

·特邀综述·

深度学习下的计算成像:现状、挑战与未来

左超^{1,2}, 冯世杰^{1,2}, 张翔宇^{1,2}, 韩静², 陈钱^{2*}

¹南京理工大学电子工程与光电技术学院,智能计算成像实验室(SCILab),江苏南京 210094;

²南京理工大学江苏省光谱成像与智能感知重点实验室,江苏南京 210094

摘要 近年来,光学成像技术已经由传统的强度、彩色成像发展进入计算光学成像时代。计算光学成像基于几何光学、波动光学等理论对场景目标经光学系统成像再到探测器采样这一完整图像生成过程建立精确的正向数学模型,再求解该正向成像模型所对应的“逆问题”,以计算重构的方式来获得场景目标的高质量图像或者传统技术无法直接获得的相位、光谱、偏振、光场、相干度、折射率、三维形貌等高维度物理信息。然而,计算成像系统的实际成像性能也同样极大地受限于“正向数学模型的准确性”以及“逆向重构算法的可靠性”,实际成像物理过程的不可预见性与高维病态逆问题求解的复杂性已成为这一领域进一步发展的瓶颈问题。近年来,人工智能与深度学习技术的飞跃式发展为计算光学成像技术开启了一扇全新的大门。不同于传统计算成像方法所依赖的物理驱动,深度学习下的计算成像是一类由数据驱动的方法,它不但解决了许多过去计算成像领域难以解决的难题,还在信息获取能力、成像的功能、核心性能指标(如成像空间分辨率、时间分辨率、灵敏度等)上都获得了显著提升。基于此,首先概括性介绍深度学习技术在计算光学成像领域的研究进展与最新成果,然后分析了当前深度学习技术在计算光学成像领域面临的主要问题与挑战,最后展望了该领域未来的发展方向与可能的研究方向。

关键词 成像系统; 计算成像; 深度学习; 光学成像; 光信息处理

中图分类号 O436

文献标志码 A

doi: 10.3788/AOS202040.0111003

Deep Learning Based Computational Imaging: Status, Challenges, and Future

Zuo Chao^{1,2}, Feng Shijie^{1,2}, Zhang Xiangyu^{1,2}, Han Jing², Qian Chen^{2*}

¹Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering,

Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China;

²Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China

Abstract In recent years, optical imaging techniques have entered into the era of computational optical imaging from the traditional intensity and color imaging. Computational optical imaging, which is based on geometric optics, wave optics, and other theoretical foundations, establishes an accurate forward mathematical model for the whole image formation process of the scene imaged through the optical system and then sampled by the digital detector. Then, the high-quality reconstruction of the image and other high dimensional information, such as phase, spectrum, polarization, light field, coherence, refractive index, and three-dimension profile, which cannot be directly accessed using traditional methods, can be obtained through computational reconstruction method. However, the actual imaging performance of the computational imaging system is also limited by the “accuracy of the forward mathematical model” and “the reliability of inverse reconstruction algorithm”. Besides, the unpredictability of real physical imaging process and the complexity of solving high dimensional ill-posed inverse problems have become the bottleneck of further development of this field. In recent years, the rapid development of artificial intelligence and deep learning for the technology opens a new door for computational optical imaging technology. Unlike “physical driven” model that traditional computational imaging method is based on,

收稿日期: 2019-11-06; 修回日期: 2019-11-22; 录用日期: 2019-12-05

基金项目: 国家自然科学基金(61722506, 61705105, 11574152)、总装“十三五”装备预研项目(30102070102)、总装“十三五”领域基金(61404150202)、国防科技项目基金(0106173)、江苏省杰出青年基金(BK20170034)、江苏省重点研发计划(BE2017162)、江苏省“333工程”科研项目资助计划(BRA2016407)、江苏省光谱成像与智能感知重点实验室开放基金(3091801410411)

* E-mail: chenqian@njust.edu.cn

computational imaging based on deep learning is a kind of “data-driven” method, which not only solves many problems considered quite challenge to be solved in this field, but also achieves remarkable improvement in information acquisition ability, imaging functions, and key performance indexes of imaging system, such as spatial resolution, temporal resolution, and detection sensitivity. This review first briefly introduces the current status and the latest progress of deep learning technology in the field of computational optical imaging. Then, the main problems and challenges faced by the current deep learning method in computational optical imaging field are discussed. Finally, the future developments and possible research directions of this field are prospected.

Key words imaging systems; computational imaging; deep learning; optical imaging; optical information processing

OCIS codes 110.0180; 100.5070; 180.6900

1 引 言

视觉是人类获得客观世界信息的主要途径,而人眼受限于视觉性能,在时间、空间、灵敏度等方面均存在局限性。光学成像技术由此应运而生,其利用各种光学成像系统,如显微镜、望远镜等,实现光信息的可视化,同时延伸扩展人眼的视觉特性。然而,一方面,传统光学成像系统因受强度成像机理、探测器技术水平、光学系统设计、成像衍射极限等因素制约,在空间分辨、时间分辨、光谱分辨、信息维度与探测灵敏度等方面仍存在一定局限性,难以满足人们对成像系统功能与性能的进一步需求,以及军民领域日益增长的高分辨、高灵敏度和多维高速成像的应用需求。采用传统光学成像系统的设计思路想要获得成像性能的少量提升,通常意味着硬件成本的急剧增加,甚至难以实现工程化应用。另一方面,光探测器规模尺寸、像元大小、响应灵敏度等均已接近物理极限,很难满足这些极具挑战性的需求。

随着成像电子学的发展,计算机数据处理能力的增强,光场调控、孔径编码、压缩感知、全息成像等光电信息处理技术取得了重大进展;此外,经过成千上万年的发展自然界已经演化出多类能够满足不同生存需求的生物视觉系统,从生物视觉系统中获得灵感无疑可以给新一代光学成像技术的发展带来有益的启示。在此背景下,20世纪90年代中期,光学成像界和图像处理界的许多研究人员不约而同地探索出了一种新型成像模式,即图像形成不再仅仅依赖于光学物理器件,还依赖于前端光学和后端探测信号处理的联合设计。这种技术就是现在广为人知的“计算成像”(Computational Imaging)技术^[1],它将光学调控与信息处理有机结合,为突破上述传统成像系统中的诸多限制性因素提供了新手段与新思路。

计算光学成像是一种通过联合优化光学系统和信号处理来实现特定成像功能与特性的新兴研究领

域。其建立在几何光学、波动光学,甚至光量子模型的基础上,采用照明与光学系统调制等方式,建立目标场景与观测图像之间的变换或调制模型,然后利用逆问题求解等数学手段,通过计算反演来进行成像。这种计算成像方法实质上就是在场景和图像之间建立某种特定的联系,这种联系可以是线性的也可以是非线性的。它突破了传统成像技术点对点一一对应的强度直接采样形式,采用了更加灵活的非直接的采样形式,更能充分发挥成像系统中各组件的特点与性能。这种灵活的设计模式可以改变光学测量的性质以获得所需的结果,并平衡物理域和计算域之间图像生成和信息提取所依赖的资源。基于信息论的概念,计算光学成像设计师不仅可以借助于传统光学设计的优势,还可以充分利用物理光学在光信号处理中的潜力来设计成像系统。这种新型的成像方式将有望改变成像系统获取信息的方式,提升其获取信息的能力,增强资源利用,赋予其诸多传统光学成像技术难以获得甚至无法获得的革命性的优势:例如,突破探测器制造工艺、工作条件、功耗成本等因素的限制,有效提高成像质量(信噪比、对比度、动态范围),简化系统(无透镜、小体积、低成本),突破光学系统与图像采集设备的分辨率限制(超像素分辨、超衍射极限),并使其功能(相位、光谱、偏振、光场、相干度、折射率、三维形貌、景深严拓、模糊复原、数字重聚焦、改变观测视角)、性能(空间分辨、时间分辨、光谱分辨、信息维度与探测灵敏度)、可靠性、可维护性等获得显著提高,有助于实现成像设备的高性能、微型化、智能化。

现如今,计算光学成像已发展为一门集几何光学、信息光学、计算光学、计算机视觉、现代信号处理等理论于一体的新兴交叉技术研究领域,成为光学成像领域的一大国际研究重点和热点。然而,隐藏在计算成像华丽外衣之下的是其所必须付出的额外成本与代价:用于进行非传统测量的物理实体器件相关的成本、多次测量产生的时间成本、数据量以及

物理模型和校准对处理性能的影响。更重要的是,计算成像技术的实际成像性能极大程度地受限于“正向数学模型的准确性”以及“逆向重构算法的可靠性”,实际物理成像过程中的不可预见性与高维病态逆问题求解的复杂性已成为这一领域进一步发展亟需解决的瓶颈问题。

近几年,DeepMind 公司研制的人工智能机器人 AlphaGo 战胜顶尖围棋棋手李世石^[2]、先进图像分类算法在具有挑战性的数据集 ImageNet 上的正确率超过人类^[3]等令人振奋消息一个接一个地传来,人工智能已经成为我们身边一个耳熟能详的词汇,国际上也开始迎来这一技术的研究热潮。当下谈到人工智能,“机器学习”、“深度学习”和“神经网络”便是经常浮现在人们脑海里的高频词汇。借助于数学中集合的概念,它们之间的关系可以理解为一种包含关系,也就是“机器学习”包含“深度学习”,“深度学习”包含“神经网络”。深度学习已经成为目前最为热门的一种机器学习方案。深度学习这一名称中的“深度”一词表示其使用的神经网络结构多于四层。一般而言,随着神经网络层数的增加,神经网络的性能会更强,学习的效果也会更佳。

互联网技术的蓬勃发展指引着大数据时代的来临,以数据推动的深度学习技术无疑是大数据时代的算法利器。相比于传统的机器学习技术:首先,深度学习技术可利用不断增多的数据不断提升其性能,而传统机器学习技术无法做到这一点;其次,有别于传统方法需要手动提取特征,深度学习技术是一项全自动的技术,它可以从海量数据中直接抽取特征,并且,对于不同的任务,不再需要设计独特的特征提取器,所有工作都可由深度学习自动完成。这是智能机器逐渐代替人工操作的一个显著体现,因此深度学习技术已成为大数据时代的一项热点技术,无论学术界还是工业界都对这项技术产生了浓厚的兴趣。特别是在计算机视觉领域,深度学习作为近年来兴起的一种“数据驱动”的技术,在图像分类、物体检测及识别等诸多应用上均取得了巨大成功。

自 2017 年初,深度学习技术逐渐走入计算成像领域研究者的视野,并在短短的两三年内已在数字全息成像^[4-9]、傅里叶叠层成像技术^[10-13]、鬼成像/单像素成像^[14-16]、超分辨显微成像技术^[17-22]、光学相干层析成像(OCT)^[23-27]、散射介质成像^[28-32]、极弱光成像^[33-34]、跨模态染色成像^[35-36]、光栅条纹分析^[37-39]与快速三维成像^[40-42]等成像体制上得以成

功应用,取得了一系列令人瞩目的开创性研究成果。令人欣喜的是,对比传统物理模型驱动的计算成像技术,样本数据驱动的深度深度学习下的计算成像技术发生了思想观念上的根本变革,它不但解决了许多过去计算成像领域难以解决的难题,还在信息获取能力、成像的功能、核心性能指标(如成像空间分辨率、时间分辨率、灵敏度等)上获得了显著提升。如今,以深度学习为主题的计算成像相关方面的论文喷井而出,呈指数式增长趋势。

在此背景下,本文概括性地介绍深度学习技术在计算光学成像领域的研究现状与最新进展。简要讨论计算成像技术与深度学习技术的基本概念,并按照深度学习技术的“目的与动机”或者说“深度学习技术为传统计算成像技术带来了哪些新的要素”进行细分,对现有深度学习计算成像技术的研究现状及其典型应用进行概述。值得注意的是,深度学习是一把“双刃剑”,它给计算成像领域研究带来了惊喜的同时也引入了一系列亟待解决的问题。本文分析了当前深度学习技术在计算光学成像领域面临的主要问题与挑战,这亦是本文重要的组成部分。最后,对深度学习在计算成像领域未来的发展方向与可能的研究方向进行讨论并展望,并给出了总结性评论。

2 深度学习下的计算成像:现状

一个典型的光学成像系统主要由光源、光学镜头组、光探测器三部分组成。其通过将三维场景中目标发出的光线聚焦在光探测器上进行“点对点”成像。然而这种“所见即所得”的成像方式因其单视角、平面投影等因素的限制,导致高维度场景信息存在缺失。除此之外,日益复杂庞大的光学成像系统也限制着其应用场景。为了解决传统光学成像系统所面临的问题,计算成像技术应运而生,其采用“先调制,再拍摄,最后解调”的成像方式。将光学系统(照明、光学器件、光探测器)与数字图像处理算法作为一个整体考虑,并在设计时一同进行综合优化,前端成像元件与后端数据处理二者相辅相成,构成一种“混合光学-数字计算成像系统”,如图 1 所示。不同于传统光学成像的“所见即所得”,计算成像建立在几何光学、波动光学,甚至光量子模型的基础上,采用照明与光学系统调制等方式,建立目标场景与观测图像之间的变换或调制模型,然后利用逆问题求解等数学手段,通过计算反演来进行成像,以获得场景目标的高质量图像与高维度物理信息。

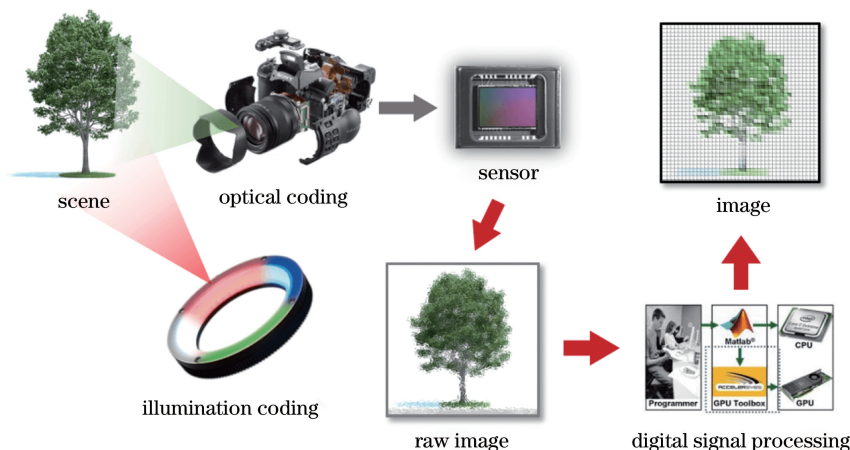


图1 计算光学成像系统的成像过程

Fig. 1 Imaging process of computational optical imaging system

俗话说“天下没有免费的午餐”，任何事物的存在都具有两面性，计算成像技术亦是如此。当设计计算成像系统时，必须权衡计算成像和传统成像相关的成本代价与预期的改进效用。计算成像技术所能带来的功能与性能上的提升往往是以复杂昂贵的系统硬件、大量额外的数据采集、复杂耗时的算法处理等为代价而换取的。更重要的是，计算光学成像华丽的外衣下还掩盖了其所依赖的两大关键问题：

