

从物理智能到微波视觉

徐丰^{1,2}, 金亚秋^{1,2}

1. 复旦大学电磁大数据与遥感智能研究所, 上海 200433
2. 复旦大学电磁波信息科学教育部重点实验室, 上海 200433

摘要 近10年来,人工智能技术得到了科技与工业界的极大的重视,预示着人类文明将进入智能时代。但是,作为智能时代基础的“智能科学”还远未成型。本文从电磁物理信息感知技术的独特视角,讨论智能科学如何发展的一些见解,指出人类智能与外在世界互为对偶问题、相互不可分割的根本属性,因此按人工智能所应对的对象及关联学科分为数学、物理、心理、意识4个阶段。其中第1阶段解决智能形成的通用学习算法的数学理论,第2阶段发展应对物理世界的物理智能。以此为基础,第3阶段发展应对智能涉及社会群体的高阶智能,第4阶段研究自由意识的本质和人工智能能否形成意识的超智能问题。结合笔者电磁信息感知专业领域,提出向物理智能发展的微波视觉新概念、相关内涵以及关键技术的建议,以笔者团队在这一方向的前期工作为例,讨论了以物理智能为基础的智能科学的研究与发展。

关键词 智能科学;物理智能;微波视觉

著名的摩尔定律成功预测了集成电路的指数次增长规律,实际上人类文明和科技的发展也大致遵循指数次增长的规律。如果把人类社会发展史绘制成一条曲线,横坐标为时间,纵坐标为文明和科技水平,那么它就是一条指数增长曲线。站在今天回望历史,不难发现人类每一次科技革命到来的时间间隔在指数次递减。例如40年前的个人计算机技术、100年前的相对论和量子力学、500年前的近代科学起源,又比如两次工业革命、农业革命、信息产业革命等,无不验证着这一规律。

有人说人工智能是下一次工业革命,亦或许是人类文明的智能时代。科技是人类文明的核心之一,文明的进步经历了资源、能量、信息、智能的4个阶段(图1)。冷兵器时代人类文明依赖于自然资源的优势,热兵器时代的制胜点在于核武器等大规模杀伤性武器,进入信息时代后体现在信息获取能力上,发达国家依靠侦察、隐身、赛博等信息技术实现对其他国家的信息不对称性优势,可以预见未来决胜的是人工智能水平的高低。

信息时代的关键是全面多类精准信息快速获取和

收稿日期:2018-04-09;修回日期:2018-04-25

基金项目:国家自然科学基金项目(61571134)

作者简介:徐丰,教授,研究方向为SAR图像解译、电磁散射建模、类脑人工智能,电子信箱:fengxu@fudan.edu.cn;金亚秋(通信作者),中国科学院院士,研究方向为复杂自然介质的电磁辐射、散射与传输,电子信箱:yqjin@fudan.edu.cn

引用格式:徐丰,金亚秋.从物理智能到微波视觉[J].科技导报,2018,36(10):30-44;doi:10.3981/j.issn.1000-7857.2018.10.004

反获取的能力,包括军事科技的侦察、隐身、对抗等技术。在大数据获取能力达到一定的高度,必须由大数据的物理信息进一步进入到其物理世界背后的意识世界中,发展人工智能与超人工智能,实现对抗式的智能感知、处理、推演、决策技术,在未来智能科技时代占据领先优势。

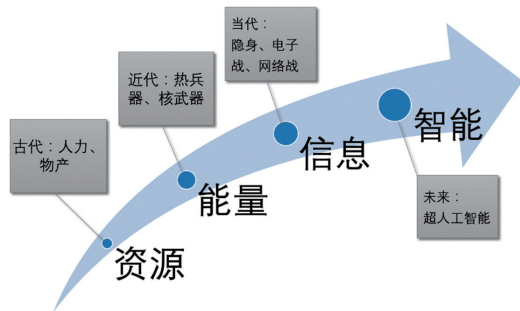


图1 人类文明发展的新阶段

Fig. 1 New era of human civilization development

10多年前神经网络研究先驱 Hinton^[1]在《Science》上发表的一篇论文在人工智能领域引发了深度学习热潮。谷歌子公司 DeepMind 开发的 AlphaGo^[2]以 4:1 战胜了韩国围棋九段棋手李世石,让人工智能得到了广泛的关注。人工智能在过去 10 年取得的重大进展,主要得益于深度学习技术,一种在强大的计算能力、海量的数据支持和改进的算法 3 个条件下成功实现突破的人工智能技术。

短短 10 年间,深度学习引发了人工智能在各行各业的研究应用热潮。在信息、生物、材料、化学等领域均开展了深度学习的人工智能应用研究。在工业界与政府部门,也纷纷推出重磅计划,为人工智能单设计划条目。深度学习技术在各领域均找到了应用,从而引起越来越多的关注。

然而,深度学习技术背后的理论积累难以支撑爆炸式增长的技术发展和应用需求。深度神经网络背后的理论基本上还是 30 多年前的人工神经网络理论,仅在深度表征和训练算法方面有局部突破。智能时代的来临亟需一门智能科学支撑。它与脑科学、认知科学等相关的诸多人脑功能学科相关联,又直接面对物理世界与社会科学界的应用需求。智能科学的举步维艰与人工智能技术应用的火爆现象的反差令人回想起前几次人工智能曾有过的寒冬。这种严峻的趋势也令多位著名学者进行深刻的思考。

本文简单讨论智能科学这一新学科方向,按人工智能所应对的对象及关联学科分为数学、物理、心理和意识 4 个阶段,并在物理智能的范畴下结合笔者前期研究提出信息感知的微波视觉新概念、相关内涵及关键技术。

1 智能科学

1.1 深度学习

深度学习属于机器学习,机器学习是人工智能的一个主要分支。深度学习的主要技术即深度神经网络,它是人工神经网络领域新兴的研究方向。人工神经网络通过将脑科学、数学、计算机科学、信息处理学、心理学等多个研究领域相互交叉结合,以建立一个可以模拟人脑智能的模型。虽然人脑智能研究仅有不到 100 年的发展历史,但已经在机器视觉、人脸识别、语音识别、智能搜索和遗传编程等领域得到了长足的发展,取得了广泛的十分有意义的应用。

人工神经网络研究起源于 20 世纪 50 年代,最早的人工神经网络模型只是对一个神经元接受刺激的响应进行简单的模拟。然而由于理论的缺乏和计算机技术的限制,人工神经网络并没能得到有效的发展。直到 20 世纪 80 年代,基于后向传播(BP)机制的神经网络得到快速发展。但是由于训练困难等问题,仍然局限于 3 层的神经网络,可以解决的问题很有限。之后,机器学习的主流方法大都是浅层结构算法(如支撑矢量机),少有人关注包含 3 层结构以上的深层算法。

2006 年, Hinton^[1]提出通过“贪婪学习”的思路,实现对多层自编码器的训练,重新引发了神经网络领域对于深层网络的研究兴趣。2012 年, Hinton 团队^[9]将深度卷积网络用于 ImageNet 图像分类挑战赛上,取得远高于以往浅层算法的成绩,引起机器学习领域的关注。从此机器学习主流开始往多层自动学习算法偏移,超过 3 层结构以上的机器学习算法被称为深度学习。2016 年,谷歌子公司 DeepMind 将深度神经网络应用于强化学习上,开发了 AlphaGo 计算机围棋程序,4:1 击败了韩国围棋九段棋手李世石,从而引发了各界人士的关注。各行各业对于人工智能的兴趣开始爆炸式增长。特别是在工业界,深度学习已被广泛的应用在各领域,如计算机视觉、语音识别、机器翻译、搜索引擎、自动驾驶、机器人等。

事实上,深度学习目前最成功的算法应该是深度卷积网络对于图像、视频类空域信息的处理,以及深度循环网络对于语音、文字类时间序列信息的处理^[4]。例如深度强化学习的一些应用,均是得益于这2个核心神经网络算法对于原始输入数据的处理。而卷积神经网络与循环神经网络这两种网络结构早在20世纪90年代就被提出。可见深度神经网络或深度学习的核心在于“深度”,即用深层结构表征原始数据。

深度学习的核心思想是层次化的特征提取结构(图2),例如深度卷积网络仿照哺乳动物视觉神经系统,利用层次化特征提取的信息处理模式。先对输入信息进行低级特征提取,在高层将低级特征组合成更高级的特征信息,经过多层特征传递,得到足够高级的特征信息,再计算最终的输出。通过训练,从海量数据中自动提取所关心的特征,将数据标签映射拟合到一个高度非线性函数中。所以,深度神经网络的本质可

