

深度学习在控制领域的研究现状与展望

段艳杰¹ 吕宜生¹ 张杰^{1,2} 赵学亮¹ 王飞跃¹

摘要 深度学习在特征提取与模型拟合方面显示了其潜力和优势. 对于存在高维数据的控制系统, 引入深度学习具有一定的意义. 近年来, 已有一些研究关注深度学习在控制领域的应用. 本文介绍了深度学习在控制领域的研究方向和现状, 包括控制目标识别、状态特征提取、系统参数辨识和控制策略计算. 并对相关的深度控制以及自适应动态规划与平行控制的方法和思想进行了描述. 总结了深度学习在控制领域研究中的主要作用和存在的问题, 展望了未来值得研究的方向.

关键词 深度学习, 控制, 特征, 自适应动态规划

引用格式 段艳杰, 吕宜生, 张杰, 赵学亮, 王飞跃. 深度学习在控制领域的研究现状与展望. 自动化学报, 2016, 42(5): 643–654

DOI 10.16383/j.aas.2016.c160019

Deep Learning for Control: The State of the Art and Prospects

DUAN Yan-Jie¹ LV Yi-Sheng¹ ZHANG Jie^{1,2} ZHAO Xue-Liang¹ WANG Fei-Yue¹

Abstract Deep learning has shown great potential and advantage in feature extraction and model fitting. It is significant to use deep learning for control problems involving high dimension data. Currently, there have been some investigations focusing on deep learning in control. This paper is a review of related work including control object recognition, state feature extraction, system parameter identification and control strategy calculation. Besides, this paper describes the approaches and ideas of deep control, adaptive dynamic programming and parallel control related to deep learning in control. Also, this paper summarizes the main functions and existing problems of deep learning in control, presents some prospects of future work.

Key words Deep learning, control, feature, adaptive dynamic programming (ADP)

Citation Duan Yan-Jie, Lv Yi-Sheng, Zhang Jie, Zhao Xue-Liang, Wang Fei-Yue. Deep learning for control: the state of the art and prospects. *Acta Automatica Sinica*, 2016, 42(5): 643–654

近年来, 深度学习 (Deep learning)^[1] 在学术界和工业界发展迅猛. 它在很多传统的识别任务上取得了识别率的显著提升^[2], 显示了其处理复杂识别任务的能力, 吸引了大批学者研究其理论和应用. 许多领域开始尝试利用深度学习解决本领域的一些问题. 在控制领域, 关于深度学习的研究已初现端倪. 本文在概述深度学习的主要思想和方法的基础上, 阐述深度学习在控制领域的研究现状, 回顾相关的深度网络与深度控制的研究内容, 介绍深度学习在控制领域常用的结合方法: 自适应动态规划及其拓展平行控制. 综合以上内容提出深度学习在控制领

域的研究展望.

1 深度学习概述

深度学习源于神经网络的研究, 可理解为深层的神经网络. 通过它可以获得深层次的特征表示, 免除人工选取特征的繁复冗杂和高维数据的维度灾难问题. 目前较为公认的深度学习的基本模型包括基于受限玻尔兹曼机 (Restricted Boltzmann machine, RBM) 的深度信念网络 (Deep belief network, DBN)^[3–4]、基于自动编码器 (Autoencoder, AE) 的堆叠自动编码器 (Stacked autoencoders, SAE)^[5]、卷积神经网络 (Convolutional neural networks, CNN)^[6]、递归神经网络 (Recurrent neural networks, RNN)^[7]. 为便于清楚地阐述深度学习在控制领域的研究现状, 本节将概述以上四种深度学习模型的主要思想和方法. 关于深度学习的更多内容, 参见深度学习领域的综述^[1, 8–11].

1.1 DBN

基于 RBM 的 DBN 由多个 RBM 堆叠而成, 其结构如图 1 所示. 网络前向运算时, 输入数据从低层

收稿日期 2015-12-26 录用日期 2016-03-26
Manuscript received December 26, 2015; accepted March 26, 2016

国家自然科学基金 (71232006, 61233001, 71402178) 资助
Supported by National Natural Science Foundation of China (71232006, 61233001, 71402178)

本文责任编辑 侯忠生

Recommended by Associate Editor HOU Zhong-Sheng

1. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室 北京 100190 2. 青岛智能产业技术研究院 山东 266000

1. The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 2. Qingdao Academy of Intelligent Industries, Shandong 266000

RBM 输入网络, 逐层向前运算, 得到网络输出. 网络训练过程, 不同于传统的人工神经网络 (Artificial neural network), 分为两个阶段: 预训练 (Pre-training) 和全局微调 (Fine tuning). 预训练阶段, 从低层开始, 每个 RBM 单独训练, 以最小化 RBM 的网络能量为训练目标. 低层 RBM 训练完成后, 其隐层输出作为高层 RBM 的输入, 继续训练高层 RBM. 以此类推, 逐层训练, 直至将所有 RBM 训练完成. 预训练阶段, 只使用了输入数据, 没有使用数据的标签, 属于无监督学习 (Unsupervised learning). 全局微调阶段, 以训练好的 RBM 之间的权重和偏置作为深度信念网络的初始权重和偏置, 以数据的标签作为监督信号计算网络误差, 利用 BP (Back propagation) 算法计算各层误差, 使用梯度下降法完成各层权重和偏置的调节. **DBN 可用于特征提取和数据分类等.**

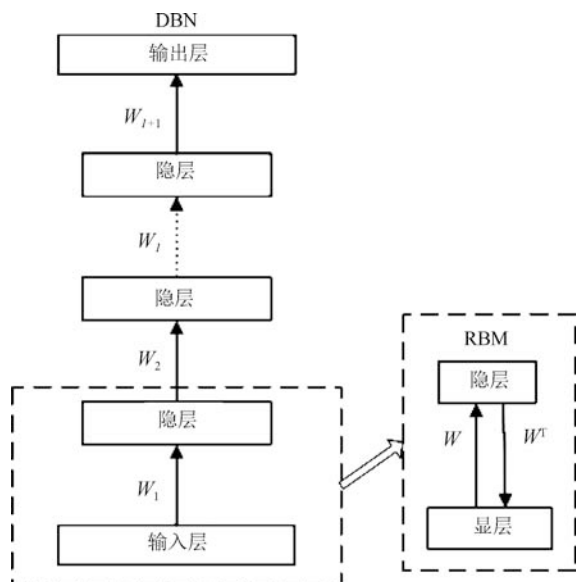


图 1 DBN 网络结构

Fig. 1 The structure of DBN

1.2 SAE

类似于 DBN, SAE 由多个 AE 堆叠而成, 其结构如图 2 所示. SAE 前向计算类似于 DBN, 其训练过程也分为预训练和全局微调两个阶段. 不同于 RBM 的是, AE 之间的连接是不对称的. 每个 AE 可视为一个单隐层的人工神经网络, 其输出目标即此 AE 的输入. 在预训练阶段, 从低层开始, 每个 AE 单独训练, 以最小化其输出与输入之间的误差为目标. 低层 AE 训练完成后, 其隐层输出作为高层 AE 的输入, 继续训练高层 AE. 以此类推, 逐层训练, 直至将所有 AE 训练完成. 同样地, SAE 的预训练阶段也只使用了输入数据, 属于无监督学习. 在全局微调阶段, 以训练好的 AE 的输入层和隐层之间的权重和偏置作为堆叠自动编码器的初始权重和偏

置, 以数据的标签作为监督信号计算网络误差, 利用 BP 算法计算各层误差, 使用梯度下降法完成各层权重和偏置的调节.

