

求解微分方程的人工智能与深度学习方法：现状及展望

卢经纬^{1,2}, 程相^{1,3}, 王飞跃^{1,3}

(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 青岛 266114; 3. 中国科学院大学人工智能学院, 北京 100049)

摘要: 随着基础理论和硬件计算能力的飞速发展, 深度学习技术在众多领域取得了令人瞩目的成绩。作为描述客观物理世界的重要工具, 长期以来微分方程是各领域研究人员关心的重点。近年来, 深度学习和微分方程的结合逐渐成了研究的热点。由于深度学习能够从大量数据中高效地提取特征, 微分方程能够反应客观的物理规律, 因此二者的结合可以有效地提升深度学习的泛化性, 同时增强深度学习的可解释性。首先, 介绍了深度学习求解微分方程的基本问题。其次, 介绍了两类深度学习求解微分方程的方法: 数据驱动和物理知情方法。然后, 讨论了微分方程深度学习求解方法在实际中的应用。与此同时, 在充分调研的基础上提出了科学智能大模型——DeDAO (微分之道), 以应对现有的挑战。最后, 对微分方程深度学习求解方法进行了简要总结。

关键词: 人工智能; 深度学习; 神经网络; 微分方程

中图分类号: TP18

文献标志码: A

doi: 10.11959/j.issn.2096-6652.202255

Artificial intelligence and deep learning methods for solving differential equations: the state of the art and prospects

LU Jingwei^{1,2}, CHENG Xiang^{1,3}, WANG Fei-Yue^{1,3}

1. The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

2. Qingdao Academy of Intelligent Industries, Qingdao 266114, China

3. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: With the rapid advancement of fundamental theories and computing capacity, deep learning techniques have made impressive achievements in many fields. Differential equations, as an important tool for describing the physical world, have long been a focus of interest for researchers in various fields. Combining the two methods has gained popularity as a study issue in recent years. Since deep learning can efficiently extract features from large amounts of data and differential equations can reflect objective physical laws, the combination of the two can effectively improve the generalization ability of deep learning and enhance the interpretability of deep learning. Firstly, the problem of solving differential equations by deep learning was briefly introduced. Then, two types of deep learning methods for solving differential equations were introduced: data-driven and physical-informed methods. Furthermore, the applications of relevant deep learning-based solving methods were discussed. Meanwhile, DeDAO (differential equations DAO), a foundation model for artificial intelligence for science, was proposed to address existing challenges. Finally, conclusions of deep learning methods for solving differential equations were presented.

Key words: artificial intelligence, deep learning, neural network, differential equation

收稿日期: 2022-11-08; 修回日期: 2022-11-18

通信作者: 王飞跃, feiyue.wang@ia.ac.cn

基金项目: 国家自然科学基金资助项目 (No.U1811463); 行动元联合研究项目伺服驱动系统建模、决策和控制算法研究

Foundation Items: The National Natural Science Foundation of China (No.U1811463), Motion G, Inc. Collaborative Research Project for Modeling, Decision and Control Algorithms of Servo Drive Systems

0 引言

随着时间的变化,客观物理世界中的事物也在变化,如山脉侵蚀、河床演变、人口迁移、经济波动、技术进步等,大量客观物理规律变化都可以由时间的函数来描述。具体来说,可由以时间和其他变量为自变量的微分方程(differential equation, DE)来描述。微分方程是表示未知函数及其导数与自变量之间关系的方程^[1],主要包括:常微分方程(ordinary differential equation, ODE)、偏微分方程(partial differential equation, PDE)、随机微分方程(stochastic differential equation, SDE)、积分微分方程(integro-differential equation, IDE)以及微分代数方程(differential algebraic equation, DAE)。在上述微分方程中,PDE作为更一般形式的微分方程,可描述的物理规律更广泛、求解更复杂且具有更大的研究价值。ODE可视为PDE的特例。广泛研究的PDE包括物理学中的泊松方程(Poisson's equation)、纳维-斯托克斯方程(Navier-Stokes equation)、麦克斯韦方程(Maxwell's equation)、薛定谔方程(Schrödinger equation)及工程应用领域的哈密顿-雅可比-贝尔曼方程(Hamilton-Jacobi-Bellman equation)等。因此,微分方程描述的物理现象广泛存在于自然科学、民生、经济和工业等领域中,其求解方法研究具有重大的理论和实际工程价值。

由微分方程描述复杂动力学系统的建模和预测仍然是具有挑战性的问题。以地球系统为例,其动力学特性受到物理、化学和生物等过程相互作用影响,这些过程发生在跨越多个数量级的时空尺度^[2]。在过去的几十年中,研究人员尝试使用有限差分法(finite difference method)^[3]、有限元法(finite element method)^[4]、谱方法(spectral method)^[5]以及无网格方法(meshfree method)^[6]以数值形式求解PDE,并在许多动力学系统应用方面取得了成功。尽管传统方法取得了不错的效果,但多数结果只适用于低维度的非线性系统,利用上述传统方法对非线性的复杂巨系统的演化进行建模和预测难以实现。此外,微分方程反问题的求解仍是棘手的问题。实际应用中数据存在缺失、裂隙或噪声边界条件等问题,这些问题都极大地影响传统方法的求解精度。因此,亟须提出新的方法来解决上述问题。