第一，如何准确地建立场景与图像之间的正向模型，并使得采样数据包含场景中所感兴趣的物理信息，这是最关键的问题；

第二，如何利用逆问题优化求解来重构图像，这也是个核心问题。它既要与非直接采样的自身特点相匹配以建立优化模型，又要在重构过程中保证重构图像的准确度。

然而通过已有的实验结果可以发现，在某些领域（特别是间接成像），目前通过计算成像算法重构的图像和基于传统透镜直接拍摄到的图像在成像质量和保真度上仍然有一些差距。其根本原因在于计算成像要求算法中采用的光学系统的数学模型能够真实全面地反映实际的物理成像过程，如果该模型不能真实地反映光学系统的复杂性，计算成像将有可能得不到理想的成像结果，其优势也可能会完全丧失。例如，如果数学模型对光学成像过程中的光的波动性质、像差或系统对温度变化的敏感度的客观参数进行忽略或者简化，则很可能出现这种情况。此外在数据采集的过程中还会受到各种环境不确定因素的影响（噪声、振动等），从而会导致所建立的成像模型并不准确。即使完全知道这些影响因素的存在，设计人员仍然会面临着这样一个难题：简单

化的模型可能无法产生精确有效的结果，而更加真实的模型可能需要大量的系统参量、很高的处理负载和很长的处理时间。实际考虑因素包括正向模型成立的条件和逆向重构算法的复杂程度、测量过程对噪声及环境扰动的敏感性，以及算法后处理引入的伪影(Artifacts)水平等。

深度学习作为近年来兴起的一种“数据驱动”的技术，其在图像分类、物体检测及识别乃至“看图作文(Image Captioning)”等诸多计算机视觉任务上均取得了巨大成功。并且由于其在“黑盒子般的盲建模”与“高维非线性特征拟合”方面的卓越表现，深度学习自2017年初也逐渐得到计算成像领域研究者的广泛关注，并在短短的两三年内取得了一系列令人瞩目的开创性研究成果。深度学习的成功不仅仅带来了人工智能相关技术的快速进步，还解决了计算成像领域许多过去被认为是难以解决甚至无法解决的难题，更重要的是它给该领域带来了思想观念上的根本变革：

1) 从“物理模型驱动”到“样本数据驱动”

在深度学习兴起之前，成像物理模型和经验驱动主宰了计算成像领域多年。一个典型的计算光学显微成像系统由照明、样品、成像系统、探测器四部分构成。照明光与样品发生作用后，成为其本质信息（如吸收、相位、光谱、三维、折射率等）的载体，通过对照明与成像系统进行光学调控使物体的本质信息转化为光强信号并由探测器离散采集，最后通过相应的重构算法对样品本质信息进行反演，获得样品的图像或其他所感兴趣的高维物理信息。为了对整个成像过程进行数学建模，通常需要基于标量衍射理论或部分相干理论对照明光产生与自由传播以

及与被测物体相互作用进行建模。例如:部分相干光场需要利用交叉谱密度/互强度或者相空间光学理论中的维格纳函数来对其进行表征;在空域利用交叉谱密度/互强度/维格纳函数所满足的传输方程,或在频域引入衍射的角谱理论去描述待测物体对照明光波的散射作用;利用 van Cittert-Zernike 定理、部分相干光学传递函数理论等去描述光学系统对成像过程的影响;最终完成从产生照明光到传感器上产生低维耦合离散光强信号的整体过程的正向数学建模。模型初步建立后,通常还需要利用 VirtualLab、Comsol 等光学仿真软件基于严格的麦克斯韦方程求解算法对成像过程进行模拟计算,并在现有成像系统上利用已知物体实测加以比对,对成像的正向模型进行进一步修正与优化。整个正向物理建模过程依赖于大量的专家知识和经验驱动,严重影响了计算成像技术的通用性和可重用性。

深度学习彻底颠覆了这种“物理模型驱动”的范式,开启了“样本数据驱动”的学习范式。具体体现在两点:第一,所谓的经验和知识也在样本数据中,在数据量足够大时无需显式的经验或知识的嵌入,直接从数据中可以学到;第二,基于深度神经网络特有的“高维特征自动提取”能力,可以直接从原始信号进行学习,而无需借助人体的特征变换或提取。数据驱动的学习范式使得科研人员无需根据经验和知识针对不同的成像问题设计不同的处理流程,从而大大提高了算法的通用性,也大大降低了解决新问题的难度。

2) 从“分步/分治”到“端到端学习”

分治或分步法,即将复杂的问题分解为若干简单子问题或子步骤,这曾经是解决复杂问题的常用思路。分步法在计算成像领域,也是被广泛采用的方法论。比如,为了解决数字全息图重构问题,过去经常将其分为预处理、相位解调、衍射计算(数值传播)、焦面判断等若干步骤。再如,为了解决非线性优化问题,可以采用分段线性方式来逼近全局的非线性。这样做的动机虽然很清晰,即子问题或子步骤变得简单、可控、更易解决,但从深度学习的视角来看,其劣势也同样明显:子问题最优未必意味着全局的最优。相反,深度学习更强调端到端的学习,即不去人为地分步骤或者划分子问题,而是完全交给神经网络直接学习从原始输入到期望输出的映射。相比分治策略,端到端的学习有协同增效的优势,有更大的可能获得全局上更优的解。当然,如果一定要把分层看作是“子步骤或者子问题”也是可以的,

但这些分层各自完成什么功能并不是预先设置好的,而是通过基于数据的全局优化来自动学习的。

3) 从“病态非线性逆问题”到“直接(伪)正向非线性建模”

计算成像中所涉及的众多复杂逆问题本质上是高度病态且非线性的,而深度学习实现了从输入到输出的非线性变换,这是深度学习在众多复杂问题上取得突破的原因之一。在深度学习出现之前,众多线性模型求解或非线性迭代优化算法是计算成像图像重构的主流技术。对于可通过近似手段线性化的逆问题(如傍轴近似下的相位恢复问题可通过光强传输方程线性化直接求解,Born 或 Rytvo 近似下的某些逆散射问题也可以实现线性化求解),相应的病态方程组求逆、反卷积与偏微分方程求解等是求解这类问题的核心算法;对于无法线性化的逆问题(如非傍轴条件或复杂照明情况下相位重构问题),可基于凸集投影与梯度搜索的优化算法进行迭代求解(尽管解空间往往是非凸的,但事实证明这些优化算法往往是奏效的)。一般而言,基于某些限制性假设的线性化求解方法所得的解可以作为更为一般条件下非线性问题求解的初值,以提高迭代算法的收敛速度与求解的稳定性。针对逆问题的病态性,通常通过引入被测物体的先验作为正则化手段限定解空间以使其良态化。这里值得一提的是压缩感知技术,它由于在解决病态逆问题方面的突出表现成为了计算成像领域中一个耳熟能详的专业词汇。压缩感知的核心假设在于已知解具有稀疏性(Sparsity),因此可以使用少量的数据来接近完美地恢复原始信号。稀疏性可以作为约束或者正则项,提供额外的先验信息。而大部分信号本身并不是稀疏的(即在自然基下的表达不是稀疏的),但是经过适当的线性变换后是稀疏的(即在另一组基下是稀疏的),如离散余弦变换与小波变换等。该领域曾经非常热门的一个研究课题是字典学习(Dictionary Learning)和变换学习(Transform Learning),通过大量的信号实例,自适应地学习最优的稀疏性表达(自组建完备字典)。为了求解稀疏约束下的最小化问题(最常用的是总变分最小化),需要进一步确定最小化能量泛函的数值求解算法(最陡下降法、非线性共轭梯度法、迭代阈值收缩法等),以及相应的自适应正则化参数的选取方法(对噪声进行统计建模,并对其局部方差进行准确估计),以获得稳定且有意义的解。

而深度学习则利用复杂的网络结构与非线性激活函数,在提取样本高维特征的同时,不断去除与目

标特征无关的信息,最终获得了足以适配足够复杂系统的非线性变换能力。实际上,深度学习同样隐含了稀疏性先验,但它认为问题本身就是稀疏的,可以不断将输入数据进行“降维”,把高维的数据空间投影到抽象但低维的认知空间。相比于传统方法“先正向建模,再求解该模型下的逆问题”的思路,深度学习技术不需要这种数据表达过程的可逆性。它直接建立了从图像到待恢复信息的“伪正向模型”——将光学系统拍摄到的图像作为“网络输入”,将待恢复的期望信息作为“网络输出”,巧妙越过“非线性病态逆问题”求解这一大障碍,直接通过高维度特征拟合实现图像与信息的提取与重建。

正是由于深度学习为计算成像技术从思想观念上带来了重大变革,且其研究内容也是极其多样与发散的,目前还没有一个比较明确的分类方法。如果按研究的问题或者成像的体制来分的话可能会较为琐碎。因此,在本节中我们将按照采用深度学习技术的“目的与动机”或者说“深度学习技术为传统

计算成像技术带来了哪些新的要素”进行细分。一般而言,引入深度学习均是为了从不同的角度、采用不同的方法来解决传统计算光学成像三方面的问题:

1) 提升传统计算成像技术的信息获取能力:突破传统计算成像技术的“信息量守恒”准则,从极少量原始图像数据中解耦并挖掘出更多场景的本质信息。

2) 降低传统计算成像技术对“正向物理模型”或“逆向重构算法”的过度依赖:绕过精确物理建模与病态非线性逆问题求解的障碍,使计算成像技术实施起来更加简单智能。

3) 突破传统计算成像技术所能够达到的功能/性能疆界:实现传统计算成像技术因物理模型所限而无法实现的功能与无法达到的性能指标。

按此方式分类的整个框架如图 2 所示。下面我们就依据此分类方式,对现有深度学习计算成像技术以及典型应用进行概述。

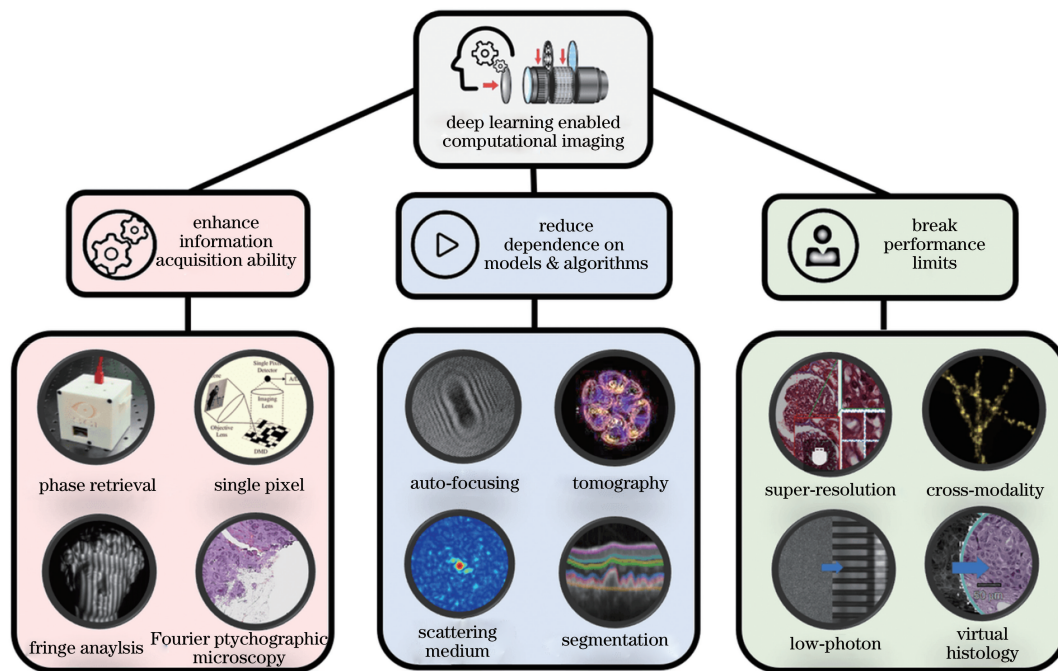


图 2 基于“目的与动机”对典型深度学习计算成像技术所作的分类

Fig. 2 Classification of typical deep learning based computational imaging techniques according to their objectives and motivations

2.1 提升传统计算成像技术的信息获取能力

从信息论的角度而言,数据不会凭空产生,但计算成像技术将光学系统的信息获取能力与计算机的信息处理能力相结合,通过光学调控与相应的信息处理技术从原始图像数据中解耦并挖掘出更多场景的本质信息。然而,对于传统的计算成像技术而言,

研究者都会或多或少地在直觉上遵守着“信息量守恒”准则,而深度学习技术的出现为突破这一准则的约束提供了可能性。

实例 1——相位恢复(无透镜显微/同轴全息)

在 Gerchberg 等^[43-44]所提出的迭代相位恢复算法(G-S 算法, Gerchberg-Saxton 算法)中,为了确保

解的存在性与唯一性(排除孪生像)^[45-47],避免因迭代算法陷于局部极小值造成的收敛停滞^[48-50]等问题,往往需要采集两幅甚至多幅不同离焦距离上的衍射图像(同轴全息图),从而利用更多测量值的约束来提高算法的收敛性和可靠性^[51]。然而,代价是需要额外获取大量的原始数据,且成像系统依赖于高精度的轴向位移装置。仅仅依靠单幅衍射图虽然也可以实现图像重构,但一般仅限于尺寸较小且分布稀疏的样本(采用空域支持域约束),对于一般的较大尺度的样品而言,由相位缺失造成的“孪生像”会在物体成像周围产生自干涉“伪影”,极大地影响成像质量。为了解决这一问题,2018年 Rivenson 等^[4]基于

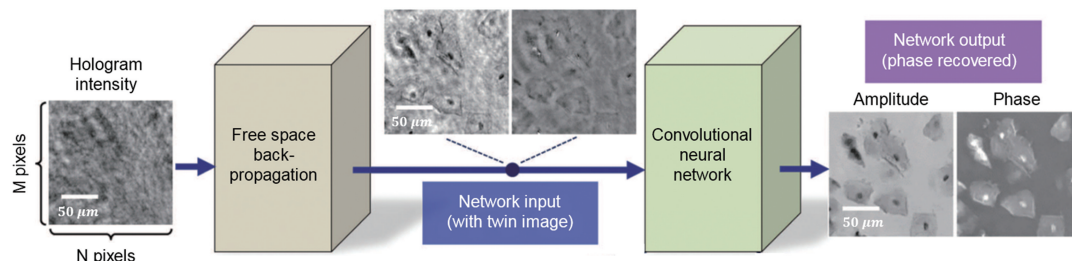


图3 使用深度学习进行单帧无透镜相位恢复^[4]

Fig. 3 Single-frame lensless phase recovery using deep learning^[4]

实例2——傅里叶叠层显微成像

2013年,Zheng等^[52]提出了一种基于相位恢复与合成孔径的计算显微成像技术——傅里叶叠层显微成像(FPM)。在该技术中,样品被不同角度的照明光束(通常是一个LED阵列)依次照射后,由低数值孔径物镜拍摄一系列低分辨率图像。由于成像系统有限孔径效应对于频域的低通滤波,改变了照明光束角度,实现了物体频谱在频域子孔径的交叠扫描。傅里叶叠层成像的核心优势在于其不仅仅能获得待测样品的相位信息,还能在基于最优化的交叠更新过程中实现频域内的合成孔径,有效促进了成像分辨率的提高^[53-55]。与传统的频域合成孔径超分辨率算法不同,傅里叶叠层成像交叠更新算法的相位恢复与频域合成孔径是同时完成的,这也正是傅里叶叠层成像技术本身的优美之处。通常情况下,低数值孔径的低倍率物镜本身具有很大的观察视场,再加之利用大角度照明光束依次照射样品,并在频域进行合成孔径,最终将成像的等效数值孔径提升到物镜与照明数值孔径之和,即保持低倍率物镜的大视场的同时,又达到很大的成像空间带宽积。然而,傅里叶叠层显微成像中空间带宽积的提升往往是以大量(数百幅)低分辨率图像数据采集与高度复杂的频域变换与空域约束反复迭代为代价的。为