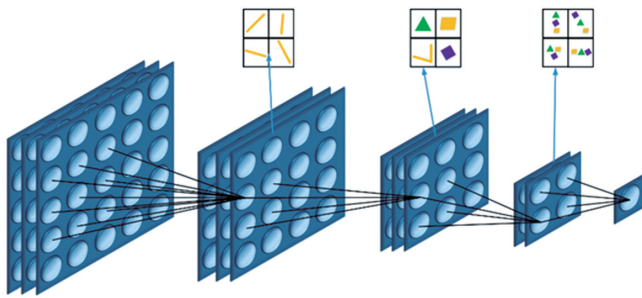


图2 深度学习的主要思想:层次化可组合的特征表征框架
Fig. 2 Key idea of deep learning: hierarchical compositional framework of feature representation

以看作是一个多层嵌套的非线性拟合函数,它巧妙地通过随机梯度下降算法将网络训练到合适的拟合精度。然而,“深度”的概念也不是最新才提出的,用超过3层的多层神经网络来拟合的方法也早就被尝试过。因此学界普遍认为深度学习爆发与3个偶然因素有关:计算能力的指数次增长、数据量的爆炸式增长和神经网络算法的改进。这3个条件的成熟使得神经网络迈入了深度学习的阶段,因而能解决更接近实际应用的复杂问题。从这个角度看,深度学习的出现对于智能科学而言仅仅是量变,还没有达到质变。

回顾历史可以发现,21世纪初出现的计算能力的增长和数据量的增长使得训练深度神经网络成为可能,而算法的改进一直到近几年来才出现的。从2006年Hinton提出的多层自编码网络直到2012年的深度卷积网络之间并没有出现很成功的应用,而从2012年之后,视觉计算的最成功应用均采用深度卷积网络,说明深度神经网络的核心算法改进发生在2006—2012年。

2006年,Hinton提出的“贪婪算法”训练多层自编码器引领了一批专家去研究深度神经网络,包括LeCun和Bengio等。深度神经网络之所以取得巨大成功,笔者认为其最核心算法改进在于采用了一类简化的激活函数,即规则化线性单元(rectified linear unit, ReLU)。由图3可知,ReLU本质上是分段线性函数 $y=\max(0, x)$ 。类似ReLU的激活函数事实上早在1975年即被提出^[5],但是在1980年代被Sigmoid等具有良好非线性、连续可导、梯度计算复杂度低等优点的激活函数全面取代。一直到2009年,LeCun团队^[6]在研究深度神经网络时,

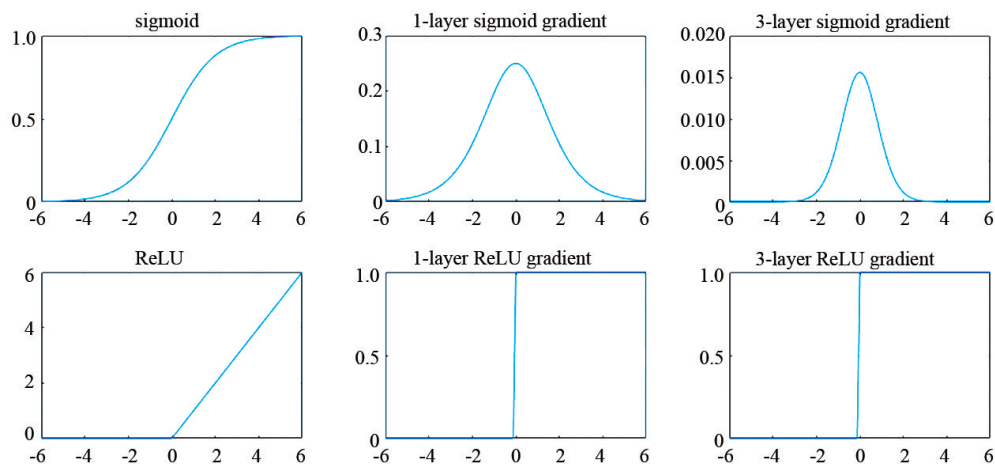


图3 2种激活函数(Sigmoid, ReLU)及其单层、多层梯度比较
Fig. 3 Two types of activation function (Sigmoid, ReLU) and their single-layer, multi-layer gradients

发现采用类 ReLU 的分段线性激活函数是实现深度神经网络端到端训练的关键因素。2011 年, Bengio 团队^[7-8]研究得出结论: 深度神经网络最适合用 ReLU 作为激活函数, 给出的理由包括 ReLU 与人脑神经元的相似性, 因为神经元对某些输入完全没有反应, 而对一些输入的反应呈单调关系, 每一时刻处在激活状态的神经元总是稀疏的。这些工作为 2012 年 Hinton 团队^[3]首次采用深度卷积网络赢得 ImageNet 挑战赛奠定了理论基础。

关于 ReLU 优点, 图 3 给出一种简单的解释, 即其梯度在多层网络后向传播时能保持稳定, 不像 Sigmoid 的梯度在累乘后容易达到饱和。关于 ReLU 的理论研究在 2012 年后还在继续, 例如 2015 年 LeCun 团队^[9]证明了深度神经网络的损失函数中的局部最小值均为较优解, 且很容易被随机梯度下降算法找到, 由此在理论上, 它解释了深度神经网络取得巨大成功的一个主要原因。

可以看出, 深度神经网络的本质是采用很多层嵌套的非线性函数来拟合海量的数据, 监督学习即拟合高维数据空间中的一个曲面, 而非监督学习如生成网络则拟合一种分布。至目前为止, 真正解决实际应用问题的深度学习技术, 无论在网络结构上有各种各样的创新, 或者在训练优化算法上有各种各样的设计, 都逃脱不了一个模式, 即用海量的训练数据拟合一个包含海量未知数的复杂多层网络。

因此, 笔者完全赞同图灵奖获得者 Judea Pearl 的观点: 无论深度学习技术取得多大成功, 各种相关研究多么五花八门, 我们还没有突破其曲线/曲面拟合的本质^[10]。不少学者已经开始反思深度学习面临的困境, 由于其背后基础理论积累不足, 深度学习技术很快就陷入一个瓶颈期。现在的深度学习相关技术过于简单, 以致可以完全被计算机掌握, 由于各种工具箱的出现使得绝大多数人可以轻易掌握深度神经网络应用技能, 绝大多数人掌握这些技能后便快速陷入调参数、调结构的瓶颈。显然, 这本身就是一份机械的工作, 可以被人工智能取代。果然, 谷歌公司于 2018 年初推出的 AutoML 工具即扮演了调参的角色, 使得一大部分应用深度神经网络的人被深度神经网络取代。谷歌公司最新开发的 AlphaGo Zero 新版围棋程序也证明了靠机械的搜索来获得较优解决方案的工作也可以被取代。可以预见, 更多的在研究网络结构和优化算法的科研人

员也将被取代。深度学习技术的出现, 预示着缺乏创造力的机械工作将会被取代。

因此, 深度学习出现的意义首先是很好地利用了数据资源优势, 实现超强拟合能力, 解决各个领域的实际应用问题。其次它引发了人们对于人工智能技术的广泛关注, 引领了一个大力研究智能科技的年代。然而, 深度学习的超强拟合能力导致人们的期望过高, 如果智能科学的基础理论迟迟没有突破, 很快深度学习或人工智能就会跟不上人们对它的期望。

1.2 智能科学的诞生

人工智能深度学习技术的广泛应用必将要求智能科学的诞生。智能科学应是一门区别于自然科学和社会科学的崭新学科。它的诞生不仅仅与深度学习技术爆发有关, 而且与脑科学、认知科学、理论神经科学、数据科学的发展密切相关, 智能科学甚至还与一般自然科学和社会科学相关, 这些学科是智能科学的前奏和基石。

智能科学是人工智能背后的基础科学, 人工智能是建立于智能科学之上的一门应用学科。首先需要关心的是人工智能需要什么样的科学支撑。社会学家将广义人工智能分为 3 类: 弱人工智能和强人工智能和超人工智能。弱人工智能是指擅长某一方面的人工智能, 现阶段发展的人工智能就是弱人工智能。强人工智能指的是与人类比肩的人工智能, 人类能做的事情它都能做。超人工智能就是科幻世界中经常出现的智力超越人类的人造机器。人工智能的目标可以归纳为 4 点——像人一样行动、像人一样思考、合理地思考、合理地行动。如果要使机器能像人一样思考的话, 必须先弄清楚人类是如何思考的。这本身就是一件复杂而有趣的工作。有 3 种办法可以完成这项任务: 通过内省——反思人类的思维方式; 通过心理实验——观察被测对象, 然后总结; 以及通过脑成像——观察大脑的活动规律。然而, 直到今天, 人类对于自己大脑的运作原理以及人类思维等的认识仍然不足, 认知科学还有诸多问题有待解决。