近年来,人工智能(artificial intelligence, AI)及相关智能技术飞速发展,其中典型代表为深度学习(deep learning)^[7-10]。深度学习作为AI的重要分支,近年来取得了重大进展。大量研究表明,深度学习在处理高维复杂结构的数据时表现优异,因此适用于处理与复杂系统相关的问题。除了在计算机视觉^[11]、自然语言^[12]以及语音识别^[13]等传统AI问题方面取得了目前最高的水平,它还在围棋博弈^[14]、预测蛋白质结构^[15]、分析粒子加速器数据^[16]以及重建大脑电路^[17]等方面取得了举世瞩目的成绩。更令人惊喜的是,基于深度学习的大模型(big/foundation model)已经在自然语言和图像的各种任务中取得了超越人类的表现,如主题分类、情感分析、问题回答、语言翻译及图片生成等,是实现通用人工智能道路上非常有潜力的手段之一^[18]。

一般来说,传统机器学习或深度学习均通过数据充分挖掘复杂系统的特征和本质,并完成特定任务。不同的是,传统机器学习需手动设计特征,再利用神经网络(neural network)、支持向量机(support vector machine, SVM)和决策树(decision tree)等方法完成任务。而深度学习则通过复杂神经网络的设计,如卷积神经网络(convolutional neural network, CNN)、循环神经网络(recurrent neural network, RNN)、Transformer等,将特征设计的任务融入神经网络,实现了端到端的学习模式,充分挖掘神经网络的潜力,极大地提升了网络性能^[7]。

尽管深度学习取得了长足的进步,但一个不可忽略的事实是,目前深度学习的成功是建立在高质量训练数据集的基础上的,如PASCAL VOC、COCO、ParallelEye-CS、English Wikipedia、Corpus等^[18-20]。但随着研究人员面对的系统复杂程度不断提升,数据量爆炸式增长,纯数据驱动的深度学习模型并不能按预期随着数据量的增长而提升性能。此外,即便纯数据驱动的深度学习模型可以在给定的数据集上出色地拟合结果,数据噪声或观测偏差等因素也会导致模型泛化性能较差,一些模型在实际应用中并不具有实用性。因此,在数据驱动的框架外,需要额外引入专家知识指导深度学习模型进行学习,为模型提供“先验/专家知识”,即在观测数据的基础上提供可靠的理论约束和归纳偏置(inductive bias)^[21]。在自然科学中,广泛存在的数学、物理

定律可被视为专家知识,这种基于物理定律或物理定律和数据双驱动的学习方式被称为物理知情机器学习(physics-informed machine learning)^[21],其典型网络为物理知情神经网络(physics-informed neural network, PINN)^[22]。在20世纪,部分研究人员就开始尝试基于神经网络或相关的机器学习方法求解(偏)微分方程,并取得了不错的效果,如王飞跃等人^[23-24]将最佳格点集与最小二乘法结合,应用于求解平板弯曲微分方程问题。Lagaris I E等人^[25]采用神经网络求解微分方程。但当时该领域并无统一的名称,直至Raissi M等人^[22]提出PINN之后,才逐渐统一为物理知情机器学习。物理知情机器学习的出发点为借助相关物理定律或约束构造可解释的神经网络模型,在数据不完美(数据量小、存在噪声或观测偏差等)的情况下,仍保证模型的泛化性和有效性,且使模型的预测符合客观物理约束。近年来,该交叉方向逐渐形成了一个新兴领域“AI for Science(科学智能)”,旨在通过AI加速数学、物理、化学等基础学科的研究。神经信息处理系统进展大会(conference on neural information processing systems, NeurIPS)、国际机器学习大会(international conference on machine learning, ICML)等知名AI会议都举行了AI for Science相关的研讨会。

为了推进微分方程深度学习求解方法研究的进一步深入,本文对常用微分方程深度学习求解方法的国内外研究进展进行简要综述,论述了求解微分方程深度学习方法的研究背景、方法和应用。在总结和调研的基础上,进一步给出了微分方程深度学习求解方法以及AI和科学发展进一步结合的可能方向,即科学智能大模型。

本文第1节描述了基于深度学习方法求解微分方程的基本问题;第2节对常用微分方程深度学习求解方法的国内外研究进展进行简要综述;第3节介绍了常用微分方程深度学习求解方法的实际应用;第4节在充分调研的基础上提出了科学智能大模型——DeDAO;第5节进行了总结。

