

神经网络在复杂系统建模中的应用

张国宁, 黄湘远

陆军指挥学院作战实验室, 南京 210045

摘要 分析了神经网络模型的特点、复杂系统的特征与建模方法论要求, 表明神经网络应用于复杂系统建模可行且适用。探讨了当前神经网络在复杂系统建模中应用的文本语言处理、语音识别、图像处理、故障诊断等, 针对所存在的不能很好地描述系统结构与演化、不能同时处理多个任务、不能很好解决非重复性问题等主要不足, 提出了解决这些问题的初步思路。

关键词 神经网络; 复杂系统; 建模

神经网络是模拟生物的大脑结构、工作模式的一种模型, 广泛应用于科研和工程中, 是当前被世界各国大力推进、重点发展的人工智能最重要的技术。“深度学习”的内核就是神经网络^[1-2]。复杂系统是有“21 世纪的科学”之称的复杂性科学中最重要的概念, 生产、社会、军事等领域的系统, 很多都属于复杂系统。研究复杂系统, 解决复杂系统的认识、优化、控制等问题, 基础在于对复杂系统进行正确、准确的建模。理清神经网络在复杂系统建模中应用的可行性、适用性及发展趋势, 对于做好复杂系统建模具有重要的指导意义。

1 神经网络的主要特点

1943 年, McCulloch 和 Pitts 共同提出 M-P 神经元模型^[3], 开创了神经网络研究的先河。1958 年, Rosenblatt 在 M-P 模型的基础之上, 集成了 Hebb、Rochester、Holland、Hodgkin、Huxley 和 Eccles 等成果, 并引入学习机制, 提出了第 1 个完整的人工神经网络模型“Perceptron

(感知机)”^[4]。此后, 经历了以 1969 年 Minsky 和 Papert 合著《Perceptron》(《感知机》)^[5]为标志的萧条, 以 1982 年 Hopfield 文章《Neural Networks and Physical Systems with emergent collective computational abilities》(《具有涌现集群计算能力的神经网络与物理系统》)^[6]为标志的复兴, 以及 20 多年的积累和沉浮, 涌现了包括 Hinton 提出的“Boltzmann Machine”(玻尔兹曼机)^[7]和“Deep Belief Network”(深度置信网络)^[8], Rumelhart 提出的“Back-propagation Algorithm”(反向传播算法)^[9], Fukushima 提出的“Neocognitron”(新认知机)^[10], LeCun 提出的“LeNet(一种卷积神经网络)”^[11]以及 Schmidhuber 提出的“Long Short Term Memory”(长短时记忆模型)^[12]等重要成果。到了 2016 年, DeepMind 开发的“AlphaGo”战胜了人类围棋冠军李世石, 神经网络以“深度学习”为名再度复兴, 成为广受公众关注的热词^[13-14]。相比较于显示的知识形式化、逻辑推理建模方法, 神经网络由于其模拟生物大脑的结构和功能而具有如下显著特点。

收稿日期: 2018-03-06; 修回日期: 2018-05-21

作者简介: 张国宁, 副教授, 研究方向为作战复杂系统建模及实验, 电子邮箱: guoning_zh@aliyun.com

引用格式: 张国宁, 黄湘远. 神经网络在复杂系统建模中的应用[J]. 科技导报, 2018, 36(12): 66-70; doi: 10.3981/j.issn.1000-7857.2018.12.009

1) 信息处理并行性。神经网络由分层的神经元连接而成,每层的神经元之间的结构是并行的,每个神经元都是一个独立的信息处理单元,整个神经网络的计算功能分布在多个处理单元上,它们处理信息的操作是同时进行的。因而神经网络信息处理的效率比串行的模型更高,同时对计算机的并行计算能力要求也更高。

2) 信息存储分布式。神经网络结构上的并行性使其信息必然以分布式的方式存储,每个神经元与其他神经元的连接上都有一个连接权值,所有这些神经元和连接权值共同构成宏观上的神经网络信息特征。单个神经元和连接权值没有多大作用,即使个别受到损坏,也不会对整体的信息特征造成太大的影响。