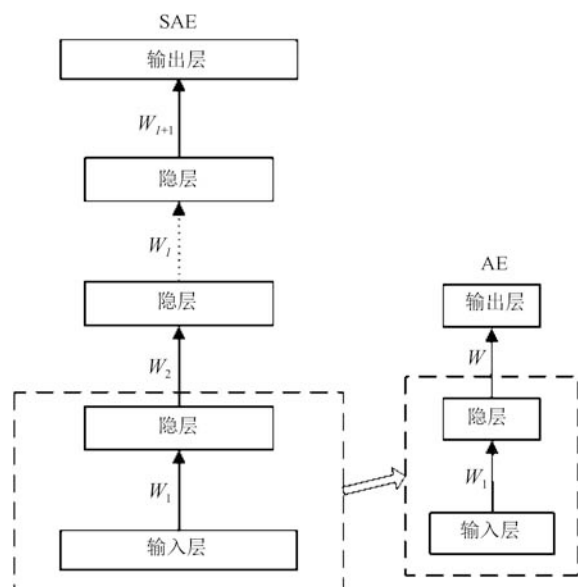


图 2 SAE 网络结构

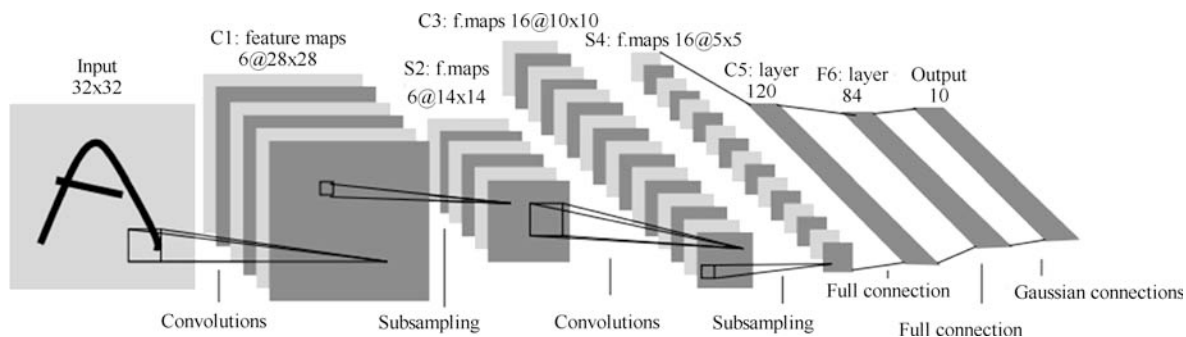
Fig. 2 The structure of SAE

1.3 CNN

CNN 由卷积层和次采样层 (也叫 Pooling 层) 交叉堆叠而成, 其结构如图 3^[6] 所示. 网络前向计算时, 在卷积层, 可同时有多个卷积核对输入进行卷积运算, 生成多个特征图, 每个特征图的维度相对于输入的维度有所降低. 在次采样层, 每个特征图经过池化 (Pooling) 得到维度进一步降低的对应图. 多个卷积层和次采样层交叉堆叠后, 经过全连接层到达网络输出. 网络的训练类似于传统的人工神经网络训练方法, 采用 BP 算法将误差逐层反向传递, 使用梯度下降法调整各层之间的参数. CNN 可提取输入数据的局部特征, 并逐层组合抽象生成高层特征, 可用于图像识别等问题.

1.4 RNN

无论是 DBN 和 SAE 还是 CNN, 都没有考虑样本之间的关联问题. RNN 则考虑了样本之间的关联关系, 将这种关联关系以神经网络之间的连接体现出来. 一般情况下, 单向 RNN^[12] 中, 如图 4 所示, 单个神经网络的隐层连接至下一个神经网络的隐层. 这种连接方式考虑了前面样本对后面样本的影响. 还有一种双向 RNN^[13] 的连接方式, 单个神经网络的隐层连接了其前后神经网络的隐层. 这种连接方式考虑了前后样本对当前样本的影响. 一般认为 RNN 的各个神经网络具有相同的权重和偏置. RNN 训练时, 可使用 RBM 或者 AE 对其进行预训

图 3 CNN 网络结构^[6]Fig. 3 The structure of CNN^[6]

练来初始化网络参数, 然后计算每个样本的输出误差, 并以累计误差训练网络参数. RNN 可用于处理时序数据或者前后关联数据. RNN 还可以和 CNN 结合使用, 处理考虑样本之间相关性的问题.

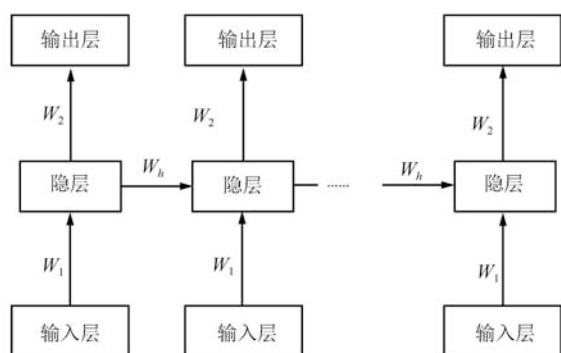


图 4 RNN 网络结构

Fig. 4 The structure of RNN

2 深度学习在控制领域的研究现状

深度学习在控制领域的研究已初现端倪. 就已有的研究报道, 目前的研究主要集中在控制目标识别、状态特征提取、系统参数辨识、控制策略计算等方面. 尤其是深度学习和强化学习的结合已经产生了令人振奋的研究成果. 如图 5 所示, 深度学习在控制系统的各环节均有应用研究, 本节从控制系统的角度分类, 从四个方面介绍深度学习在控制领域的研究现状.

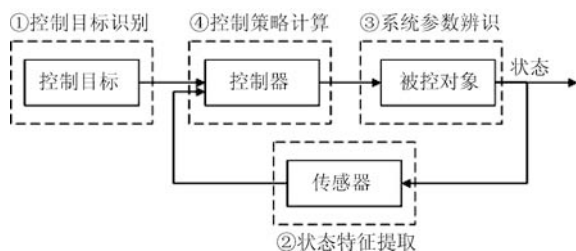


图 5 深度学习在控制系统各环节的应用

Fig. 5 The application of deep learning in control system

2.1 控制目标识别

在传统的控制系统中, 控制目标一般以给定量的形式直接给出. 而在智能控制系统中, 控制目标有时并不明确或者不能直接得到. 就已有研究深度学习进行控制目标识别的报道, 在先进机械手抓取系统^[14]中, 系统首先需要检测被抓取物的位置、识别被抓取物并估计其姿态以确定机械手需要到达的位置和角度. 本节以此为例说明深度学习在控制目标识别中的作用.

图 6^[14]显示了先进机械手抓取系统的流程. 输入数据是颜色深度图. 在物体检测过程中, 系统利用前景背景分割技术检测出物体位置. 然后从图片中分割出包含物体的统一大小图片. 在物体识别和姿态估计过程中, 获取的统一大小图片输入 CNN 网络, 经过前向计算在网络输出端获得物体的类别. 该类别信息包含了物体的种类和姿态. 最后, 机械手按照规划的路径到达物体上方, 调整夹持器的姿态, 使得其姿态符合物体的姿态, 然后执行抓取动作.

