

## 人工智能技术发展概述<sup>1</sup>

俞祝良 华南理工大学自动化科学与工程学院

近年来，以深度学习为核心的人工智能技术，取得了一系列的重大突破。本文将就人工智能的产业化热潮，主要研究流派及发展历史，以深度学习为核心的成功应用，以及存在的一些问题和今后的可能研究方向进行介绍。

### 一、前言

从1956年达特茅斯会议首次定义“人工智能”（Artificial Intelligence, AI）开始，AI研究经历了几次历史浮沉。在一次又一次的高潮和低谷的交替中，不可否认，AI无论是在理论和实践上都取得了扎实的进步，人类对于智能的理解进一步加深。尤其是近期以深度学习（Deep Learning, DL）为代表的AI技术取得了突破性的进展，从而在全世界范围内掀起了又一个AI研究热潮。与以往不同的是，这次的研究热潮同时伴随着AI商业化浪潮，实验室成果很快就进入工业界，甚至工业界在这股热潮中也站在了学术研究的前沿，这在以往的技术发展史上是非常罕见的。

2015年7月，人工智能被写入《国务院关于

积极推进“互联网+”行动的指导意见》；2016年3月，人工智能一词被写入“十三五”规划纲要；2016年5月，国家发改委等四部门联合下发《“互联网+”人工智能三年行动实施方案》。李克强总理的政府工作报告中也提到了人工智能产业发展。中国科学技术部“科技创新2030—重大项目”近期或将新增“人工智能2.0”，人工智能将进一步上升为国家战略。这可以充分体现我国对AI的重视程度。2017年，中国工程院院刊信息与电子工程学部分刊《信息与电子工程前沿（英文）》出版了“Artificial Intelligence 2.0”专题，潘云鹤等多位院士及专家学者对AI 2.0所涉及的大数据智能、群体智能、跨媒体智能、混合增强智能和自主智能等进行了深度阐述。

<sup>1</sup> 本研究受广州市科技计划项目产学研协同创新重大专项，“金融智能服务系统关键技术攻关”（201604016113）和“面向未来银行的智慧金融平台关键技术研发及产业化”（201604046018）资助。

面对人工智能热潮，我们该如何理解、看待其进步，如何了解其功能和限制呢？虽然现在已经有不少书籍<sup>[1]</sup>和论文<sup>[2]</sup>讨论了上述问题，本文将从人工智能的产业化浪潮、学术流派和研究方法，以深度神经网络为主的AI发展历史，近期成果和问题等诸多方面对人工智能进行介绍。希望能对读者了解AI有所帮助。

注：达特茅斯会议上定义的人工智能是指用计算机模拟人的逻辑思维，实际上这个定义比较适合基于符号逻辑的演绎系统（符号学派），如专家系统等。实际上人类还有归纳总结能力（联接学派）。严格讲，这不包括在狭义的人工智能当中，所以神经网络、模糊逻辑和遗传算法等结合构成了“计算智能”的范畴。为了便于表述，我们文中采用广义的人工智能，这个概念和“机器智能”<sup>[1]</sup>一致，包含一切机器具有的智能。

## 二、人工智能商业化浪潮

20世纪末，当以神经网络为主流的AI研究又一次跌入低谷的时候，加拿大多伦多大学的Hinton教授还是坚守阵地，辛勤耕耘，并在2006年获得了突破<sup>[3,4]</sup>。2012年他和两位学生成立“深度神经网络研究”（DNN Research）公司，数月后被Google收购，从此Hinton教授身兼多伦多大学教授和Google研究者的双重身份。Google随后斥资4亿美元收购初创的人工智能前沿企业DeepMind。另外，Google还收购了乌克兰面部识别技术开发商Viewdle。

紧随Hinton教授的步伐，纽约大学Yann LeCun教授2013年底被聘请为Facebook人工智能研究院的主任。斯坦福大学吴恩达（Andrew Ng）教授2014年被百度聘任为首席科学家，负责百度大脑的计划（2017年已经辞职）。斯坦福大学李飞飞教授（Fei-Fei Li）成为谷歌云计算部门的负责人之一。这些现象一方面说明了人工智能现在

