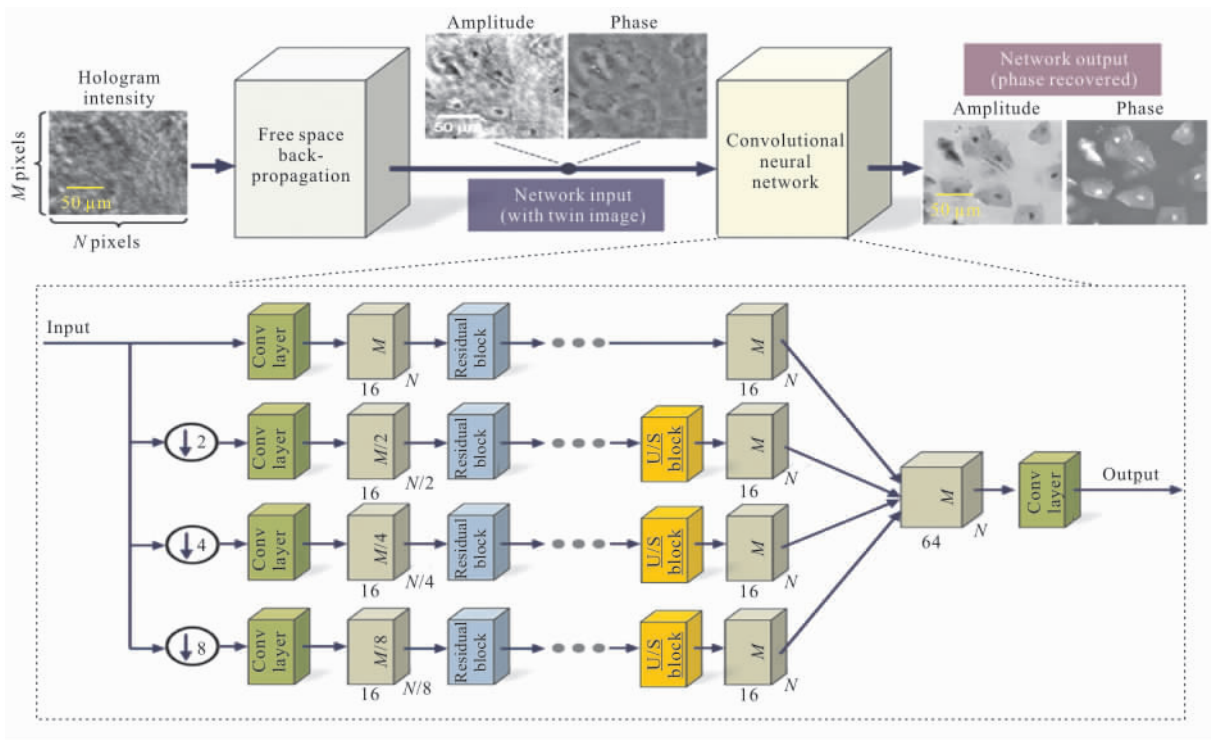


图 11 卷积神经网络自动实现全息图无孪生像重建^[70]

Fig.11 Convolutional neural network for hologram reconstruction without twins-image^[70]

根据不同的设计目标,还可以训练并优化系统,使其能够自动获取全息重建距离进行成像而无需人工干预。从图 12 中可以看出,由于记录过程是同轴全息的,原始单一强度全息图的重建会受到孪生像的干扰,而经过训练的卷积神经网络输出重建能有效抑制孪生像的干扰,获得清晰的重建结果。相比于多高度图像迭代重建的传统方法,该方法具有更高

的计算效率和更好的图像质量。

超表面全息是近几年新兴起的方向,它结合了超表面二维平面阵列排列方式灵活、按需编码等优点,拥有对光场相位、振幅、偏振、角动量、频率等进行调控的能力,以及全息技术的成像能够再现三维物体和复振幅特点,具有多自由度编码成像、主动亚波长调控等优势^[71-74]。结合深度学习设计所需要的

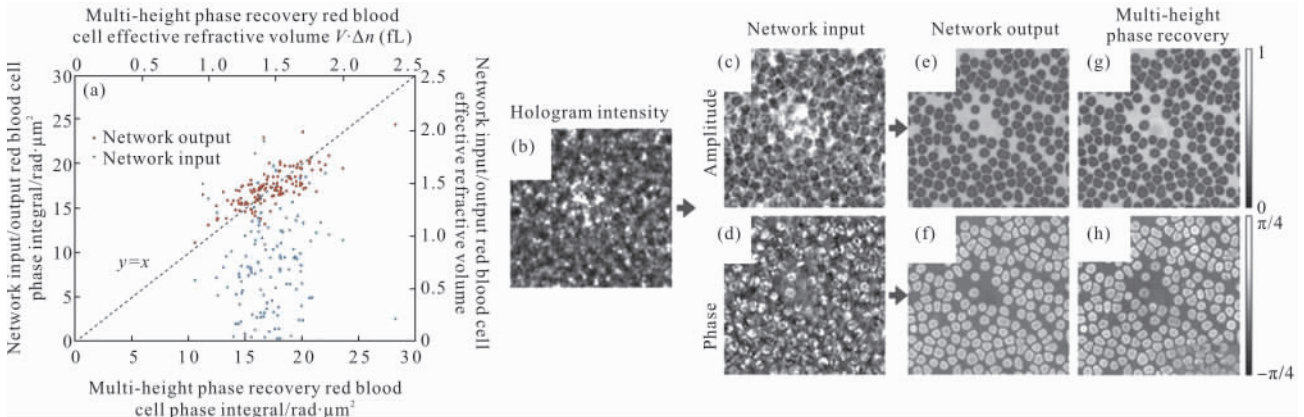


图 12 采用深度学习相位复原方法实现红细胞预测。(a)多高度图像相位复原重建与神经网络输出的红细胞折射率体积参数比;(b)单幅全息图;(c)、(d)从单幅全息图中得到振幅和相位图作为神经网络输入;(e)、(f)深度学习神经网络输出振幅和相位图;(g)、(h)采用多幅图像相位复原算法得到振幅和相位图^[70]

Fig.12 Using deep neural network phase retrieval method to achieve red blood cell estimation. (a) Ratio of phase retrieval and reconstruction of multi-height image and red blood cell index volume parameter output from deep neural network; (b) Single hologram; (c), (d) Amplitude and phase map obtained from single hologram as input of neural network; (e), (f) Amplitude and phase images output by the deep neural network from input (c), (d); (g), (h) Using multiple images phase retrieval algorithms to obtain the amplitude and phase images^[70]