图 6 机械手抓取系统^[14]Fig. 6 Robotic grasping system^[14]

从系统流程可以看出, 深度学习作用于物体识别和姿态估计过程. 具体地, 使用了 CNN 网络进行分类. 网络的输出类别按照选取的 n 种物体不同角度的姿态进行区分, 第 i 种物体有 c_i 种姿态, 那么输出的类别总数为 $C = \sum_{i=1}^n c_i$, 即 CNN 网络的输出层节点个数. 网络的输入是统一大小的图片, 以二维矩阵的方式输入网络. 所使用的 CNN 通过卷积层和次采样层交叉堆叠的方式建立特征提取的网络, 然后通过全连接层连接到输出层建立分类器. 网络结构建立后, 收集一定数量具有类别标记的图片样本, 一部分用于训练网络参数, 剩余部分用于测试网络性能. 经过迭代训练, 获取训练完成的 CNN 网络. 当有新的图片样本输入网络时, 网络经过前向计算,

输出图片中物体的类别,包括物体的种类和姿态.机械手获取物体类别后,在执行抓取动作之前,将夹持器调整到和物体的姿态相一致的状态,以完成抓取动作.在这个系统中,机械手抓取之前需要达到的姿态为机械手控制系统的控制目标,此控制目标由物体的种类和姿态决定,所以采用深度学习进行物体分类实质上在控制系统中起到了控制目标识别的作用.至于机械手如何达到控制目标则由系统的控制器决定.

2.2 状态特征提取

现代控制理论引入了状态空间的概念,以系统的状态为基础进行分析与综合.在基于视觉的控制系统中,系统状态由图片体现,其维度高,不容易进行人工特征抽取,不利于系统的分析综合与控制.而深度学习十分适合对图片数据进行降维与特征抽取,所以深度学习在基于视觉的控制系统中研究较多.基于视觉的控制系统中,深度学习多与强化学习方法结合,使得系统能够根据深度学习提取的特征给出控制策略.在深度学习用于控制系统状态特征提取的研究中,Lange 等^[15-16]利用 SAE 进行特征提取;Mnih 等^[17]则利用 CNN 进行特征提取.无论采用哪种深度学习模型,其思想基本一致,以下以 Atari 游戏系统^[17-18]为例介绍深度学习在状态特征提取中的作用.

Atari 是一款计算机游戏,玩家通过计算机屏幕显示的游戏场景,决策执行的游戏动作,获得游戏得分的变化.在使用深度学习进行 Atari 游戏的系统中,游戏场景以图片的形式输入 CNN 网络,CNN 网络和 Q 网络连接,游戏动作的决策是最终的输出,系统结构如图 7 所示.其中, Q 网络是用于获得系统预期回报函数的网络.在这个系统中,可执行的动作集合包含有限个游戏动作元素,而状态集合包含无限个场景状态元素.评价网络的输出层节点个数为可执行的游戏动作个数,每个节点的输出即在输入场景下采用该节点所代表动作的预期回报.系统选择评价网络输出节点中预期回报最高的节点所代表的游戏动作执行. Atari 游戏执行了游戏动作后,更新游戏场景图像并得到一个强化信号.

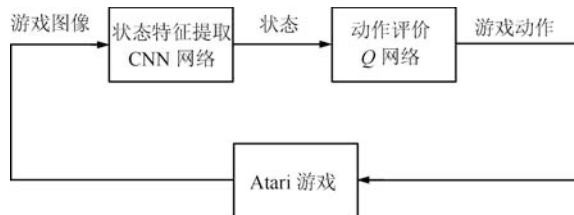


图 7 使用深度学习进行 Atari 游戏

Fig. 7 Playing Atari with deep learning

整个网络的权重参数根据强化学习的方法进行

训练,具体的训练方法和技巧不在本文介绍的范围,这里仅说明其主要思想.当前时刻的预期回报一方面是当前时刻来自于游戏的强化信号与下一时刻来自于网络的最高预期回报之和,另一方面也是当前时刻来自于网络的最高预期回报,二者之间的差距即整个网络的损失函数,训练的目标是最小化该损失函数.由此,网络获得了可传递的误差用于网络参数的训练.据报道,该系统在实验中的多个游戏任务上超越了人类的水平.

2.3 系统参数辨识

系统辨识是现代控制理论的重要研究内容,它根据系统输入输出来确定描述系统的模型.对于复杂的非线性动态系统,其模型难以用线性函数或者根据先验知识建立.而神经网络具有拟合复杂非线性函数的能力,可以用于系统辨识.但浅层神经网络在训练中容易受到局部最优等问题的影响,有时并不能准确描述动态系统.深度学习的发展给解决这一问题带来了启发.已经有部分研究专注于利用深度学习的方法进行系统辨识.由于系统模型由深度神经网络代替,系统辨识任务就转变成深度神经网络的参数优化,故而称这类任务为系统参数辨识.

使用深度学习进行系统参数辨识的控制系统一般涉及模型预测控制.模型预测控制的主要思想是在每个时刻基于系统当前状态和预测模型,计算出系统未来一段时间的最优控制序列,并执行序列的首个控制策略.其中,预测模型使用深度学习实现.这类研究中,直升机动态模型^[19]使用深度 ReLU 网络模型拟合,主要思想是利用历史一段时间的数据预测未来时刻的加速度.切割机器人^[20]考虑了系统状态在时间序列上的相关性,使用 RNN 构建系统动态模型,利用历史时间窗的系统状态预测未来时间窗的系统状态.车杆摆动系统^[21]使用深度神经网络拟合系统动态模型,利用当前时刻的状态和控制输入预测下一时刻状态与当前时刻状态的差值.

值得一提的是,在车杆摆动系统中,深度网络经过训练能够拟合系统动态模型后,在输出端引入一个新的节点代表预期回报函数.这样,深度网络同时表达了系统动态模型和动作评价函数.如图 8^[20]所示, s_t 表示系统在 t 时刻的状态, a_t 表示系统在 t 时刻采取的动作, $Q(s_t, a_t)$ 表示系统在状态 s_t 下,执行动作 a_t 所得的预期回报函数.网络训练思路是,先忽略输出端的 Q 函数,采用随机初始化的状态和任意动作产生一系列 $(s_t, a_t, s_{t+1} - s_t)$ 用于拟合系统动态模型.然后忽略输出端的 $s_{t+1} - s_t$,按照强化学习的方法,每一步执行最优的动作并且根据每一步产生的回报训练 Q 网络.这种训练方法基于一个基本假设,那就是对于状态预测有效的隐层特征有

助于预期回报函数的预测. 用深度网络拟合系统动态模型的步骤相当于对 Q 网络的预训练, 这一步不需要考虑回报. 实验证明, 经过预训练的 Q 网络能够比直接使用强化学习训练的 Q 网络更快达到较好优化水平.

2.4 控制策略计算

在单输出控制系统中, 控制策略指的是单一的控制量. 而在智能控制系统中, 控制策略除了指单一控制量之外, 也可以指一串动作或一个决策. 对于传统的控制器而言, 需要获取给定量和系统输出之间的差, 从而计算控制策略. 而对于智能控制系统而言, 控制器获取系统状态进行控制策略计算. 经典的 PID 控制从系统机理出发设计控制器, 而深度神经网络是一个需要监督信号进行训练的模型. 所以深度学习在控制策略计算方面的研究也主要集中在有监督信号的情形.