受工业界的欢迎程度，同时也说明了人工智能目前的发展趋势是学术研究和企业开发快速深度结合。

为了迎合AI的发展热潮，大量的开源学习平台不断问世，贾扬清的Caffe、Google的TensorFlow、FaceBook的FBLearner Flow、Tesla领衔的OpenAI、百度深度机器学习开源平台等等，都为AI的研究和产业开发起到了巨大的推动作用。当然还有很多其他工业巨头的产品，例如IBM的沃森系统、微软的同声翻译等等。

科技发展，人才为本。AI也不例外，目前企业对于AI人才的抢夺更是趋于白热化。目前AI领域，尤其是深度学习领域面临着缺乏专家的困境。这个领域刚刚开始发展，所以专家、即使是博士生都特别少。吴恩达教授曾总结深度学习领域人才匮乏的几个原因：首先是数据，获取解决某些领域的问题的数据常常非常困难；其次是计算基础和架构工具，包括计算机硬件和软件，入门不易；最后是这个领域的工程师培养时间长。为了解决上述问题，工业界的科技巨头，如Google、Facebook、Twitter、百度等纷纷通过收购深度学习领域的初创公司来招揽人才。其中最典型的是Google，它通过不断收购深度学习领域的公司，抢到一批世界一流专家。总而言之，人工智能产业的发展，使得其相关领域的人才成为稀缺之宝，对该领域的研究人员来讲，既是机遇，也是挑战。

## 三、人工智能的主要研究学派

人类的智能主要包括归纳总结和逻辑演绎两大类。我们大量的感知处理，如视听觉、身体感知处理等都是下意识的，属于归纳总结智能。而数学推导，逻辑推理等都是基于公理系统的符号演绎方法。由于在发展过程中对于智能的理



解不同，渐渐形成了几个经典学派。每个学派从不同的角度看待问题，提出解决方案。比如最为主要的两个学派，符号学派和联结学派，前者从哲学、逻辑学和心理学出发，将学习视为逆向演绎，使用预先存在的知识来解决问题，大多数专家系统使用符号学派的方法；后者专注于通过神经元之间的连接来推导表示知识，该学派聚焦于物理学和神经科学，并相信大脑的逆向工程，他们用反向传播算法来训练人工神经网络以获取结果<sup>[5]</sup>。其他学派，如进化学派在遗传学和进化生物学的基础上得出结论，贝叶斯学派注重统计学和概率推理，类推学派更多是关注心理学和数学优化来推断相似性判断。

虽然上述主流学派各自都取得了很大的成就，但是其各自采用的研究方法都遇到了诸多困

难。而且这些学派对于AI的研究思路和方法难以形成一个统一的框架。为了更好地理解AI的本质。本文拟采用<sup>[1,5]</sup>中的观点来介绍人工智能的主流研究方法。

基于对“机器智能由什么决定”这个问题的回答，人工智能领域七十多年的发展形成三大主流研究方法：结构模拟、功能模拟，和行为模拟。这三种研究方法之间缺乏一种内在的联系，使得人们普遍认为AI理论不成体系。为了解决上述问题，[1,5]提出了第四类方法，即机制模拟。

近代科学强调“结构决定论”，认为只要系统的结构清楚了，功能也就认识清楚了。最先提出来的智能模拟就是结构模拟的思路，其代表性成果有神经元的MP模型<sup>[6]</sup>，多层感知器MLP模型<sup>[7]</sup>和人工神经网络<sup>[8]</sup>等等。在机械系统的研究中非

常有用的结构决定论，在智能系统的研究中，却存在很多问题。结构只是硬件基础，不能完全确定系统的智能行为。正如现在对于单个神经元的生理活动机制已经研究得非常深入，但是对大量的神经元连接在一起所呈现的智能行为，目前的理解还是非常肤浅。作为结构模拟的主流成功方法，人工神经网络具有非常突出的优点：人工神经网络具有比较规范的结构；系统具有大量可以调节的参数，自由度大，可以实现非常多样的系统；网络具有并行处理的机制；其信息是分布存储，提供了优良的记忆和联想能力；系统具有很强的自适应能力和高度的容错能力。当然人工神经网络也具有很多弱点和局限性：人的智能系统不仅结构复杂，而且机制深奥，不是现在的人工神经网络所能完全模拟的；人工神经网络能做到的规模和人脑相比还是太小；神经元的工作机理过度简化，难以保证系统能最终模拟人脑。还有人工神经网络将一切智能都归为数值计算，在未能回答任何问题是否都可以形式化或者数值化这个问题之前，这种数值化计算能否完全模拟人类智能还是一个问题。