样品特性和全息图分布,超表面全息便能更加有效地根据人为需求设计实现相应的器件和功能。深度学习辅助的超表面全息是个热点问题^[75]。超表面结构设计和选取往往需要大量的时间和计算能力。利用深度学习进行多种结构和等离子共振模式类型的训练,使神经网络能够根据简单的任务目标给出候选结构,这样就能大大地降低时间成本,并且准确率上也有很高的保证。这将对推动超表面全息的发展具有极其深刻的意义。

2.3 全息图的产生与再现

全息术是指能够记录三维物体振幅和相位全部信息的技术。随着计算机水平的日益提高,传统的光学全息与计算相结合成为必然趋势。传统的数字全息图大多是基于两束相干光束的干涉得到的,利用数码探测器件 (Charge Coupled Device, CCD/ Complementary Metal Oxide Semiconductor,

CMOS)探测到全息图,然后利用计算机模拟再现过程,得到重建像。这种全息称为数字全息(Digital Holography, DH)^[76-77]。

如图 13 所示,全息图的记录采用典型的马赫曾德离轴干涉的方法^[78]。再现过程则利用计算机模拟

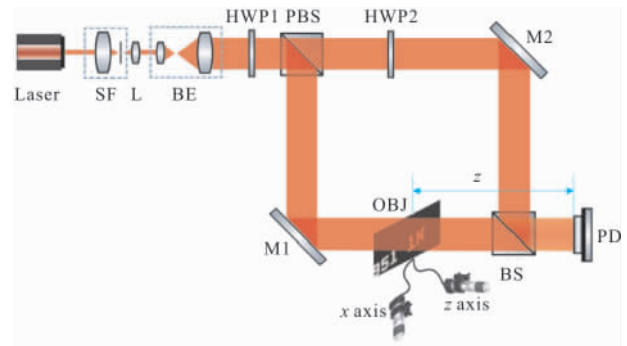


图 13 离轴数字全息记录光路图^[78]
Fig.13 Optical path of off-axis digital holography recording^[78]

菲涅尔传播得到重建像。重建距离与全息图的记录距离有关,由于某些不确定因素可能导致距离参数误差过大,将很难重建清晰的像,这就给图像的后处理带来了很大困难。结合深度学习的优点,计算

机可以通过先期的训练网络学习全息图的重建过程,如图 14 所示,进而可在全息图的采集过程中实时完成全息图的重建,无需模拟全息图的非涅尔域的传播距离等参数,从而大大减少数据后处理的时

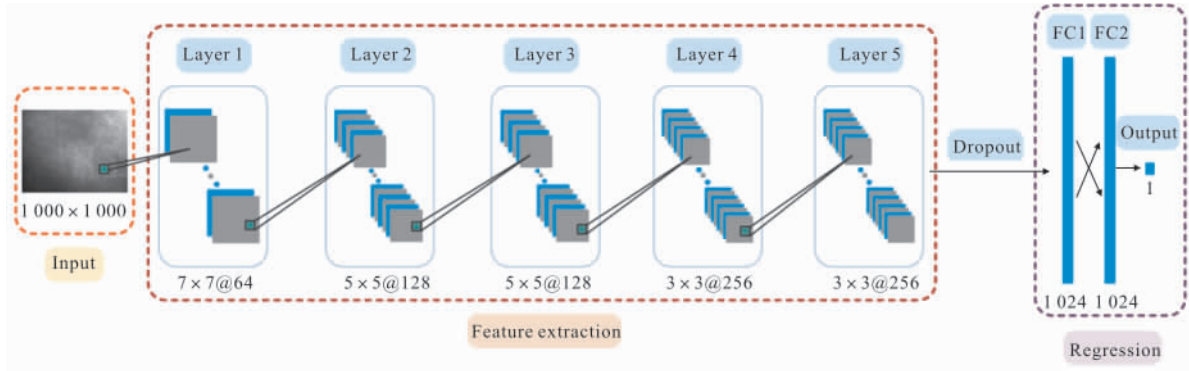


图 14 网络模型训练示意图^[78]

Fig.14 Schematic of the neural network training^[78]

间,提高实验效率^[78]。同时,经过适当训练后的网络模型能够重建经过平移、缩放、倾斜的振幅、相位和复振幅全息图。深度学习的使用有效地提高了数字全息实验效率,对于类似的科研工作很有借鉴意义。

另一种不同于数字全息的计算全息技术(Computer Generated Hologram, CGH)^[79],即通过一定的算法利用计算机产生全息图。计算全息的再现过程是将生成的全息图置于光路中,利用光的衍射实现重建。深度学习通过学习全息图生成的算法规则,能够实时、快速产生基于 Gerchberg-Saxton (GS)算法的全息图^[80]。借助于空间光调制器(Spatial Light Modulator, SLM)的实时刷新、动态调制特点,可以将生成的全息图加载到 SLM 上,便可观测到动态的全息图重建像。

如图 15 所示,结合深度卷积残差网络(Deep Convolutional Residual Network, ResNet),去学习计算全息图的重建过程规则^[80]。经过大量的训练,使得该网络模型能够具有重建任意二维灰度 GS 全息图的功能。测试的重建图像与传统计算全息算法的重建图做了比较,神经网络的输出具备更清晰的成像效果。将该方法推广到不同类型全息图的产生和重建中是非常有必要的。注意某个神经网络模型不能够仅仅简单地局限于较窄的问题,而应该泛化(或容易迁移学习)到解决同一类问题,才更具有价值。

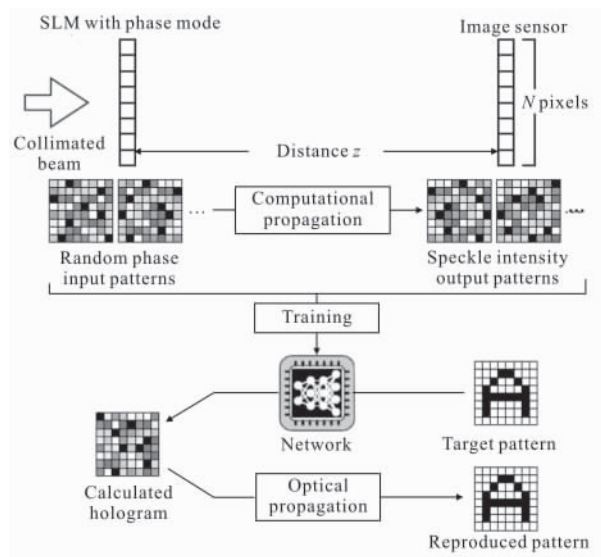


图 15 神经网络生成计算全息图^[80]

Fig.15 Schematic of the neural network computer generated hologram^[80]