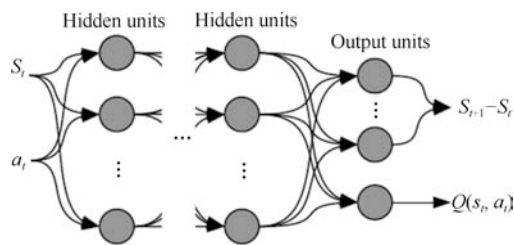


图 8 进行状态预测和学习 Q 函数的深度网络^[20]

Fig. 8 Neural network for learning state prediction and Q function^[20]

此类研究中, 有利用深度学习模型拟合 PID 控制器的报道^[22]. 其主要思想是对于一个线性定常系统, 首先设计一个能够满足控制要求的 PID 控制器. 然后以 DBN 网络替代 PID 控制器. DBN 网络的参数由 PID 控制器的输入输出数据训练. 训练后的 DBN 网络的输入是原来 PID 控制器的输入, 经过前向计算的网路输出替代原来 PID 的输出作为控制量. 这项工作虽然没有说明深度学习用于此处的必要性和优势, 但是却证明了深度学习可以用于控制策略计算. 有研究比较单层网络、多层网络和 RNN 网络分别用于粗糙地形运动系统^[23]. 在这项研究中, 网络使用状态作为输入, 动作作为输出. 而训练则采用引导策略搜索法, 需要的监督信号来源于已有的范例. 网络从已有的成功的运动范例中采样其状态-动作样本, 用于网络训练. 此研究的实验结果表明多层网络或 RNN 网络需要更多的训练样本才能达到比较好的性能, 而随着样本数目的增加, 二者的表现性能都有超越单层网络的趋势. 训练好的网络对于新的状态输入, 能够通过前向计算给出需要执行的动作, 即本系统的控制策略. 在运动控制函数^[24]的研究中, 使用深度网络计算从初始状态到

目标状态的控制指令序列和状态轨迹序列. 其中, 初始状态和目标状态作为网络的输入, 控制指令序列和状态轨迹序列作为网络的输出.

比较特别的是, 在运动控制函数的研究中, 输入数据的维度远小于输出数据的维度. 研究中首先使用 SAE 对输出数据进行降维. 降维过程中逐层训练 AE 的阶段, 不同于一般的从底层开始逐层训练的方法, 这项研究从最顶层开始逐层训练 AE. 如图 9 所示, 高层 AE 训练完后, 其隐层输出作为低层 AE 的输入, 继续训练低层 AE. SAE 训练完成后, 得到从低维特征到输出控制和轨迹序列的网络, 这里需要注意的是此网络取自 SAE 的上半部分. 然后使用神经网络将系统输入端连接至已经训练好的网络. 最后使用随机的到达样本对整个网络进行训练, 每个到达样本包含运动的初始状态、目标状态和从初始状态到目标状态的轨迹. 训练好的网络对于一个输入的初始状态和目标状态, 经过前向计算输出控制指令和运动轨迹序列.

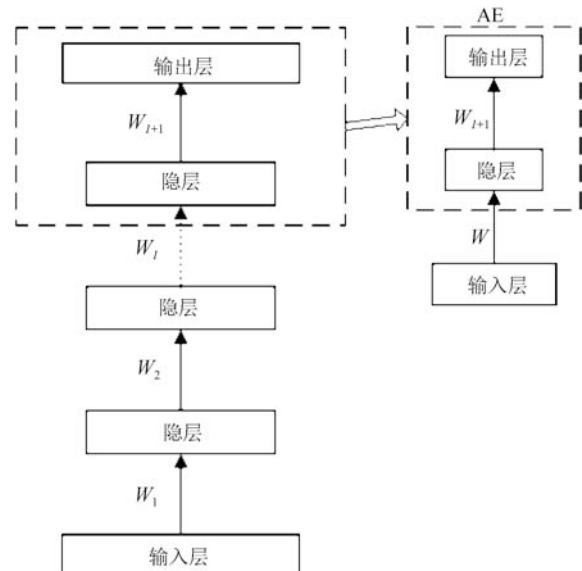


图 9 进行运动控制函数研究的深度网络

Fig. 9 Deep neural network for motor control function

3 深度网络与深度控制

早在深度学习出现之前, 就有学者曾利用多层神经网络实现一种自适应模糊逻辑控制系统^[25]. 从今天的角度看, 可以称为深度模糊控制网络. 另外, Saridis 及合作者在研究智能机控制理论的过程中, 曾采用玻尔兹曼机的结构和最小化网络能量的方法用于智能机的“组织”阶段, 即获取最优任务(动作)序列^[26-27]. 这些方法、思想和当前的深度学习方法具有一定的相似性和不同之处. 期望本节介绍能够拓宽深度学习在控制领域的研究思路, 激发深度学习在控制领域的研究灵感.

3.1 深度模糊控制网络

深度模糊控制网络系统的目标是开发一种基于模糊逻辑和神经网络实现拟人控制的方法,该方法主要针对这样一类系统:构建数学模型不可行或难以实行;熟练操作人员知识可用于实现控制目的.因此人类控制涉及的知识表示和知识进化在此系统中均有体现.如图10所示,该系统包含3类子网络:SN、RN和CN,分别用于模式识别、模糊推理和控制合成.

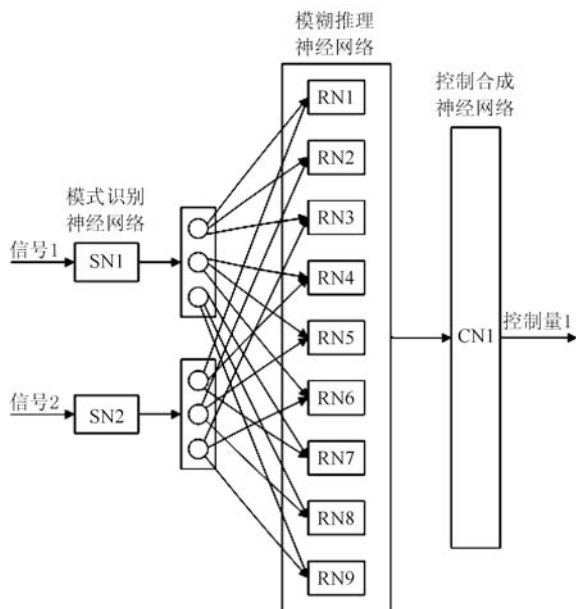


图10 深度模糊控制网络

Fig. 10 Neuro-fuzzy network

模式识别神经网络实现输入信号的模糊化.具体地,将输入信号通过隶属度函数映射为模糊语义项的隶属度.图10的示例中,有两个输入信号,每个信号分别对应三个模糊语义项.此处的模糊语义项描述了信号模式.语义项的隶属度在0到1之间,表示输入信号符合语义项描述的程度.SN被训练以替代隶属度函数.输入信号的模糊化完成后,语义项及其隶属度作为输入进入模糊推理神经网络.该子网络运用规则集(知识库)进行模糊推理.具体地,利用满足输入信号模糊语义项的规则进行推理,计算出每条决策规则的耦合强度.图10的示例中,有九条决策规则,每条规则遵循if-then的决策模式.每条规则对应一个RN.对于一个输入信号,每条规则的条件只含有其一个模糊语义项.所以对于图10的示例,每个RN的输入是两个模糊语义项,分别来源于两个不同的SN.每个RN被训练以替代规则集中的决策规则.RN输出规则的耦合度作为控制合成神经网络的输入,此网络利用一系列步骤和控制模糊语义项的隶属度函数输出最终的控制量.这些步骤

包括单个规则的推理,产生模糊控制量和解模糊.这些步骤均使用CN网络替代完成.SN、RN和CN网络构造完成后,可以连接成如图10所示的深度模糊控制网络,在3个子网络分别训练的基础上进行全局训练,进一步优化控制效果.深度模糊控制网络能够实现从采集状态信号到输出控制变量的全过程,更具体的过程参见文献[25],此处不再赘述.

文献[28]曾试图利用Memetic进化计算推广Neuro-fuzzy网络,并将中间节点值抽取出来,组成基于规则的决策系统,通过语言动力学系统进行描述^[29-30],实现从数计算到词计算之间的转换.这种方法可用于“当地简单,远程复杂”的网络化深度学习的硬件实现.能有效结合在线的实时决策与远程的云计算,从而低成本地提高系统性能.