面对结构模拟所存在的问题，另外一类观点认为，人工智能的研究无需去理会智能的具体结构，只要能够模拟智力功能即可，这就是“功能主导论”下的功能模拟思路。实际上，功能模拟的最典型代表就是传统的人工智能，比如专家系统<sup>[9]</sup>就是其最成功的应用。功能模拟也常常被称为符号主义、逻辑主义、心理学派。符号主义认为人工智能源于数理逻辑。数理逻辑在20世纪30年代开始应用于描述智能行为，并在计算机上实现逻辑演绎系统。后来由符号主义者进一步发展为启发式算法→专家系统→知识工程理论和技术。这方面的研究一开始取得了不少成绩，但是一直被批评为难以解决实际问题。直到专家系统出

现，为工业界、经济和社会领域带来了成功的方案，如第一个专家系统DENDRAL<sup>[9]</sup>用于质谱仪分析有机化合物的分子结构，MYCIN医疗专家系统<sup>[10]</sup>用于抗生素药物治疗等等。符号主义曾经在人工智能领域中一枝独秀，为人工智能发展做出极大贡献。我国吴文俊院士关于几何定理机械化证明的工作就是其中一项非常重要的成果。当然，功能模拟也具有明显的缺点：系统的智能水平与获得的知识水平有很大的关系。而且很多知识获取困难，专家知识充满矛盾和片面，再加上现有逻辑理论的局限性等，使得功能模拟在发展过程中也困难重重。

在功能模拟和结构模拟都暴露了各自的缺陷后，20世纪90年代开始出现了行为模拟的思路，即“行为表现论”。该观点认为，无论采用什么样的结构和具有什么样的功能，只要系统能表现出智能行为（在外界刺激时能够产生恰当的行为响应），就等于模拟了智能系统。在这个研究方法中，首先是机器感知，然后针对感知信息做模式分类，最后是对感知判断结果做出的模拟智能的行为，这也被称为感知-动作系统。这方面的典型应用是Brooks完成的模拟六脚虫的爬行机器人等系统<sup>[11]</sup>。感知-动作系统涉及到一个重要问题：随着任务环境的变化，系统如何能自主学习并扩充从感知到动作之间的映射知识。这当然就是一个机器学习问题。行为模拟也具有明显的缺点：只有那些能用行为表现智能才能被模拟，可是很多智能过程无法用行为直接表达。

综上所述，结构模拟、功能模拟和行为模拟都具有先天不足，而且这三大方法之间缺乏理论上的统一性。后来的研究发现，智能的生成机制才是智能系统的核心。机制模拟方法认为，无论对什么问题、环境和目标，智能系统的生成机制必然要获得“问题、约束条件、预设目标”等信



息，然后提取和建立相关知识，进而在目标控制下，利用上述信息和知识演绎出求解问题的策略，并转化为相应的智能行为作用于问题，并解决问题<sup>[1,5]</sup>。这可以概括为一种“信息-知识-智能”转换过程。根据这个观点，结构模拟可以认为是“信息-经验知识-经验策略”转换过程，功能模拟是“信息-规范知识-规范策略”转换过程，行为模拟可以认为是“信息-常识知识-常识策略”转换过程。因此，结构模拟、功能模拟、行为模拟三者都是平行的，而机制模拟和谐地统一了上述三种模拟方法，成为了一个统一的理论。

在AI发展过程中，上述多个方法各自都出现过自己的发展巅峰和低谷时期。目前的AI热潮则源于结构模拟方法方面的突破，即由于解决了深度神经网络的训练问题，加上大数据的高性能计算平台（云计算，GPU等）变成现实，使得深度