2.4 微纳结构设计与其共振曲线预测

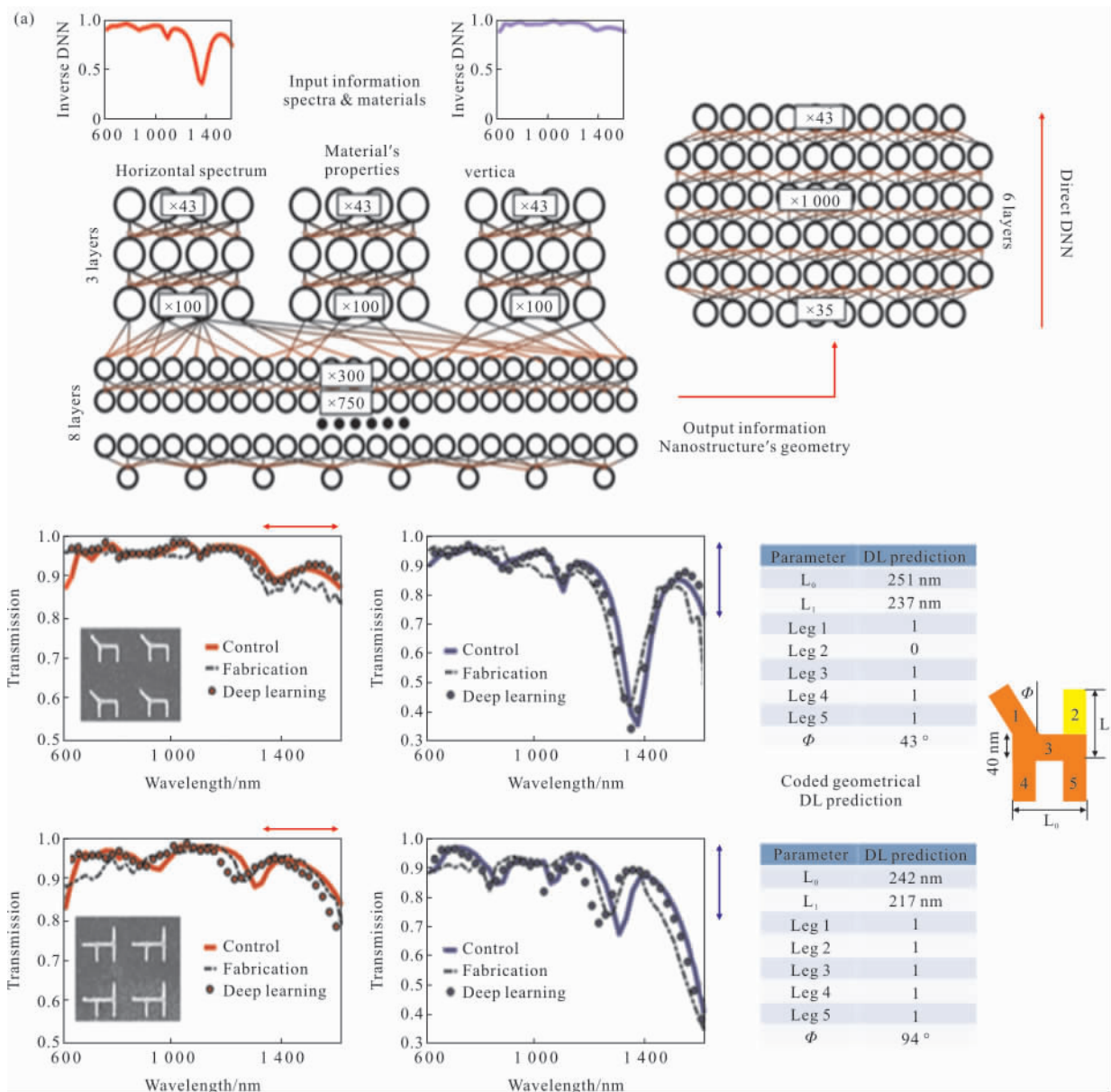
超表面包含微纳米尺度的图案化结构阵列。基于超表面对光波前调控特性的多样性,若要实现特定的波前调控特性,如光束整形,非线性光学等都需要利用矢量电磁场分析方法,如有限元模型、严格耦合波分析、时域有限差分法进行长时间的建模和扫描,得到需要的结构分布或共振响应曲线^[81-84]。时间成本和计算力水平都是限制其发展的重要因素。

如图 16 所示,深度学习以其特有的特征学习和建模的能力将为解决这个问题提供便利。利用深度

学习模型将微纳结构图案与共振响应曲线相映射，实现微纳结构与共振曲线的一一对应。经过适当的训练之后，模型能够泛化在某种条件下的一类问题。这样在已知微纳结构时，便能够通过网络模型得到对应的真实的共振曲线，反之亦然^[85-88]。参考文献[85]中将深度神经网络应用到微纳结构的设计上，获得了很大的成功。如图 16(a)所示，该网络模型深度已经超过 10 层，且每层的相连接的神经元最高甚至达到 1 000 个。模型已经达到相对复杂的程度，但是该模型对结构的光谱曲线的拟合程度仍比较高。该模型是一个双向模型，既能够通过结构预测光谱曲线，又能够根据曲线和必要条件得到结构图

案。经过训练好的模型能够在毫秒量级时间内得到输出结果，相比于 FDTD 或 FEM 计算求解 Maxwell's 方程组时间大大缩短。同时需要注意的是，该模型的逆向结构的预测是基于特定的图案的变化。例如，文中^[85]采用的微纳结构图案的变化是基于上图 16(a)中“H”型结构。DNN 模型的逆向预测给出了特定的自由度，如臂长，“1”臂的旋转角等。该网络并不能够自由地预测和逆向设计任意形式的微纳结构，在一定程度上限制了模型的推广应用。另外，提高模型的深度和逆向结构预测的自由度将会给未来推向实际微纳光学应用提供一个方法。