深度学习提出了单帧相位恢复技术,该方法的思想在于仅使用一幅相机拍摄到的离焦强度图像进行相位恢复。将拍摄到的离焦图像直接反向传播至焦平面,即将传播得到的复振幅作为网络输入,以使用8幅离焦图像通过传统G-S迭代算法得到的样品清晰相位为目标,利用深度神经网络来模拟相位提取算法的过程,成功地在孪生像与物体伪影的干扰中从单幅同轴全息图提取并分离出了待测样品的真实信息,获得了准确的振幅与相位信息(图3)。相比于传统的G-S迭代算法,不仅避免了复杂的迭代优化过程,还大大降低了成像所需的图像数目,单幅重建也使得该系统不再需要复杂的轴向机械位移装置。

为了解决这一问题,2018年,Nguyen等^[10]提出将深度学习技术应用于FPM领域。该方法思想在于将使用传统FPM技术获得的相位作为目标真值,将在自行设计的照明模式下拍摄到的5幅图像作为输入图像,利用深度神经网络模拟FPM中图像信息的提取与合成,从输入图像中提取物体的相位信息,整个过程如图4所示。经过训练的网络在保证重构成像质量的前提下,将FPM成像所需的图像数目大大降低,并在减少数据量的需求的同时,避免了传统重构算法繁琐的迭代优化过程。

实例3——条纹相位分析(相位测量)

在条纹相位分析领域,光学相位测量技术已被广泛应用于光学干涉测量、数字全息、电子散斑干涉、莫尔轮廓术及条纹投影轮廓术等。这些方法的一大共性在于通过干涉或者投影的方式在物体表面形成周期性的结构条纹,从而使所测量物体的相关物理量直接或间接地反映在条纹的相位信息中。所以,从根本上而言,这些光学技术的测量精度直接取决于条纹图案的相位解调精度。因此,条纹图案分析是光学相位测量技术中最核心的步骤,也通常是最困难的部分。经典的条纹分析技术大致可分为两类:

1) 时域相移解调法:采用多幅具有相对相位差的条纹图像进行相位提取。该方法能够实现像素级

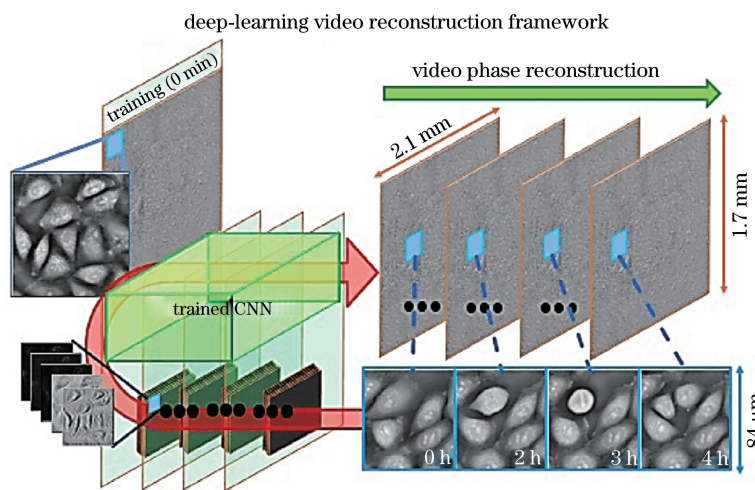


图 4 使用深度学习技术进行少图快速 FPM 成像^[10]

Fig. 4 Fast FPM imaging with few images using deep learning technology^[10]

的高分辨率相位测量,但需要采集多幅条纹图像,容易受到物体运动/环境振动等外界干扰的影响,通常难以应用于动态测量。

2) 空域相位解调法:仅采用单幅条纹图像(通常是包含载频的高频条纹)实现相位信息的提取,如傅里叶变换法(FT)、加窗傅里叶变换法(WFT)等。但对条纹陡变、不连续以及物体细节丰富的区域较为敏感,难以实现高精度、高分辨率的相位测量。且算法一般具有较多的参数(如滤波窗尺寸等)需手动调节,难以实现全自动化操作。

针对这一问题,本课题组首次将深度学习技术应用在条纹分析中,并有效提高了条纹投影轮廓术的三维测量精度^[37]。该方法的思想在于仅采用一幅条纹图像作为输入,利用深度神经网络来模拟相

移法的相位解调过程。如图 5 所示,构建两个卷积神经网络(CNN1 和 CNN2)。CNN1 负责从输入条纹图像(I)中提取背景信息(A);随后,CNN2 利用提取的背景图像和原始输入图像生成所需相位的正弦部分(M)与余弦部分(D);最后,将该输出的正余弦结果代入反正切函数计算得到最终的相位分布。相比于傅里叶变换法与加窗傅里叶变换法,该方法能够更为准确地提取相位信息,特别是针对具有丰富细节的物体表面,相位精度可提升 50% 以上,仅采用一幅输入条纹图像但总体测量效果接近于 12 步相移法[如图 5(a)~(d)所示]。该技术目前已被成功应用于高速三维成像,实现了速度高达 20000 frame/s 的高精度三维面型测量^[40]。

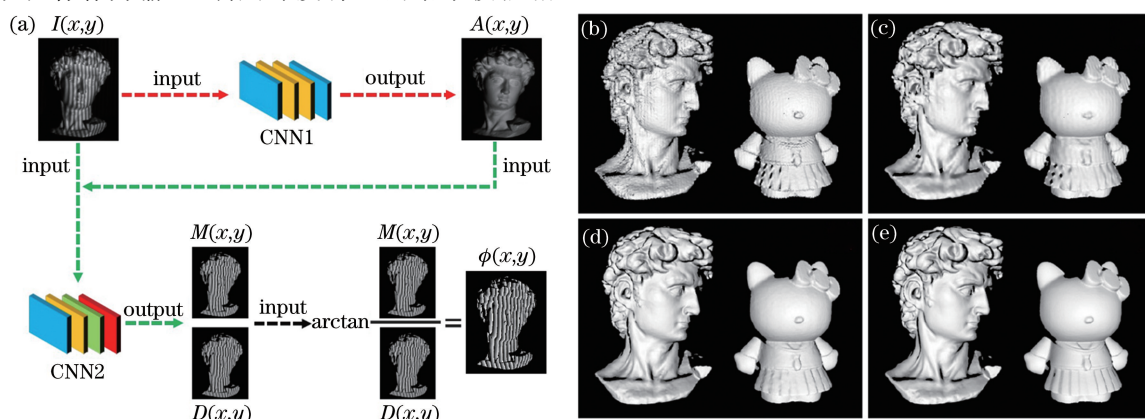


图 5 基于深度学习的条纹分析方法原理与相位重构结果对比^[37]。(a)基于深度学习的条纹分析方法原理图; (b)傅里叶变换法重构结果;(c)加窗傅里叶变换法重构结果;(d)深度学习法重构结果;(e) 12 步相移法重构结果

Fig. 5 Principle of fringe analysis method based on deep learning and comparison of phase reconstruction results^[37]. (a) Principle of fringe analysis method based on deep learning; (b) reconstruction result of FT; (c) reconstruction result of WFT; (d) reconstruction result of proposed deep-learning method; (e) reconstruction result of 12-step phase-shifting profilometry

实例 4——单像素成像(计算“鬼”成像/“对偶摄影”)

在很多重要的成像领域,如极弱光成像、远红外成像、深紫外成像等,很难制造出具有高空间采样率并且成本低廉的阵列探测器。而制造满足相同技术指标的单像素探测器,尤其是制作非可见光波段的单像素探测器,要容易且成本低廉得多。因此,单像素成像成为了简化成像系统、降低成本的一个良好选择。单像素成像最早起源于双光子纠缠鬼成像^[56],利用纠缠态光子对的空间信息相关性来探测目标物体的空间信息,从而实现物体图像的重建,随后该项技术被拓展到了热光源^[57]与赝热光源^[58]。而“计算鬼成像”技术通过空间光调制器(SLM)产生随机散斑以模拟光子的随机性,因此无需再使用面阵探测器来探测散斑图案,即仅需要使用一个单像素探测器(作为唯一的探测器),就可以实现真正意义上的“单像素成像”技术。2005年,Sen等^[59]提出了“对偶摄影”(dual photography),利用投影仪与摄像机的互换性实现了一系列新奇的成像功能,如场景渲染与绕墙成像。或许很多光学成像领域的研究人员现在还不知道:当下十分热门的单像素成像技术^[60-61](起源于2006年)与计算鬼

成像技术^[62](起源于2012年)其实就是对偶摄影的一种特殊形式。不论是单像素成像、计算“鬼”成像还是“对偶摄影”,都需要对场景进行多次图案投影并用单像素探测器收集散射光场,故往往需要上万次甚至数十万次的原始数据测量,十分繁琐耗时,难以实现动态成像。压缩感知技术^[61]利用“先压缩,后采样”的压缩感知思想来得到物体信息在空间域的欠采样数据,并以稀疏性先验为约束,以较少的测量数据重建出物体的图像(实验中仅使用1500次测量,就重建出像素数为原图像像素数的2%的图像)。该方法虽然大大降低了采样数据量,运算却极其复杂耗时,并且很难准确恢复出图像细节部分。针对此问题,Lyu等^[14]于2017年首次提出了一种基于深度学习的单像素技术(如图6所示)。该方法利用空间光调制器显示图像数据集,并获取了相机拍摄得到的原始图像。之后以空间光调制器上显示的图像作为训练目标,分别在以原始图像和传统鬼成像法得到的相位图作为输入的情况下对神经网络进行训练,从而在5%的信息采样率下获得了显著超过传统压缩感知鬼成像的重构结果,并且该方法在噪声鲁棒性方面也得到了明显提升。

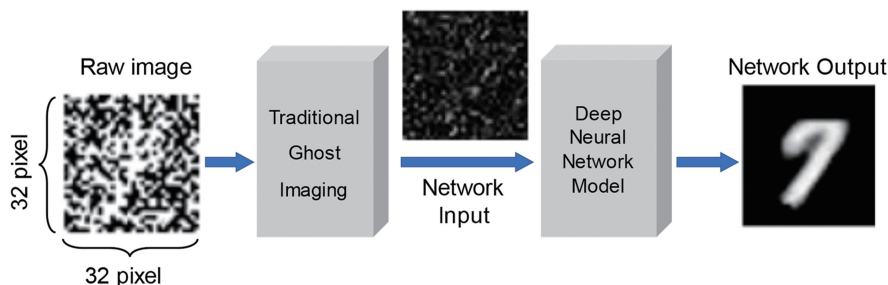


图6 使用深度神经网络的单像素技术框架^[14]

Fig. 6 Framework of single-pixel technique using deep neural network^[14]

2.2 降低传统计算成像技术对“正向物理模型”或“逆向重构算法”的过度依赖

正如前文所言,计算成像的两大核心内容之一正是如何准确地建立场景与图像之间的正向模型,并精确地设计成像系统,使得采样数据能够包含全部的场景信息。这就要求所涉及的正向物理模型能够真实地反映真实世界成像的物理过程。然而,由于实际成像过程的复杂性与不可预知性,精确的正向物理模型通常难以获得,从而就会使得我们使用计算成像技术获得的图像信息与真实情况相差甚远。而深度学习则利用复杂的网络结构与非线性激活函数,直接建立了从原始图像到待恢复信息的“伪

正向模型”,打破了传统计算成像技术中“正向物理模型精确可知”这一限制性条件。作为一个强大的高维特征提取工具,神经网络是直接从大量样本数据中学习输入与输出之间的复杂高维关联,难以预知的不确定因素(如噪声、像差等)也自然而然地被纳入其中,最终获得足以适配足够复杂的真实成像系统的非线性变换能力。

除了如何设计精确可知的正向物理模型外,计算成像技术另一大核心内容便是如何利用逆问题优化求解来重构图像。这要求在重构过程中保证重构图像的准确度。然而,传统的数值优化算法与物理模型中大量的“中间参数”往往需要相关从业人员进

行手动调节选取,最终的图像重构质量很大程度上依赖于参数的人为选取,难以实现无人工干预下的全自动处理。而深度学习技术具有“端到端”的特殊映射机制及高维特征“自行提取”的特点,模型一旦训练完成后就没有任何自由参数需要调节,从而可完全实现“无参数”与“全自动”。

实例 1——(穿透)散射介质成像

光在均匀介质中是沿直线传播的。然而当经过浑浊媒介、生物组织等介质时,光会在这些介质内发生多重散射,出射后的光场将变为散斑场。这是一种不可逆的扩散过程,严重影响了目标的可见性。在传统计算光学成像范畴,实现(穿透)散射介质成像的方法包括反馈波前调制^[63]、传输矩阵^[64]、相位共轭技术^[65-66]、散斑相关^[67-68]等。然而这些方法的有效性往往被限制在光学记忆效应区,即介质的散射作用不可过强并在一定入射角内可以被视为一个线性不变系统。而复杂强散射介质形成的散斑空间分布是散射体微观排列和入射光场波前的复杂函数,难以对其建立全面精确的物理模型并给出简单直接的逆散射解决方案,且算法对于不同类型的散

射介质的可迁移性较差。

深度学习技术为解决这些问题提供了很多新的思路。Li 等^[28]提出了一种具有统计特征的“一对多”深度学习技术(如图 7),该技术从大型数据集中识别出隐藏的统计不变性,其封装了多个微观结构不同的散射介质系统的一系列统计变化,使神经网络模型能够适应散斑的去相关(decorrelation)。卷积神经网络能够学习在具有相同宏观参数的散射体上捕获的散斑强度图案中包含的统计信息(如图 7 所示)。经过训练后,该网络能够迁移至未经训练的散射介质环境中进行成像,且可对不同类型物体生成高质量的目标预测,在数据类型、系统结构等方面表现出了良好的泛化性。Lü 等^[69]构建了混合神经网络(HNN)模型,在强散射情形下实现了隐藏物体的恢复。实验中使用的散射介质是 3 mm 厚的白色聚苯乙烯平板,其记忆效应范围小于 0.01° ,光学厚度为 13.4,远超记忆效应区。说明了基于深度学习的散射成像方法可以不受“正向物理模型精确可知”的约束,突破传统技术中依赖的记忆效应视场角的限制。

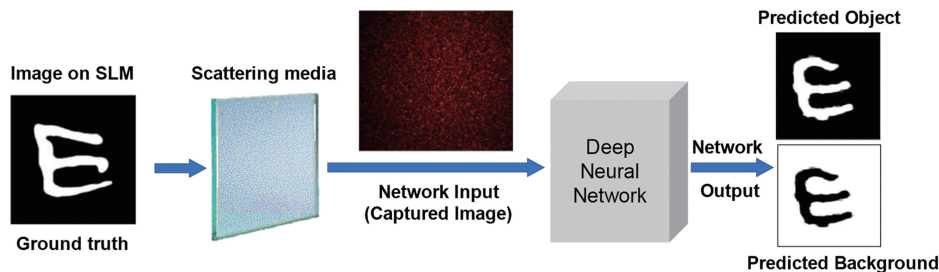


图 7 基于深度学习进行散射介质成像的网络原理图^[28]

Fig. 7 Network of deep learning based imaging through scattering medium^[28]