神经网络的表达能力得到了充分的发挥，对AI的发展起到了推波助澜的作用。本文将进一步以深度学习为主介绍其发展和成功案例。

#### 四、神经网络发展过程及深度学习的兴起

让机器具有智能，是人类一直的梦想，但是实质性的进展却是在20世纪50年代开始的。在McCulloch和Pitts神经元计算模型工作<sup>[6]</sup>的基础上，康内尔大学的Rosenblatt提出了感知器（Perceptron）模型<sup>[7]</sup>。在感知器研究中，Rosenblatt受到Hebb工作的启发，提出一套算法来精确定义网络的学习规则，这也是首个具有自组织自学习能力的模型。Hebb认为知识和学习在大脑中主要是通过神经元间突触的形成与变化来实现。感知器通过调整针对输入值的权重，利用一个非常简单直观的学习方法，从输入数据上实现学习功能。Rosenblatt还用定制硬件的方法实现了感知器，可以用来学习并对 $20 \times 20$ 像素输入中

的简单形状进行正确分类。自此，机器学习问世了。感知器不仅是日后许多新的神经网络模型的始祖，还为20世纪60年代带来了人工智能的第一个热潮。

这股热潮同时迎来了激烈的批评。1969年，Minsky和Papert在名为《感知器》的书中提出了强烈的批判。他们认为单层的感知器网络无法解决非线性可分问题（如异或门，XOR问题）。另外，网络模型所需的计算量也超出了当时计算机的能力。学术界普遍认为这本书对人工智能步入第一个冬天起到了推波助澜的作用——人工智能进入泡沫幻灭期，相关资助和出版都遭冻结。

Minsky和Papert关于感知器的分析证明用单个感知器无法解决XOR问题，指出必需要多层感知器网络（所谓的多层神经网络）才可以完成任务，而且Rosenblatt的学习算法对多层网络无用。这个问题最终的解决方案就是著名的反向传播算法<sup>[12,13]</sup>，该方法让研究者广泛理解了应该如何训练多层神经网络来解决复杂学习问题，其中包括非线性可分问题。通过在神经网络里增加一个或者多个隐层，可以使得多层神经网络具有非常强的解决复杂问题的能力。更有趣的发现是数学能证明，多层前馈神经网络是普适模拟器（Universal Approximator），本质上，多层结构使得神经网络能够在理论上执行任何函数表达，当然包括XOR（异或）问题<sup>[14,15]</sup>。Minsky提出的计算量问题也很快得到了解决，传统的感知器用所谓“梯度下降”的算法纠错时，其运算量和神经元数目的平方成正比，Hinton和Rumelhart在合作论文<sup>[13]</sup>中，系统地提出了应用反向传播算法，把纠错的运算量下降到只和神经元数目成正比的方法。Hinton和其博士后Yann LeCun于1989年采用美国邮政系统提供的近万个手写数字的样本来训练神经网络系统，在独立的测试样本中错误率低至5%，达到

了实用水准<sup>[16]</sup>。随后LeCun进一步运用“卷积神经网络”（convoluted neural networks）的技术<sup>[17]</sup>，开发出商业软件，用于读取银行支票上的手写数字，获得了巨大的成功。神经网络掀起了第二次热潮。

但是很快研究者发现了反向传播算法具有本质缺陷——梯度消失（Vanishing Gradient Problem）<sup>[19]</sup>，也就是说，神经网络的代价函数（Cost Function）的误差从输出层向输入层反向传播时，梯度衰减极快，学习速度变得极慢，甚至无法学习，神经网络很容易停滞于局部最优解。这使得理论上可以学习任意函数的多层神经网络在实用中无法实现。同时，多层网络由于具有较多的参数，学习自由度大，算法训练时会出现过度拟合（Overfitting）问题，使得学习过程中表现良好的网络的泛化误差很大，无法真正应用。在多层网络遭遇上述重大问题的时候，贝尔实验室的Vapnik提出了支持向量机（Support Vector Machine, SVM）的算法<sup>[19]</sup>，通过使用所谓“核机制”（Kernel Trick）的非线性映射，使得本来线性不可分的样本映射到线性可分的高维特征空间中。从20世纪90年代初开始，SVM在图像和语音识别上找到了广泛的应用。由于其理论完备、机理简单受到了研究者的追捧。上述多方面的作用下，神经网络的研究再一次进入了冬天。