1 问题描述

本节采用数学的语言描述深度学习方法求解微分方程的问题。作为描述物理系统动态特性的重要数学工具之一,广泛研究的二阶微分方程可表示

为式(1)~式(2):

$$F(\mathbf{x}, u(\mathbf{x}), \nabla u(\mathbf{x}), \nabla^2 u(\mathbf{x}), \lambda) = 0, \mathbf{x} \in \Omega \quad (1)$$

$$B(u, \mathbf{x}) = 0, \mathbf{x} \in \partial\Omega \quad (2)$$

其中, $\mathbf{x} = [x_1, \dots, x_n]^\top \in \Omega$, $\Omega \subset \mathbb{R}^n$ 表示定义域。 $u(\mathbf{x})$ 为未知解。 ∇ 为微分算子,可以是ODE算子、PDE算子或者是IDE算子等。 λ 为方程中的参数。 $B(u, \mathbf{x})$ 为边界条件,如狄里克雷(Dirichlet)条件、诺依曼(Neumann)条件、洛平(Robin)条件或者一些周期边界条件。

微分方程的求解问题是指,在有限信息条件下,寻找满足式(1)与式(2)的解 u 与参数 λ 的组合。通常将微分方程的求解问题分为两类:正问题和反问题。正问题是在参数 λ 和边界条件已知的情况下,根据方程求出解析或者数值解,预测系统运动,产生观测数据的过程。反问题是指参数 λ 未知,利用有限的观测数据反推出最佳的参数 λ 。

由于偏微分方程具有较好的实用价值,大量的研究人员从事求解微分方程方法的相关研究工作。采用深度学习方法求解常微分方程和偏微分方程,近几年受到了大量学者的关注。本文以抛物形方程为例,介绍这类求解方法的基本框架,其发展方程如式(3)所示:

$$u_t + F(\mathbf{x}, u(\mathbf{x}), \nabla u(\mathbf{x}), \nabla^2 u(\mathbf{x}), \lambda) = 0 \quad (3)$$

其中, $t \in [0, T]$ 。该方程表示曲面沿其法线以与平均曲率成比例的速度运动的演化过程。采用深度神经网络来近似式(3)的解 $u(\mathbf{x}, t)$,并定义近似 $u(\mathbf{x}, t)$ 的神经网络为 $\hat{u}(\mathbf{x}, t)$,一个简单的3层前馈神经网络(feed forward network, FNN)的表示形式如式(4)所示:

$$\hat{u}(\mathbf{x}, t) = \theta_2^\top \sigma(\theta_1^\top [\mathbf{x}^\top, t]^\top + b_1) + b_2 \quad (4)$$

其中, $\theta = [\theta_1^\top, \theta_2^\top]^\top$ 为神经网络权重矩阵, $\mathbf{b} = [b_1, b_2]^\top$ 为偏置向量, $\sigma(\cdot)$ 为激活函数。进一步可得到偏微分方程的表示形式 $f(\mathbf{x}, t)$ 如式(5)所示:

$$f(\mathbf{x}, t) := \hat{u}_t + F(\mathbf{x}, \hat{u}, \nabla \hat{u}, \nabla^2 \hat{u}, \lambda) \quad (5)$$

可见 $f(\mathbf{x}, t)$ 由物理方程和深度神经网络组合表示,参考文献[22]称其为PINN。将初始和边界条件看作软约束处理,则可得到如式(6)所示的损失函数:

$$\text{Loss} = \text{MSE}_u + \text{MSE}_f \quad (6)$$

其中, MSE_u 表达式如式 (7) 所示, MSE_f 表达式如式 (8) 所示:

$$MSE_u = \frac{1}{N_u} \sum_{i=1}^{N_u} |u(t_u^i, x_u^i) - u^i|^2 \quad (7)$$

$$MSE_f = \frac{1}{N_f} \sum_{i=1}^{N_f} |f(t_f^i, x_f^i)|^2 \quad (8)$$

其中, $\{t_u^i, x_u^i, u^i\}_{i=1}^{N_u}$ 表示初始条件和边界条件产生的训练数据, 也可从实际应用中观测得到的数值, 可视为数据驱动的部分。 $\{t_f^i, x_f^i\}_{i=1}^{N_f}$ 表示采样数据, 用于计算控制方程的残差, 可视为物理定律驱动的部分。

显然, $f(\mathbf{x}, t)$ 的真实值为 0, 调整 $\hat{u}(\mathbf{x}, t)$ 的网络参数以最小化 $f(\mathbf{x}, t)$ 的残差, 同时满足初始条件和边界条件, 能够获得方程解的一种可行的数据驱动估计。深度学习求解微分方程的基本框架如图 1 所示。