当前深度网络的巨大应用成功仅仅发源于人们对于视觉神经网络的理解和仿生。大脑和智能作为至今人类还没有研究透彻的领域, 不断引起了科技界, 乃至工业界和政府部门的重视。系统的脑科学研究主要分为自底向上的生物学研究和自顶向下的理论神经科学或认知科学研究, 神经生物学通过生物观测实验手段

从微观层面开始,不断向上积累基础知识,试图理解整个大脑的工作机制,例如通过神经元细胞切片的观察,一直到宏观尺度的大脑活动功能核磁共振(fMRI)成像等,神经生物学研究目前无法彻底理解大脑工作机制,主要局限于对活体大脑实时活动的微观观察手段的缺乏,近年来发展的荧光标记高分辨率光学成像技术将是该领域下一个突破口。理论神经科学或认知科学研究通过对人类宏观行为的理解建立数学理论模型以图仿制大脑智能。理论神经科学仿照牛顿物理学首先从数学上对大脑工作机制进行建模,通过假设的基本规律来逐步解释更复杂的大脑工作机制,与生物学手段比起来,理论神经科学的研究一直停滞不前。实验观测和理论建模必须有效结合,才能得到可靠的进展。例如通过观测发现的脉冲时间相关可塑性(SDTP)机制被认为是记忆和学习的神经化学基础,对于神经网络模型的构建具有重要推动作用。除了系统的脑科学研究,目前提及最多的是更接近应用层面的类脑人工智能算法(即深度学习)及其实现(即神经网络芯片),以及机器人外围技术(如机械骨架、仿生材料)。当前类脑人工智能算法与真正意义上的神经生物学的结合不够紧密,这将是下一代人工智能发展的方向之一。而机器人外围技术由于其在工业生产、家庭护理、国家安全等领域的重要应用得到重视,在人工智能技术的推动下,机器人可以实现的功能已快速接近以前只有在科幻世界中才出现的场景。

美国、欧盟、中国均已经或即将推出各自的大脑研究计划和人工智能战略,在神经生物学、类脑人工智能技术等方面开展研究。解开大脑的奥秘是开启未来智能世界的关键,是未来科技发展的战略制高点,谁掌握人工智能,谁将成为未来核心技术的掌控者。

人工智能及其相关技术日新月异,每周都有相关突破性进展报道。但是如前论述,深度学习本质即为超强函数拟合,其技术红利可能将很快被用尽。显然,目前的人工智能依旧处于弱人工智能范畴,走向强人工智能和超人工智能的道路依旧一团迷雾。在这个过

程中,需要先解开大脑思维的奥秘,发展系统的智能科学。就像看到鸟的展翅飞翔,有了人类飞行的萌芽,但要发明飞机,则要发展现代航空力学与航空工业。

多名学者针对深度学习的火爆现象进行了冷静思考,并提出智能科学的研究建议。深度学习的引领人物也不停在反思,Hinton等^[11]在深度学习进入火热的现象后,一直在思考,最近提出要推翻后向传播梯度下降的算法,认为人脑不是靠梯度下降学习,提出“胶囊网络”的新概念。LeCun^[12]针对机器学习的发展也提出无监督学习与预测学习是机器学习的一个大方向。

此外,深度学习也受到来自机器学习其他领域学者的质疑,比如谷歌公司的Rahimi^[13]在第31届神经信息处理系统大会(NIPS 2017)上公开将当前深度学习的研究比喻为“炼金术”,其本意就是批评当前大量的深度学习研究仅仅浮于网络结构和参数的不断调整尝试上,而没有试图建立一门基于严谨、周密、可验证的理论之上的科学。

正如人人会做的酿酒术,并不能容易产生近代化学一样。LeCun认为:工程总是先于理论,就像望远镜先于光学、蒸汽机先于热动力学、计算机先于计算科学等。对于此,麻省理工学院Poggio^[14]则认为当前深度学习太依赖于数据,应该回归到理解人类智能上来,即脑科学和神经科学,比如首先解释人脑智能是如何产生的这样的根本科学问题。

2011年图灵奖获得者、贝叶斯网络的前驱Judea Pearl^[10]最近发表论文阐述了机器学习理论的可能发展方向,认为当前深度学习的本质即函数拟合,其所关联的数据科学也只跟数据有关并没有科学。Pearl认为强人工智能的正确途径是引入因果关系,他将其分为关联(association)、交互(intervention)和想象(counterfactual)3个层面(图4^[10])。当前深度学习仅停留在关联的层面,在交互的层面需要引入互动并观察其因果关系,而在想象层面则需要具备逻辑推理能力。最后,他提出应采用他本人发明的用于描述因果关系的贝叶斯网络等数学工具来研究下一步的强人工智能。

因果层次	人类动作	思考问题	具体例子
1. 关联	观察	是什么?	从某种症状推断某种疾病
2. 交互	动手	怎么做?	吃了某种药, 疾病痊愈了
3. 想象	思考	会怎样?	如果不吃药的话现在会怎样

图4 Pearl的3层因果关系

Fig. 4 Pearl's three layers of casual relationship

加州大学洛杉矶分校朱崧纯^[15]对人工智能从现状、任务到构架与统一进行了系统的阐述。他将人工智能学科分为6个方向:计算机视觉、自然语言理解与交流、认知与推理和机器人学、博弈与伦理和机器学习。他对于人工智能的一个观点也是与数据依赖性有关,认为人工智能不应该依赖大量数据,更应该关注人脑内在因素。认为智能科学是牛顿与达尔文理论体系的统一,提出应该把面向应用的人工智能变成智能科学。他提出一个很好的观点:“物理学把生物的意志排除在研究之外,而这正好是智能科学要研究的对象。智能科学要研究的是一个物理与生物混合的复杂系统。智能作为一种现象,就表现在个体与自然、社会群体的相互作用和行为过程中。我个人相信这些行为和现象必然有统一的力、相互作用、基本元素描述。”这是笔者所了解到的对智能科学最深邃的见解之一。

1.3 智能科学与物理智能

智能科学是智能时代所必须的支撑基础科学,它的内涵远不止当前的深度学习技术。智能科学应以人脑为支点,研究人与自然世界、人与人类社会的多体问题。但智能科学不能是孤立的研究人脑的科学。人脑是进化和后天学习的产物,生物进化给予人类一套可塑的神经组织,通过人与世界和社会的交互才形成了人类智能。人脑本身的结构组成和工作机制离不开外界对它的影响。可以想象一个初生的婴儿若一直孤独地生存在太空中,不与任何物质或信息交互,则不可能具备人类智能。所以说,人类智能是外在世界对人脑施加作用后所产生的反作用,外在世界与人脑智能是一双对偶问题。现有科学已经对外在世界建立较为成熟的理论体系。

因此智能科学应是一门以人脑为中心,研究人脑与自然世界相互作用、人脑与人脑相互作用,人脑与人类社会相互作用的学科。它不仅依赖于一套能描述人脑原理的数学理论,还需结合用于描述自然世界单体问题的自然科学和用于描述人类社会单体问题的社会科学。

按这一观点将人工智能分为下面4个阶段。

第1个阶段研究对象为人类智能的大脑原生组织结构,即自然进化赐予人类的可塑神经网络,借助于脑科学与计算神经科学等,发展一套可以建模大脑原生组织结构的数学工具,可以称之为数学智能。这一阶段的目的是用数学和计算工具描述和模拟初生婴儿大

脑的通用组织结构和通用学习算法。

第2个阶段研究对象为外在物理世界加到人脑的“外力”导致在人脑中产生的“应力”,即人脑与物理世界交互后学会的适应物理世界并在其中生存的能力,可以称之为物理智能。物理智能即能适应客观世界的人工智能,如机器人。构造物理智能必须依赖数学智能和物理学。人类建立起来的强大的物理学理论体系有助于构造强于人类的物理智能,即比人类更适应物理世界的智能体,本文第2节讨论的微波视觉即属于这一范畴。物理智能可以推广到广义的自然智能,即研究能够适应各种自然科学现象的人工智能。

第3个阶段研究对象为他人或社会对于人脑的“外力”与所产生的“应力”,即人脑如何学会与他人相处、如何在人类社会中适应和生存的能力,可以称之为心理智能,或广义的社会智能。这个阶段研究智能与智能之间的高阶相互作用,也可称为高阶智能;则物理智能属于一阶智能,因为其研究范畴只涉及单个智能体。