佐治亚理工大学的 Wenshan Cai 课题组提出采



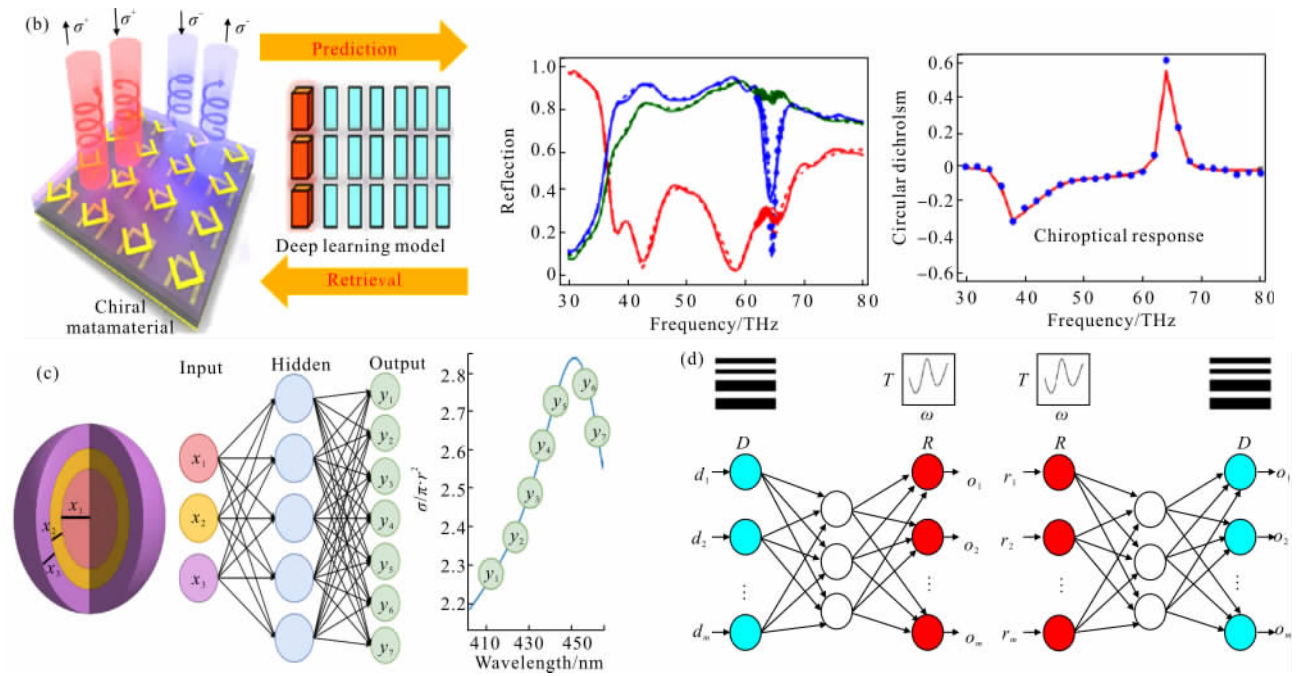


图 16 四种纳米结构光谱预测和逆向纳米参数的神经网络示意图:(a)~(d)四种不同双向网络模型^[85-88]

Fig.16 Four schematics of neural networks for the prediction of nanostructure spectrum and inverse nanostructure parameters:

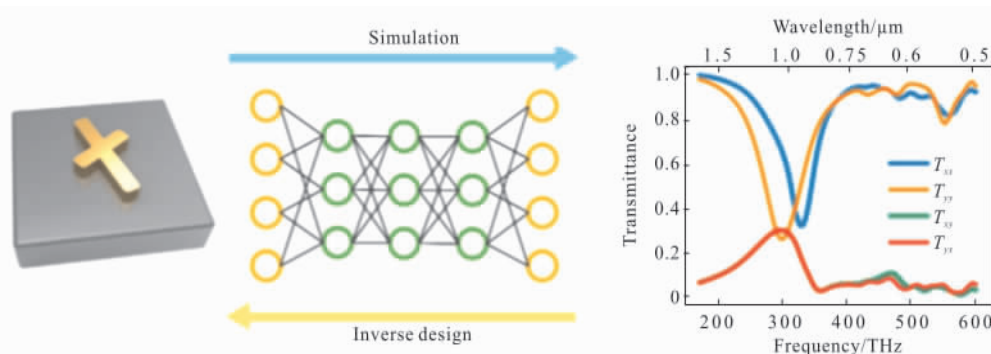
(a)~(d) Four different bi-directional network models^[85-88]

用生成对抗式神经网络进行不同结构的光谱曲线预测和逆向设计^[89]。对抗神经网络能够随机生成图案模型,输入到 GAN 模型中。网络的输出端为该结构的光谱曲线。其逆向设计为从网络的输出端到输入端,根据光谱曲线输出结构图,如图 17 所示。

从图 18 预测结果可以看出,对比网络的输出和真实值,二者存在一定的误差。如图 18(a)中 C 型和扇形结构呈翻转形式;图 18(d)~(e)中预测的曲线虽然接近,但结构却不同。这就说明该模型存在非唯一解。模型预测的精度也是需要衡量的关键。但是, GAN 网络能够解决传统的神经网络只能针对特定结构训练的问题,能够多自由度的搜索可能的解空

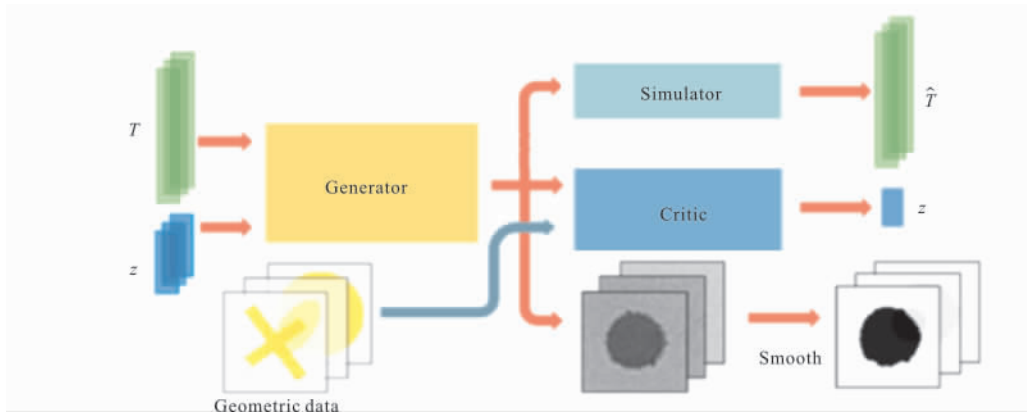
间,大大提高了网络模型的工作效率和泛化能力。但该文并没有分析在预测和训练学习过程中从曲线到图案预测可能会出现非唯一解的情况。这可能需要对网络结构和参数进一步改进和优化。

除此之外,利用深度学习生成光能量分束器和波长选择器等都实现了较好的效果^[87-91]。如图 19(a)所示,利用深度学习逆向设计生成的 Z 型波导弯曲器件和分光器。Z 型波导弯曲器可以实现波导传输的方向改变,并且能量损失低。图 19(b)分光器将入射多波长的光束分成低通与高通两路。图 19(c)则实现不同偏振光束(TE 和 TM 波)的分离。这个设计是基于直接二值搜索算法,设计大小为 $2.4 \mu\text{m} \times 2.4 \mu\text{m}$ 。



(a) 双向神经网络模型

(a) Bi-directional neural network model



(b) GAN model^[89]

(b) GAN 模型^[89]