从上述分析可以看出，神经网络研究再一次进入冬天有其本质性的问题，如梯度消失、过度拟合和计算量大等等。在这个冬天中，有些研究者如Hinton等依然相信问题能获得解决，并坚守这块阵地。经过多年的努力，最终迎来了神经网络的复兴，并且开启了深度学习的大门。在神经网络的训练上，无监督学习被引入网络的初始化中。在论文<sup>[3]</sup>中，Hinton利用限制玻尔兹曼机（RBM）对神经网络实现无监督训练

(Unsupervised Training)。RBM从输入数据中通过无监督训练发现重要特征，对神经网络的权重进行有效的初始化。然后将多层RBM叠加在一起形成深度网络，再对整体网络用反向传播算法进行微调，取得了很好的效果。另外在论文[20]中提出了用一种“修正线性单元”(Rectified Linear Unit, RELU)的转换函数来替代传统神经单元的非线性转换函数。RELU函数简单，而且其导数为常数，输入小于零时为0，大于零时为1，不存在传统转换函数在反向传播计算中的梯度消失问题。基于无监督学习的初始化和RELU函数的采用，很好地解决了多层网络的训练问题。

为了解决多层网络训练中的过拟合问题，Hinton在[21]中提出了一种称为丢弃(Dropout)的算法。在每次训练中，以一定的比例，在训练中忽略这部分神经元。该算法可以使得网络变得更鲁棒，避免过度拟合。另外，由于大数据变成了现实，将大数据用于训练，也是解决过度拟合的一个重要途径。例如现在的ImageNet数据库，就可以提供高达千万张图片用于训练。大数据成了深度学习发展的一个重要基础。

当多层神经网络的训练和过拟合得到了解决后，理论上我们可以利用非常深的网络来学习实现任意复杂的函数。现在的深度网络往往具有百万到数亿的学习参数，而且为了实现这样的网络训练，往往需要海量的数据，计算量自然更是个问题。幸运的是，随着云计算的普及，我们有大量的计算资源可以使用。而且斯坦福大学的吴恩达发表论文<sup>[22]</sup>，通过用GPU实现大规模学习，实现了数十倍到数百倍的速度提升。从此人类迎来了神经网络的大复兴，并且打开了深度学习的大门。关于深度学习更完整的论文资料请参考[23]。

如果我们认为深度学习的提出仅仅是由于算

法的改进和计算能力的增强，那对深度学习的本质还是缺乏认识。实际上，深度学习的提出和现在取得的突飞猛进的进展，有其更加重要而深刻的思想变化。在脑科学研究中发现，大脑不同的功能区域专门负责同一类任务，例如视觉图像识别、语音信号处理和文字处理等等。因此科学家为不同的任务开发不同的算法，如Gabor滤波器、SIFT特征提取算子、马尔科夫随机场等等来提取信号的特征，最终用于模仿大脑功能。但[24, 25]中的研究表明，大脑实际上是一台通用学习机器(Universal Learning Machine)，同样的学习机制可以用于完全不同的任务，不同的脑功能区可以转换，而且转换过程中能自动学习特征，大脑的神经网络具有极强的可塑性。大脑学习算法的普适性和可塑性一直激励着计算机科学家不懈地努力探索。历史性的突破发生在2006年<sup>[4, 26]</sup>，Hinton等突破深度学习的技术瓶颈，进而引领了深度学习的热潮。更有意思的是通过深度学习获得的特征居然和大脑的视觉处理过程非常相似，从而证明深度学习从某个方面已经对大脑的学习机制做了很好的模仿，从数据中学习到功能。

## 五、深度学习推动的人工智能发展

深度学习提出至今，已经在各类应用上取得了巨大的进展。尤其是基于深度学习的AlphaGo系统一举击败韩国棋手李世石以后，人们都对以深度学习为主的AI研究充满了期待。事实上，现在几乎每天都可以看到AI取得各类突破的报道。感兴趣的读者可以在网上找到各种最新的进展报道。本文将选择其中几个主要进展进行简单介绍。

深度学习在实际应用中的成功案例首推图像识别。2009年，普林斯顿大学建立第一个超大型图像数据库供计算机视觉研究者使用<sup>[27]</sup>，随后在



以ImageNet为基础的大型图像识别竞赛ImageNet Large Scale Visual Recognition Challenge 2012中，Hinton团队将深度学习应用到ImageNet图像识别问题上，正确率稳居第一，并且性能遥遥领先第二名团队。这标志着深度学习在图像识别领域大幅度超越其他技术，成为AI技术突破点。随后以深度学习为主的图像分析处理方法层出不穷，比如深度残余学习（Deep Residual Learning）方法等。目前基于深度学习的图像识别功能已经超越了人类。