每个阶段还应该研究对应的人脑原生的驱动力,如生理需求、心理需求等,这些需求是驱动通用人工智能算法进行正确学习的必要源动力。

第4个阶段研究对象为意识的本质,意识如何形成是一个根本科学问题,人工智能是否能产生意识更是一个哲学问题,这一根本问题的研究有助于解答人类一直寻求的答案:人是从哪里来的。有一点可以肯定是:没有自我意识的人工智能不可能超越人类。

这种分类方法(图5)体现了智能在进化生长过程不同阶段中所应对的对象不同。当前智能科学相关的研究往往不特别区分这些阶段,而这个分阶段的生长过程却定义了智能本质。人类智能就是循序渐进的从先适应客观世界的物理智能,发展到适应人类社会的心理智能。人工智能的发展也应遵循这一规律。先建立一种通用的人工智能算法,能够建模人类大脑的基本机能,如学习、记忆等。这种算法与物理世界交互即可产生物理智能,例如机器人能够应对牛顿力学。不同阶段的研究可以并行进行,如在数学智能没有完全研究成熟之前可以进行相应的物理智能研究。例如基于现阶段的深度神经网络理论,可以结合计算物理进行基于深度学习的物理智能研究。数学智能与物理智能可以相互补充、相互促进。人工智能来源于人类智能但完全有可能在某些方面超越人类,就像人类学鸟飞发明飞机,而飞机却超越鸟类,当然也超越了不会飞

的人类的能力;显然,已经建立的物理学理论体系可以帮助人们发明在某一方面超越人类的物理智能。

通过2个途径,人们可能构建超越人类的物理智能:1) 借助物理学对于物理世界的严谨理论表述,我们可以将物理原理设计在人工智能中,使其具备比人类更强的适应或战胜物理世界的能力;2) 采用计算物理模拟器来训练通用人工智能算法,使其具备人类不可能拥有的超能力。第1点类似于在人工智能大脑内部用计算物理引擎构建1个模型来帮助预测外部世界。关于第2点,谷歌 AlphaGo Zero 算法的成功给人们一定的信心,通过在围棋模拟器上学习3天即超过人类3000

年的集体智慧,说明模拟器训练是一种可行的途径。

因此,计算物理模拟器是物理智能的关键。采用模拟器构建物理智能的途径利用了人类发展史上的2次革命成果:1) 科学革命建立了人类对于物理世界规律精确表述的理论和技术创新;2) 信息革命建立了人类利用计算机对物理世界规律进行计算仿真的能力。这相当于将人类千年的智慧汇集到一个人工智能大脑中。物理智能将超越人类智能,因为物理学描述的现象超越人类感官范畴。例如,物理学涵盖的尺度范围和速度范围远超过人类能适应的范围,电磁波涵盖的频谱远超过人类能感知的光谱范围(图6)。

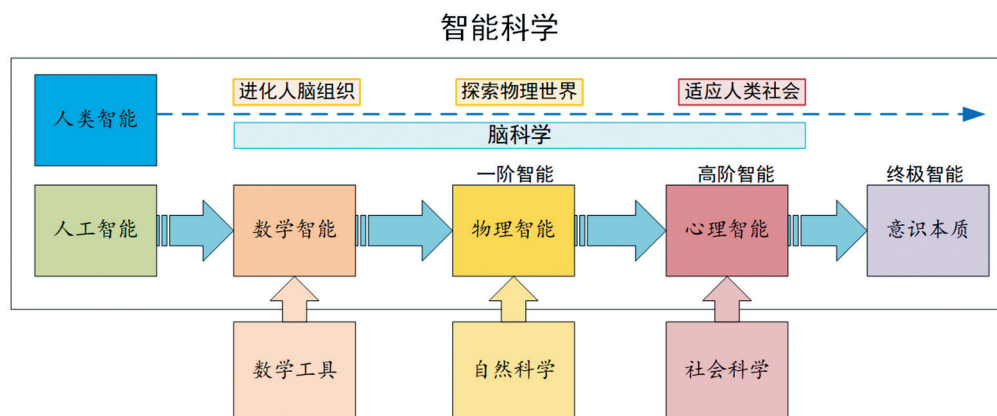


图5 未来智能科学体系

Fig. 5 Architecture of future intelligent science

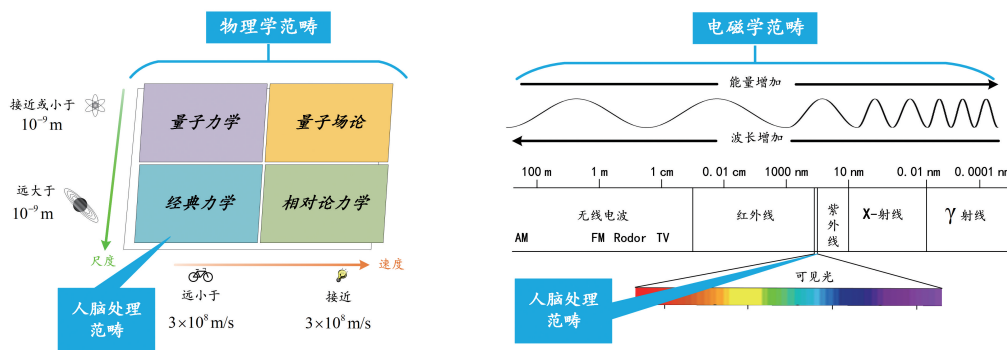


图6 物理智能在适应物理世界的能力上将超越人类智能

Fig. 6 Physical intelligence will excel over human being in adapting to the physical world

作为物理智能一个典型的例子,通过力学模型构建的人工智能可以完美地控制机器人的运动。另一个

例子就是“微波视觉”,一种基于计算电磁学引擎的物理智能,像人类处理光信息一样处理微波信息。

2 微波视觉

2.1 内涵

人的视觉能看见光,但看不见微波。若有外星人的眼睛能看见微波,那他就具有“微波眼睛”,他的视觉神经中枢必与人类视觉不同,可称为“微波视觉”。第二次世界大战期间发明的雷达等技术已使人类感受到微波的信息,是否也能发展基于人工智能的信息感知与理解的“微波视觉”?

在阳光下历经亿年进化,人类形成与生俱来的适应光学信息的大脑,初生婴儿的大脑再通过其与光学世界交互进行学习,最终形成能感知世界、认识自我的人类视觉。人类视觉能高效理解光学信息、实现自主定位导航,同时还是产生自我意识的关键因素。

深度学习前沿技术 AlphaGo Zero 给予人们启发,智能算法与虚拟围棋世界的交互和学习可以获得超越人类的智能。早期 AlphaGo 版本采用人类棋谱进行学习与训练,AlphaGo Zero 版本则完全通过在围棋游戏模拟器里面相互对弈,然后相互学习优化,经过3天的学习就达到了人类冠军的水平。AlphaGo Zero 达到这一成绩的核心前提条件在于模拟器精确的模拟了围棋游戏世界的规则,也就是说如果有一个能精确模拟物理世界的模拟器,相信人们也能用现有深度学习技术训练出实用的人工智能。

显然,物理智能的其中一种途径(或者说当前可行的途径)就是用精确高效的计算物理模拟器训练基于深度学习的人工智能,而这样的思想已经在国际上得以推行。例如在自动驾驶领域的绝大多数企业都在使用虚拟现实模拟器和真实采集的数据一起训练自动驾驶算法。2017年底,美国美国国防高级研究计划局(DARPA)启动“进攻性蜂群使能战术”OFFSET研究项目,旨在研发能自动控制上百架无人机进行协同作战的人工智能技术。该项目的主要技术途径即依赖虚拟现实飞行模拟器 AirSim 训练这一人工智能技术。

基于现有深度学习技术,可以发展适应电磁信息的认知智能算法,并用超算模拟电磁世界,通过与电磁模拟器的交互进行演化学习,实现人造微波视觉。研究微波视觉能自动解译雷达图像、复杂电磁环境等微波信息,实现微波视觉自主定位导航,为探索意识本质提供基础。

图7阐述了微波视觉的主要构思,光视觉是人在光

世界中进化和学习获得,而我们提出通过模拟电磁(微)波世界,实现相类似光视觉但又截然不同电磁信息感知认知的微波视觉。主要技术途径是借鉴 AlphaGo/OFFSET、基于电磁学理论,构建微波世界模拟器,用于演化和训练微波视觉。