实例 2——三维衍射层析成像

在讨论相位成像技术时,通常都会假设大部分待测物体属于二维(薄)物体,可以将其表示为由吸收与相位构成的二维的复透射率分布,透射光场的复振幅分布即为入射光场复振幅与物体复透射率的乘积。然而,相位延迟其实是样品三维折射率在一个二维平面上的投影(俗称 2.5D 成像),这是一个沿光传播方向的从入射表面到出射表面各个平面的相位延迟的积分变化量,并不是“真三维”的立体信息^[70-73]。三维衍射层析技术^[70,72,74-75]可以有效解决这一问题,其可以对三维样品内部各点的折射率实现全方位(横向+轴向)高分辨率成像,从而获取样品三维折射率分布。该项技术通常需将相位测量技术(数字全息或相位恢复技术)与计算机断层扫描技术相结合,通过旋转物体^[76-77]或改变照明方向^[78-81]

等方式得到多组定量相位信息,然后结合反投影滤波^[82]、逆 Radon 变换(忽略衍射效应)^[78,83]或是 Wolf 的衍射层析理论(考虑衍射效应)^[70,79-80],重建出物体的空间三维折射率分布。图 8 给出基于深度学习进行三维衍射层析重建的基本框图。近年来,“强度衍射层析技术”——一种基于非干涉强度测量原理的衍射层析技术逐渐崭露头角。相比于传统光学衍射层析技术,该方法只需要直接拍摄物体不同焦面或者不同照明角度的强度图像,再利用图像重构算法就可以反演出物体的三维折射率分布,这有效避免了传统衍射层析技术干涉测量与光束机械扫描的难题。强度衍射层析成像技术主要分为两类:基于轴向扫描的三维光强传输技术^[84-86]与基于角度扫描的三维傅里叶叠层成像技术^[87-89]。不管是哪种技术都依赖对光与三维物体相互作用最终形成图像

的物理过程的准确数学模型:这通常需要假设样品满足纯/弱/缓变相位近似、Born 或 Rytvo 近似的弱散射近似、多层叠加近似(multi-slice)或非负折射率近似等。然而,不论哪种近似都存在一定的局限性,特别地,在强散射、多次散射、后向散射、大数值孔径

照明情况下的普适性普遍较差,因此目前文献中大多数衍射层析成像的实验结果并不十分理想。另一方面,三维层析成像往往还需要巨量的原始数据,这也对其后续算法的高效重建(速度与存储上)提出了巨大挑战。

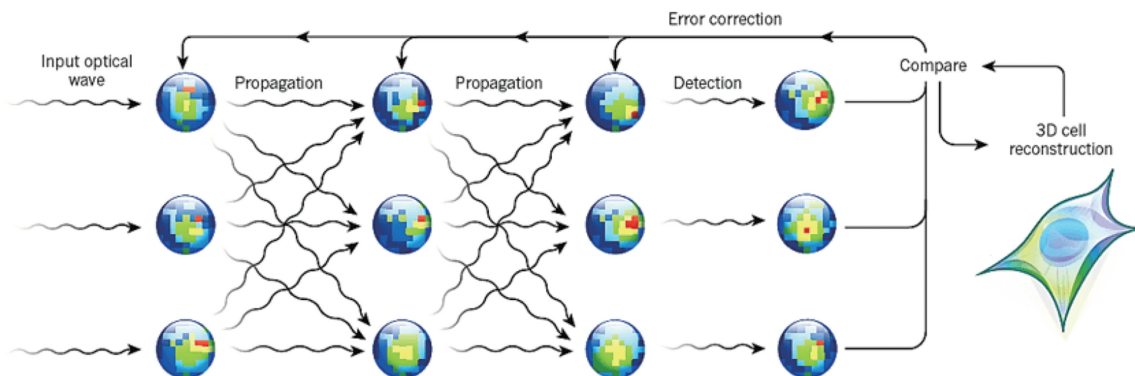


图 8 基于深度学习进行三维衍射层析重建的基本框图^[26]

Fig. 8 Basic framework of 3D diffraction tomography reconstruction based on deep learning^[26]

借助于机器学习算法则有望巧妙地“规避”上述问题。Kamilov 等^[90]将人工神经网络应用于衍射层析成像,他们利用神经网络构建了一个类似于 multi-slice 的三维衍射传播模型,将目标对象按一组薄片切片建模:每个切片由一个网络层表示,三维对象的每个像素由网络节点表示。因为入射光与三维物体的散射过程及其复杂,难以通过理论推导得到完善的数学模型,所以通过神经网络强大的高阶拟合特性去“学习”是一种巧妙且有效的替代方案(注意,由于采用的网络结构较为简单、层数较少,他们的工作严格来说并不算深度学习,但其实核心思想别无二致)。神经网络的训练数据由一组从不同角度捕获的三维物体的二维全息图组成,使用“反向传播”最小化训练数据和模型解之间的差异来预测物体的三维折射率。他们根据实验获得的数据直接通过该方法训练后的神经网络成功恢复出了 HeLa 细胞的三维折射率结构。2018 年,Nguyen 等^[27]利用深度学习技术和衍射强度分布直接对样品三维折

射率进行重建。为了获取相应的训练数据集,他们首先仿真了一个具有不同空间折射率分布的物体,并生成了一系列不同角度投影的二维的相位图像。再将二维图像显示在空间光调制器上并通过相机拍摄到了一组原始的强度图,利用这种“仿真与实验结合”的方式完成了数据集的构建。然后将相机拍摄到的一组原始光强信息经过逆 Radon 变换后作为网络输入,通过深度学习网络使其直接匹配至样品的三维折射率分布(图 9)。尽管这种处理方式的合理性与最终实验结果的准确性仍然有待商榷,但这的确为三维衍射层析数据重建提供了一种新颖的思路。

实例 3——数字全息成像自动聚焦

在数字全息图的重建过程中,具体样品位置具有不可预知性,一种通常的做法是分步处理并循环搜索:如首先对全息图进行预处理去除噪声提升信噪比与对比度(可选),然后通过相位求解算法获取衍射场的相位分布,再经过衍射计算(数值传播)实现不同焦面图像的重建,最后经过经典的清晰度判据(如梯度、

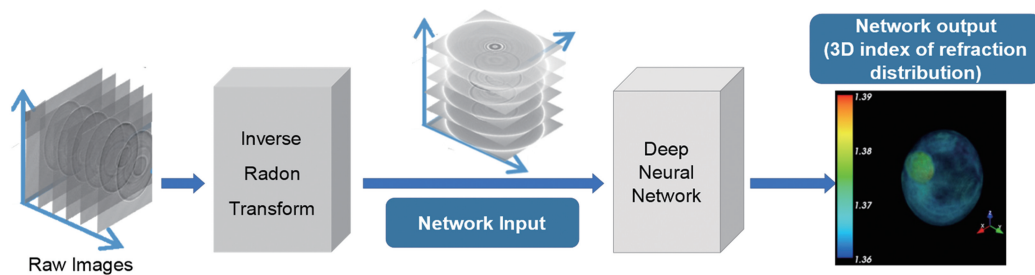


图 9 基于深度学习进行光学衍射层析的网络原理图^[27]

Fig. 9 Schematic of network of optical diffraction tomography based on deep learning^[27]

图像熵等)判断当前面是否是真实物体所在的平面,对上述过程反复进行迭代搜索以获取焦面的最优估计。然而这一过程不仅繁琐费事,且依据传统聚焦判断得到的结果还有可能不够准确。为了简化这一复杂过程,2018年Ren等^[6]提出利用深度学习网络对数字全息图的离焦距离进行预测。该方法主要思想在于利用高精度位移台控制物体位置以在多个不同离焦距离下拍摄到相应的数字全息图,从而构建了数字全息图与其相应离焦距离的数据集,并利用深度神经网络进行训练,直接建立离焦距离和衍射图之间的对应关系,输出参数只有一个数字,即对应了样品的离焦距离。如图10所示,经过训练的深度神经网络无需迭代搜索,可以直接根据输入的数字全息图输出相应的离焦距离。更进一步,Zhang等^[9]利用深度学习技术从离焦的离轴干涉全息图中直接恢复得到聚焦状态下的相位与振幅,极大简化了传统数字全息技术重建过程中对于物理模型参数的调整及获取过程。

实例4——光学相干层析成像图像分割

光学相干层析成像因可以获得微米级分辨率的人体组织三维截面图像,被广泛应用于医学与工业

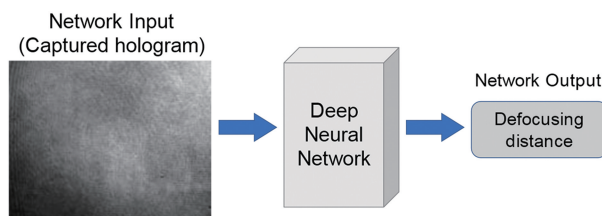


图10 使用深度神经网络的数字全息离焦距离计算框架^[6]
Fig. 10 Framework of defocusing distance calculation in digital holography based on deep neural network^[6]

成像领域中。在许多视网膜疾病的研究中,光学相干层析成像的图像信息的准确量化(如视网膜图像的边界分割)对于提高病灶识别及致病过程等因素的分析至关重要。然而,光学相干断层图像中视网膜的边界分割往往依赖于医生的经验,难以实现全自动处理。Fang等^[23]提出了一种结合卷积神经网络和图形搜索方法的视网膜光学相干断层图像边界自动分割框架,原理如图11所示。得益于神经网络对特殊视网膜层特征的准确提取,该方法可对九层视网膜边界进行准确分割,有效避免了人工分割时的主观性和时间成本。

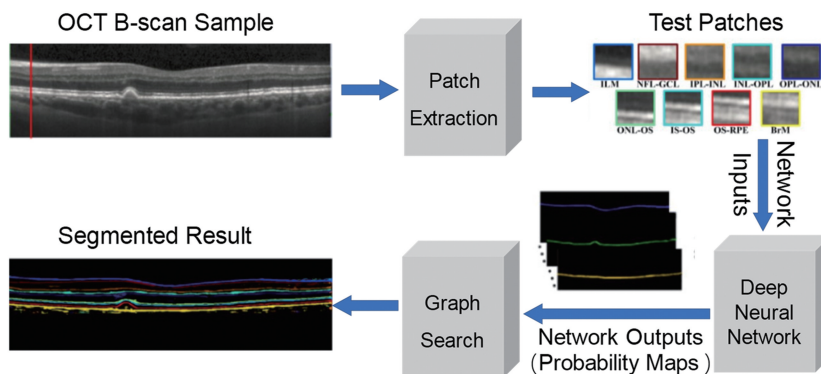


图11 针对视网膜光学相干断层图像的边界自动分割原理图^[23]

Fig. 11 Schematic of automatic boundary segmentation framework for retinal OCT image^[23]

2.3 突破传统计算成像技术所能够达到的功能/性能疆界

计算成像技术通过将光学系统的信息获取能力与计算机的信息处理能力相结合,突破了传统光学成像系统对于成像器件硬件的过度依赖。然而“物理模型驱动”的计算成像技术不可避免地受到物理模型的制约,这些制约既包括光学系统与成像条件的制约,如衍射极限、空间带宽积、成像光子数、器件灵敏度等“先天不足”,也包括重构物体信息与光信息为载体之间的关联性制约,如相位、光谱、三维等与所获取的强度信息间需要具有“显式”关联。由于信息不会凭空产生,因此当所获取的图像数据因受

这些制约而无法直接与目标信息相关联时,传统计算成像技术是无能为力的。而在深度学习技术中,由于最终的输出不仅仅取决于输入数据,还和神经网络从大量训练数据中学习到的成百上千万个权值参数紧密相关,这些参数不仅为图像重建提供了强大而完备的“先验数据库”,还建立了输入与输出间难以通过简单公式表示的“隐性高维”关联,从而有望实现数据“无中生有”与“点石成金”,为突破传统计算成像技术的物理模型限制和拓宽其功能/性能疆界提供了可能。

实例1——超分辨率成像(突破衍射极限)

由于成像系统具有有限孔径效应,一个理想

物点发出的光在图像平面并不会形成一个理想的几何点,而是会形成一个弥散斑(艾里斑)。对于非相干的衍射受限系统而言,艾里斑的半径为 $1.22\lambda/NA$,其中 λ 为成像光波的波长, NA 为成像系统的数值孔径,这被称为“阿贝衍射极限”^[91]。为了提升光学显微镜的分辨率,往往需要采用高数值孔径的油浸物镜,使用起来非常不便。为了解决这一问题,Rivenson等^[20]于2017年提出了基于深度学习的显微成像超分辨算法。该方法首先利用40倍物镜(干)和100倍物镜(油)下拍摄的同一样品图像生成数据集,而后利用深度神经网络

学习低分辨率图像到高分辨率图像之间的“映射”关系。从傅里叶光学理论上,这种映射关系是没有科学依据的。因为低分辨率图像在频域是没有高频细节信息的,不论如何处理,数据也不会“无中生有”。但深度学习的能力的确让人惊叹,经过大量样本训练后的神经网络的确只需输入一幅低分辨率图像(图12左图)即可成功地突破像素分辨率及衍射极限的限制,生成相应的高倍物镜下的高分辨图像(图12右图)。即利用40倍干镜获得了100倍油镜的成像效果,省去复杂油浸物镜观测时的诸多不便。

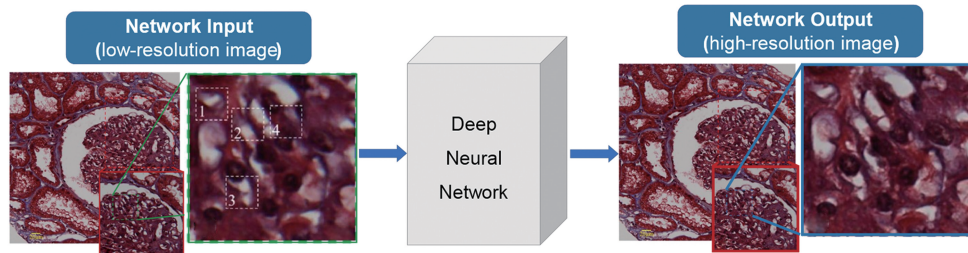


图12 基于深度学习进行超分辨率成像的网络框架示意图^[20]

Fig. 12 Network framework of super-resolution imaging based on deep learning^[20]

高数值孔径油浸物镜虽然能够提升成像分辨率,但受阿贝衍射极限所限,最终能达到的分辨率也不会超过光波波长的一半。随着人类对微观世界的探索逐步深入,需要观测的微观尺度越来越小,传统光学显微镜的分辨率已无法满足科学研究的需要,人们迫切需要分辨率更高的显微技术。2014年,诺贝尔化学奖的三位得主使用荧光分子和特殊的光物理原理,巧妙地突破了普通光学显微镜无法突破的“阿贝极限”,其开创性的成就使得人们能够窥探纳米世界,这些技术包括受激发射损耗(STED)技术^[92]、光激活定

位显微技术(PALM)^[93]、随机光学重建显微技术(STORM)^[94]等。但这些技术依赖于复杂昂贵的硬件系统,且实际操作和使用起来非常复杂不便,STED技术逐点扫描的成像机理使其对环境扰动非常敏感,难以实现动态成像,PALM和STORM还依赖于特殊荧光分子标记,成像过程需要成千上万次的图像采集,复杂耗时。针对这一问题,Wang等^[17]直接通过深度神经网络实现了传统聚焦显微镜图像的超分辨,不借助于任何额外的物理硬件获得了与STED技术相当的成像分辨率,结果如图13所示。