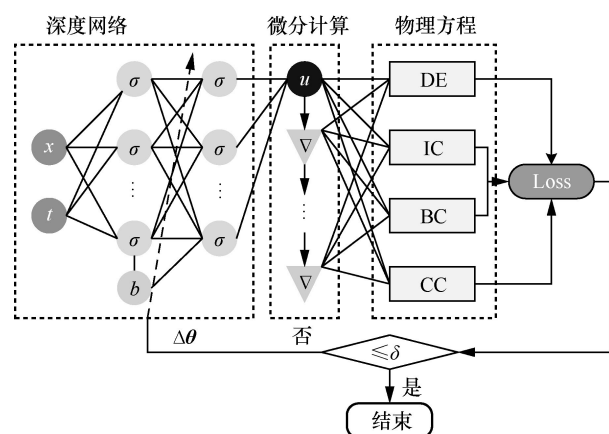


图1 深度学习求解微分方程基本框架

图1中, DE表示微分方程, IC表示初值条件(initial condition), BC表示边界条件(boundary condition), CC表示约束条件(constraint condition), δ 为收敛精度。在训练中, 可根据具体情况选择合适条件进行训练, 并非所有条件都必须采用。由图1可看出, 深度学习求解物理方程的方法将神经网络作为基本逼近单元, 调整网络参数以最小化原微分方程的残差。得益于深度神经网络强大的函数逼近能力, 该方法能够进行无网格求解, 得到封闭形式的解。网络训练基于最优化理论, 反向传播方法是最常用的方法。在训练过程中, 通常需要计算误差函数相对于网络参数的梯度, 采用自动微分技

术将微分方程的微分形式融入神经网络训练流程。基于神经网络近似解构造的微分方程 $f(\mathbf{x}, t)$ 被称为PINN^[22]。训练中, 损失函数可被写成两部分的组合。第一部分由初始或边界条件组成; 第二部分由微分算子和方程表达式组成, 根据正反问题需求, 方程系数可以是已知的或者是待估计的。最后, 训练网络参数以满足微分方程。在AI模型中嵌入物理方程, 能够加速训练, 增强模型的泛化能力。

2 微分方程求解的深度学习方法

2.1 方法概述及分类

由于其强大的学习和函数近似能力, 深度学习在很多应用领域, 如图像处理和机器翻译等, 取得了显著的成功。含多个隐藏层的多层前馈神经网络是可描述多种深度学习的网络结构, 式(4)表示的3层前馈神经网络就是一种简单的网络结构, 其基本结构被称为感知机, 它能够对输入进行线性和非线性变换。在过去的几十年中, 已经发展了很多种不同的神经网络, 如CNN、RNN和Transformer等。基于FNN的深度学习方法是解决大多数微分方程求解问题^[26]。将3层前馈神经网络推广成一般形式, 可得到一个L层深度FNN递归计算式如式(9)所示:

$$N_l(\mathbf{x}) = \sigma_l(\theta_l N_{l-1}(\mathbf{x}) + b_l), \quad 2 \leq l \leq L \quad (9)$$

其中, $N_l(\mathbf{x})$ 表示第l层神经网络的输出, θ_l 和 b_l 为权重向量和偏置, $\sigma_l(\cdot)$ 为第l层的激活函数。网络规模需根据方程解的复杂性来选择。

更复杂的深度学习网络, 已有相关研究者展开研究。参考文献[27]提出了ConvPDE-UQ框架, 采用卷积神经网络构造方程的数值求解器, 并建立了神经网络在不同域上近似PDE解的理论证明。Li Z J等人^[28]指出注意力机制能够提供一种灵活的方式利用输入中的隐藏模式, 及查询点和输入之间的隐式关系, 提出了一种基于注意力的数据驱动算子学习框架。

目前常见的微分计算方法主要有4种: 手动求导、符号微分、数值微分及自动微分。在深度学习框架中, 自动微分已经成为标配。自动微分基于链式法则递推的计算网络输出对参数的偏导数, 在输出变量维度较高、网络层数较多、参数数量庞大时, 计算效率更高, 更加适合深度神经网络的大规模计

算需求。除此之外,微分与卷积存在某些联系,对此 Dong B 等人^[29]建立了卷积核/滤波器的微分阶与微分算子的阶之间的对应关系。基于卷积与微分的对应关系,Long Z C 等人^[30]采用可学习的卷积近似微分算子,用深度网络表示偏微分方程,数据驱动地发现未知偏微分方程。