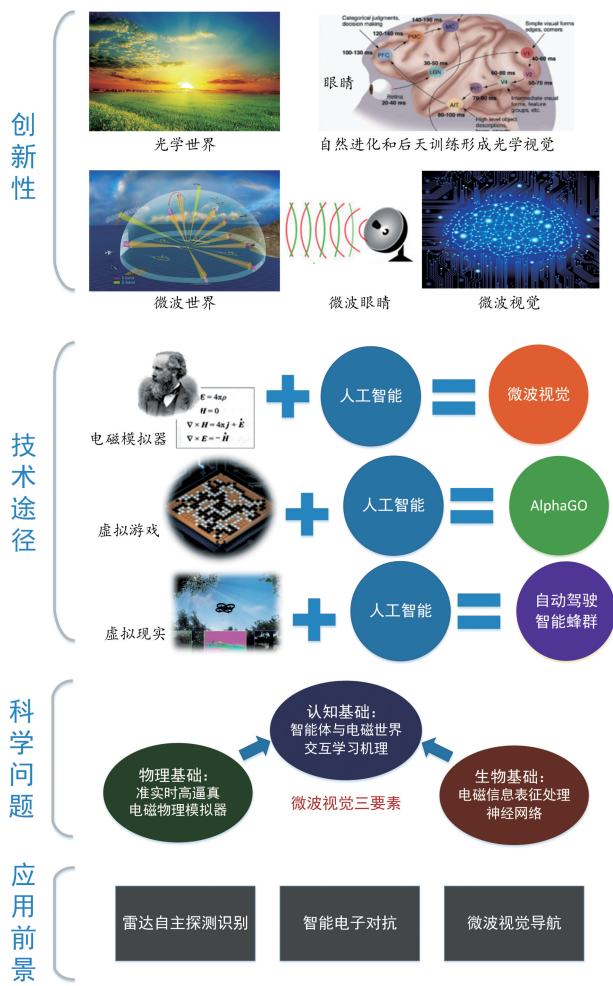


图7 微波视觉的概念与内涵
Fig. 7 Microwave vision

借鉴人脑的光视觉发展人工智能的微波视觉,人脑产生光视觉的3个必要条件是:外在光学世界的存在;自然进化产生的光视觉神经网络;人与光世界交互学习演化的能力。因此发展微波视觉需要研究对应3个问题:高性能电磁物理模拟器、处理电磁信息的微波视觉神经网络、智能体与电磁模拟器的交互学习机制,这3个问题分别构成了微波视觉的电磁物理基础、神经生物基础和认知学习基础,称之为微波视觉三要素。

1) 微波视觉的电磁物理基础——准实时高逼真的电磁物理模拟器。

首先解决微波视觉的虚拟训练环境问题,为微波视觉的进化和学习提供虚拟环境。这一问题的关键体现在既要实现高逼真度的模拟效果,又要在资源约束条件下满足准实时交互的要求,这是采用虚拟环境训练人工智能的可行性条件。若采用精确几何物理建模和现有计算电磁学方法,其模型精细度难以在有限人力条件下做到高逼真要求,算法复杂度无法达到准实时要求。突破这一瓶颈的思路是将观测大数据同化到计算物理模型中,研究数据驱动的电磁模拟器,通过大数据挖掘提取半经验底层散射模型,进一步发展混合多尺度计算电磁学方法,实现既能提高仿真逼真度、又能降低算法复杂度、满足准实时交互的需求。

2) 微波视觉的神经生物基础——处理电磁信息的微波视觉神经网络。

人类进化的视觉神经网络适应于光信息处理,需要设计专门表征处理电磁信息的微波视觉神经网络。以视觉神经网络底层组织、环路和网络结构为启发,对比光信息与微波信息,仿照设计构建电磁信息神经网络,体现从以粒子效应为主到波场效应为主的物理属性转变。解决神经元如何表征电磁信息的时、频、相、极化等多维度特征,神经信息传递机制如何体现电磁波相干散射与传播规律,神经网络结构如何契合电磁散射逆问题求解模式等微波视觉神经网络基础问题。设计用于结构化表征的深度生成网络和用于反演的深度鉴别网络等基本组件。

3) 认知学习基础——智能体与电磁模拟器交互学习机制。

光视觉是人脑视觉神经网络与外在光世界交互学习演化的产物,基于交互的学习机制是突破通用人工智能须解决的根本问题。尽管脑科学这方面研究相对滞后,现阶段的微波视觉研究可参考已知人类光视觉的学习机制和运作机制,如视觉注意、视觉印象、视角关联、预测推断等,受它们启发研究如何建立具备注意、记忆、关联、预测等功能的微波视觉强化学习架构。借鉴视觉神经认知机制,探求如何设计交互式迭代学习算法,按训练和测试2个阶段分别研究不同交互机制,在训练阶段的交互以采样为目的,测试阶段的交互以实时预测反演为目的。

“微波视觉”与人类长期依靠的“光视觉”不同,它的产生将颠覆传统的雷达信号处理与以光视觉为基础的图像处理技术,将使目前似乎看不见摸不着的微波

监测成为“微波视觉”中看得清理解透的新的技术形态,如自主定位导航、电子侦查对抗等电磁波技术。微波视觉将实现天地海目标雷达回波图像在线自动解译,改变目前地面站人工判读业务形态;实现雷达回波直接重构目标与场景的可视化表征,颠覆传统电磁信号处理、雷达回波图像解译的思维与方法;实现直接对散射辐射场的视觉语义概念生成、推理、决策和交互,改变现有雷达、侦查、干扰对抗的技术形态;实现基于微波视觉的智能自主定位导航,补充传统视觉导航技术。

微波视觉将从基本理论到多种关键技术再到广泛应用的体系化研究,面临多个维度的挑战。理论上需解决微波视觉的根本科学问题,包括如何理解智能与物理世界交互的机制,如何实现高效电磁模拟器,如何使用模拟器构建微波视觉。在技术上需解决技术难题包括大规模场景与目标的快速计算仿真、无限解空间的搜索与学习、视觉的交互学习算法以及而向应用的各种方法等。

2.2 前期研究

复旦大学电磁波信息科学教育部重点实验室长期从事电磁散射机制、微波遥感图像解译和参数反演等研究,前期在微波视觉相关领域特别是合成孔径雷达(synthetic aperture radar, SAR)图像快速仿真、智能解译和目标识别方面做了相关研究工作。

合成孔径雷达能够全天时、全天候、高分辨率成像,因此适用于目标的侦察、监视和识别,同时也已经被广泛应用于地球科学、天气变化与环境系统监测、海洋资源利用、行星探测、战场侦察等领域,是空间遥感的前沿技术。由于其微波波段成像和相位相干处理的特性,SAR图像与光学图像表现不同,难以直观解读。事实上,SAR图像包含了丰富的目标信息,例如几何形状、材质、结构等。SAR图像解译和信息获取是在轨雷达卫星实现成功应用的最后决定性的一步。

遥感大数据时代下的SAR图像解译是一个极大的科学应用挑战。大数据时代的智能方法如计算机视觉技术目前已经取得很大的成功,但是需要发展先进的SAR微波雷达智能信息获取方法。每天都有海量的数据从太空返回而等待处理,需要强大的解译和信息获取方法。

微波视觉的物理基础在于准实时高逼真的电磁物理模拟平台,复旦大学电磁波信息科学教育部重点实验室开发的POLARS EYES仿真平台(图8)采用双向解

析射线追踪算法,并在通用图形处理器上大规模并行实现,可实现准实时的大规模地物场景与目标的SAR

成像仿真,包括上百平方公里的地形起伏、覆盖地物如植被等以及人工目标如建筑物、船只等。

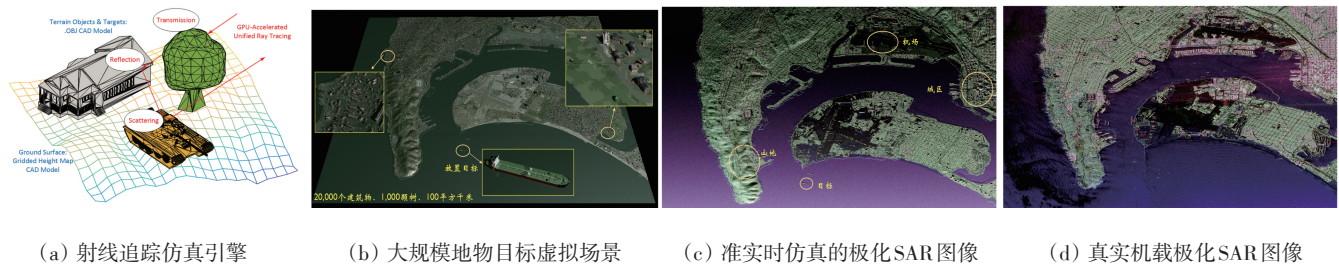


图8 POLSAR EYES 准实时大规模地物场景与地物的SAR仿真

Fig. 8 POLSAR EYES: Near-real time SAR simulation platform for large-scale terrain scene and targets