通过对以上方法的介绍,我们可以发现,此方法和现在的深度学习方法既有共通之处又有不同特点.二者从整体上看都是多层神经网络结构,均是由不同的神经网络模块组合而成.二者都可以采用预训练和全局微调的训练过程.不同的是,前者可采用并行的方式对3个子网分别训练,而后者需采用串行的方式逐层训练.前者将人类知识以规则的形式嵌入网络,子网训练过程使用了这些规则提供的启发式信息.后者在子网训练过程中一般采用自动训练的方法,减少人为干预.对于网络结构和训练结果,前者可解释其含义,后者则不可解释.深度模糊控制网络最大的特点是包含知识结构,网络结构融合了知识表示,训练学习过程蕴含了知识进化.

3.2 智能机的组织

Saridis开发了一种分层递阶智能控制方法,将智能机的控制分为3个阶段:组织、协调和执行^[31].其中,组织阶段负责智能机的任务规划,主要问题是对期望的动作构建最小化熵的正确任务序列.Saridis及合作者利用玻尔兹曼机建立了熵的解析方法来衡量任务序列的正确性.

玻尔兹曼机的构建方法:输入节点表示期望的目标;输出节点表示为满足目标智能机需要执行的任务序列;隐藏节点使得输入和输出节点之间能够发展复杂的交互.节点的状态是1或0,表示任务执行或不执行.熵定义成任务序列正确概率的函数,任务序列正确概率定义成网络能量的函数,而网络能量由节点状态和节点之间的权重计算得到.在一定的约束条件下,网络能量越小,任务序列的正确概率越大,熵越小.玻尔兹曼机通常使用网络能量作为代价函数,通过最小化网络能量找到节点最优状态.类似地,这里以最小化熵为目标,利用最优化方法,例如遗传算法等,搜索出最优的网络权重,计算出对应的节点状态.从而获得最优节点状态对应的最优任

务序列.

4 自适应动态规划与平行控制

从深度学习在控制领域的研究现状可以看出, 强化学习方法中的自适应动态规划方法在深度控制系统的实现中起到重要的作用. 本节主要对自适应动态规划方法进行介绍, 进而引出其在复杂系统中的扩展: 平行控制.

为实现最优控制, 自适应动态规划方法和深度学习的结合是研究深度控制的关键领域. 最优控制以性能指标最优为目的, 是在上世纪 50 年代起发展起来的重要控制领域. 其不但在工程控制中具有重要意义, 在社会经济系统的管理中也有着重要的应用. 由 Bellman 提出的动态规划方法^[32-33] 以倒推的思想, 求解最优控制的 Bellman 方程:

$$J^*(x(k)) = \min_{u(k)} \{l(x(k), u(k)) + J^*(x(k+1))\} \quad (1)$$

$$J(x(k)) = \sum_{i=k}^{\infty} l(x(i), u(i)) \quad (2)$$

$$\text{s.t. } x(k+1) = F(x(k), u(k)), \quad k = 0, 1, \dots \quad (3)$$

需要存储不同状态下的最优性能指标 $J^*(x(k))$, 进而通过求解 Bellman 方程 (1) 中的最优性能条件得到针对任意状态的最优控制. 动态规划方法可用于求解闭环形式的最优控制, 且广泛地适用于线性系统、非线性系统、离散系统和连续系统等各种情况. 然而, 在使用动态规划方法求解复杂最优控制问题时, 会面临所谓“维数灾难”^[33-34], 即, 算法所需存储量和计算量随状态空间的维数增长而迅速提高. 1977 年, Werbos^[35] 首次提出使用多层神经网络逼近动态规划方法中的性能指标函数和控制策略以近似地满足 Bellman 方程, 从而时间向前 (Forward-in-time) 获得最优控制和最优性能指标函数. 近年来, 基于这种自适应动态规划 (Adaptive dynamic programming, ADP) 思想的研究层出不穷, 但使用的名称不尽相同, 如 Adaptive critic designs、Approximate dynamic programming、Asymptotic dynamic programming、Relaxed dynamic programming、Neurodynamic programming、Neural dynamic programming 等, 因其思路相同或相近, 在本文中统称为自适应动态规划^[36]. 近年来, 自适应动态规划成果斐然, 学者结合神经网络和强化学习思想实现了对不确定系统的鲁棒控制^[37]、大型系统的分散式控制^[38]、对自适应系统的可持续激励控制^[39] 等.

4.1 自适应动态规划的神经网络结构

启发式动态规划 (Heuristic dynamic programming, HDP)^[40] 是由 Werbos 提出的一种自适应动态规划近似结构.

如图 11 所示, 在这种自适应动态规划的近似结构中, 为了近似性能指标函数, 神经网络扮演了重要的角色. 该结构由模型网络、评价网络和执行网络三个多层神经网络组成. 模型网络用于近似状态方程 (3). 评价网络根据 Bellman 方程中最优性的满足程度, 对系统输入的控制信号进行评价, 给出评价信号即最小化的误差

$$E_c = |J(k) - l(k) - J(k+1)|^2 \quad (4)$$

执行网络则根据评价网络的评价信号更新并获得近似的最优控制, 最小化性能指标对控制的偏导

$$E_a = \left\| \frac{\partial l(x_k, u_k)}{\partial u_k} + \frac{\partial J(x_{k+1})}{\partial x_{k+1}} \frac{\partial F(x_k, u_k)}{\partial u_k} \right\| \quad (5)$$

由此可见, 早期的自适应动态规划研究中已经包含了深层神经网络的思想. 然而自适应动态规划中的深层神经网络并未运用预训练等现代深度学习技术, 无法利用未标记数据, 对于过深的结构还不能摆脱陷入局部收敛的困境.

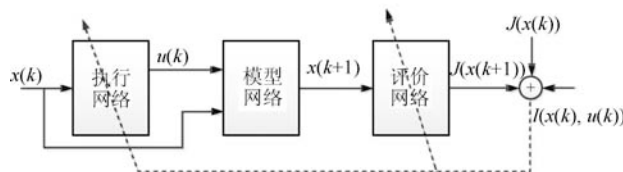


图 11 自适应动态规划的神经网络结构

Fig. 11 The network structure of adaptive dynamic programming

4.2 自适应动态规划的迭代型算法

2002 年, Murray 等率先提出了对连续状态方程的自适应动态规划迭代型算法^[41]. 迭代型 ADP 依然可以使用神经网络近似性能指标函数, 在每次迭代中都更新现有的最优性能指标及最优控制律. 迭代型 ADP 最显著的优势是从理论上可以证明迭代算法的稳定性和收敛性, 是自适应动态规划研究中的重大进展. 目前主要有值迭代 ADP 和策略迭代 ADP 两类算法.

依然以离散系统为例, 值迭代自适应动态规划方法^[42-44] 首先以 $J_0^*(x_k) = 0$ 作为最优性能指标函数的初值. 对每次迭代 $i = 0, 1, \dots$, 求解迭代的近似最优控制

$$u_i(x_k) = \arg \min_{u_k} \{l(x_k, u_k) + J_i^*(x_{k+1})\} \quad (6)$$

随后求得迭代的近似最优性能指标函数

$$J_{i+1}^*(x_k) = l(x_k, u_i(x_k)) + J_i^*(F(x_k, u_i(x_k))) \quad (7)$$

Wei 等^[45] 给出了不同初值情况下的值迭代 ADP 的收敛性证明.