另外，Kaggle网站举办了一场在卫星图像上进行场景特征检测的比赛，数据集由英国国防科学与技术实验室（DSTL）提供。卫星领域产生的大量图像数据，非常适合用深度学习方法去处

理，而且最终的竞赛结果发现，深度学习方法取得了非常优异的性能。在图像理解方法这方面，深度学习方法结合递归神经网络（RNN），就可以实现图像的自然语言层面理解。斯坦福大学李飞飞团队结合卷积网络和递归网络实现了图片标题的自动生成。

语言是人机交流的一种重要途径，攻克语音识别是AI必须面对的问题之一。最先开始在语音识别上取得成功的深度学习方法是Hinton、邓力等发表在[28]中的方法，该方法用RBM对神经网络进行预训练，再用深度神经网络模型(DNN)识别语音。在Google的一个基准测试中，单词错误率降低到12.3%。文献[29]使用RNN/LSTM等技术，在音位错误率测试中的表现优于同期的所有其他



技术。AI在语音识别上的成功是继图像识别之后的又一个技术突破点。

疾病诊断，一直是医生的专利，也是一个神秘的专业领域。现在通过深度学习，我们可以让机器到达医学专家的诊断水平。论文[30]中报道了利用深度学习诊断皮肤癌的工作。我们知道皮肤癌是人类最常见的恶性肿瘤之一，目前主要是通过视觉诊断的。论文使用深度卷积神经网络进行皮肤病变分类，该网络仅使用图像和疾病标签作为输入，学会正确分类。他们在两个二分类任务：角质形成细胞癌（Keratinocyte Carcinomas）和良性脂溢性角化病（Benign Seborrheic Keratoses），就是恶性黑色素瘤和普通的痣上进行了测试，发现深度卷积神经网络在这两个任务上的表现都达到了所有测试中的专家的水平，这证明了在皮肤癌诊断问题上，机器达到了皮肤科专业医生的水平。最新论文[31]中报道，科学家使用能够自学习的AI技术，让计算机在预测心脏病的发作上击败了人类医生。该技术一旦投入使用，这一新的诊疗手段每年将拯救数以千计甚至百万计的生命。

除此之外，深度学习在游戏方面也取得了突破。最为世人所称道的是AlphaGo对局李世石的比赛<sup>[32]</sup>，AlphaGO以压倒性的胜利赢了人类顶级棋手。其中很多精妙的招式让人叹为观止。到底是深度学习算法已经像人类一样具有了创造力还是神经元参数、适当的算法结合CPU的计算蛮力下的成功，值得人们深思。Google的DeepMind团队开发的深度Q网络（DQN）在49种Atari像素游戏中，29种达到乃至超过人类职业选手的水平<sup>[33]</sup>。

机器在计算能力、存储空间和检索速度方面优势巨大，如果说机器在逻辑思维上能模仿并击败人类那是合情合理的。但是如果说机器能学会艺术鉴赏，那将令人难以置信。实际上，论文[34]

中报道了深度学习已经可以分辨不同艺术作品的内容和风格，并且学会把一幅作品的艺术风格应用到其他作品中去。Google的AI实验室成功实现让机器学习小说的叙述和文字用法，甚至句子的构造。这意味着深度学习可以量化很多人类独有的一些模糊的判决能力，有些感觉人类本身也不清楚如何解释，甚至只是停留在感觉层面，如艺术风格、文采特色、棋类游戏棋中的棋面优势等。这些本来只可意会，不可言传的概念居然被机器学会了。到底是机器学习厉害，还是这些虚幻的概念后面其实有个虽然非常复杂，但是可以获得的函数在支配，而深度学习只是用其蛮力去把这些函数反映到了数亿冰冷的参数中去？

## 六、人工智能发展面临的主要问题和新的研究方向

AI概念提出以后，迅速发展成为一门广受关注的交叉和前沿学科，在很多方面取得了丰硕成果。尤其是近年来深度学习的发展，更是将AI研究推向一个新的高潮。深度学习的成功故事每天都在上演，似乎深度学习是无所不能的。所以，人们对AI产生了浓厚兴趣，产业界首先布局，大量资本与并购的涌入并加速AI技术与应用的结合，使其蔓延升温。那么，抛开商业因素，AI在技术上是否还存在什么问题？它今后的发展方向是什么？