3) 自组织与自学习。神经网络模型需要利用样本数据进行训练,训练过程就是神经网络模型自组织、自学习的过程,神经网络依据样本数据,按一定规则不断调整神经元之间的连接权值,从而构建起具有实际应用能力的神经网络。

4) 问题求解概率化。神经网络模拟生物大脑的神经网络,在处理问题时强调对综合输入的训练和处理,不拘泥于明确的特征参数的选择和计算,对于问题的解,也是以概率的形式给出,而不像严格的逻辑推理和知识工程那样给出精确的、明确的解。这让神经网络在处理需要进行综合的联想、比较、推理、归纳等复杂问题时更加有效。

2 复杂系统的特征和建模方法论要求

复杂性科学是20世纪80年代提出的,它确立了一种有别于“老三论”“新三论”等理论方法解决系统的认识、建设、控制等新理念、新视角。按照复杂性科学的观点,复杂系统指的是一类组成关系复杂、不能用还原论的方法处理的系统,具有自组织适应性、不确定性和涌现性等显著特征。自组织适应性,是指复杂系统可以自适应、自学习、自组织、自聚集,根据环境条件改变自身的结构、行为、能力,以更好地适应环境的变化。不确定性,是指复杂系统是不确定的系统,通常不可能对复杂系统进行形式化的分析,不能重复再现复杂系统的行为。涌现性,是指复杂系统中的每个个体都遵从局部规则,但通过交互会形成新的整体效应,即“1+1>2”,体现了一种质变,强调个体之间的相互关

系。复杂系统的这些特征,对建模方法论提出了“超越还原论”的要求。

“超越还原论”是复杂性科学的方法论,它坚持还原论与整体论的辩证统一。在复杂性科学中,还原论与整体论既是对立、矛盾和对抗的,又是互补、共生、合作的。一方面,要还原到元素层次,了解局部的精细结构,才能避免猜测性、笼统地认识系统整体;而如果没有整体的观点,对事物的认识就极可能是零碎的、片面的、局部的,只见树木,不见森林,不能从整体上把握事物、解决问题。另一方面,哥德尔的不完备性原理表明,复杂系统存在着内在逻辑认识的盲点,要想清楚地认识它,不能仅仅在局部上,而要站在以这个系统为部分的整体上来观察和研究它。指导复杂系统建模的复杂性科学方法论坚持整体与部分、多样性与统一性、有序性与无序性的辩证统一观点。

1) 整体与部分辩证统一。整体的各个部分是有机地、有组织地相互联系、相互作用着的,这种组织性使得各个部分所不具有的新特征和属性涌现出来;同时,部分在成为整体的成员后,它作为部分所独有的特征和属性也会受到整体的限制和束缚。这要求复杂系统建模必须注重同时考察部分和整体。

2) 多样性与统一性辩证统一。复杂系统由多个属性、功能、行为各不相同的部分所构成,系统的统一性产生并依赖于其组成部分的多样性,这种统一性必须要有一定数量、一定规模的多样性来支撑;同时,部分的多样性也受到整体的统一性的限制、维持和保护,没有整体的统一性,部分的多样性就没有意义。这要求复杂系统建模要注重多样的输入输出、分布式和并行处理。

3) 有序性与无序性辩证统一。在复杂系统中,无序逐渐自组织为有序,形成系统的组织结构,且具有相对的稳定性;新的环境条件变化又会影响和破坏系统的有序,就会进入下一个无序自组织、自学习,进而又成为有序的过程。这要求复杂系统建模要考虑适应性、学习和自组织等行为模式。

3 神经网络在复杂系统建模中的应用

将神经网络主要特点与复杂系统特征及建模方法论的要求进行比较可以发现,神经网络的信息处理并行性、信息存储分布式、自组织与自学习、问题求解概

率化,与复杂系统建模“注重同时考察部分和整体,注重多样的输入输出、分布式和并行处理,考虑适应性、学习和自组织”的要求相一致。对于模型反映复杂系统的适应性、不确定性、涌现性具有显著的适用性。