早期的 AI 方法主要采用数据驱动的方式,即通过学习“标签数据”来训练神经网络以表征物理系统输出的解。网络的输入可以是参数、空间、时间等,可根据需要选择。然而,对于一些工业系统,如油田和电力行业,尽管有许多运行数据,但可用于复杂物理系统建模的实验数据量是有限的。其中唯一可用的数据是边界条件和初始条件,而具体的控制偏微分方程和相关参数是可用或者可估计的。对此,图 2 描绘了实际研究中可能面临的部分情况。易于产生大数据的系统,如自然语言、绘画艺术等,存在少量的物理规律,基于数据驱动寻找其内在规律,是当前最有效的手段之一。对于物理知识丰富,而数据信息较少的情况,仅有初始条件和边界条件以及部分系数的数据,基于物理定律的分析和预测是唯一的途径。现实中还存在一种中间情况,即有部分数据和一些物理知识,基于物理知情机器学习方法整合数据和物理定律,弥补观测能力的不足和物理模型的缺失^[21]。综上,根据面对数据种类和数量的不同,求解偏微分方程的深度学习方法可以划分为数据驱动和物理知情两种方法。

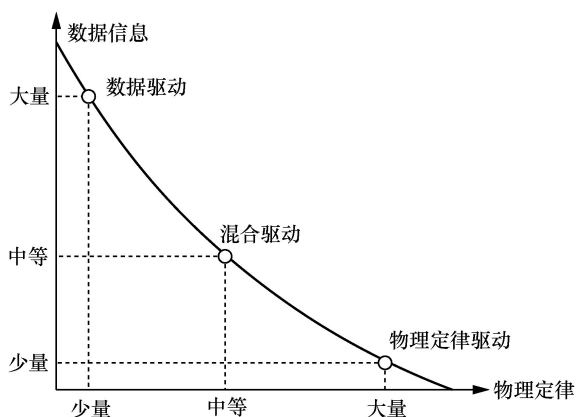


图2 数据和物理信息量不同的3种情形

2.2 数据驱动求解方法

不同于基于网格的求解方法,基于深度学习的驱动求解方法是用神经网络近似求解高维微分方程,而不是基函数的线性组合。寻求未知微分

方程或其算子的函数近似结构表示,通过学习相关数据找出背后蕴藏的微分方程模型,进而预测系统动力学特性。近年来,基于深度学习方法设计微分方程的数据驱动求解方法受到了大量关注。其中神经网络算子是当前学术界的研究热点之一,其设计的思想是期望采用神经网络表示各种显式与隐式算子,学习函数到泛函的映射能力,而不是表示传统意义上的函数,即变量到变量的映射关系,进而在输入空间上任意取一点,网络都可以返回相应的函数值。

针对 PDE 算子的近似问题,研究者期望神经网络学习的解算子是一个以函数为输入,以函数为输出的函数^[31-32]。采用神经网络逼近算子,有以下一般近似定理。假设 φ 是一个连续的非多项式函数, X 表示一个巴拿赫空间, K_1, K_2 为两个紧集,且有 $K_1 \subset X, K_2 \subset \mathbb{R}^d, V$ 为 $C(K_1)$ 中的紧集, G 为一个非线性连续算子表示,表示从 V 到 $C(K_1)$ 的映射。接下来,对于 $\varepsilon > 0$, 有正整数 p, q 与 r , 常数 $c_i^k, \xi_{ij}^k, \theta_i^k, \zeta_k \in \mathbb{R}, w_k \in \mathbb{R}^d, x_j \in K_1, i=1, \dots, n, k=1, \dots, p, j=1, \dots, r$, 则有式 (10):

$$G(u)(y) = \sum_{k=1}^p \sum_{i=1}^q c_i^k \varphi \left(\sum_{j=1}^m \xi_{ij}^k u(x_j) + \theta_i^k \right) \varphi(w_k y + \zeta_k) \quad (10)$$

对所有 $u \in V$ 和 $y \in K_2$ 都成立。该定理表明对于任意的非线性连续算子,总存在包含两个子网络的算子网络,其中分支网络用于编码输入函数,主干网络用于编码输出函数在特定位置上的值。因此,对于任意的非线性连续算子,总能找到满足式 (10) 的神经网络形式无限逼近。基于非线性算子的神经网络近似理论, Lu L 等人^[31]采用两个深度神经网络分别编码离散的输入函数空间和输出函数域,提出了一种具有较小泛化误差的深度算子网络: DeepONet, 并给出了该网络能够基于数据有效近似各种显式算子、各种确定和随机微分方程隐式算子的数学证明。神经算子的网络结构设计,基于神经网络关于算子的一般近似定理,能够以自监督的方式有效地学习微分方程的解算子。

随着微分方程维度的增长,高维微分方程的求解变得异常困难。采用深度学习,以数据驱动的方式实现高维微分方程的数值计算,同样受到了大量的关注。针对高维问题, Han J 等人^[33]设计了基于深度学习的高维抛物形偏微分方程求解方法,该方法类似于深度强化学习,将梯度作为策略函数,用