图 17 GAN 模型训练示意图

Fig.17 Schematic of GAN model training

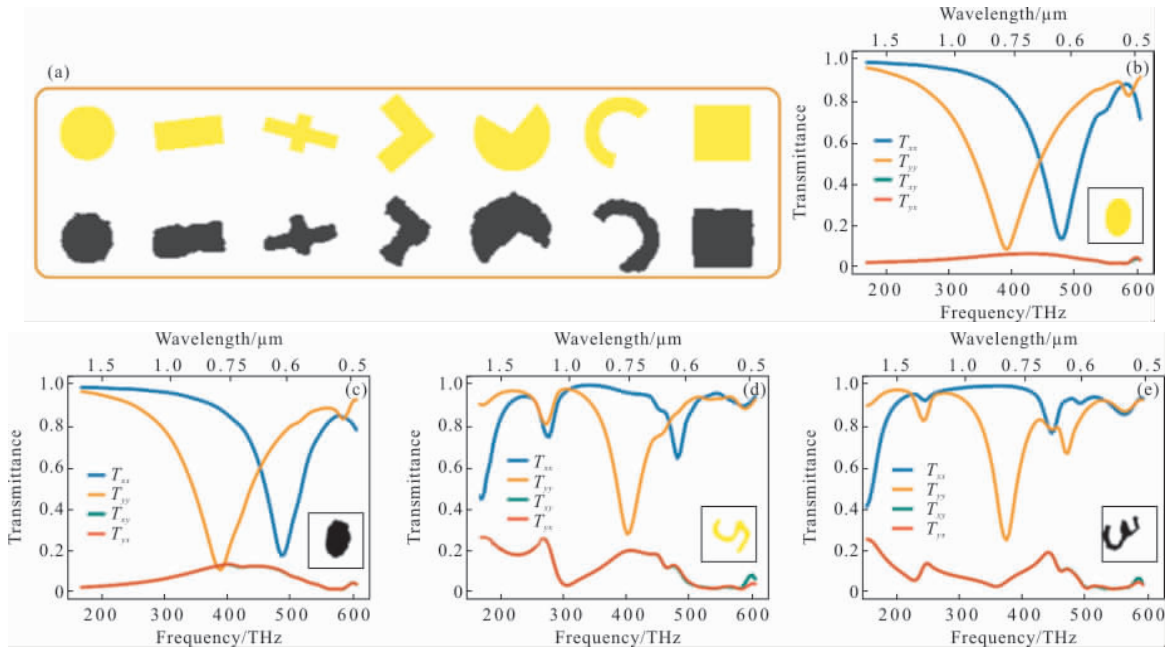
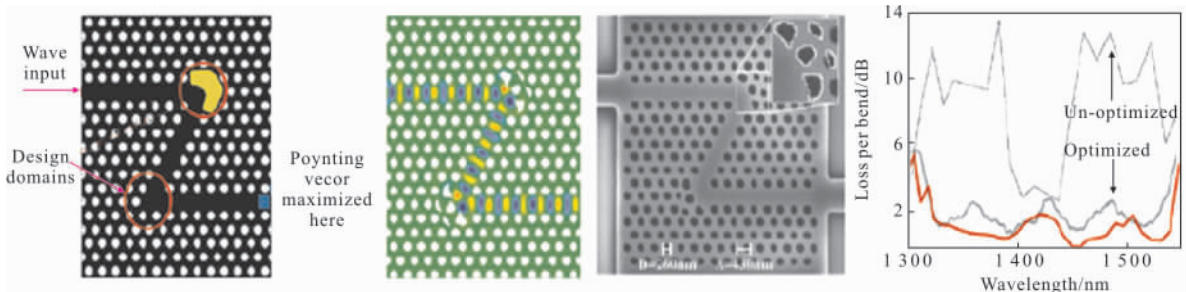


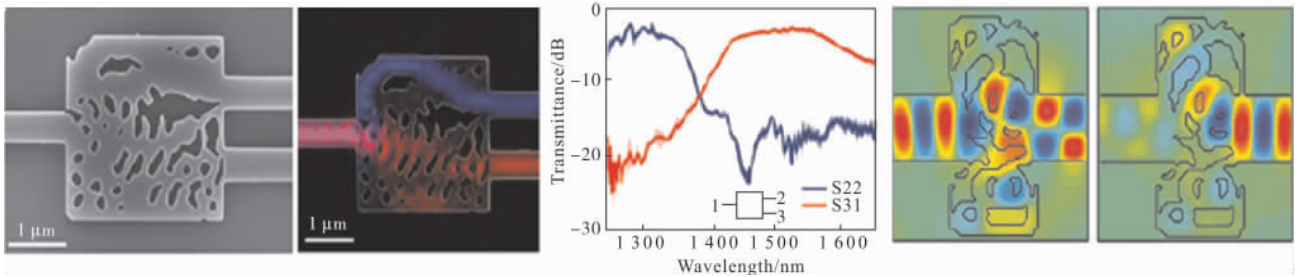
图 18 GAN 纳米结构光谱及逆向纳米结构预测。(a)测试和预测结构形状;(b)、(d)实际光谱曲线和对应结构;(c)、(e)模型输出的光谱曲线和对应结构^[89]

Fig.18 GAN nanostructure spectrum and inverse nanostructure prediction. (a) Testing and predicting the shape of the structure; (b), (d) Actual spectral curve and corresponding structure; (c), (e) Spectral curve of the network output and corresponding structure^[89]



(a) 光束偏折器

(a) Beam deflector

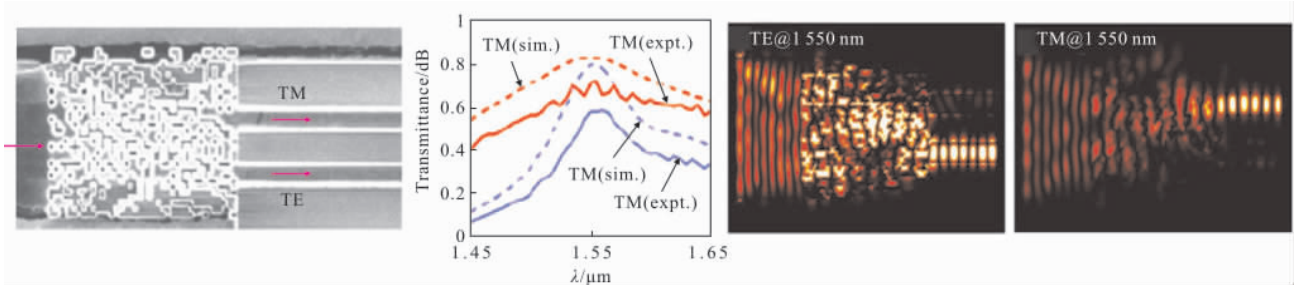


(b) 波长分束器

(c) 光束隔离器

(b) Wavelength beam splitter

(c) Beam isolator



(d) TE/TM波分束器^[90-93]

(d) TE/TM wave splitter^[90-93]

图 19 深度学习逆向设计结构示例

Fig.19 Inverse design applications based on deep learning

这类算法一般是基于目标函数的优化。从优化的结果来看(图 19 (d)),逆向设计的结构都是无序的非均匀非圆孔分布。这样设计的结构具有较大的自由度和非线性空间,因此能够较为容易地完成复杂任务的逆向设计。