3.1 神经网络应用于复杂系统建模的领域

经过多年的理论研究和实践应用,神经网络在复杂系统建模中的应用较多,如在文本语言处理、语音识别、图像处理、故障诊断等应用领域大展神通,取得了很多突破性进展。

文本语言处理任务中,如何提取典型、有效的特征是解决问题的关键。基于神经网络模型的特征学习方法可以快速、有效地学习短文本的向量表示,解决了短文本语义匮乏和快速检索难等问题,对文本分类、信息检索等任务有着积极作用^[15]。

语音识别领域中,在深层神经网络框架下,从后验概率计算、假设检验模型和有监督分类多个不同的角度对发音错误检测算法进行改进,可增强对中文声调和英文重读的判别能力,有效修正非标准发音和标准声学模型之间的不一致,大幅度降低语音识别的单词识别错误率,从而突破某些实际应用场景下对语音识别性能要求的瓶颈,使语音识别技术走向真正实用化。基于深度神经网络的隐马尔可夫模型可将语音到词语的转换错误率从 27.4%降低到 18.5%^[16-17]。

在图像处理领域中,利用深度学习的层次结构提取图像的潜在复杂结构规则,然后利用神经网络构建从视觉特征到高层语义的映射模型,可实现视觉内容与高层语义之间的有效映射,获取图像的高层语义内容,有效解决基于视觉认知机理的图像语义内容及图像理解问题。例如,谷歌公司的 Schroff 等利用三元组损失函数进行网络训练以直接学习人脸特征,并将这些特征用于人脸验证和人脸聚类中,算法取得了 99.63% 的精度^[18]。

在故障诊断领域,神经网络是模拟人类思维的学习以及推理过程,通过有效的特征提取、选择和分类识别处理诊断信息,模拟人类专家,以灵活的诊断策略对监测对象的运行状态和故障作出智能判断和决策。例如,东南大学孙文郡等采用具有无监督特征提取优势的稀疏自编码模型来学习故障特征,利用基于深度神经网络方法的稀疏自编码算法实现异步电机故障分类^[19]。

3.2 神经网络复杂系统建模存在的问题

神经网络是复杂系统的模型,但又不是一个完备的模型,主要体现在以下 3 个方面。

1) 神经网络长于处理信息而短于描述系统的结构和演化。神经网络主要用于复杂系统的信息处理,是一个信息处理模型,而复杂系统建模,不仅仅是一个信息处理问题,还包括对复杂系统结构和演化的建模问题的研究。神经网络是对复杂系统信息处理功能的建模,在复杂系统的结构和演化建模问题上还有待完善。

2) 神经网络长于处理单一任务而短于同时处理多个任务。大多数复杂系统都可以同时处理多个任务。而目前的神经网络应用,有视觉处理、文字处理,可以胜任分类、推理、预测等任务,但都限于特定的单一领域,尚未有能够同时处理多个不同领域任务的通用神经网络模型,既就是处理相近领域的任务也需要重新进行训练。

3) 神经网络长于处理重复性工作而短于解决创新性问题。神经网络模型能够较好的完成重复性的工作,训练好的神经网络,可以用来解决一些重复性的工作,能够以比人更快的速度、更高的精度完成任务。但在面对相同领域中的新问题时,要在缺乏相关数据的条件下找出解决问题的办法,仍具有挑战性。

4 结论

神经网络是复杂系统建模的一种重要的手段,与复杂系统建模注重同时考察部分和整体、注重多样的输入输出与分布式和并行处理、考虑适应性与学习和自组织等行为模式的要求相一致,能够很好地描述复杂系统的信息处理。但也存在着不能建模复杂系统的结构与演化、多任务处理、创新能力的不完备性问题。

破解神经网络在复杂系统建模中的应用问题,主要就是破解其不完备性问题。本文提出 2 种解决思路。

一是通过多模块综合实现多任务处理。将神经网络的神经元区分为不同的群,分别描述和解决问题的不同方面,执行不同的任务,从而使整体的网络具有多任务处理功能。Hinton 近期实现了“CapsNets”(胶囊网络)模型^[20],在视觉处理中将一些已经学习到的特征相对固定下来,从而在新的场景下更快地发现这些特征,