策略迭代自适应动态规划方法多数讨论连续系统^[41, 46-48]. 对于离散系统, Liu 等^[49] 提出一种策略迭代 ADP 并分析了该算法的稳定性和收敛性. 策略迭代 ADP 初始于一个给定的容许控制律 $u_0(x_k)$. 对每次迭代 $i = 0, 1, \dots$, 求解近似的最优性能指标函数 V_i 以满足

$$J_i^*(x_k) = l(x_k, u_i(x_k)) + J_i^*(F(x_k, u_i(x_k))) \quad (8)$$

随后求得迭代的近似最优控制

$$u_{i+1}(x_k) = \arg \min_{u_k} \{l(x_k, u_k) + J_i^*(x_{k+1})\} \quad (9)$$

4.3 平行控制

平行控制^[50] 是一种反馈控制, 是 ACP (Artificial societies, Computational experiments, Parallel execution) 方法^[51-55] 在控制领域的具体应用, 其核心是利用人工系统进行建模和表示, 通过计算实验进行分析和评估, 最后以平行执行实现对复杂系统的控制. 平行控制是自适应控制的自然扩展, 在图 11 中, 被控对象往往并无自主行为的能力, 使用神经网络模型可以得到近似的微分方程或差分方程以描述系统状态变化. 在控制包含具有自主行为代理的复杂系统时, 无法使用状态方程近似被控对象状态对控制输入的变化. 控制与管理、试验与评估、学习与培训代替了经典控制系统中单一的控制变量, 复杂的人工系统则代替了被控对象的模型网络, 如图 12^[50] 所示.

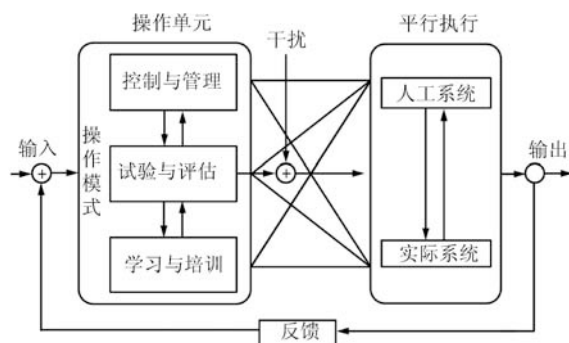


图 12 平行控制系统^[50]

Fig. 12 Parallel control systems^[50]

神经网络在平行控制中依然扮演重要的角色. 一方面, 尽管无法使用神经网络模型近似系统状态方程 (3), 我们却可以使用神经网络模型近似剥离了

代理模型自主行为的方程

$$x(k+1) = F(x(k), u(k), u_1(k), \dots, u_n(k)) \quad (10)$$

其中 u_1, \dots, u_n 为代理 $1, \dots, n$ 的主动行为, 通过在实际系统中充分激励代理行为和系统控制, 可获得方程 (10) 的近似. 另一方面, 自适应动态规划中通过迭代求解的近似最优性能指标, 在平行控制中则拓展为以计算实验和涌现观测获得的性能指标估计, 在对控制策略的评估中, 评价网络依然可以作为这一性能指标估计的近似手段.

平行控制是自适应控制方法在复杂系统中的扩展, 以自适应动态规划和基于代理的控制 (Agent-based control, ABC) 作为主要计算方法, 已经成功应用于社会计算^[56-58]、智能交通^[59-60]、健康^[61-62]等领域.

5 总结与展望

本文首先概述了当前较为公认的四种深度学习基本模型, 在此基础上, 对深度学习在控制领域的研究现状进行分类描述. 为打破深度学习在控制领域的研究局限, 对深度学习概念出现之前的深度控制相关研究进行介绍. 对当前和深度控制联系十分紧密的自适应动态规划方法的神经网络结构以及算法进行了详细描述, 并对其在复杂系统中的扩展平行控制的概念进行了介绍. 以下针对深度学习在控制领域的研究现状进行总结. 由于在特征提取以及模型拟合等方面显示出的潜力和优势, 深度学习已经被用于控制领域的相关研究, 并且在一些控制系统中表现出了较好的性能. 但是, 从上文的研究现状可以看出, 深度学习主要在涉及视觉的控制系统中具有应用的优势. 对于控制目标识别, 其识别的对象是图片中物体的种类和姿态. 对于状态特征提取, 其提取的是场景图片的抽象特征. 对于系统参数辨识, 其本质也是从系统状态中提取特征, 然后拟合特征和输出之间的关系. 对于控制策略计算, 虽然有直接拟合 PID 控制器的研究, 但是没有体现出深度学习的优势甚至引入的必要. 输出控制策略的深度网络需要从已有的范例中学习如何计算控制策略, 网络对系统的控制性能在很大程度上会受训练范例的影响, 不能估计其对系统出现的新情况的控制效果. 综上, 目前深度学习在控制领域的研究虽然已有以上介绍的一些实例, 但是相关研究的报道仍然相对较少, 研究的广度和深度都略显不足.

虽然深度学习在控制领域的研究存在一些问题, 不可否认的是, 深度学习在控制领域的研究前景仍然十分广阔. 就在本文投稿后一个月, 谷歌在 *Nature* 上发表了利用深度神经网络和树搜索方法进行围棋博弈的论文^[63], 其围棋程序 AlphaGo 在

随后进行的人机围棋比赛中以 4:1 的战绩打败了专业围棋选手. 这无疑给深度学习在控制领域的研究带来新的动力和想象力. 以下对未来值得研究的内容进行展望.

5.1 仿人感知与控制

深度学习之所以受到重视, 最早是其在语音识别和图像识别方面取得的重大突破. 而语音和图像识别的目的是能够达到人对语音和图像的识别水平, 所以深度学习从本质上是模拟人的听觉和视觉. 听觉和视觉是人感知世界的能力. 对于控制来讲, 例如机器人的机械手抓取控制系统的目的是能够达到人对物体抓取的速度和精度, 所以深度学习是否能够模拟人的运动控制系统是其用于控制领域的关键问题. 模拟系统以及实现该系统的硬件都是值得研究的问题. 深度学习需要监督信号才能训练出有效的模型. 引导策略搜索法是一种利用已经观察到的范例样本训练网络的方法. 引导策略搜索方法中, 对某些控制系统例如机器人行走, 人的行走模式可以当作训练中的范例样本. 所以对于智能控制领域的问题, 如何充分利用人的智能这一资源是十分有意义的. 而深度学习具有表达人类某些智能的潜力, 所以挖掘深度学习对于仿人型机器人等系统中存在的智能控制问题具有十分重要的意义.

5.2 无人系统

自动控制领域的研究使得无人系统的软硬件研究得到了快速发展, 例如无人机 (Unmanned aerial vehicle, UAV)、无人驾驶车等. 对于这类复杂的无人系统, 它们的动态模型一般具有复杂非线性特点, 利用传统的建模方法较难表示其动态特性. 而深度学习擅长对复杂非线性系统进行建模, 所以深度学习对这类无人系统的建模是值得进一步研究的问题. 无人系统需要尽量减少人的参与或者没有人的参与, 深度学习自动学习状态特征的能力使得深度学习在无人系统的研究中具有先天的优势. 如何充分利用和发挥深度学习在无人系统中的优势并发展深度学习在环的无人系统控制是下一步研究的方向.

5.3 强化自适应控制

强化学习模拟人对世界认知的过程, 利用从外界反馈的回报信号和接收或采集的环境状态不断调整系统参数, 改变控制策略, 达到自适应控制的效果. 对于环境状态维度较高的情况, 例如视觉环境, 深度学习能够自动提取识别状态特征, 为强化学习提供源数据. 深度学习与强化学习的结合能够使感知环境和系统控制无缝连接, 其所在的系统能够自动感知不确定的环境状态并完成自适应控制. 这样的系统值得进一步研究.