SAR 图像中自动目标识别是一个关键应用,也是最适合利用机器学习方法实现的。采用卷积神经网络(CNN)对10类地面车辆SAR数据集进行分类,取得较好效果(图9)^[16]。针对SAR图像数据比较少、对观测条件敏感等特点,直接用SAR数据训练CNN很容易出现过拟合的问题。因为CNN自由参数太多,但训练样本不充足,这就导致了严重的过拟合。CNN中绝大部分的可训练参数都包含在全连接层,通过用卷积层取代全连接层,而不是大量地减少网络的层数,减少需要训练的参数。尽管这种改变降低了网络的表示能力,但是也大大减小了过拟合。在标准测试集(观测条件、目标配置几乎不变)上取得99%的分类精度,而在扩展测

试集(观测条件、目标配置发生变化)上分类精度有一定下降但仍比传统方法具有优势。由此可以明显看到现有深度神经网络过度依赖训练数据的缺陷。

将CNN用于多极化SAR地物分类中,发现可以取得比传统方法更好的效果^[17]。对于地物分类而言,其应用模式与图像分割非常类似,因此可以按像素选择局部窗口进行训练和测试,训练样本数量远大于目标识别的情形,但要求训练样本与测试样本的观测条件和地物类型保持一致。同时也测试了CNN地表分类器的通用性,图10显示了2个不同地点不同时间采集的极化SAR图像经过同一分类器得到的结果性能一致,说明了CNN地表分类器具有较好的泛化性能。

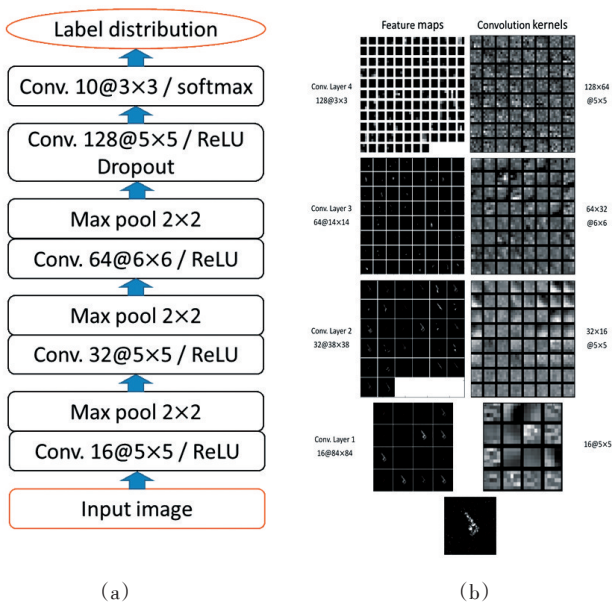


图9 全卷积网络(a)及其在MSTAR数据集训练得到的卷积核和特征图(b)

Fig. 9 (a) AConvNet; (b) Convolutional kernel and feature maps trained on MSTAR dataset

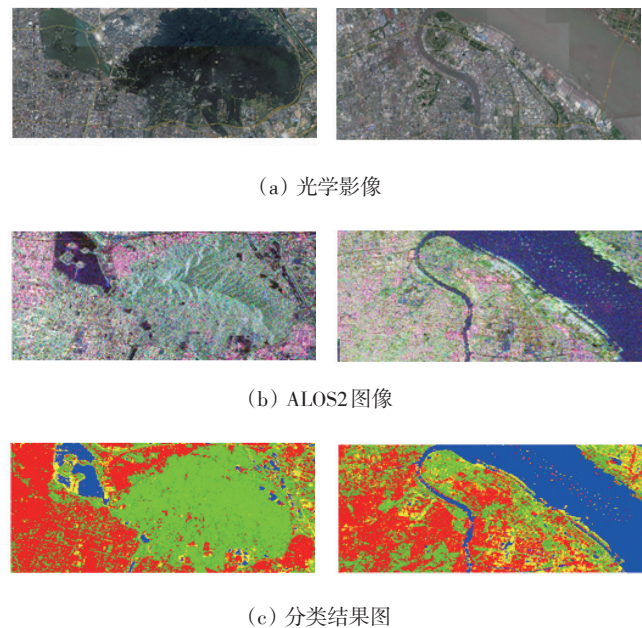


图10 用CNN分类的ALOS极化SAR图像及其分类结果和对比光学影像

Fig. 10 Classification of ALOS POLSAR image with CNN

对于多通道相干的SAR图像,如极化SAR或干涉SAR,其通道间相干相位差带有重要信息。因此针对带相位信息的SAR图像,将CNN推广到了复数域,称为复数卷积网络(complex-valued-CNN, CV-CNN)^[18]。为了区分,实数神经网络可称为real-valued CNN (RV-CNN)。特别是如极化SAR和干涉SAR,均包含了不同通道之间的相位差,该相位信息不再是完全随机的,往往隐含了散射地物的位置、形状、散射机制的信息,这

些信息对地物分类和识别至关重要。CV-CNN不仅将复数数据作为输入,同时对各层传递了相位信息,也即每层的神经元信息和权重均由复数表示,神经元操作以及学习算法均推广至复数域。对比实验结果表明,在相同未知数规模下,采用复数输入按复数计算的CV-CNN和将实部虚部作为2个独立实数输入的RV-CNN相比, CV-CNN能有效提高网络性能,对于极化SAR地表分类中,其错误率可以降低2/3以上(图11)^[18]。

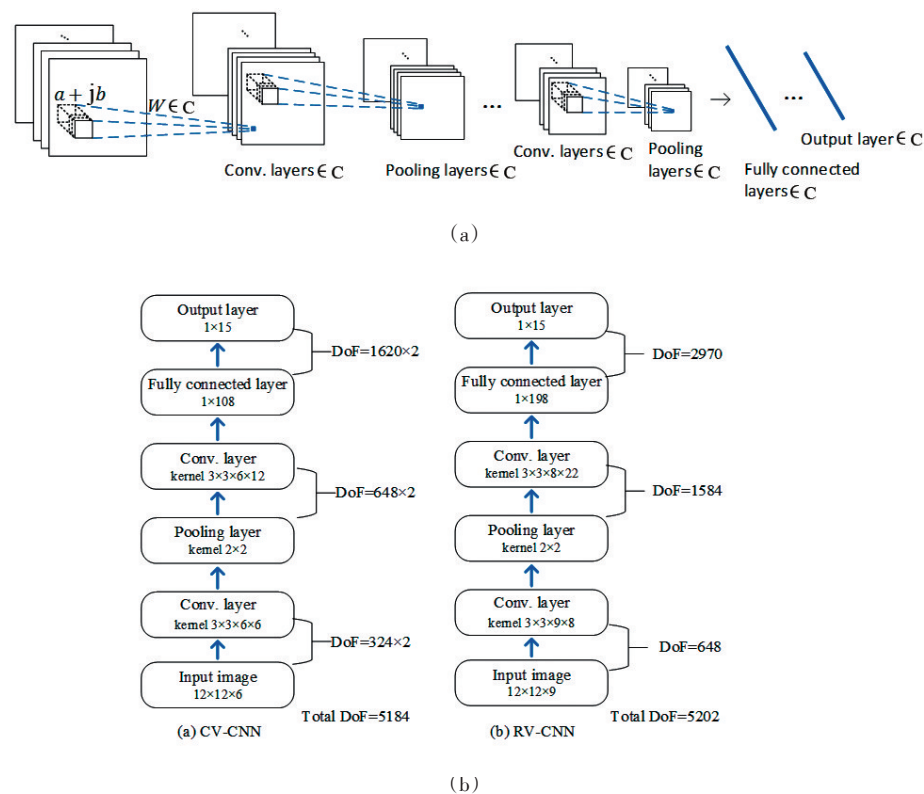


图11 CV-CNN架构(a)及用于等效对比实验的CV-CNN和RV-CNN(b)

Fig. 11 CV-CNN architecture (a) and the comparison between CV-CNN and RV-CNN with similar degrees of freedom(b)

采用神经网络的方法还可以由单极化或双极化SAR图像重构全极化SAR图像(图12)。采用多尺度卷积网络提取单极化SAR图像上的空域特征,通过训练将其转化到极化特征,然后重构出对应的多极化SAR图像,称为对灰度雷达图像填上颜色^[19]。该方法首先利用预训练的CNN从单极化灰度雷达图像中提取多尺度空域特征矢量,然后训练一个全连接深度网络将空域特征空间映射到极化域特征空间,由此得到每个像素的极化散射矩阵。