神经网络逼近未知解的梯度。参考文献[34]探索伽辽金方法 (Galerkin method) 和机器学习相结合的高维微分方程无网格求解方法, 并称该方法为“深度伽辽金方法 (deep Galerkin method, DGM)”, 同时该文指出随着隐藏层数的增加, 深度网络能更精确地收敛于偏微分方程的解。利用网格方法求解高维 PDE 几乎不具备可行性, 而深度学习方法能够通过随机采样的方式, 将 PDE 问题转化为机器学习问题, 避免了网格计算, 利用随机梯度下降等方法, 在随机采样的空间点上训练深度神经网络满足微分算子、初始条件和边界条件。因此, 在处理高维微分方程时, 基于深度学习的微分方程求解方法, 比传统方法更具有优势。

为了解决时变微分方程的参数估计问题, 参考文献[35]将测量数据作为节点值, 给出了一种未知时变偏微分方程的深度神经网络建模框架。参考文献[36]采用受限卷积核表示微分算子, 将二维受限卷积核推广到三维受限卷积核, 提出了三维偏微分方程智能求解方法。Long Z C 等人^[30]基于卷积核逼近微分算子的方法, 提出了 PDE-Net 2.0, 从观察数据中学习时间相关 PDE 的未知参数, 数值仿真结果表明该方法只需少量假设和先验知识, 即可具备高精度的长期预测能力。综上, 基于深度学习的数据驱动微分方程求解方法, 核心思想是在物理空间中进行数据驱动学习和建模, 实现方程及其中各项微分算子的智能表示形式, 进而得到方程的近似解。

2.3 物理知情求解方法

如前所述, 纯数据驱动的深度学习方法不能随着数据量的增长而提升性能, 且泛化性能不能得到保证。因此, 在机器学习中引入物理定律作为“知识/先验经验”的物理知情机器学习受到了广泛关注。

物理知情求解方法的基石工作为参考文献[22], 但类似思想的相关工作在 20 世纪末已有学者研究。王飞跃等人^[23]利用数论近似分析中的最佳格点集为最小二乘法提供一种简洁、公式化的布点格式, 并应用于求解平板弯曲问题, 取得了令人满意的效果。Meade A J 等人^[37]利用神经网络求解 OED, 将分段线性映射作为激活函数, 将 L_2 范数作为网络近似误差, 证明网络逼近误差范数随隐藏层神经元数量的增加而单调递减。Van Milligen B P 等人^[38]利用单层神经网络求解 PDE, 该方法较为简单, 且无须使用

有限差分法和坐标变换。Lagaris I E 等人^[25]将微分方程的解写成两部分: 第一部分满足初始/边界条件, 不包含可调参数; 第二部分的构造不影响初始/边界条件, 并涉及一个包含可调参数的 FNN。Mai-Duy N 等人^[39]采用径向基函数 (radial basis function, RBF) 和区域分解法 (domain decomposition method) 设计了无网格的数值求解方法, 并在泊松方程上实现了高精度的求解。Reynaldi A 等人^[40]将有限元法和神经网络结合, 采用反向传播算法求解正问题, 采用 Levenberg-Marquardt 算法求解逆问题, 且该方法中的计算均不涉及矩阵求逆。

回到近些年的工作, 参考文献[22]提出 PINN, 引起了科研界和工业界的广泛关注。之后 DeepMind 提出的 AlphaFold 2 成功预测了 2 亿多种蛋白质结构^[41], 唤起了研究者利用 AI 加速科学研究的热情, 催生了“AI for Science”。在参考文献[22]中, 网络训练的损失函数由两部分构成: PDE 的控制方程、初始/边界条件和网络实际输出之差。前者可视为物理定律驱动, 后者可视为数据驱动。这种使用客观物理方程作为损失函数的设计理念有效地规范了训练过程中的优化^[42]。实际上, 若中间数据带有标签, 同样可融入网络损失函数直接进行监督学习。PINN 将初始/边界条件均融入网络训练, 但并不一定严格精确地满足, 可视为一种“软约束”。为了进一步提升高维微分方程的求解精度, Bar-Sinai Y 等人^[43]尝试设计方程在网格尺度下逼近其动力学特性, 其基本思想为将 PINN 和有限差分法结合, 用 PINN 代替多项式函数, 该方法称为数据驱动离散化 (data-driven discretization), 可视为离散版的 PINN。类似地, Pang G F 等人^[44]采用时间离散化技术提出了 fPINN (fractional PINN), 用于求解时间分数阶偏微分方程。Gao H 等人^[45]提出了物理知情的总卷积神经网络: PhyGeoNet, 用于在不规则域上求解无标注数据的偏微分方程, 该方法引入一个椭圆映射将不规则物理域转换为规则参考域, 使非矩形几何形状和非均匀网格可直接使用 CNN。Zhu Y H 等人^[46]将微分方程控制方程融入损失函数, 得到了物理约束的 CNN 模型, 在无数据标签的情况下通过 K-L 散度 (Kullback-Leibler divergence) 进行训练。Sun L N 等人^[47]提出了一种物理约束的深度学习方法用于流体建模, 该方法通过结构化的网络设计使初始/边界条件严格满足, 且无须仿真数据。与参考文献[22]的方法相比, 参考文献[47]的方法可