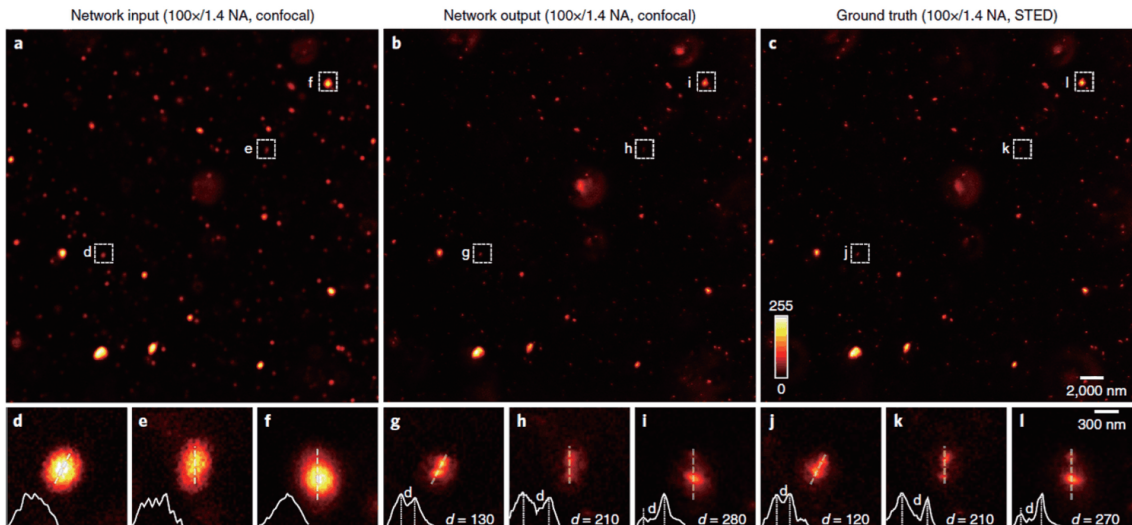


图13 基于深度学习进行STED超分辨率成像的实验结果^[17]

Fig. 13 Experimental results of STED super-resolution imaging based on deep learning^[17]

实例 2——高灵敏度(低照度、低光子数)成像

随着电子设备的快速发展和专业摄像器材应用的普及,人们能拍摄到质量越来越高的图片。但是在实际摄影过程中,总会存在各种不可控因素,致使获得的图片存在各种缺陷。尤其是拍摄环境较暗或光照条件不足等会导致图像信噪比较差,随后的转换、传输、存储等操作进一步降低了低照度图像的质量。然而,低照度图像被广泛存在于诸如地质勘测、水下探测与生物医学等领域中,如何提高低照度条件下的成像质量就成为了当前一大研究热点。当前,主要从两方面着手:提升探测器元件的响应灵敏度;对探测器获得的信号进行图像增强。然而,目前光探测器响应灵敏度指标正逐渐逼近物理极限(已实现单光子探测),难以进一步满足低照度成像的要求。而简单的低照度图像增强处理算法(如直方图均衡化等)仅能简单提升视觉效果,难以应用于图像保真度要求较高的医疗或科研领域。2018年,Chen等^[34]将深度学习技术应用于极弱光成像领域。其主要思想为在使用相同相机的情况下,先在极低照度下拍摄到短曝光(约 1/30 s)图像,而后将在长曝光时间下拍摄到的图像作为深度神经网络的匹配目标。经过训练的深度学习网络可以在照度低于 0.1 lx 的情况下,仅根据一幅极弱光条件下拍摄到的短曝光时间(约 1/30 s)图像[如图 14(a)所示]恢复得到一张细节清晰的正常图像[如图 14(c)所示],相比于图 14(b)所示的使用高感光灵敏度 CCD

拍摄到的图像,深度学习所得结果无论是色彩、细节还是阴影中的背景均得到了更好的还原。

实例 3——跨模态成像

对组织标本进行显微成像观察是对临床上多种疾病进行诊断的基本工具,也是组织病理学与生物科学的必备工具。通过临床手段获得组织切片的标准染色图像通常需要一系列复杂工序:福尔马林固定和石蜡包埋(FFPE)、切片(通常为 2~10 μm)、标记染色、风干封片等多个步骤,整个过程极其繁琐耗时。为了简化切片染色流程、降低染色成本,Rivenson等^[35-36]于 2018 年利用深度学习技术对虚拟组织染色技术进行了研究。该方法的主要思想在于利用组织切片染色前后的图像构建训练数据集,利用对抗神经网络学习未染色切片与染色切片之间的映射关系。虽然,从现有认知范畴而言,这种映射关系是没有科学依据的,因为未染色样品本身不具备化学染料所存在的生化反应过程,更不会具备组织样本各组分特异性。但深度学习却似乎能够从大量测试样本数据中发现这些看似无关的数据集的隐性复杂关联,而这种关联是无法利用我们现有知识体系来建立甚至理解的。实验发现,经过训练的神经网络可以根据一张未染色的切片图像的“自发荧光图像”或者“定量相位图像”直接生成其染色后的结果(如图 15 所示),使得组织学切片分析“绕开”了切片染色这一复杂繁琐的过程,有望为“即时病理诊断”打开一扇新的大门。

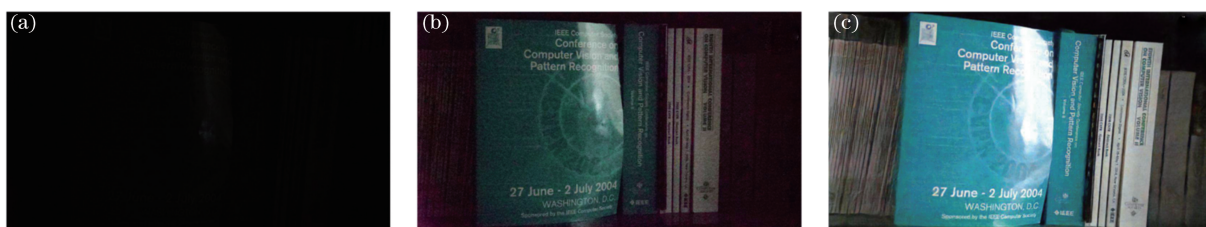


图 14 基于深度学习进行极弱光成像的结果^[34]。(a)摄像机输出(ISO 8000);(b)摄像机输出(ISO 409600);(c)由原始数据(a)恢复得到的结果

Fig. 14 Results of imaging using very weak light based on deep learning^[34]. (a) Camera output with ISO 8000; (b) Camera output with ISO 409600; (c) recovered result from raw data of Fig. 14(a)

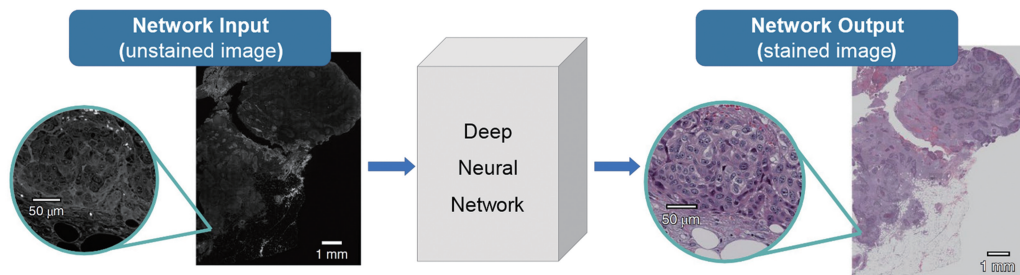


图 15 基于深度学习进行虚拟染色成像的网络框架示意图^[35]

Fig. 15 Network framework of virtual staining imaging based on deep learning^[35]

3 深度学习下的计算成像:挑战

第二章从提升信息获取能力、降低对“正向物理模型”与“逆向重构算法”的依赖,以及突破传统计算成像技术的性能极限这三个方面简述了深度学习下计算成像的研究现状与最新进展。由于篇幅限制,仅列举了一些该领域内的代表性案例,但从这些典型案例中已经可以看出,深度学习下的计算成像领域取得了许多令人振奋的进展,这些定将为计算成像领域的进一步发展注入新的活力。本章中,将把注意力进一步向前转移,探讨深度学习在计算成像领域所面临的诸多挑战,这将有助于为我们未来的下一步研究制定大胆的战略构想。

3.1 (实测)训练数据的获取与标注成本高

近年来,深度学习技术的发展及其相关应用呈现爆炸式增长。值得注意的是,神经网络的想法并非最近几年才提出,它已有几十年的历史。但是直到最近它才受到如此广泛的关注。这其中重要的因素在于神经网络是一种以数据为导向的算法,卓越的网络输出取决于大量/巨量的数据训练。正是近年来互联网时代积累的大量数据和算力,释放了深度学习神经网络的潜力,也为人工智能应用的飞跃式发展打下了坚实基础。但就目前而言,大部分在计算成像领域中取得成功应用的深度学习技术的案例中,往往需要借助实际成像系统,通过实验获得大量的实测训练数据来进行标注(获取所期待的真值)。这里就存在两个关键问题:

1) 耗时费力:大部分光学成像实验的数据采集过程都是极其复杂且耗时的,而深度学习所依赖的大规模训练数据的获取与正确标注无疑让这一问题雪上加霜,因此需要耗费大量的人力和物力成本。加之光学成像领域的公开数据集稀少,这又增加了深度学习计算成像技术的实施难度。虽然有一些方法可以减少其对数据的依赖,比如迁移学习、少样本学习、无监督学习和弱监督学习,但是到目前为止,它们的性能还没法与大样本的监督学习相比。

2) 真值难知:数据采集后无法获得目标准确可靠的理想真实值,是数据集构建的另一大限制性因素。例如,进行物体相位信息的获取时,往往需要使用传统算法进行信息提取。然而这不仅制约了神经网络的表现,即网络的输出质量很难超越用于训练的数据,还使得多种类、大数据量的信息获取十分困难、繁琐。因此,在进行深度网络训练时,如何获取真实、有效并且具有良好代表性的数据集仍是

计算成像领域面临的一大问题。

值得注意的是,对于某些正向物理模型精确可知的光学成像应用而言,一种可能的做法是直接通过物理建模并在计算机中进行仿真来获取网络训练所需的大量训练数据。这种方式虽然规避了上述两个问题,但也丧失了深度学习技术的最大优势:深度学习能够从大量实验样本数据中学习实际成像系统输入与输出之间的复杂高维关联,获得足以适配足够复杂的真实成像系统的非线性变换能力。而计算机仿真是无法准确还原实际系统中难以预知的不确定因素(如噪声、像差等),所以所学习到的模型也不一定能够真实全面地反映实际的物理成像过程,从而将有可能得不到理想中的成像结果。

3.2 目标合作度与环境稳定性要求高

在计算成像领域,为了能够获得足以适配足够复杂的真实成像系统的映射能力,深度学习训练数据往往通过实际的成像系统获取。这往往需要复杂耗时的数据采集过程,不仅使得训练集构建繁杂耗时,更为深度学习在一些特定领域的实际应用,如穿透散射介质成像(军事、生物成像等),带来了一定的困难。这主要是由以下两方面问题导致的:

1) 目标合作度要求高:传统的监督学习方法往往需要由大量具备不同特征目标的样本组成的数据集进行训练。然而方法本身隐含着目标侧灵活可控的假设,只有满足这一假设才能稳定地获得对应样本的真值信息。然而,在某些实际情况下,满足这一条件是十分困难的。例如,穿透散射介质成像中通常需获得大量已知标准样本的散斑场图像,这就需要在物体侧频繁替换目标物。而在实际军事(如穿透雾霾、遮障等)成像环境中,很难直接对目标侧进行自由操控。生物医学成像(深层穿透成像)应用也存在类似的问题,在人体皮下或者脏器内置放(大量)合作目标通常也是不切实际的。

2) 环境稳定性要求高:在构建深度神经网络训练所需的数据集时,往往需要假设成像系统的物理模型是趋于稳定甚至不随时间变化的,这样所采集到的大量实验样本数据才能够集中体现成像系统真实复杂的物理成像过程。然而样本采集过程往往十分繁杂耗时(可能长达数小时甚至数天),这就必须保证数据采集过程中的系统环境要尽可能地保持稳定一致。然而,在许多实际应用中,该条件往往难以满足。仍以穿透散射介质成像为例:雾霾、遮障、大气湍流、水下浑浊等外界环境往往都会随时间而改变,从而使得获取的数据难以真实有效地反映实际

成像时散射介质对成像系统的影响。生物医学成像应用中也存在类似的问题:人体的呼吸、血流及新陈代谢活动都是永不停息的,这也意味着散射介质并不会在较长的时间内保持一成不变,这种系统与环境的稳定性给深度学习技术的实际应用带来了相当大的挑战。

3.3 网络结构的选取趋于经验主义

针对特定的成像需求,到底选择什么样结构的神经网络合适?这是初次尝试深度学习的研究人员经常面对的一个问题。尽管从前人的相同或者相似工作中能找到网络结构设计的灵感,但是在神经网络后期的调试与优化过程中,如何调整超参数(Hyper Parameters,如神经网络的层数、CNN中滤波器的大小、特征的数量等)使得该网络能够在既定的应用中表现出色仍是一个难以回答的问题。通过试错法进行超参的调整尽管有一定效果,但这一手段依赖于从业者本身对深度学习调参的理解,且试错过程中时间成本过高。

此外,深度神经网络的规模同样也是研究人员在进行网络设计时需要考量的因素。随着神经网络层数的增多,其非线性拟合能力也就越强,往往训练结果的精度也会得到提升,但是当网络层数、参数数量达到一定规模时,不仅训练过程会变得复杂,还会给网络能否快速输出结果提出了硬件性能上的挑战。快速输出神经网络的运算对于算力强大的服务器与工作站而言可能算不上太大的负担,但对于移动终端或穿戴设备(如手机、平板等)而言,其往往难以承担规模过大的神经网络的部署,这时需要在设计阶段考虑对网络结构、尺寸进行合理限制,合理权衡成像性能与运算资源。

3.4 “调参好比炼丹”式的试错法训练机制

对于尝试使用深度学习技术的许多人来说,深度学习方法预测最终结果的过程往往是难以理解的。虽然深度神经网络所基于的基本运算过程,如卷积运算、激活函数运算、梯度求解等,十分简单易懂,但是随着深度神经网络规模的扩大与参数的增加,大量乃至巨量的参数迭代过程、梯度优化过程互相耦合,使得整个训练过程难以理解,导致研究人员在很多情况下只能通过最终的测试结果来对神经网络性能的优劣进行判断,使得优化和提升神经网络性能这一目标充斥着大量的试错过程。但是类似“试错式”的研究方法往往意味着大量的计算资源与时间成本。尤其是对于大规模的深度神经网络,巨量的参数使得完成一次训练甚至可能需要占用数十

乃至上百块高性能TPU、耗费数个昼夜甚至更久的时间,这种多次且无明确目标方向的试错极易带来时间与算力的大量浪费。

近年来,越来越多的研究人员也逐步意识到这个问题的重要性与严重性。为了解释神经网络的学习过程,Zeiler等^[95]提出了一种针对卷积神经网络的可视化方法。该方法对神经网络学习的特征进行了可视化,为优化网络结构、提升预测的准确性提供了思路。而Shwartz-Ziv等^[96]于2017年也尝试使用“信息瓶颈理论”解释深度学习的训练过程,其发现了深度学习训练过程存在的“特征拟合”和“特征压缩”两个阶段,并进行了相应的可视化分析。