深度学习技术和早年的多层神经网络相比，理论上的突破其实不大。深度学习存在一个根本的缺陷，就是缺乏完善的理论支撑，对于其结果缺乏解释，包括网络的设计、参数的设置和取得的效果。寻找理论上的深入解释，是深度学习发展必须解决的一个重要问题。

目前深度学习一般都需要大数据支持，但不是所有的应用都具备大数据条件。结合传统知识表达和数据驱动知识学习，可以解决很多迫切的

现实问题。这也是今后发展的一个重要方向。

还有，深度学习在训练网络中需要大量有标记数据去学习输入和输出的映射关系。这样获得的模型往往无法将其泛化到与训练时不同条件的数据集上。而现实应用中，我们遇到的数据集常常会包含很多新场景，许多数据是模型在训练过程中没遇到过的，因此学习得到的模型可能无法很好地预测结果。就将学习得到的知识迁移到新的条件和环境的能力通常被称为迁移学习，这是今后一个重要的研究方向。如果我们将迁移学习做到极限，仅仅从少数几个甚至零个样本中学习（如一次和零次学习），就将能解决更多实际问题。执行一次和零次学习的模型是机器学习中最难的问题之一。可这对我们人类而言却不是那么困难。这是AI发展一个值得深入研究的问题。

另外还有一个非常值得进一步思考的问题：是否非得要采用深度神经网络进行学习？如果所考虑问题的数据和目标之间的函数关系比较简单，那么我们完全可以用浅度的网络进行建模学习。但是如果这个函数的确比较复杂，是否一定要用深度网络呢？针对这个问题，南京大学周志华教授提出一种基于树的方法，叫深度森林<sup>[35]</sup>，来挑战深度学习。在设置可类比的情况下，深度森林取得了和深度神经网络相当甚至更好的结果，而且更容易训练，小数据也能运行。更重要的是相比深度神经网络，基于树的方法不会存在那么困难的理论分析问题。该方法为在许多任务中使用深度神经网络之外的方法打开了一扇门。

## 七、结论

到目前为止，人工智能的研究依然还处于初级阶段，距离最终的目标还有很长的路要走。深度学习方法取得了巨大的进展，但是没有坚实的理论基础，无法实现对系统和性能的透彻理解和预测。还有很多的问题摆在我们面前，如迁移学

习、小样本学习、增强学习等等，这些都是人工智能研究急需解决的问题。虽然人工智能在很多方面已经赶超人类，但是本质上离真正的智能还是有很大的距离，这也是我们进一步期待和努力的方向。