视为“硬约束”。Meng X H 等人^[48]针对长时问题设计了 PPINN (parareal PINN), PPINN 将长时问题分解为多个短时问题,并采用粗粒度 (coarse-grained) 的求解器进行监督,取得了不错的效果。Yang L 等人^[49]提出了 B-PINN (Bayesian PINN) 以解决存在噪声数据时偏微分方程的正和逆问题,其中 PDE 的 PINN 和贝叶斯网络用作先验估计,哈密顿蒙特卡罗 (Hamiltonian Monte Carlo) 或变分推理 (variational inference) 用作后验估计。Wang Y Z 等人^[50]提出了 DSP-PIGAN 求解偏微分方程,该方法受启发于生成对抗网络 (generative adversarial network, GAN), 包括生成器、PINN 后处理模块和判别器,其基本思想为使用判别器判定输入是 PINN 后处理模块还是微分方程实际值,当判别器无法判定真伪时,则认为生成器和 PINN 后处理模块训练完成。基于 DeepONets, Wang S 等人^[51]进一步提出了物理知情的深度算子网络 (physics-informed DeepONets), 利用物理方程、边界条件等信息作为网络参数学习的约束,实现无任何配对输入输出训练数据情况下的高精度求解。值的指出的是,目前大部分物理知情求解方法均采用单网络,神经网络的各阶时间偏导数和空间偏导数利用 PyTorch^[52]和 Tensorflow^[53]中的自动微分技术处理获得,参考文献[54]为了减少对自动微分的依赖及计算冗余,提出了多网络的物理知情机器学习,其每个网络分别用于近似未知解函数及其各阶时间偏导数和空间偏导数,且互相独立。Han J H 等人^[55]基于采样布朗运动,针对拟线性椭圆形偏微分方程提出了一种深度学习数值方法,该方法先重构偏微分方程再利用神经网络近似解的梯度,因此无须显式地计算神经网络对输入变量的导数。

3 应用案例

近些年,已有国内外学者致力于尝试深度学习方法来求解工程应用中出现的复杂微分方程,并取得了相当丰富的成果。

首先是工业领域。参考文献[56]开展了深度算子网络在光热发电系统中的应用研究,使用预测值的历史数据替换主干网络中的时间坐标,提出了一种数据驱动的机器学习建模框架,案例分析表明,该方法能够以更低的计算成本实现对光热发电系统状态的高精度预测。Franklin T S 等人^[57]采用物理知情神经网络构造的虚拟传感器取代油井系统中

的传统传感器,能够结合系统动力学相关的先验知识,训练长短期记忆(long short term memory, LSTM)神经网络,而不需要测量所有状态。Shi H J 等人^[58]提出了一个端到端深度学习框架,同时学习库普曼 (Koopman) 内嵌函数和库普曼算子,以解决非线性系统库普曼函数的设计困难。参考文献[59]采用深度神经网络修正 Verruijt-Booker 解的围岩位移因子,以建立地表沉降与隧道开挖面空间位置的关联,通过构建数据-物理定律双驱动的 PINN 模型,实现深度网络在满足物理机理约束的空间中进行训练,缓解了地表沉降预测中对训练样本需求量较大的问题,实际应用结果表明,该策略能够提升模型的泛化性能。万鹏等人^[60]提出了一种基于元学习 (meta learning) 和 PINN 的刀具磨损预测方法,首先建立基于 PINN 的刀具磨损融合预测模型,然后采用元学习方法优化刀具磨损融合预测模型的损失函数,提高模型的鲁棒性,实验结果表明,该方法能够获得一个快速适应新工况的预测模型,在变工况条件下刀具磨损预测精度更加稳定。郑素佩等人^[61]提出了基于黏性耗散机制的正则化物理知情神经网络,将考虑黏性正则化的浅水波方程作为网络构建中的物理约束,训练深度网络用正则化方程的光滑解逼近原方程的间断解,并对满足不同初始条件的一维、二维浅水问题进行数值模拟,数值结果表明算法泛化能力强、可预测任意时刻的解、分辨率高、不会出现抹平和伪振荡现象。Goswami S 等人^[62]研究基于相场法的脆性断裂预测问题,不同于其他 PINN 算法最小化控制方程的残差,将系统的变分能量作为损失函数,修改神经网络输出,使损失函数中不存在边界条件损失分量,从而使网络能够完全满足边界条件。Zhang E R 等人^[63]采用深度神经网络分别近似正向问题的解和未知材料参数,通过深度学习方法构造物理知情深度神经网络,求解非均匀材料的辨识问题。