深度学习在光学中的应用广泛,如拓扑优化实现调控光束波形,隐身斗篷设计,通信集中器逆向设计等^[94-100],诸多功能有待进一步挖掘。

3 结论与展望

深度学习算法的思想在计算机科学,特别是图像分类与识别、自动驾驶、语义分割、自然语言处理等领域有着广泛的应用,并在光学领域崭露头角,显示出巨大潜力。在大数据时代的今天,能够有效快速的分析和处理图像及更复杂的数据至关重要。但是深度学习算法及其与光学的结合仍然具有改善和提升的空间^[101]:

(1) 数据集的获取及处理。目前主流的机器学习算法大多是监督学习机制,即训练的数据需要包含数据标签。在正文中提到的全息成像,光谱预测等

方面需要对实验数据进行预处理和标签。因此,在大多数文章中结构相对简单、数据量少、模型参数少,限制了人工神经网络,特别是神经网络应对复杂光学问题的能力。发展和训练无监督学习算法是非常有意义和使用价值的。然而越复杂的网络,越大的数据量则会增加预处理时间。既要保证神经网络模型的精度和泛化能力,又要降低训练模型的时间是深度学习落地应用的重要挑战之一。(2) 神经网络与光学机制的契合。光的衍射传播遵从一定规律,试图让神经网络去理解光的衍射干涉等规律,摒弃传播噪声和误差等从而实现准确的预期目标具有一定的难度。(3) 目前应用深度学习实现的功能多是单一网络模型的作用。倘若能够结合多种深度学习模型,发挥不同机器学习算法的优势,往往能扬长避短,提高神经网络在解决具体问题时的表现。(4) 提高计算能力和完善自学习的机制。普通的光学系统一般无法实现在线的实时计算。同时,当光学系统发生非预期变化时,探测器记录的图像数据往往偏离训练数据的分布规则,这就导致深度学习的模型可能会失效,需要重新学习和计算,鲁棒性还不高。因此发

展和推动主动自学习的深度学习算法和快速计算也是非常必要的。深度学习算法用于辅助和推动光计算与通信、芯片结构设计、全息图产生与再现、图像处理、超分辨成像、动态智能显示、逆向微纳结构设计等领域,具有重要研究意义^[102]。

参考文献:

- [1] Simon H. A Comprehensive Foundation [M]. 2nd ed. New Jersey: Prentice Hall International, Inc, 1999.
- [2] Simon H. Machines, Neural Network and Learning[M]. 3rd ed. New Jersey: Pearson Education, Inc, 2009.
- [3] Gardner M W, Dorling S R. Artificial neural networks (the multilayer perceptron)–a review of application in the atmospheric sciences[J]. *Atmospheric Environment*. 1998, 32(14): 2627–2636.
- [4] Hornik K. Multilayer feedforward networks are universal approximators [J]. *Neural Network*, 1989, 2 (5): 359–366.
- [5] Arel I, Rose D C, Karnowski T P. Deep machine learning –a new frontier in artificial intelligence research [J]. *IEEE Computational Intelligence Magazine*, 2010, 5(4): 13–18.
- [6] Qiao Fengjuan, Guo Hongli, Li Wei, et al. Research on deep learning classification based on SVM: a review[J]. *Journal of Qilu University of Technology*, 2018, 32(5): 39–44. (in Chinese)
- [7] Cortes C, Vapnik V. Support –vector networks [J]. *Machine Learning*, 1995, 20(3): 273–297.
- [8] Lecun Y, Bengio Y, Hinton G. Deep learning [J]. *Nature*, 2015, 521(7553): 436–444.
- [9] Lecun Y, Boser B E, Denker J S, et al. Handwritten digit recognition with a back–propagation network[C]// Proc Advances in Neural Information Processing Systems, 1990: 396–404.
- [10] Lecun Y, Bottou L, Bengio Y, et al. Gradient –based learning applied to document recognition [C]// Proceedings of IEEE. 1998, 86(11): 2278–2324.
- [11] Hochreiter S, Schmidhuber J. Long short–term memory [J]. *Neural Computation*, 1997, 9(8): 1735–1780.
- [12] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, et al. Learning phrase representations using RNN encoder –decoder for statistical machine translation [J]. *arXiv Preprint*, 2014, arXiv: 1406.1078.
- [13] Goodfellow I, Pouget –Abadie J, Mirza M, et al. Generative adversarial nets [C]//Advances in Neural Information Processing Systems, 2014: 2672–2680.
- [14] Yoav F, David H. Unsupervised learning of distributions of binary vectors using 2 layer networks[C]//Advances in Neural Information Processing Systems, 1991, 4: 912–918.
- [15] Hinton G E, Osindero S, Teh Y. A fast learning algorithm for deep belief nets [J]. *Neural Computation*, 2006, 18(7): 1527–1554.
- [16] Wang Bingrui, Lan Huiying, Chen Yunji. Programming frameworks for deep learning algorithms [J]. *Big Data Research*, 2018, 4(4): 56–63. (in Chinese)
- [17] Jake B. Notes on convolutional neural networks [D]. Cambridge: MIT Press, 2006.
- [18] Lecun Y, Bengio Y. Convolutional Networks for Images, Speech, and Time Series [M]. Cambridge: MIT Press, 1995.
- [19] Waibel A, Hanazawa T, Hinton G, et al. Phoneme recognition using time–delay neural networks [J]. *IEEE Transactions on Acoustics Speech and Signal Processing*, 1989, 37(3): 328–339.
- [20] Bottou L, Fogelman –Soubié F, Blanchet P, et al. Experiments with time delay networks and dynamic time warping for speaker independent isolated digit recognition[C]//Euro Speech: 1989, 89: 537–540.
- [21] Simard P Y, Steinkraus D, Platt J C. Best practices for convolutional neural networks applied to visual document analysis[C]//Seventh International Conference on Document Analysis and Recognition, 2003: 958–962.
- [22] Vaillant R, Monroq C, Lecun Y. Original approach for the localization of objects in images [J]. *IEEE Proceedings –Vision Image and Signal Proceeding*, 1994, 141(4): 245–250.
- [23] Nowlan S, Platt J. A convolutional neural network hand tracker [C]//Advances Neural Information Processing Systems, 1995: 901–908.
- [24] Lawrence S, Giles C L, Tsoi A C, et al. Face recognition: A convolutional neural –network approach [J]. *IEEE Transactions on Neural Networks*, 1997, 8 (1): 98–113.
- [25] Goodfellow Ian, Bengio Yoshua, Courville Aaron. Deep Learning [M]. Translated by Zhao S, Li Y, Fu T, et al. Beijing: Post and Telecom Press, 2017: 274 –298. (in Chinese)