3.5 特定样本训练后的网络缺乏泛化能力

泛化能力评价的是一个神经网络完成训练后,在处理“从未遇见过”的输入数据时的表现。虽然传统计算成像技术的实际成像性能受限于“正向数学模型的准确性”以及“逆向重构算法的可靠性”,但只要“模型全面准确”且“算法稳定可靠”,对于不同的观测对象都可获得较为理想的成像结果。但对于基于数据驱动的深度计算成像,神经网络输入与输出之间关系的建立主要依赖于对大量样本数据的反复训练过程。显而易见,对于训练数据中常见的图像特征,神经网络更容易学习到从该特征到输出结果的映射。而对于训练过程中出现较少的图像特征或者实际成像中遇到的一个区别于训练集的全新样本,神经网络一般难以给出正确的输出。因此泛化能力通常与训练数据样本的规模和多样性密切相关。而正如前文所说,在计算光学成像领域,深度学习技术所依赖的数据集通常获取起来较为困难,仿真得到的数据集与真实的成像过程总是存在偏差,而多类样本的大量实验数据又难以获取,这一困难成为了制约深度神经网络泛化能力的一大因素。

值得注意的是,对于神经网络泛化能力本身需要一分为二来看待。这好比是关于“通才”与“专才”的思考。通才的知识面广,但深度较为欠缺。而专才尽管知识面相对较窄,但能精通一到两项专业特长。对于社会的发展而言,通才和专才都是不可或缺的。回到光学成像的范畴,由于不少应用面向的对象本身就较为单一,因此不断增加同类型的训练数据,可对面向特定应用的成像系统性能提升起到积极的作用。因此,从实际成像需求出发,辩证地看待深度学习辅助下光学成像方法的泛化能力,在某些场合缺乏泛化能力反倒不是一件坏事。

3.6 “深度学习下的计算机视觉”≠“深度学习下的计算成像”

在过去的7年中,深度学习技术以其特有的“高维特征自动提取”的功能避免了传统机器学习极其依赖的人工“特征工程”选取工作,迅速成为了机器学习领域的主流。而在这其中,最为令人瞩目的便是其在计算机视觉领域的重大进展:目标识别、三维视频渲染、图像去模糊、图像超分辨等技术都因为深度学习技术的出现彻底改头换面。然而,当关注到其在计算机视觉领域的成功并为之欣喜时,也同样要认识到其潜在的缺陷:深度学习并不是魔术,数据也并不会“无中生有”,由大量训练数据中学习“先验数据”并不能与某个待测样品所获取“真实信息”画上等号。对于“深度学习所获得的结果是真实有效且准确可靠的”,相信目前谁也没有办法拍胸脯保证,至少现在还无法做到。

可能有人会提出质疑,这真的那么重要吗?不管白猫黑猫,能抓到老鼠的就是好猫。的确在某些领域,这并不算是个问题。计算机视觉所面向的很大一部分应用就是满足人类观测的需要(如消费电子、影视娱乐),人眼是最终的受体也是评价者,“看起来”好看、“看起来”清楚、“看起来”真实足矣(一个更为恰当的英文单词叫 photorealistic)! 而“它归根到底是不是真的”,其实(似乎)并不是那么重要。例如,深度学习技术可基于单张图像实现图像(像素)超分辨率重建,这种数据的无中生有显然是违背经典的信息论的。换言之,无法保证经过深度神经网络所“长出来的”图像细节与真实高分辨率场景中的完全一致。但是“who cares”? 只要知道图像的确变清楚了,马赛克的确消失不见了,这就足矣!

然而,在光学范畴的计算成像领域,上述深度学习技术的潜在缺陷或许是“致命的”。不仅仅是为了满足视觉观测的需要,计算成像技术往往还和工业测量、医疗诊断、科学发现等领域密不可分,这就意味着不仅需要得到一个“看起来还不错”的结果,更需要确保它们的“准确、可靠、可重复、可溯源”。而这些要素都很不幸的是深度学习技术的“软肋”。“如果我都不能保证结果是真的,我还要它有什么意义?”这要求看似有些苛刻,但在某些特定领域的确也是必须的。因为谁都不希望自己的产品在质检阶段被深度学习算法的“存在某个瑕疵”而打上不合格的标签,更不希望自己的体检报告中由于深度学习算法的“长出了某个病灶”而被诊断为得了不治之症。更进一步说,深度学习技术的成功所依赖的是

从大量训练样本中学习并提取的“共性”信息(特征),这恰恰导致其在面对“罕见样本”(与训练数据集差异较显著)时,所得结果的准确性往往并不理想。而对这类“罕见样本”的正确检测与可靠识别往往也是工业检测与医学诊断领域最有意义且最具挑战性的一部分。毫不夸张地说,“异常”更是一切科学新发现的起源。因此,在当前深度学习技术所获得的巨大成功面前,还是应当对其在计算成像领域的应用保持清醒且理性的态度。

3.7 “深度学习”缺乏“深入理解”的能力

目前,深度学习技术仍然在很大程度上依赖大量数据进行特征信息提取,换言之,当深度神经网络面对一项截然不同的任务时,需要使用新的数据进行相应训练。正如图灵奖得主、贝叶斯网络之父 Judea Pearl 所说,当前的深度学习不过只是“曲线拟合”,清华大学的张钹院士也曾指出现代的人工智能基本方法有缺陷,而我们必须走向具有理解能力的 AI,这才是真正的人工智能。需要明确的是,现有的深度学习缺乏理解和推理能力的原因在于它缺乏常识信息。举例来说,利用深度学习来进行条纹分析以计算条纹图中蕴含的包裹相位信息,目前的方法是两步走:先利用深度学习技术求解条纹的正弦和余弦部分,然后将它们带入反正切函数计算包裹相位。由于缺乏推理能力,深度学习技术无法预知包裹相位具有不连续的空间跳变这一常识(而陡变区域的高精度拟合往往对于卷积神经网络是非常具有挑战性的),使得难以训练出准确可靠的端对端(条纹到相位的直接映射)的神经网络。为了改善这一问题,需要建立常识库,将常识信息引入到深度学习,使神经网络在预测时既考虑已看到的样本又与有关真实世界的常识相联系。

4 深度学习下的计算成像:未来

4.1 搭上深度学习技术发展的顺风车

当了解到计算成像系统的性能、功能与成像能力因物理模型(如衍射极限、逆问题模型)等受限时,深度学习技术为计算成像所带来的性能优势就显而易见了。毫无疑问,深度学习技术与深度神经网络模型仍会在接下来的若干年不停地向前发展,而计算光学成像技术也必定将搭着这列顺风车继续快速前行。

4.1.1 对抗学习(GAN)

GAN 是从 Goodfellow 等^[97]的研究工作里演化出来的一个深度学习分支。图灵奖获得者 Yann

LeCun 曾评价对抗学习为“Adversarial training is the coolest thing since sliced bread(对抗训练是自切片面包以来最酷的事情)”。这个由博弈论启发而产生的技术包含两个算法,一个是生成器算法,一个是鉴别器算法,它们的目标是在训练的过程中欺骗对方。当这两者的博弈达到平衡时,模型训练结束。此时利用生成器即可输出最终的结果。

GAN 的优势在于它是一种以半监督方式训练的方法,适合标签较少的训练数据。而且 GAN 模型只用到了反向传播,不需要复杂的马尔可夫链。此外,GAN 不仅可用作图像生成,还可以用于图像分类。从计算机视觉顶会 CVPR 2018 年的论文统计数据来看,以 GAN 为关键词的论文数量占比已接近论文总量的 10%。且纵观 CVPR 2014 至 CVPR 2018,与 GAN 有关的论文数量呈现逐年翻倍的情况。这足以说明 GAN 这项技术正不断地影响着深度学习技术未来的发展。

4.1.2 迁移学习与少样本学习

迁移学习专注于利用已有问题的解决模型求解其他不同但相关问题。比如说,用辨识轿车的模型来提升识别卡车的能力。迁移学习的初衷是节省人工标注样本的时间,让模型可以通过已有的标记数据向未标记数据迁移,从而训练出适用于未标记数据的运算模型。具体来说,迁移学习算法先在一个拥有更大的数据集的任务(源任务)上训练,然后再被迁移为学习另一个只有较少数据集的任务(目标任务)。如果存在一个与目标任务有相关性的任务,且该任务具有丰富的数据,那么可先训练一个针对该任务的模型,然后在我们的目标任务中重用这个模型,或者将这个模型作为我们目标任务模型的训练起始点。这将有利于加速训练的过程,提升神经网络性能。

少样本学习是迁移学习的一个分支,它的产生依赖于人类非常擅长通过少量的样本识别一个新物体。比如,小孩只需要学习几幅图片就能辨别“狗”、“猫”、“牛”等动物。受人类这种快速学习能力的启发,少样本学习在机器学习一定类别的大量数据后,对于新的物体,只需要少量的样本就能迅速完成学习。

4.1.3 自动化机器学习(AutoML)

深度学习算法的性能受许多决策的影响。对于没有丰富计算机技术背景的研究人员,深度学习神经网络的设计总是给他们带来不小的困扰。研究人员需要选择相应的神经网络架构、正则化方法、超参

数等。所有的这些操作对神经网络的性能都有很大的影响。自动机器学习的目标就是使用自动化的方式做出上述的决策。使用者只需提供训练数据,自动机器学习系统就能自动地决定最佳的训练方案。让不同领域的研究人员不必苦恼于学习各种机器学习的算法。

目前,自动化机器学习的实现方式包括:超参数优化(Hyper-parameter Optimization)、元学习(Meta Learning)、神经网络架构搜索(Neural Architecture Search)等。对于超参数优化,常用的方法有网格搜索(Grid Search)、随机搜索(Random Search)和贝叶斯优化。对于元学习,它的主要任务是让机器学习“如何学习”。通过对现有的学习任务之间的性能差异进行系统观测,然后让机器学习已有的经验和元数据,用于更好地执行新的学习任务。从某种意义上来说,元学习的过程蕴含了超参数优化。因为它学习了超参数、流水线构成、神经网络架构、模型构成与元特征等。对于神经网络架构搜索,伴随着深度学习的流行,神经网络的架构变得越来越复杂。利用主观经验来确定合适的神经网络架构难度也越来越大,神经网络架构搜索就是为了解决这个问题。通过定义搜索空间(Search Space)、确定搜索策略(Search Strategy)、性能评价(Performance Estimation Strategy)这三个阶段,机器可根据反馈进行每一轮的架构搜索。自动化机器学习将大幅降低机器学习技术的使用门槛,进一步推动其在光学成像领域中的应用。

4.2 物理模型驱动数据与数据驱动物理模型

4.2.1 物理模型驱动数据

物理模型驱动是当前深度学习发展的一个重要方向,即在深度学习中嵌入或内蕴特征规则先验,代替单一的纯数据驱动式学习。众所周知,深度学习是一种标准的数据驱动型方法,它将深度网络作为黑箱,依赖于大量数据解决现实问题;而模型驱动方法则是从目标、机理、先验出发,首先形成学习的一个代价函数,然后通过极小化代价函数来解决问题。模型驱动方法的最大优点是只要模型足够精确,解的质量可预期甚至能达到最优,而且求解方法是确定的;模型驱动方法的缺陷是在应用中难以精确建模。模型驱动深度学习方法有效结合了模型驱动和数据驱动方法的优势。

2018年,Xu等^[98]提出一种模型驱动与数据驱动相结合的深度学习方法(图16),给出了模型驱动深度学习标准流程:1)根据问题,建立模型族

(Family of Models); 2) 根据模型族, 设计算法族 (Family of Algorithms) 并建立算法族的收敛性理论; 3) 将算法族展开 (unfold) 成深度网络并实施深度学习。这种方法将物理模型、逻辑规则作为先验

引入到深层神经网络中, 利用人类意图和领域知识对神经网络模型进行引导, 包括特征规则约束、网络架构设计等, 可有效提升网络大样本学习效率、小样本/零样本学习能力、数据泛化能力。



图 16 模型驱动的深度学习方法^[98]

Fig. 16 Model-driven deep-learning approach^[98]

4.2.2 数据驱动物理模型

深度学习方法一般来说只能学习到数据集中已有的知识, 比较擅长于归纳, 而不擅长演绎, 也就是说网络难以推演物理定律, 但是可以在一定约束/条件下拟合物理定律。例如, 概率生成模型^[97]可用于自然图像的生成, 训练 1000 万张图片后生成的模型可以自动学习到其内部分布, 能够解释给定的训练图片并同时生成新的图片。与庞大的真实数据相比, 概率生成模型的参数量远远小于数据量, 在训练过程中生成模型会被迫去发现数据背后更为简单的统计规律, 从而能够生成这些数据。2017 年, Lin 等^[99]给出深度学习定性的物理解释(图 17): 1) 基本

的物理学定律都是 2 到 4 阶, 而且拥有对称性等, 这些约束使得解空间变小, 因此 DNN 可以近似得到这个解; 2) 所有物质由简单的基本单元构成, 这个分层结构与 DNN 相似, DNN 的层数越多, 生成的结果越复杂。

然而未知定律必定隐含在数据中, 既然深度学习能够对已知分布/规律进行可靠判别筛选, 那么可以只利用深度学习分析数据并试图找到突出/奇异点, 而不一定要找到特定规则或者新规律。在检测器物理学中, NOvA 中微子实验的研究人员将 CNN 用于粒子识别和分类、粒子轨迹重建、粒子的相互作用分析等^[100]; 在天体物理学方面, CNN 被用来发现

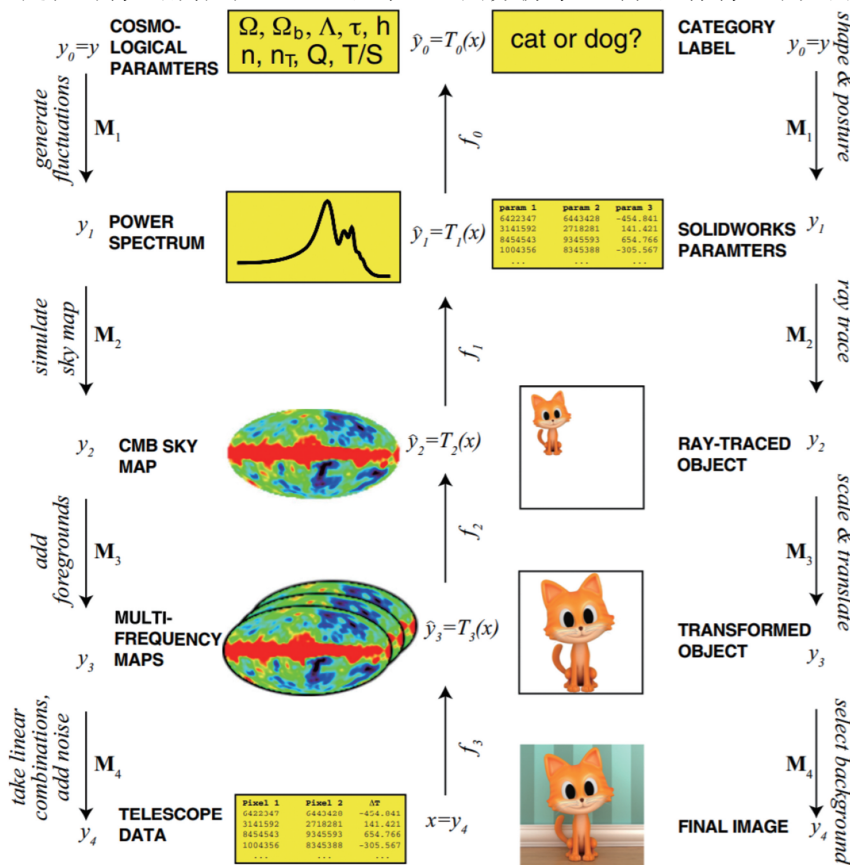


图 17 物理(左)和图像分类(右)关联的因果层次结构^[99]

Fig. 17 Causal hierarchy structure relevant to physics (left) and image classification (right)^[99]