SAR图像中的相干斑抑制是理解SAR图像的基本处理步骤。笔者还提出了基于卷积网络的SAR相干斑

滤波网络^[20]。如图13所示,将SAR图像邻域窗口输入CNN训练其预测当前像素的真实散射系数的分布,而训练该网络所需要的真实散射系数分布的真值标签难以获得,因此基于相干斑乘积模型,设计了可以非监督训练的目标函数,即将预测散射系数的对数分布与理想相干斑的对数分布进行卷积再使其等于输入SAR图像块像素值的对数分布。该网络架构对于仿真和真实SAR图像进行测试均得到理想效果。

用于训练样本的SAR图像往往并不充分,笔者提出基于生成网络的零样本学习方法,采用生成网络来层次化表征SAR图像特征,由此构建连续的可解释的

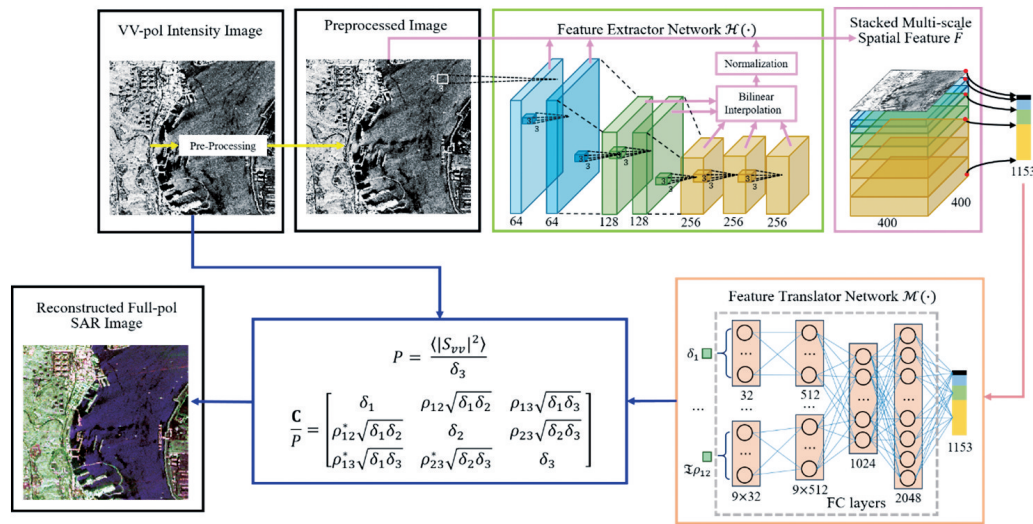


图 12 从单极化 SAR 图像中重构全极化散射矩阵的深度网络架构

Fig. 12 Deep neural networks used to reconstruct full polarimetric scattering matrix from single-pol SAR image

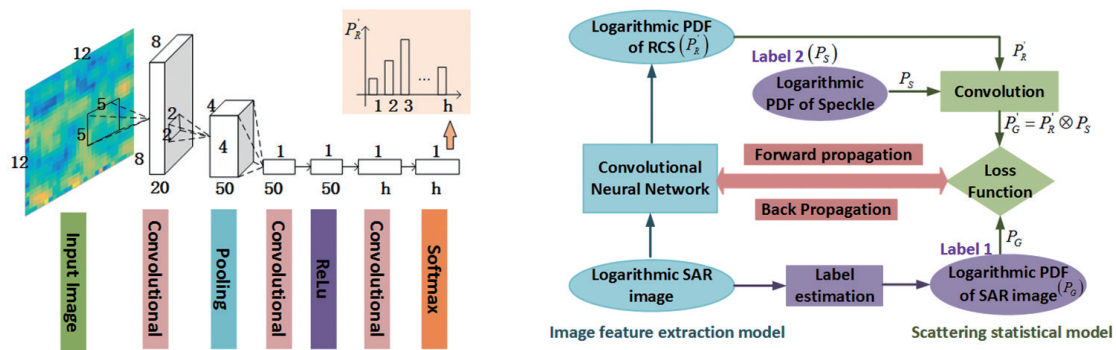


图 13 用于 SAR 图像相干斑抑制的深度卷积神经网络架构

Fig. 13 Deep convolutional neural network for SAR speckle reduction

特征空间,通过将 SAR 图像映射到该特征空间来实现对于未知样本的解译^[21]。图 14(a)为深度生成网络架构,由构造网络、生成网络和解译网络 3 部分组成,分别进行从已知目标标签到特征空间的映射、从特征空间到图像空间的生成,以及从未知目标图像空间到特征空间的映射;图 14(b)给出的即是对 MSTAR 中 7 类已知目标经过生成网络学习后构造的二维特征空间及其中的目标分布。

基于 SAR 目标重构传统的方法,将被观察目标重构成人类可以直接理解的表征,即三维几何形状和一些可识别的特点,重构的几何结构应保留目标识别用

到的关键特征。相对于 SAR 自动目标识别的, SAR 目标重构的 2 个关键优势在于:不仅局限对已知目标库的识别,应对所有目标都适用;因此也不需要对应的训练数据进行训练。这样的方法必须直接从物理模型中提炼出来,因为只有物理规律才具有普适性。提出一个分层框架实现 SAR 目标重构^[22],如图 15 所示,该框架由 2 层组成:基础层,首先提取和识别原始散射机制,包括模型参数的估计;高级层,组合孤立散射体而进一步重构得到完整几何结构。

如何利用人工智能技术进一步发展 SAR 图像智能解译是亟需研究的课题。图 16 给出一种数据驱动和模

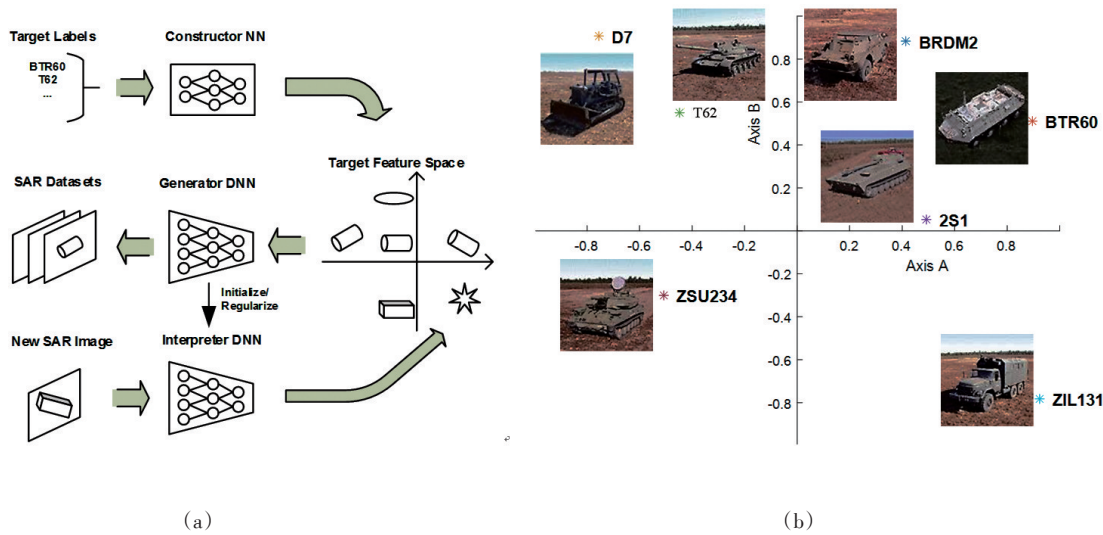


图 14 用于零样本学习 SAR 目标特征空间的深度生成网络架构
 Fig. 14 Deep generative neural network for zero-shot learning of SAR target feature space