- Chinese)
- [26] Shin H, Roth H R, Gao M, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning [J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5SI): 1285-1298.
- [27] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [28] Litjens G, Kooi T, Bejnordi B E, et al. A survey on deep learning in medical image analysis [J]. *Medical Image Analysis*, 2017, 42: 60-88.
- [29] Esteva A, Kuprel B, Novoa R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. *Nature*, 2017, 542(7639): 115.
- [30] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651.
- [31] Murtagh P, Tsoi A C. Implementation issues of sigmoid function and its derivative for VLSI digital neural networks [J]. *IEEE Proceedings -E Computers and Digital Techniques*, 1992, 139(3): 207-214.
- [32] Xu G, Wu H, Shi Y. Structural design of convolutional neural networks for steganalysis [J]. *IEEE Signal Processing Letters*, 2016, 23(5): 708-712.
- [33] Han J, Cho G, Kwak K. A design of convolutional neural network using ReLU-based ELM classifier and its application [C]//Proceedings of the 9th International Conference on Machine Learning and Computing, 2017.
- [34] Mirza M O S. Conditional generative adversarial nets [J]. *arXiv preprint*, 2014, arXiv: 1411.1784.
- [35] Radford A M L C S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. *arXiv preprint*, 2015, arXiv: 1511.06434.
- [36] Reed S, Akata Z, Yan X, et al. Generative adversarial text to image synthesis [J]. *arXiv preprint*, 2016, arXiv: 1605.05396.
- [37] Wu J, Zhang C, Xue T, et al. Learning a probabilistic latent space of object shapes via 3D generative-adversarial modelin [J]. *arXiv preprint*, 2016, arXiv: 1610.07584.
- [38] Smith E, Meger D. Improved adversarial systems for 3D object generation and reconstruction[J]. *arXiv preprint*. 2017, arXiv: 1707.09557.
- [39] Denton E, Gross S, Fergus R. Semi-Supervised Learning with Context-Conditional Generative Adversarial Networks [J]. *arXiv preprint*, 2016, arXiv: 1611.06430.
- [40] Yi Z, Zhang H, Tan P, et al. DualGAN: Unsupervised dual learning for image-to-image translation [J]. *arXiv preprint*, 2017, arXiv: 1704.02510.
- [41] Ledig C, Theis L, Huszar F, et al. Photo-realistic single image super-resolution using a generative adversarial network[J]. *arXiv preprint*, 2017, arXiv: 1609.04802v5.
- [42] Vondrick C, Pirsaviash H, Torralba A. Generating videos with scene dynamics [J]. *arXiv preprint*, 2016, arXiv: 1609.02612v3.
- [43] Perarnau G, van de Weijer J, Raducanu B, et al. Invertible conditional GANs for image editing[J]. *arXiv preprint*, 2016, arXiv: 1611.06355.
- [44] Antonia Creswell A A B. Inverting the generator of a generative adversarial network [J]. *arXiv preprint*, 2016, arXiv: 1611.05644.
- [45] Zhou S, Xiao T, Yang Y, et al. GeneGAN: learning object transfiguration and attribute subspace from unpaired data[J]. *arXiv preprint*, arXiv:1705.04932.
- [46] Kim T, Cha M, Kim H, et al. Learning to discover cross-domain relations with generative adversarial networks[J]. *arXiv preprint*, 2017, arXiv: 1703.05192.
- [47] Wang C, Wang C, Xu C, et al. Tag disentangled generative adversarial network for object image re-rendering [C]//The Twenty-Sixth International Joint Conference on Artificial Intelligence, 2017: 2901-2907.
- [48] Antipov G, Dugelay M B J. Face aging with conditional generative adversarial networks [J]. *arXiv preprint*, 2017, arXiv: 1702.01983.
- [49] Arjovsky M, Chintala S, Bottou L. Wasserstein GAN [J]. *arXiv preprint*, 2017, arXiv: 1701.07875.
- [50] Ando T, Horisaki R, Tanida J. Speckle-learning-based object recognition through scattering media [J]. *Optics Express*, 2015, 23(26): 33902-33910.
- [51] Horisaki R, Takagi R, Tanida J. Learning based imaging through scattering media [J]. *Optics Express*, 2016, 24(13): 13738-13743.
- [52] Takagi R, Horisaki R, Tanida J. Object recognition through a multi-mode fiber [J]. *Optical Review*, 2017,