引力透镜(引力透镜是指可以扭曲来自它们后面的遥远星系的光的大型天体),加速对望远镜数据扫描以寻找引力透镜扭曲现象的过程^[101];在机器视觉应用中也提出了诸多新奇检测方法,采用深度学习对数据中新奇的或未观测到的数据进行检测识别^[102-103]。

4.3 深度学习的可解释性有待进一步探究

所谓可解释性是指在我们需要了解或解决一件事情的时候,可以获得所需要的足够的可以理解的信息。反过来说,如果在一些情境中我们无法得到相应的足够的信息,那么这些事情对人们来说都是不可解释的。正如前文所述其依赖于多层简单的线形运算的组合,最终实现了高度非线性化的高维特征提取功能并获得了极高的模型表现能力。但是,虽然人们创造了准确度极高的网络模型,但最后只得到了一堆看起来“毫无意义”的模型参数

与匹配度非常高的判定结果。尽管如此,但模型本身也意味着知识,我们希望知道模型究竟从数据中“学”到了哪些知识从而支撑该模型进行了最终的决策。

当准备将神经网络应用于某些特定成像领域时,除了获得最终的理想的成像结果,还希望能够了解到神经网络究竟提取了原始输入中的哪些信息,是基于什么形式的运算得到了这个结果。除此之外,不可解释同样意味着模型的“危险性”,图 18 所示为一个非常经典的关于对抗样本的例子,对于一个 CNN 模型,在图片中添加了轻微随机噪声之后熊猫却被判定为长臂猿^[104]。因此,如何进一步推动深度学习的数据解释性,如何进一步提高神经网络结果的可溯性、稳定性,都是目前迫切需要解决的问题,这些问题的解决也将为深度学习在计算成像领域打开更广阔的应用空间。

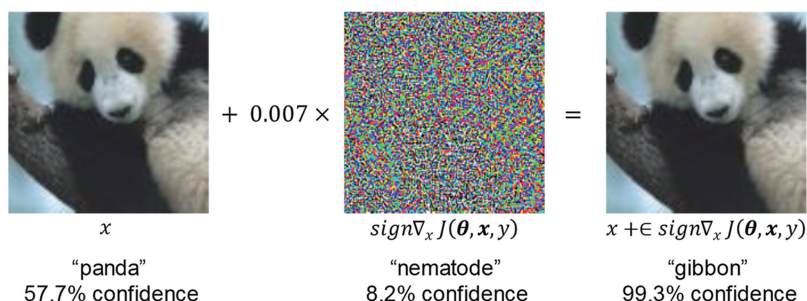


图 18 在熊猫图片中加入轻微随机噪声,CNN 模型将图片识别为长臂猿^[104]

Fig. 18 After adding slight noise into Panda image, CNN model recognizes image as Gibbon^[104]

4.4 脑神经科学启发的思路值得更多的重视

作为多层神经网络,深度学习是受脑神经科学启发而发展起来的。特别是卷积神经网络,其根源为 Fukushima 在 1980 年提出的认知机模型,而该模型的提出动机就是模拟哺乳动物视觉神经系统,通过逐层提取由简及繁的特征,实现语义逐级抽象的视觉神经通路。在诺贝尔奖获得者 Hubel 和 Wiesel 的共同努力下,该通路从 20 世纪 60 年代起逐渐清晰,为 CNN 的诞生提供了良好的参考。但值得注意的是,生物视觉神经通路极其复杂,神经科学家对初级视觉皮层中简单神经细胞的边缘提取功能是清晰的,对通路后面部分越来越复杂的神经细胞的功能也有一些探索,但对更高层级上的超复杂细胞的功能及其作用机制尚不清晰。这意味着 CNN 等深度模型是否真的能够模拟生物视觉通路还是不得而知的。但可以确定的是,生物神经系统的链接极为复杂,不仅仅有自下而上的前馈和同层递归,更有大量的自上而下的反馈,以及来自其他神

经子系统的外部链接,这些都是目前的深度模型尚未建模的。但无论如何,脑神经科学的进步可以为深度模型的发展提供更多的可能性,这是非常值得关注的。例如,最近越来越多的神经科学研究表明,曾一度被认为功能极为特异化的神经细胞其实具有良好的可塑性,例如,视觉皮层的大量神经细胞在失去视觉处理需求后不久即被重塑,转而处理触觉或其他模态的数据。神经系统的这种可塑性使其面向不同的智能处理任务时具有良好的通用性,这为通用人工智能的发展提供了参照。因此我们可以大胆展望,未来很有可能会出现更加智能的可塑化模型来代替现在的固定结构的深度学习模型。

4.5 既要“深度”又要“深入”

深度学习的出现似乎一度改变了计算机视觉与光学成像领域,有计算机视觉领域的研究者曾开玩笑地说道:“深度学习让我感觉到之前一切的所学似乎都白学了。”大家似乎再也不用绞尽脑汁自己去研究背后的数学物理机理、推导模型并进行求解预测。

只要统一地甩给计算机一张网络就可以让它自己去学习具体的模型了,这大大降低了知识学习的成本,减少了手动建模的工作量。理论研究的最终目的不就是建立真实物理世界的数学模型,然后利用模型再去造福世界的吗?现在不需要研究理论也可以建立一个近乎准确的模型了,为什么还要费尽力气深入研究理论?

这看似很有道理,但我们不妨换一个思路去思考一下?一项工作是选择由深度学习去完成还是经典理论算法去完成,归根到底还是在于二者谁能够完成得“更出色”。如果深度学习真的能够“保证”预测出一个近乎准确的模型,输出最为理想的重建结果,我们定能够安心地把这件事交给它而不必再劳

神费力。但谁能保证做到这一点?即使你是这么认为,你是否真的深入了解并实现过那些经典的物理模型驱动方法,并保证能对此结果做出一个不偏不倚的公正评价?你会为让经典算法得到一个最优的结果而通宵调参吗?你是否真的保证你所实现的深度学习算法在训练时没有因数据有意无意的泄露而得到一个过度美化的结果?你会客观而随机地选择测试数据并以此去评价最终的实验结果吗?因此,个人之见是“真正的”深度学习其实并没有(至少现阶段)把科学研究变得更简单,反而拉高了科学研究的门槛,因为他要求研究者不但既要公正且有效地利用“深度”学习这一工具,又要对此领域的研究足够“深入”,以保证真正得到客观而准确的结论(图 19)。

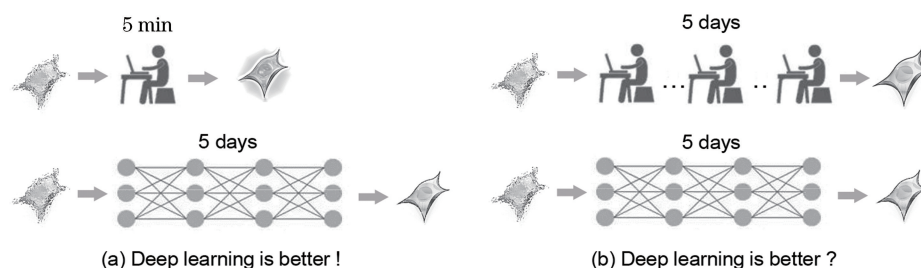


图 19 深度学习与经典理论算法之间的客观公证对比

Fig. 19 Comparison between deep learning and classical theoretical algorithm should be objective

4.6 既要“有所为”又要“有所不为”

现阶段深度学习技术在很多领域(包括计算成像)已经步入了一个爆炸式增长的时代,它似乎成为了该领域研究的未来潮流。一个不能用经典模型解释的现象,只要套上“深度学习”似乎就变得“高大上”起来。这真的让人细思恐极!很多初出茅庐的研究者们似乎发现其实不用扎实的基础理论也可以发表一篇高档次的论文,他们当然会很快地成为该领域的拥护者,并幻想着似乎自己已经走在了该领域研究的最前沿。但熟不知,这大量形形色色的披着深度学习华丽外衣的研究论文背后只不过是一套“模板化”的菜谱,自己学着去做其实未必能做出真正漂亮且有营养的一道美食。

深度学习终究不过是一种基于大量样本数据的统计类方法,统计类方法在推理性的任务面前是不可靠的。在物理学中,想要建立一个模型通常需要三个步骤:1)学习数学物理理论;2)观测真实实验数据;3)基于数学物理理论对观测数据进行建模。看似最重要的建模只是三道工序中的最后一环,实际上大多数困难费事且重要的工作都花在前两步(这也就是为什么大部分人不喜欢花了那么多年时间才读到博士,即使读到了博士还要继续花大量时间读

文章做实验)。深度学习的优势是省去了第一步,简化了第二步,直接到了第三步,这听上去似乎有些难以置信。因为神经网络的参数选择没有理论基础,仅用数据驱动。显然只通过有限(局部的)数据得到的符合这些数据的函数可能是不唯一的,但真实的函数却只有一个,这就注定了仅依赖局部数据驱动预测的模型类似于一种赌博。诚然在计算成像领域某些很难公式化建模的任务上,神经网络取得了巨大的成功,但想要使用神经网络从大量衍射图样中学出菲涅耳衍射定律这样的通用的数学公式却不太现实。

最后值得一提的是,如果你真的对深度学习那么情有独钟,非深度学习不可,那请你在最后的时刻冷静下来思考一下这个问题:“这到底值不值得、不适合用深度学习去做?”深度学习切记不可乱用。在某些问题前面,传统基于物理模型的方法已经能够给出足够简单而精确的解决方案,并不需要深度学习。这就好比一个矩形的面积明明通过长乘宽就可以得到,却非要拿微积分算一遍一样。在有的时候,这种“不必要”有时候其实并不是那么容易被“意识到”。去年一篇来自哈佛和谷歌的用深度学习预测余震位置的 Nature 论文^[105]遭受了地震一般地猛

烈质疑,被封为“深度学习的错误用法”(图 20)。这其实并不是单单因为利用深度学习预测地震从常识上讲多么不靠谱,而是有人质疑如果采用文中(具有泄露嫌疑)的训练数据,传统任意一个简单模型,如支持向量机(SVM)、广义相加模型(GAM)等,只需要 1500 行数据都可以得出和原本 470 万行数据的神经网络相当的结果^[106]。另一方面,你必须清楚地认清深度学习技术“它最好能做到多好”很大程度上取决于“你给它的数据有多可靠”。例如,在某些场合下,采集到的真实数据对应的真值无从得知。此时你若想偷个懒,直接用经典算法处理的结果去标注并让网络去建立关联,那么就算再理想再强大的神经网络也至多能和经典算法打个平手,那还为何去劳驾深度学习呢?恰恰是因为这个原因,很多问题并不适合直接用深度学习去解决。例如:干涉测量中低质量包裹相位图的空间解包裹(对应实测数据的正确绝对相位信息难以甚至无法获得),条纹投影中直接从相位到深度的端到端映射(真实世界物体的绝对三维坐标难以通过仪器量测)等。

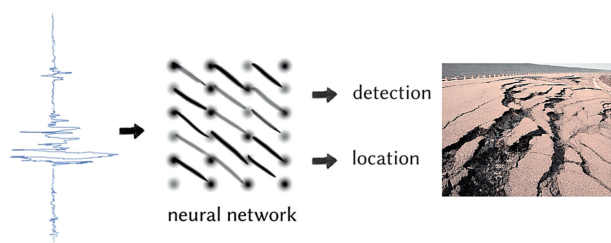


图 20 深度学习被用于预测地震遭到了质疑

Fig. 20 Forecasting earthquake using deep learning hit with rebuttals has been questioned

5 结束语

当下快速发展的深度学习技术为计算光学成像的发展打开了一扇新的窗户,它有效提升了传统计算成像技术的信息获取能力,降低了传统计算成像技术对“正向物理模型”或“逆向重构算法”的过度依赖,突破了传统计算成像技术所能够达到的功能/性能疆界,并为这个领域带来很多令人瞩目的开创性研究成果。但同样,深度学习技术在光学成像领域的应用还面临着巨大的挑战。这不仅需要依赖深度学习的专家们进一步去完善这一工具,还需要光学成像的专家们更加理性地去借鉴与使用。相信大家都期望看到今后越来越多的文章不再是各类网络结构与各类计算成像体制简单排列组合般的堆砌,而能够真正大胆地把深度学习下的计算成像所面临的这些挑战毫不避讳地拿到台面上去讨论,甚至有勇

气去挑战它们!只有这样,“深度学习下的计算成像”这一研究领域才不会“昙花一现”,而是能够真正地走得更远……

参 考 文 献

- [1] Wikipedia. Computational imaging [EB/OL]. (2019-10-15) [2019-11-20]. https://en.wikipedia.org/wiki/Computational_imaging.
- [2] Wikipedia. AlphaGo versus Lee Sedol [EB/OL]. (2019-11-06) [2019-11-20]. https://en.wikipedia.org/wiki/AlphaGo_versus_Lee_Sedol.
- [3] Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge [J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [4] Rivenson Y, Zhang Y B, Günaydn H, et al. Phase recovery and holographic image reconstruction using deep learning in neural networks[J]. Light: Science & Applications, 2018, 7(2): 17141.
- [5] Wu Y C, Rivenson Y, Zhang Y B, et al. Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery[J]. Optica, 2018, 5(6): 704-710.
- [6] Ren Z B, Xu Z M, Lam E Y. Learning-based nonparametric autofocusing for digital holography [J]. Optica, 2018, 5(4): 337-344.
- [7] Wang K Q, Dou J Z, Qian K M, et al. Y-Net: a one-to-two deep learning framework for digital holographic reconstruction [J]. Optics Letters, 2019, 44(19): 4765-4768.
- [8] Nguyen T, Bui V, Lam V, et al. Automatic phase aberration compensation for digital holographic microscopy based on deep learning background detection [J]. Optics Express, 2017, 25 (13): 15043-15057.
- [9] Zhang G, Guan T, Shen Z Y, et al. Fast phase retrieval in off-axis digital holographic microscopy through deep learning [J]. Optics Express, 2018, 26(15): 19388-19405.
- [10] Nguyen T, Xue Y J, Li Y Z, et al. Deep learning approach for Fourier ptychography microscopy [J]. Optics Express, 2018, 26(20): 26470-26484.
- [11] Kappeler A, Ghosh S, Holloway J, et al. Ptychnet: CNN based Fourier ptychography [C] // 2017 IEEE International Conference on Image Processing (ICIP), September 17-20, 2017, Beijing, China. New York: IEEE, 2017: 1712-1716.
- [12] Jiang S W, Guo K K, Liao J, et al. Solving Fourier ptychographic imaging problems via neural network modeling and TensorFlow [J]. Biomedical Optics