## 参考文献

- [1] 李蕾与王小捷. 机器智能. 清华大学出版社, 2016.
- [2] 顾险峰. 人工智能的历史回顾和发展现状. 自然杂志, 2016(03): 第157–166页.
- [3] Hinton G E, Osindero S and Teh Y. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, 18(7): 1527–1554.
- [4] Hinton G E and Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, 313(5786): 504–507.
- [5] 钟义信. 高等人工智能原理—观念、方法、模型、理论. 2014: 科学出版社.
- [6] McCulloch, et al. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 1990, 52(1): 99–115.
- [7] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 1958, 65(6): 386–408.
- [8] [8]. Haykin S. *Neural Networks: A comprehensive Foundation*. Prince Hall, 2004.
- [9] Lindsay R K, et al., DENDRAL: A case study of the first expert system for scientific hypothesis formation. *Artificial Intelligence*, 1993, 61(2): 209–261.
- [10] Shortliffe E H. Mycin: a rule-based computer program for advising physicians regarding antimicrobial therapy selection. *Stanford University*, 739–739, 1975.
- [11] Brooks R A. Intelligence without representation. *Artificial Intelligence*, 1991. 47(1–3): 139–159.
- [12] Rumelhart D E, Hinton G E and Williams R J. Learning internal representations by error propagation. *Readings in Cognitive Science*, 1988, 399–421.
- [13] [13]. Rumelhart D E, Hinton G E and Williams R J. Learning representations by back-propagating errors. *Nature*, 1986, 323: 533–536.
- [14] [14]. Hornik K. Approximation Capabilities of Multilayer Feedforward Networks. *Neural Networks*, 1991, 2(4): 251–257.
- [15] Csaj B C. Approximation with Artificial Neural Networks. Eotvos Lorud University, Hungary, 2001..
- [16] Lecun Y et al. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1989, 1(4): 541–551.
- [17] LeCun Y et al. Gradient-based learning applied to document recognition. *Proc. IEEE*, 1998, 86(11): 2278–2324.
- [18] Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 1998, 6(2): 107–116.
- [19] Vapnik V. *The Nature of Statistical Learning Theory*. Springer, 1999.
- [20] Glorot X, Bordes A and Bengio Y. Deep Sparse Rectifier Networks. *Journal of Machine Learning Research*, 2011, (15): 315–323.
- [21] Hinton G E, et al. Improving neural networks by preventing co-adaptation of feature detectors. *Computer Science*, 2012, 3(4): 212–223.
- [22] Raina R, Madhavan A and Ng A Y. Large-scale deep unsupervised learning using graphics processors. in *International Conference on Machine Learning, ICML 2009*. Montreal, Quebec, Canada.
- [23] 深度学习论文阅读路线图. <https://github.com/songrotek/Deep-Learning-Papers-Reading-Roadmap>
- [24] Sharma Angelucci J A and Sur M. Induction of visual orientation modules in auditory cortex. *Nature*, 2000, 404(6780): 841–847.
- [25] Vuillerme N and Cuisinier R. Sensory supplementation through tongue electrotactile stimulation to preserve head stabilization in space in the absence of vision. *Investigative Ophthalmology & Visual Science*, 2009, 50(1): 476–481.
- [26] LeCun Y, Bengio Y and Hinton G. Deep learning. *Nature*, 2015, 521(7553): 436–444.
- [27] Deng J, et al. ImageNet: A large-scale hierarchical image database. in *IEEE Conference on Computer Vision and Pattern Recognition, 2009, CVPR 2009*.
- [28] Hinton G, et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine*, 2012, 29(6): 82–97.
- [29] Graves A, Mohamed A and Hinton G. Speech

- Recognition with Deep Recurrent Neural Networks. arXiv, 2013, 38(2003): 6645–6649.
- [30] Esteva A., et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 2017. 542(7639): 115.
- [31] Hutson M. Self-taught artificial intelligence beats doctors at predicting heart attacks. Science, 2017.
- [32] Silver D, et al. Mastering the game of Go with deep neural networks and tree search. Nature, 2016, 529(7587): 484–489.
- [33] Mnih V, et al. Human-level control through deep reinforcement learning. Nature, 2015, 518(7540): 529–533.
- [34] Gatys L A, Ecker A S and Bethge M. A Neural Algorithm of Artistic Style. Computer Science, 2015.
- [35] Zhou Z H and Feng J. Deep Forest: Towards An Alternative to Deep Neural Networks. arXiv, 2017.

本文转载自《南京信息工程大学学报（自然科学版）》2017年第9卷第3期，297–304.



【作者简介】俞祝良，工学博士，教授，博士生导师，华南理工大学自动化学院副院长，教育部新世纪优秀人才，广东省“千百十”省级培养对象。俞祝良分别于1995年和1998年从南京航空航天大学电子工程系获得本科和硕士学位。2006年从新加坡南洋理工大学获得博士学位。2000年至2008年在新加坡南洋理工大学信号处理中心从事科研工作，任研究组组长。2008年加入华南理工大学自动化科学与工程学院，现主要从事信号处理、机器学习、大数据处理及在脑信号处理、智能机器人、医学数据处理中的应用研究工作。在国际期刊及会议上发表论文约100余篇，其中SCI检索论文60余篇，EI检索论文50余篇。出版英文著作（章节）6部。先后承担国家自然科学基金，广东省自然科学基金重点项目，各类人才项目和企业合作项目多项。主要学术兼职：国际SCI期刊MSSP，副主编；中国生物医学工程学会医学神经工程分会委员会委员；全国高等院校智能科学与技术专业系列规划教材编委会委员；中国人工智能学会教育工作委员会第二届委员会，常务委员等。