- 24(2): 117–120.
- [53] Horisaki R, Takagi R, Tanida J. Learning based focusing through scattering media [J]. *Applied Optics*, 2017, 56(15): 4358–4362.
- [54] Horisaki R, Takagi R, Tanida J. Learning based single-shot superresolution in diffractive imaging [J]. *Applied Optics*, 2017, 56(32): 8896–8901.
- [55] Sinha A, Lee J, Li S, et al. Lensless computational imaging through deep learning [J]. *Optica*, 2017, 4(9): 1117–1125.
- [56] Li R, Zeng T, Peng H, et al. Deep learning segmentation of optical microscopy images improves 3-D neuron reconstruction [J]. *IEEE Transactions on Medical Imaging*, 2017, 36(7): 1533–1541.
- [57] Lyu M, Wang W, Wang H, et al. Deep-learning-based ghost imaging [J]. *Scientific Reports*, 2017, 7: 17865.
- [58] Jo Y, Park S, Jung J, et al. Holographic deep learning for rapid optical screening of anthrax spores[J]. *Science Advances*, 2017, 3(8): e1700606.
- [59] Lin X, Rivenson Y, Yardimeir N T, et al. All-optical machine learning using diffractive deep neural networks [J]. *Science*, 2018, 361(6406): 1004.
- [60] Xiao H, Rasul K, Vollgraf R, et al. fashion-mnist[EB/OL]. San Francisco: GitHub database site, 2017 [2019–12–17]. <https://github.com/zalando-research/fashion-mnist#fashion-mnist.htm>.
- [61] Lecun Y, Cortes C, Burges C J C. The MNIST database of handwritten digits [EB/OL]. New York: New York University, 1998 [2019–12–17]. <http://yann.lecun.com/exdb/mnist.htm>.
- [62] Li Jingxi, Mengu Deniz, Luo Yi, et al. Class-specific differential detection in diffractive optical neural networks improves inference accuracy [J]. *Advanced Photonics*, 2019, 1(4): 046001.
- [63] Huang B, Bates M, Zhuang X. Super resolution fluorescence microscopy [J]. *Annual Review of Biochemistry*, 2009, 78: 993–1016.
- [64] Willig K I, Rizzoli S O, Westphal V, et al. STED microscopy reveals that synaptotagmin remains clustered after synaptic vesicle exocytosis [J]. *Nature*, 2006, 440(7086): 935–939.
- [65] Rittweger E, Han K Y, Irvine S E, et al. STED microscopy reveals crystal colour centres with nanometric resolution[J]. *Nature Photonics*, 2009, 3(3): 144–147.
- [66] Rust M J, Bates M, Zhuang X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)[J]. *Nature Methods*, 2006, 3(10): 793–795.
- [67] Dong C, Loy C C, He K, et al. Image super-resolution using deep convolutional networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295–307.
- [68] Rivenson Y, Gorocs Z, Gunaydin H, et al. Deep learning microscopy[J]. *Optica*, 2017, 4(11): 1437–1443.
- [69] Wu Y, Rivenson Y, Zhang Y, et al. Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery[J]. *Optica*, 2018, 5(6): 704–710.
- [70] Rivenson Y, Zhang Y, Gunaydin H, et al. Phase recovery and holographic image reconstruction using deep learning in neural networks [J]. *Light-Science & Applications*, 2018, 7: e17141.
- [71] Huang L, Muehlenbernd H, Li X, et al. Broadband hybrid holographic multiplexing with geometric metasurfaces [J]. *Advanced Materials*, 2015, 27(41): 6444.
- [72] Huang L, Chen X, Muehlenbernd H, et al. Three-dimensional optical holography using a plasmonic metasurface[J]. *Nature Communications*, 2013, 4: 2808.
- [73] Zijlstra P, Chon J W M, Gu M. Five dimensional optical recording mediated by surface plasmons in gold nanorods[J]. *Nature*, 2009, 459(7245): 410–413.
- [74] Zhao R, Sain B, Wei Q, et al. Multichannel vectorial holographic display and encryption[J]. *Light-Science & Applications*, 2018, 7: 95.
- [75] Jiang Q, Jin G, Cao L. When metasurface meets hologram: principle and advances [J]. *Advanced Optics and Photonics*, 2019, 11(3): 000518.
- [76] Goodman J W. Introduction to Fourier Optics [M]. 2nd ed. Singapore: McGraw-Hill, 1996: 66–67.
- [77] Schnars U, Falldorf C, Watson J, et al. Digital Holography and Wavefront Sensing: Principles, Techniques and Applications [M]. Berlin: Springer, 2015.
- [78] Ren Z, Xu Z, Lam E Y. Learning-based nonparametric autofocusing or digital holography [J]. *Optica*, 2018, 5(4): 337–344.
- [79] Slinger C, Cameron C, Stanley M. Computer-generated holography as a generic display technology [J].

- Computer*, 2005, 38(8): 46–53.
- [80] Horisaki R, Takagi R, Tanida J. Deep learning – generated holography [J]. *Applied Optics*, 2018, 57(14): 3859–3863.
- [81] Pfeiffer C, Grbic A. Metamaterial Huygens’ surfaces: tailoring wave fronts with reflectionless sheets [J]. *Physical Review Letters*, 2013, 110(19): 197401.
- [82] Arbabi A, Horie Y, Bagheri M, et al. Dielectric metasurfaces for complete control of phase and polarization with subwavelength spatial resolution and high transmission[J]. *Nature Nanotechnology*, 2015, 10(11): 190–197.
- [83] Butet J, Brevet P, Martin O J F. Optical second harmonic generation in plasmonic nanostructures: from fundamental principles to advanced applications [J]. *ACS Nano*, 2015, 9(11): 10545–10562.
- [84] Li G, Chen S, Pholchai N, et al. Continuous control of the nonlinearity phase for harmonic generations [J]. *Nature Materials*, 2015, 14(6): 607–612.
- [85] Malkiel I, Mrejen M, Nagler A, et al. Plasmonic nanostructure design and characterization via deep learning[J]. *Light–Science & Applications*, 2018, 7: 60.
- [86] Liu D, Tan Y, Khoram E, et al. Training deep neural networks for the inverse design of nanophotonic structures[J]. *ACS Photonics*, 2018, 5(4): 1365–1369.
- [87] Ma W, Cheng F, Liu Y. Deep learning enabled on–demand design of chiral metamaterials [J]. *ACS Nano*, 2018, 12(6): 6326–6334.
- [88] Peurifoy J, Shen Y, Jing L, et al. Nanophotonic particle simulation and inverse design using artificial neural networks[J]. *Science Advances*, 2018, 4(6): aar4206.
- [89] 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.
- [90] Tahersima M H, Kojima K, Koike–Akino T, et al. Deep neural network inverse design of integrated photonic power splitters[J]. *Scientific Reports*, 2019, 9: 1368.
- [91] Yao K, Unni R, Zheng Y. Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale[J]. *Nanophotonics*, 2019, 8(3): 339–366.
- [92] Su L, Piggott A Y, Sapra N V, et al. Inverse design and demonstration of a compact on–chip narrowband three–channel wavelength demultiplexer [J]. *ACS Photonics*, 2018, 5(2): 301–305.
- [93] Piggott A Y, Lu J, Babinec T M, et al. Inverse design and implementation of a wavelength demultiplexing grating coupler[J]. *Scientific Reports*, 2014, 4: 7210.
- [94] Molesky S, Lin Z, Piggott A Y, et al. Inverse design in nanophotonics[J]. *Nature Photonics*, 2018, 12(11): 659–670.
- [95] Lan L, Sun F, Liu Y, et al. Experimentally demonstrated a unidirectional electromagnetic cloak designed by topology optimization [J]. *Applied Physics Letters*, 2013, 103(12): 121113.
- [96] Fujii G, Watanabe H, Yamada T, et al. Level set based topology optimization for optical cloaks [J]. *Applied Physics Letters*, 2013, 102(25): 251106.
- [97] Sell D, Yang J, Doshay S, et al. Periodic dielectric metasurfaces with high –efficiency, multiwavelength functionalities [J]. *Advanced Optical Materials*, 2017, 5(23): 00645.
- [98] Deng Y, Korvink J G. Topology optimization for three–dimensional electromagnetic waves using an edge element –based finite –element method [C]//Proceedings of the Royal Society A –mathematical physical and engineering sciences, 2016, 472(2189): 20150835.
- [99] Andkjaer J, Sigmund O. Topology optimized low –contrast all–dielectric optical cloak[J]. *Applied Physics Letters*, 2011, 98(2): 021112.
- [100] Lin Z, Groever B, Capasso F, et al. Topology –optimized multilayered metaoptics [J]. *Physical Review Applied*, 2018, 9(4): 044030.
- [101] Yin Baocai, Wang Wentong, Wang Litong. Review of deep learning [J]. *Journal of Beijing University of Technology*, 2015, 41(1): 48–59. (in Chinese)
- [102] Barbastathis George, Ozcan Aydogan, Situ Guohai, On the use of deep learning for computational imaging[J]. *Optica*, 2019, 6(8): 921–943.