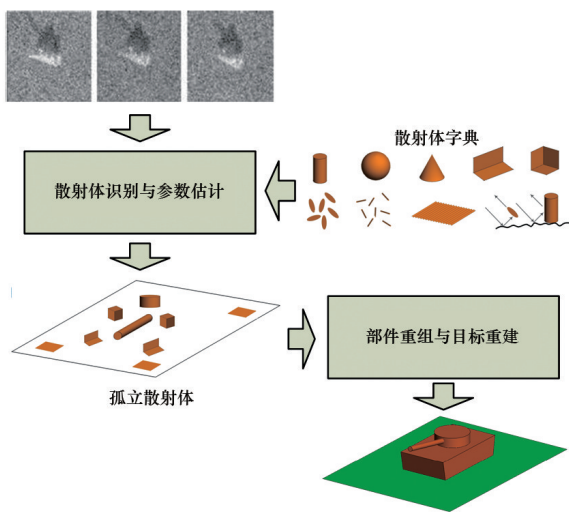


图 15 SAR 目标重构框架
 Fig. 15 Framework for SAR target reconstruction

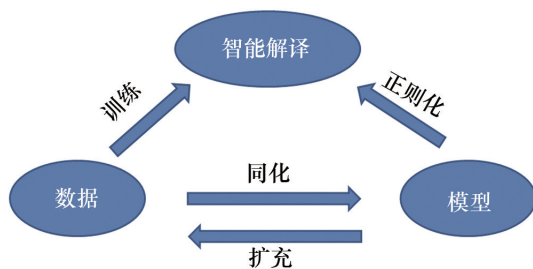


图 16 数据驱动-模型约束下的 SAR 智能解译
 Fig. 16 SAR intelligent interpretation with data driven-model constraint

型约束下的 SAR 智能解译框架^[23],深度学习技术本身必须由海量数据驱动,但是往往人们忽视了模型的作用,也就是将先验知识融合到智能算法中。特别是针对 SAR 图像这样的电磁波散射物理过程的产物,需要融合电磁散射理论和机器学习理论,将物理规律等先验知识体现在智能解译算法中。这一目的可由模型、数据和算法三者协作达成:模型可以产生模拟数据,实测数据可以同化模型中,同时模型可以直接对机器学习算法进行正则化约束。这一思路即是物理智能/微波视觉的一个例子。

3 结论

深度学习引发人类社会进入智能时代,智能时代迫切需要智能科学,然而完整的智能科学体系还不明朗。笔者认为智能科学是与自然科学和社会科学并列的大领域,并不局限于研究人类智能的脑科学。由于智能形成的因素和存在的必要条件,智能科学的范畴应该与自然科学和社会科学是对偶问题,因此三者应该相结合进行交叉研究。例如物理智能是一类典型的智能科学研究,即建立人工智能能适应甚至战胜物理世界。

物理智能是一个重要的智能科学研究领域,通过物理学与智能科学交叉研究能应对物理世界的人工智能。物理智能应与现有物理学紧密结合进行研究,特

别是某些超越人类感官范畴、能力范围的领域,依赖于人类已建立的描述物理世界的精确理论体系,物理智能将超越人类感官范畴、超越人脑计算能力,其中微波视觉技术即旨在研发处理微波频段电磁信息的人工智能技术。物理智能的研究应该与智能科学的进展同时开展,例如在现阶段就可以利用虚拟世界模拟器来训练深度学习技术,从而实现如自动驾驶、智能蜂群或微波视觉的实用技术。

参考文献 (References)

- [1] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507.
- [2] Silver D, Huang A, Maddison C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. *Nature*, 2016, 529(7587): 484-489.
- [3] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[J]. *Advances in Neural Information Processing Systems*, 2012, 25(2): 1097-1105.
- [4] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. *Nature*, 2015, 521(7553): 436-444.
- [5] Fukushima K. Cognitron: A self-organizing multilayered neural network[J]. *Biological Cybernetics*, 1975, 20(3-4): 121-136.
- [6] Jarrett K, Kavukcuoglu K, Ranzato M, et al. What is the best multi-stage architecture for object recognition?[C]//*IEEE International Conference on Computer Vision*. Piscataway NJ: IEEE, 2010: 2146-2153.
- [7] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks[J]. *Journal of Machine Learning Research*, 2010, 9: 249-256.
- [8] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks[J]. *Journal of Machine Learning Research*, 2012, 15: 315-323.
- [9] Choromanska A, Henaff M, Mathieu M, et al. The loss surface of multilayer networks[J]. *Arxiv.org*, 2014, arXiv:1412.0233
- [10] Pearl J. Theoretical impediments to machine learning with seven sparks from the causal revolution[J/OL]. [2018-03-20]. http://www.cse-lab.ethz.ch/wp-content/uploads/2014/09/Judea_Pearl_GREAT.pdf.
- [11] Sabour S, Frosst N, Hinton G E. Dynamic routing between capsules[J/OL]. [2018-03-20]. <http://papers.nips.cc/paper/6975-dynamic-routing-between-capsules.pdf>.
- [12] LeCun Y. Predictive learning[EB/OL]. [2018-03-20]. <https://www.youtube.com/watch?v=Qi1Yry33TQE>.
- [13] Ali Rahimi's talk at NIPS [EB/OL]. [2018-03-20]. <https://www.youtube.com/watch?v=Ount2Y4qxQo>.
- [14] 周翔. Tomaso Poggio: 深度学习需要从炼金术走向化学[EB/OL]. (2018-01-28) [2018-03-20]. http://www.sohu.com/a/219443091_697750.
Zhou Xiang, Tomaso Poggio: Deep learning from alchemy to chemistry[EB/OL]. (2018-01-28)[2018-03-20]. http://www.sohu.com/a/219443091_697750.
- [15] 朱崧纯. 浅谈人工智能: 现状、任务、构架与统一[EB/OL]. (2017-11-02) [2018-03-20]. https://mp.weixin.qq.com/s/-wSYLu-XvOrsST8_KEUa-Q.
- [16] Chen S, Wang H, Xu F, et al. Target classification using the deep convolutional networks for SAR images[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2016, 54(8): 4806-4817.
- [17] Zhou Y, Wang H, Xu F, et al. Polarimetric SAR image classification using deep convolutional neural networks[J]. *IEEE Geoscience & Remote Sensing Letters*, 2017, 13(12): 1935-1939.
- [18] Zhang Z, Wang H, Xu F, et al. Complex-valued convolutional neural network and its application in polarimetric SAR image classification[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2017, 55(12): 7177-7188.
- [19] Song Q, Xu F, Jin Y Q. Radar image colorization: Converting single-polarization to fully polarimetric using deep neural networks[J]. *arXiv.org*, 2017, arXiv:1707.07225.
- [20] Yue D, Xu F, Jin Y Q. Deep despeckling neural network for SAR images[J]. *International Journal of Remote Sensing*, 2018, in press.
- [21] Song Q, Xu F. Zero-shot learning of SAR target feature space with deep generative neural networks[J]. *IEEE Geoscience & Remote Sensing Letters*, 2017, 14(12): 2245-2249.
- [22] Xu F, Jin Y Q, Moreira A. A preliminary study on SAR advanced information retrieval and scene reconstruction[J]. *IEEE Geoscience & Remote Sensing Letters*, 2016, 13(10): 1443-1447.
- [23] 徐丰, 王海鹏, 金亚秋. 深度学习在 SAR 目标识别与地物分类中的应用[J]. *雷达学报*, 2017, 6(2): 136-148.
Xu Feng, Wang Haipeng, Jin Yaqiu. Deep learning as applied in sar target recognition and terrain classification[J]. *Journal of Radars*, 2017, 6(2): 136-148.

From the emergence of intelligent science to the research of microwave vision

XU Feng^{1,2}, JIN Yaqiu^{1,2}

1. Institute of Electromagnetic Data Science and Remote Sensing Intelligence, Fudan University, Shanghai 200433, China

2. Key Laboratory for Information Science of Electromagnetic Waves (MoE), Fudan University, Shanghai 200433, China

Abstract In the past decade, artificial intelligence has attracted a great attention from both academics and industry, suggesting the transition of the human civilization into a new intelligence era. However, the "intelligent science" has not yet developed into the science foundation of the intelligent age. This paper, from a specific viewpoint of electromagnetic information sensing, discusses how the intelligent science might be developed in future, and points out that the human intelligence and the external world form a pair of inseparable dual problems, as the fundamental nature of the intelligent science. Therefore, the development of the artificial intelligence goes through four stages according to the objects and the related disciplines, namely, the mathematical intelligence, the physical intelligence, the psychological intelligence, and the free consciousness. In the first stage, the mathematical theory of a general learning algorithm is developed to form the basis of the intelligence. In the second stage, the physical intelligence is developed to deal with the physical world. Based on this, the third stage of development responds to high-order intelligence involving social groups, while in the fourth stage, the nature of the free consciousness and the question whether the artificial intelligence can finally become a conscious super-intelligence are studied. Within the author's own research field of the electromagnetic information sensing, a new concept of the microwave vision is proposed along the direction of the physical intelligence to explain its key ideas and essential problems and to paint the roadmap of its technological development. Taking the previous studies of the authors' team in this direction as an example, we further discuss the research and development of the physical intelligence as one of the specific directions of a general intelligent science.

Keywords intelligent science; physical intelligence; microwave vision ●



(责任编辑 刘志远)