5.4 复杂系统

平行控制针对复杂系统, 以大数据和 ACP 方法为基础, 采用人工系统作为建模工具, 通过虚实互动的平行执行方式来完成. 平行控制是一种数据驱动的计算控制方法, 与深度学习有天然的契合. 平行控制作为自适应动态规划在复杂系统的自然扩展, 鉴于自适应动态规划在深度控制系统发挥的重要作用, 平行控制与深度学习的结合为解决复杂系统的管理与控制问题提供了重要的解决思路.

5.5 控制性能评估

最后, 我们应该看到深度学习用于控制系统中的理论研究仍然欠缺. 目前没有理论能够评估使用了深度学习的控制系统的稳、准、快等性能. 虽然深度学习在控制系统中能够表现出一定的控制效果, 但是仅仅通过试错等方式不能保证控制性能. 这方面理论上的欠缺有可能阻碍深度学习在控制系统中的研究发展. 所以理论方面的研究也是一个非常值得注意的方向.

References

- 1 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015, **521**(7553): 436–444
- 2 Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. In: *Proceedings of the 2012 Advances in Neural Information Processing Systems*. Lake Tahoe, Nevada, USA: Curran Associates, Inc., 2012. 1097–1105
- 3 Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, **313**(5786): 504–507
- 4 Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, **18**(7): 1527–1554
- 5 Bengio Y, Lamblin P, Popovici D, Larochelle H. Greedy layer-wise training of deep networks. In: *Proceedings of the 2007 Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2007. 153–160
- 6 Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, **86**(11): 2278–2324
- 7 Sutskever I. Training recurrent neural networks [Ph.D. dissertation], University of Toronto, Canada, 2013
- 8 Bengio Y. Learning deep architectures for AI. *Foundations and Trends® in Machine Learning*, 2009, **2**(1): 1–127
- 9 Arel I, Rose D C, Karnowski T P. Deep machine learning — a new frontier in artificial intelligence research. *IEEE Computational Intelligence Magazine*, 2010, **5**(4): 13–18
- 10 Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(8): 1798–1828

- 11 Schmidhuber J. Deep learning in neural networks: an overview. *Neural Networks*, 2015, **61**: 85–117
- 12 Boulanger-Lewandowski N, Bengio Y, Vincent P. Modeling temporal dependencies in high-dimensional sequences: application to polyphonic music generation and transcription. arXiv: 12066392, 2012.
- 13 Schuster M, Paliwal K K. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 1997, **45**(11): 2673–2681
- 14 Yu J, Weng K, Liang G, Xie G. A vision-based robotic grasping system using deep learning for 3D object recognition and pose estimation. In: Proceedings of the 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO). Shenzhen, China: IEEE, 2013. 1175–1180
- 15 Lange S, Riedmiller M. Deep auto-encoder neural networks in reinforcement learning. In: Proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN). Barcelona: IEEE, 2010. 1–8
- 16 Mattner J, Lange S, Riedmiller M. Learn to swing up and balance a real pole based on raw visual input data. In: Proceedings of the 19th International Conference on Neural Information Processing. Doha, Qatar: Springer, 2012. 126–133
- 17 Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M. Playing Atari with deep reinforcement learning. arXiv: 1312.5602, 2013.
- 18 Mnih V, Kavukcuoglu K, Silver D, Rusu A A, Veness J, Bellemare M G, Graves A, Riedmiller M, Fidjeland A K, Ostrovski G, Petersen S, Beattie C, Sadik A, Antonoglou I, King H, Kumaran D, Wierstra D, Legg S, Hassabis D. Human-level control through deep reinforcement learning. *Nature*, 2015, **518**(7540): 529–533
- 19 Punjani A, Abbeel P. Deep learning helicopter dynamics models. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation. Seattle, WA: IEEE, 2015. 3223–3230
- 20 Lenz I, Knepper R, Saxena A. DeepMPC: learning deep latent features for model predictive control. In: Proceedings of Robotics: Science and Systems (RSS). Rome, Italy, 2015.
- 21 Anderson C W, Lee M, Elliott D L. Faster reinforcement learning after pretraining deep networks to predict state dynamics. In: Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN). Killarney: IEEE, 2015. 1–7
- 22 Cheon K, Kim J, Hamadache M, Lee D. On replacing PID controller with deep learning controller for DC motor system. *Journal of Automation and Control Engineering*, 2015, **3**(6): 452–456
- 23 Levine S. Exploring deep and recurrent architectures for optimal control. In NIPS (Neural Information Processing Systems) 2013 Workshop on Deep Learning, 2013. arXiv: 1311.1761, 2013.
- 24 Berniker M, Kording K P. Deep networks for motor control functions. *Frontiers in Computational Neuroscience*, 2015, **9**: 32
- 25 Wang F-Y, Kim H-M. Implementing adaptive fuzzy logic controllers with neural networks: a design paradigm. *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology*, 1995, **3**(2): 165–180
- 26 Saridis G N, Moed M C. Analytic formulation of intelligent machines as neural nets. In: Proceedings of the 1988 IEEE International Symposium on Intelligent Control. Arlington, VA: IEEE, 1988. 22–27
- 27 Moed M C, Saridis G N. A Boltzmann machine for the organization of intelligent machines. *IEEE Transactions on Systems, Man, and Cybernetics*, 1990, **20**(5): 1094–1102
- 28 Wang F-Y. Evolutionary Neuro-fuzzy Networks for Analysis of Complex Systems: a Memetic Approach, Technical report#03-09-99, Program for advanced research of complex systems, the University of Arizona, 1999
- 29 Wang F-Y. Modeling, analysis and synthesis of linguistic dynamic systems: a computational theory. In: Proceedings of the 1995 IEEE International Workshop on Architecture for Semiotic Modeling and Situation Control in Large Complex Systems. Monterey, CA: IEEE Press, 1995. 173–178
- 30 Wang Fei-Yue. Fundamental issues in research of computing with words and linguistic dynamic systems. *Acta Automatica Sinica*, 2005, **31**(6): 844–852
(王飞跃. 词计算和语言动力学系统的基本问题和研究. 自动化学报, 2005, **31**(6): 844–852)
- 31 Saridis G N, Stephanou H E. A hierarchical approach to the control of a prosthetic arm. *IEEE Transactions on Systems, Man, and Cybernetics*, 1977, **7**(6): 407–420
- 32 Bellman R. On the theory of dynamic programming. *Proceedings of the National Academy of Sciences of the United States of America*, 1952, **38**(8): 716–719
- 33 Bellman R. *Dynamic Programming*. Princeton: Princeton University Press, 1957.
- 34 Dreyfus S E, Law A M. *The Art and Theory of Dynamic Programming*. New York: Academic Press, 1977.
- 35 Werbos P J. Advanced forecasting methods for global crisis warning and models of intelligence. *General Systems Yearbook*, 1977, **22**(12): 25–38
- 36 Wang F-Y, Zhang H G, Liu D R. Adaptive dynamic programming: an introduction. *IEEE Computational Intelligence Magazine*, 2009, **4**(2): 39–47
- 37 Liu D R, Wang D, Wang F-Y, Li H L, Yang X. Neural-network-based online HJB solution for optimal robust guaranteed cost control of continuous-time uncertain nonlinear systems. *IEEE Transactions on Cybernetics*, 2014, **44**(12): 2834–2847
- 38 Liu D, Wang D, Li H. Decentralized stabilization for a class of continuous-time nonlinear interconnected systems using online learning optimal control approach. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, **25**(2): 418–428
- 39 Xu B, Yang C, Shi Z. Reinforcement learning output feedback NN control using deterministic learning technique. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, **25**(3): 635–641
- 40 Werbos P J. Approximate dynamic programming for real-time control and neural modeling. *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*. New York: Van Nostrand Reinhold, 1992. 493–525