也减少了对于数据量和计算能力的依赖。这个模型虽然还达不到多模块综合实现多任务处理,但提供了一个可以继续研究的思路。

二是建立由一个“管理者神经网络”与多个“执行者神经网络”构成的“知识应用神经网络”,解决初级的神经网络创新的问题。即遵循哥德尔的不完备性原理,增加具有高一层次视角的管理者,借鉴深度迁移学习的方法,将在某些场景下训练好的神经网络作为初始执行者网络,利用变化了的数据进行训练,辅以强化学习方法,以较小的数据量实现已有知识的改进,由管理者神经网络负责组合若干个执行者网络类似条件下类似的任务,实现“综合集成”的创新。

参考文献(References)

- [1] Hinton G, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507.
- [2] Vos N, Meijden H, Denessen E. Effects of constructing versus playing an educational game on student motivation and deep learning strategy use[J]. *Computers & Education*, 2011, 56(1): 127-137.
- [3] McCulloch W, Pitts W. A logical calculus of the ideas immanent in nervous activity[J]. *Bulletin of Mathematical Biophysics*, 1943, 5(4): 115-133.
- [4] Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain[J]. *Psychological Review*, 1958, 65(6): 386-408.
- [5] Minsky M, Paypert S. *Perceptrons*[M]. Cambridge: MIT Press, 1969.
- [6] Hopfield J. Neural networks and physical systems with emergent collective computational abilities[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1982, 79(8): 2554-2558.
- [7] Ackley D, Hinton G, Sejnowski T J. A learning algorithm for Boltzmann machines[J]. *Cognitive Science*, 1985, 9: 147-169.
- [8] Hinton G, Osindero S, Teh Y. A fast learning algorithm for deep belief nets[J]. *Neural Computation*, 2006, 18: 1527-1554.
- [9] Rumelhart D, Hinton G, Williams R. Learning representations by back-propagating errors[J]. *Nature*, 1986, 323(6088): 533-536.
- [10] Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. *Biological Cybernetics*, 1980, 36: 193-202.
- [11] LeCun Y, Bottou L, Bengio Y, et al. Gradient based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [12] Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [13] Silver D, Huang A, Maddison C, et al. Mastering the game of Go with deep neural networks and tree search[J]. *Nature*, 2016, 529(7587): 484-489.
- [14] Tan T, Qian Y, Yin M, et al. Cluster adaptive training for deep neural network[C]//International Conference on Acoustics, Speech and Signal Processing. South Brisbane, Queensland, Australia: IEEE, 2015: 4325-4329.
- [15] Yu Z, Wang H X, Lin X M, et al. Understanding short texts through semantic enrichment and hashing[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2016, 28(2): 566-579.
- [16] Hu W P, Qian Y, Soong F K, et al. Improved mispronunciation detection with deep neural network trained acoustic models and transfer learning based logistic regression classifiers [J]. *Speech Communication*, 2015, 67: 154-166.
- [17] Seide F, Li G, Yu D. Conversational speech transcription using context-dependent deep neural networks[C]//Proceedings of the Annual Conference of the International Speech Communication Association. New York: Curran Associates, 2011: 437-440.
- [18] Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015: 815-823.
- [19] Sun W J, Shao S Y, Zhao R, et al. A sparse auto-encoder-based deep neural network approach for induction motor faults classification[J]. *Measurement*, 2016, 89: 171-178.
- [20] Sabour S, Hinton G. Dynamic routing between capsules [EB/OL]. [2018-03-06]. <http://arxiv.org/pdf/1710.09829v1.pdf>.

Applications of neural network in complex system modeling

ZHANG Guoning, HUANG Xiangyuan

Combat Laboratory, Army Command College, Nanjing 210045, China

Abstract By analyzing and comparing the characteristics of neural network models and complex systems and the requirements of their modeling methodology, it is shown that the applications of the neural network in the complex system modeling are feasible. The main application fields of the neural network in the complex system modeling are discussed, such as the natural language processing, the speech recognition, the visual processing, and the fault diagnosis; and the main problems are pointed out, including the facts that the neural network can not describe the structure and the evolution of the system, it can not handle multiple tasks at the same time, and it can not solve the non repetitive problems well. Also the solutions of these problems are discussed preliminarily.

Keywords neural network; complex system; modeling ●



(责任编辑 王丽娜)