- Express, 2018, 9(7): 3306-3319.
- [13] Cheng Y F, Strachan M, Weiss Z, et al. Illumination pattern design with deep learning for single-shot Fourier ptychographic microscopy [J]. Optics Express, 2019, 27(2): 644-656.
- [14] Lü M, Wang W, Wang H, et al. Deep-learning-based ghost imaging [J]. Scientific Reports, 2017, 7: 17865.
- [15] He Y C, Wang G, Dong G X, et al. Ghost imaging based on deep learning [J]. Scientific Reports, 2018, 8: 6469.
- [16] Shimobaba T, Endo Y, Nishitsuji T, et al. Computational ghost imaging using deep learning [J]. Optics Communications, 2018, 413: 147-151.
- [17] Wang H D, Rivenson Y, Jin Y Y, et al. Deep learning enables cross-modality super-resolution in fluorescence microscopy [J]. Nature Methods, 2019, 16(1): 103-110.
- [18] Nehme E, Weiss L E, Michaeli T, et al. Deep-STORM: super-resolution single-molecule microscopy by deep learning [J]. Optica, 2018, 5(4): 458-464.
- [19] Ouyang W, Aristov A, Lelek M, et al. Deep learning massively accelerates super-resolution localization microscopy [J]. Nature Biotechnology, 2018, 36(5): 460-468.
- [20] Rivenson Y, Göröcs Z, Günaydin H, et al. Deep learning microscopy [J]. Optica, 2017, 4(11): 1437-1443.
- [21] Heinrich L, Bogovic J A, Saalfeld S. Deep learning for isotropic super-resolution from non-isotropic 3D electron microscopy [M] // Descoteaux M, Maier-Hein L, Franz A, et al. Medical image computing and computer-assisted intervention-MICCAI 2017. Lecture notes in computer science. Cham: Springer, 2017, 10434: 135-143.
- [22] Wang H, Rivenson Y, Jin Y, et al. Deep learning achieves super-resolution in fluorescence microscopy [J]. Biorxiv, 2018: 309641.
- [23] Fang L Y, Cunefare D, Wang C, et al. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search [J]. Biomedical Optics Express, 2017, 8(5): 2732-2744.
- [24] Lee C S, Baughman D M, Lee A Y. Deep learning is effective for classifying normal versus age-related macular degeneration OCT images [J]. Ophthalmology Retina, 2017, 1(4): 322-327.
- [25] Schlegl T, Waldstein S M, Bogunovic H, et al. Fully automated detection and quantification of macular fluid in OCT using deep learning [J]. Ophthalmology, 2018, 125(4): 549-558.
- [26] Waller L, Tian L. Machine learning for 3D microscopy [J]. Nature, 2015, 523(7561): 416-417.
- [27] Nguyen T, Bui V, Nehmetallah G. Computational optical tomography using 3-D deep convolutional neural networks [J]. Optical Engineering, 2018, 57(4): 043111.
- [28] Lü M, Wang H, Li G, et al. Learning-based lensless imaging through optically thick scattering media [J]. Advanced Photonics, 2019, 1(3): 036002.
- [29] Li S, Deng M, Lee J, et al. Imaging through glass diffusers using densely connected convolutional networks [J]. Optica, 2018, 5(7): 803-813.
- [30] Horisaki R, Takagi R, Tanida J. Learning-based imaging through scattering media [J]. Optics Express, 2016, 24(13): 13738-13743.
- [31] Satat G, Tancik M, Gupta O, et al. Object classification through scattering media with deep learning on time resolved measurement [J]. Optics Express, 2017, 25(15): 17466-17479.
- [32] Cheng S F, Li H H, Luo Y Q, et al. Artificial intelligence-assisted light control and computational imaging through scattering media [J]. Journal of Innovative Optical Health Sciences, 2019, 12(4): 1930006.
- [33] Goy A, Arthur K, Li S, et al. Low photon count phase retrieval using deep learning [J]. Physical Review Letters, 2018, 121(24): 243902.
- [34] Chen C, Chen Q F, Xu J, et al. Learning to see in the dark [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 3291-3300.
- [35] Rivenson Y, Wang H D, Wei Z S, et al. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning [J]. Nature Biomedical Engineering, 2019, 3(6): 466-477.
- [36] Rivenson Y, Liu T R, Wei Z S, et al. PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning [J]. Light: Science & Applications, 2019, 8: 23.
- [37] Feng S J, Chen Q, Gu G H, et al. Fringe pattern analysis using deep learning [J]. Advanced Photonics, 2019, 1(2): 025001.
- [38] Yan K T, Yu Y J, Huang C T, et al. Fringe pattern denoising based on deep learning [J]. Optics Communications, 2019, 437: 148-152.
- [39] Wang K Q, Li Y, Qian K M, et al. One-step

- robust deep learning phase unwrapping[J]. *Optics Express*, 2019, 27(10): 15100-15115.
- [40] Feng S J, Zuo C, Yin W, et al. Micro deep learning profilometry for high-speed 3D surface imaging[J]. *Optics and Lasers in Engineering*, 2019, 121: 416-427.
- [41] Luo W, Schwing A G, Urtasun R. Efficient deep learning for stereo matching[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 26-July 1, 2016, Las Vegas, Nevada. New York: IEEE, 2016: 5695-5703.
- [42] Kuznietsov Y, Stuckler J, Leibe B. Semi-supervised deep learning for monocular depth map prediction [C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 2215-2223.
- [43] Gerchberg R W, Saxton W O. A practical algorithm for the determination of phase from image and diffraction plane pictures[J]. *Optik*, 1972, 35(2): 237-250.
- [44] Gerchberg R W. Phase determination from image and diffraction plane pictures in the electron microscope[J]. *Optik*, 1971, 34(3): 275-284.
- [45] Fienup J R, Wackerman C C. Phase-retrieval stagnation problems and solutions[J]. *Journal of the Optical Society of America A*, 1986, 3(11): 1897-1907.
- [46] Seldin J H, Fienup J R. Numerical investigation of the uniqueness of phase retrieval[J]. *Journal of the Optical Society of America A*, 1990, 7(3): 412-427.
- [47] Guizar-Sicairos M, Fienup J R. Understanding the twin-image problem in phase retrieval[J]. *Journal of the Optical Society of America A*, 2012, 29(11): 2367-2375.
- [48] Wackerman C C, Yagle A E. Use of Fourier domain real-plane zeros to overcome a phase retrieval stagnation[J]. *Journal of the Optical Society of America A*, 1991, 8(12): 1898-1904.
- [49] Lu G, Zhang Z, Yu F T S, et al. Pendulum iterative algorithm for phase retrieval from modulus data[J]. *Optical Engineering*, 1994, 33(2): 548-555.
- [50] Takajo H, Takahashi T, Kawanami H, et al. Numerical investigation of the iterative phase-retrieval stagnation problem: territories of convergence objects and holes in their boundaries [J]. *Journal of the Optical Society of America A*, 1997, 14(12): 3175-3187.
- [51] Misell D L. A method for the solution of the phase problem in electron microscopy [J]. *Journal of Physics D: Applied Physics*, 1973, 6(1): L6-L9.
- [52] Zheng G A, Horstmeyer R, Yang C. Wide-field, high-resolution Fourier ptychographic microscopy [J]. *Nature Photonics*, 2013, 7(9): 739-745.
- [53] Horstmeyer R, Chen R Y, Ou X Z, et al. Solving ptychography with a convex relaxation [J]. *New Journal of Physics*, 2015, 17(5): 053044.
- [54] Yeh L H, Dong J, Zhong J S, et al. Experimental robustness of Fourier ptychography phase retrieval algorithms [J]. *Optics Express*, 2015, 23(26): 33214-33240.
- [55] Zuo C, Sun J S, Chen Q. Adaptive step-size strategy for noise-robust Fourier ptychographic microscopy [J]. *Optics Express*, 2016, 24(18): 20724-20744.
- [56] Pittman T B, Shih Y H, Strekalov D V, et al. Optical imaging by means of two-photon quantum entanglement[J]. *Physical Review A*, 1995, 52(5): R3429-R3432.
- [57] Bennink R S, Bentley S J, Boyd R W. "Two-photon" coincidence imaging with a classical source [J]. *Physical Review Letters*, 2002, 89(11): 113601.
- [58] Gatti A, Brambilla E, Bache M, et al. Ghost imaging with thermal light: comparing entanglement and classical correlation[J]. *Physical Review Letters*, 2004, 93(9): 093602.
- [59] Sen P, Chen B, Garg G, et al. Dual photography [C]//*ACM SIGGRAPH 2005 Papers on SIGGRAPH'05*, July 31-August 4, 2005, Los Angeles, California. New York: ACM, 2005: 745-755.
- [60] Dharmpal T, Laska J N, Michael B, et al. A new compressive imaging camera architecture using optical-domain compression [J]. *Proceedings of SPIE*, 2006, 6065: 606509.
- [61] 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.
- [62] Sun B, Edgar M P, Bowman R, et al. 3D computational imaging with single-pixel detectors [J]. *Science*, 2013, 340(6134): 844-847.
- [63] Vellekoop I M, Mosk A P. Focusing coherent light through opaque strongly scattering media [J]. *Optics Letters*, 2007, 32(16): 2309-2311.
- [64] Popoff S M, Lerosey G, Carminati R, et al. Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media [J]. *Physical*

- Review Letters, 2010, 104(10): 100601.
- [65] Leith E N, Upatnieks J. Holographic imagery through diffusing media[J]. Journal of the Optical Society of America, 1966, 56(4): 523.
- [66] 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.
- [67] Bertolotti J, van Putten E G, Blum C, et al. Non-invasive imaging through opaque scattering layers [J]. Nature, 2012, 491(7423): 232-234.
- [68] Yang W Q, Li G W, Situ G H. Imaging through scattering media with the auxiliary of a known reference object [J]. Scientific Reports, 2018, 8: 9614.
- [69] Lü M, Wang H, Li G, et al. Learning-based lensless imaging through optically thick scattering media [J]. Advanced Photonics, 2019, 1(3): 036002.
- [70] Wolf E. Three-dimensional structure determination of semi-transparent objects from holographic data [J]. Optics Communications, 1969, 1(4): 153-156.
- [71] Kak A C, Slaney M. Principles of computerized tomographic imaging[M]. New York: IEEE Press, 2001.
- [72] Haeberlé O, Belkebir K, Giovaninni H, et al. Tomographic diffractive microscopy: basics, techniques and perspectives[J]. Journal of Modern Optics, 2010, 57(9): 686-699.
- [73] Rappaz B, Marquet P, Cuche E, et al. Measurement of the integral refractive index and dynamic cell morphometry of living cells with digital holographic microscopy[J]. Optics Express, 2005, 13(23): 9361-9373.
- [74] Lauer V. New approach to optical diffraction tomography yielding a vector equation of diffraction tomography and a novel tomographic microscope [J]. Journal of Microscopy, 2002, 205(2): 165-176.
- [75] Choi W. Tomographic phase microscopy and its biological applications [J]. 3D Research, 2012, 3(4): 5.
- [76] Charrière F, Marian A, Montfort F, et al. Cell refractive index tomography by digital holographic microscopy[J]. Optics Letters, 2006, 31(2): 178-180.
- [77] Charrière F, Pavillon N, Colomb T, et al. Living specimen tomography by digital holographic microscopy: morphometry of testate amoeba [J]. Optics Express, 2006, 14(16): 7005-7013.
- [78] Choi W, Fang-Yen C, Badizadegan K, et al. Tomographic phase microscopy [J]. Nature Methods, 2007, 4(9): 717-719.
- [79] Sung Y, Choi W, Fang-Yen C, et al. Optical diffraction tomography for high resolution live cell imaging[J]. Optics Express, 2009, 17(1): 266-277.
- [80] Kim K, Yoon H, Diez-Silva M, et al. High-resolution three-dimensional imaging of red blood cells parasitized by *Plasmodium falciparum* and *in situ* hemozoin crystals using optical diffraction tomography [J]. Journal of Biomedical Optics, 2014, 19(1): 011005.
- [81] Cotte Y, Toy F, Jourdain P, et al. Marker-free phase nanoscopy [J]. Nature Photonics, 2013, 7(2): 113-117.
- [82] Devaney A. A filtered backpropagation algorithm for diffraction tomography[J]. Ultrasonic Imaging, 1982, 4(4): 336-350.
- [83] Barty A, Nugent K A, Roberts A, et al. Quantitative phase tomography [J]. Optics Communications, 2000, 175(4/5/6): 329-336.
- [84] Soto J M, Rodrigo J A, Alieva T. Label-free quantitative 3D tomographic imaging for partially coherent light microscopy [J]. Optics Express, 2017, 25(14): 15699-15712.
- [85] Li J, Chen Q, Sun J, et al. Three-dimensional tomographic microscopy technique with multi-frequency combination with partially coherent illuminations[J]. Biomedical Optics Express, 2018, 9(6): 2526-2542.
- [86] Soto J M, Rodrigo J A, Alieva T. Optical diffraction tomography with fully and partially coherent illumination in high numerical aperture label-free microscopy [Invited] [J]. Applied Optics, 2018, 57(1): A205-A214.
- [87] Horstmeyer R, Chung J, Ou X, et al. Diffraction tomography with Fourier ptychography[J]. Optica, 2016, 3(8): 827-835.
- [88] Tian L, Waller L. 3D intensity and phase imaging from light field measurements in an LED array microscope[J]. Optica, 2015, 2(2): 104-111.
- [89] Zuo C, Sun J, Li J, et al. Wide-field high-resolution 3D microscopy with Fourier ptychographic diffraction tomography[J/OL]. (2019-05-26)[2019-11-20]. <https://arxiv.org/abs/1904.09386>.
- [90] Kamilov U S, Papadopoulos I N, Shoreh M H, et al. Learning approach to optical tomography[J]. Optica, 2015, 2(6): 517-522.
- [91] Abbe E. Beiträge zur theorie des Mikroskops und der Mikroskopischen Wahrnehmung[J]. Archiv Für

- Mikroskopische Anatomie, 1873, 9(1): 413-468.
- [92] Hell S W, Wichmann J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy [J]. Optics Letters, 1994, 19(11): 780-782.
- [93] Betzig E, Patterson G H, Sougrat R, et al. Imaging intracellular fluorescent proteins at nanometer resolution[J]. Science, 2006, 313 (5793): 1642-1645.
- [94] Rust M J, Bates M, Zhuang X W. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)[J]. Nature Methods, 2006, 3(10): 793-796.
- [95] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks [M] // Fleet D, Pajdla T, Schiele B, et al. Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8689: 818-833.
- [96] Shwartz-Ziv R, Tishby N. Opening the black box of deep neural networks via information [J/OL]. (2017-04-29)[2019-11-20]. <https://arxiv.org/abs/1703.00810>.
- [97] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Advances in Neural Information Processing Systems, December 8-13, 2014, Montreal, Quebec, Canada. Canada: NIPS, 2014: 2672-2680.
- [98] Xu Z B, Sun J. Model-driven deep-learning [J]. National Science Review, 2018, 5(1): 22-24.
- [99] Lin H W, Tegmark M, Rolnick D. Why does deep and cheap learning work so well? [J]. Journal of Statistical Physics, 2017, 168(6): 1223-1247.
- [100] Behera B. Status of a deep learning based measurement of the inclusive muon neutrino charged-current cross section in the NOvA near detector [J/OL]. (2017-10-10) [2019-11-20]. <https://arxiv.org/abs/1710.03766>.
- [101] Nguyen T Q, Weitekamp D, Anderson D, et al. Topology classification with deep learning to improve real-time event selection at the LHC[J]. Computing and Software for Big Science, 2019, 3 (1): 12.
- [102] Perera P, Patel V M. Deep transfer learning for multiple class novelty detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 16-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 11544-11552.
- [103] Perera P, Nallapati R, Xiang B. OCGAN: one-class novelty detection using GANs with constrained latent representations[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 16-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 2898-2906.
- [104] Goodfellow I J, Shlens J, Szegedy C. Explaining and harnessing adversarial examples[J/OL]. (2015-03-20)[2019-11-20]. <https://arxiv.org/abs/1412.6572>.
- [105] DeVries P M R, Viégas F, Wattenberg M, et al. Deep learning of aftershock patterns following large earthquakes [J]. Nature, 2018, 560 (7720): 632-634.
- [106] Machine Learning. Deep learning of aftershock patterns following large earthquakes [R/OL]. [2019-11-20]. https://www.reddit.com/r/MachineLearning/comments/9bo9i9/r_deep_learning_of_aftershock_patterns_following/.