- 41 Murray J J, Cox C J, Lendaris G G, Saeks R. Adaptive dynamic programming. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 2002, **32**(2): 140–153
- 42 Bertsekas D P. *Dynamic Programming and Optimal Control*. Massachusetts: Athena Scientific Belmont, 1996.
- 43 Sutton R S, Barto A G. *Reinforcement Learning: An Introduction*. Cambridge: MIT Press, 1998.
- 44 Al-Tamimi A, Lewis F L, Abu-Khalaf M. Discrete-time nonlinear HJB solution using approximate dynamic programming: convergence proof. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2008, **38**(4): 943–949
- 45 Wei Q, Liu D, Lin H. Value iteration adaptive dynamic programming for optimal control of discrete-time nonlinear systems. *IEEE Transactions on Cybernetics*, 2015, **46**(3): 840–853
- 46 Abu-Khalaf M, Lewis F L. Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. *Automatica*, 2005, **41**(5): 779–791
- 47 Zhang H G, Wei Q L, Liu D R. An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games. *Automatica*, 2011, **47**(1): 207–214
- 48 Bhasin S, Kamalapurkar R, Johnson M, Vamvoudakis K G, Lewis F L, Dixon W E. A novel actor-critic-identifier architecture for approximate optimal control of uncertain nonlinear systems. *Automatica*, 2013, **49**(1): 82–92
- 49 Liu D, Wei Q. Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, **25**(3): 621–634
- 50 Wang Fei-Yue. Parallel control: a method for data-driven and computational control. *Acta Automatica Sinica*, 2013, **39**(4): 293–302
(王飞跃. 平行控制: 数据驱动的计算控制方法. 自动化学报, 2013, **39**(4): 293–302)
- 51 Wang Fei-Yue. Parallel system methods for management and control of complex systems. *Control and Decision*, 2004, **19**(5): 485–489
(王飞跃. 平行系统方法与复杂系统的管理和控制. 控制与决策, 2004, **19**(5): 485–489)
- 52 Wang Fei-Yue. Computational theory and method on complex system. *China Basic Science*, 2004, **6**(5): 3–10
(王飞跃. 关于复杂系统研究的计算理论与方法. 中国基础科学, 2004, **6**(5): 3–10)
- 53 Wang Fei-Yue, Lansing J S. From artificial life to artificial societies — new methods for studies of complex social systems. *Complex Systems and Complexity Science*, 2004, **1**(1): 33–41
(王飞跃, 史帝夫·兰森. 从人工生命到人工社会 — 复杂社会系统研究的现状和展望. 复杂系统与复杂性科学, 2004, **1**(1): 33–41)
- 54 Wang Fei-Yue. On the modeling, analysis, control and management of complex systems. *Complex Systems and Complexity Science*, 2006, **3**(2): 26–34
(王飞跃. 关于复杂系统的建模、分析、控制和管理. 复杂系统与复杂性科学, 2006, **3**(2): 26–34)
- 55 Wang Fei-Yue. Artificial societies, computational experiments, and parallel systems: a discussion on computational theory of complex social-economic systems. *Complex Systems and Complexity Science*, 2004, **1**(4): 25–35
(王飞跃. 人工社会、计算实验、平行系统 — 关于复杂社会经济系统计算研究的讨论. 复杂系统与复杂性科学, 2004, **1**(4): 25–35)
- 56 Wang F-Y. Toward a paradigm shift in social computing: the ACP approach. *IEEE Intelligent Systems*, 2007, **22**(5): 65–67
- 57 Wang F-Y, Carley K M, Zeng D, Mao W. Social computing: from social informatics to social intelligence. *IEEE Intelligent Systems*, 2007, **22**(2): 79–83
- 58 Wang Fei-Yue. Study on cyber-enabled social movement organizations based on social computing and parallel systems. *Journal of University of Shanghai for Science and Technology*, 2011, **33**(1): 8–17
(王飞跃. 基于社会计算和平行系统的动态网民群体研究. 上海理工大学学报, 2011, **33**(1): 8–17)
- 59 Wang F-Y. Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications. *IEEE Transactions on Intelligent Transportation Systems*, 2010, **11**(3): 630–638
- 60 Zhu F, Wen D, Chen S. Computational traffic experiments based on artificial transportation systems: an application of ACP approach. *IEEE Transactions on Intelligent Transportation Systems*, 2013, **14**(1): 189–198
- 61 Wang F-Y, Wong P K. Intelligent systems and technology for integrative and predictive medicine: an ACP approach. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2013, **4**(2): 32
- 62 Duan W, Cao Z D, Wang Y Z, Zhu B, Zeng D, Wang F-Y, Qiu X G, Song H B, Wang Y. An ACP approach to public health emergency management: using a campus outbreak of H1N1 influenza as a case study. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2013, **43**(5): 1028–1041
- 63 Silver D, Huang A, Maddison C J, Guez A, Sifre L, Van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D. Mastering the game of Go with deep neural networks and tree search. *Nature*, 2016, **529**(7587): 484–489



段艳杰 中国科学院自动化研究所复杂系统管理与控制国家重点实验室博士研究生。主要研究方向为智能交通系统, 机器学习及应用。

E-mail: duanyanjie2012@ia.ac.cn

(DUAN Yan-Jie Ph.D. candidate at The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. Her research interest covers intelligent transportation systems, machine learning and its application.)



吕宜生 中国科学院自动化研究所复杂系统管理与控制国家重点实验室助理研究员. 主要研究方向为交通数据分析, 动态交通建模, 平行交通管理与控制系统.

E-mail: yisheng.lv@ia.ac.cn

(**LV Yi-Sheng** Assistant professor at The State Key Laboratory of Management and Control for Complex Systems,

Institute of Automation, Chinese Academy of Sciences. His research interest covers traffic data analysis, dynamic traffic modeling, and parallel traffic management and control systems.)

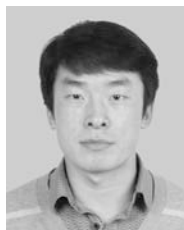


张 杰 中国科学院自动化研究所复杂系统管理与控制国家重点实验室助理研究员. 主要研究方向为拍卖机制, 最优控制与博弈论.

E-mail: jie.zhang@ia.ac.cn

(**ZHANG Jie** Assistant professor at The State Key Laboratory of Management and Control for Complex Systems,

Institute of Automation, Chinese Academy of Sciences. His research interest covers online auctions, optimal control and game theory.)



赵学亮 中国科学院自动化研究所复杂系统管理与控制国家重点实验室博士研究生, 中国自动化学会工程师. 主要研究方向为社会计算, 智能信息处理.

E-mail: xueliang.zhao@ia.ac.cn

(**ZHAO Xue-Liang** Ph. D. candidate at The State Key Laboratory of Management and Control for Complex

Systems, Institute of Automation, Chinese Academy of Sciences, engineer at Chinese Association of Automation. His research interest covers social computing and intelligent information processing.)



王飞跃 中国科学院自动化研究所复杂系统管理与控制国家重点实验室研究员. 主要研究方向为智能系统和复杂系统的建模, 分析与控制. 本文通信作者.

E-mail: feiyue.wang@ia.ac.cn

(**WANG Fei-Yue** Professor at The State Key Laboratory of Management and Control for Complex Systems, In-

stitute of Automation, Chinese Academy of Sciences. His research interest covers modeling, analysis, and control of intelligent systems and complex systems. Corresponding author of this paper.)