SDE 和柯尔莫哥洛夫 (Kolmogorov) 偏微分方程能够描述经济系统的动态变化,被广泛应用于金融系统的建模中,然而大多数柯尔莫哥洛夫偏微分方程的近似方法都会遇到维数灾难的问题,或者只能在单个固定时空点上实现偏微分方程解的近似。为了克服上述两个困难,Beck C 等人^[64]将柯尔莫哥洛夫偏微分方程求解问题转化为无限维随机优化问题,通过全连接深度神经网络近似时空离散化后的方程,进而提出了求解柯尔莫哥洛夫偏微分方程

的深度学习方法,并通过数值仿真验证了采用深度学习方法求解该类微分方程的可行性。Glau K 等人^[65]研究深度学习求解 PDE 方法在期权定价中的应用,采用无监督的方式训练深度,因而无须在解空间采样,所提出的深度参数偏微分方程方法(deep parametric PDE method)分为离线和在线两个阶段,离线阶段对神经网络进行训练,在线阶段进行状态、参数值和灵敏度的评估。

在其他科学领域,复杂微分方程的求解同样取得了大量的成果。参考文献[66]利用 PINN 分析了2020年新冠疫情的传播情况,并做出了以周为单位的短期预测,为疫情防控提供了科学分析基础。Grimm V 等人^[67]同样基于 PINN 分析了新冠疫情的感染率,所采用的方法以经典的 SIR(susceptible-infected-removed)和 SEIR(susceptible-exposed-infected-removed)模型为基础。Cavanagh H 等人^[68]利用 PINN 表示了亚洲大豆锈病的形态动力学特性,基于 PINN 以图像描述细胞-药物相互作用的形状变化(形态动力学),将 PINN 扩展到多模态数据的应用中。参考文献[69]基于物理知情机器学习提出了预测海洋环境中局部悬浮泥沙浓度的统计学习框架,可实现6h内的高精度预测,评价相对误差为5.80%~9.44%。Gross M R 等人^[70]开发了物理知情的机器学习自动工作流程用于预测裂缝马塞勒斯页岩储层的产量,该工作流程将快速降阶模型(reduced order model)与高保真油藏模型相结合,以匹配生产历史,并根据压力管理提供生产的实时预测。Kashinath K 等人^[71]将物理知情机器学习应用于天气预测,通过10个案例研究和分析展示了该方法如何成功用于模拟和预测天气过程。Pombo D V 等人^[72]探讨了基于堆叠机器学习的模型的实用性,使用物理知情机器学习方法预测同一地点潜在的风能和太阳能。参考文献[73]使用物理信息机器学习模型评估台风期间洪水预报中输入特征的有效空间特征。Bukhari A H 等人^[74]针对巴基斯坦拉合尔市,通过基于分数阶洛伦兹(Lorenz)的物理知情混合计算范式 SARFIMA-NARX,预测未来两天内每小时的 PM2.5 浓度值和空气质量。Rice J L 等人^[75]设计了基于物理知情机器学习的 Koopman 方法以预测长期的海面温度。

综上,微分方程深度学习求解方法在实际应用中的研究已经展开。同时需注意的是,实际应用中

多模态数据的引入为求解微分方程提供了大量值得挖掘的信息,但也给微分方程深度学习求解方法的实际应用带来了不小的挑战。同时,客观物理事物、参与其中的人、探索和改进这些客观事物和人类复杂系统的研究人员是一个有机的整体,不应忽略某些要素,或将其中的要素割裂开分析,尤其是其中难以预测的复杂社会和人的因素,应予以充分考虑。

4 迈向科学智能大模型: DeDAO

基于上述调研,本文认为微分方程的深度学习方法求解及科学智能的未来研究将围绕以下5个问题开展。

(1) 复杂高维微分方程(动力学系统)求解及预测。尽管相关深度方法在一些典型微分方程求解问题上取得了令人满意的效果,但面对高维复杂问题,尚无广泛认可的方法。参考文献[33,44]尝试基于时间、数据离散的方法分段逼近,以求解高维偏微分方程。但理论限制较多,且处理的数据相对单一。随着应用需求的增大,这一问题将日益凸显。

(2) 求解/预测精度和训练效率的提升。由于深度学习方法采用非线性函数近似的数值方法,而非解析方法,理论和实际应用中,网络的近似误差总是存在^[76]。因此降低近似误差以符合不同的应用场景始终是研究的重点。提升网络复杂度是一条可行的思路^[34,37],但随着网络复杂程度的提升,训练的时间和空间复杂度也随之提升,且过于庞大的神经网络在实际应用中并不现实。

(3) 神经网络近似解的泛化性。在给定定义域上,当训练数据分布合理且量足够的时候,通过适当的设计和训练,神经网络具有良好的拟合效果。但在给定数据集之外,神经网络的表现往往难以符合预期。在某些情况下,网络训练集上表现很好,而测试集上表现不如人意,该现象称为“过拟合”。泛化性是神经网络应用研究的重点。提升网络复杂度和提升数据覆盖面是有效的方法,但同样会带来巨大的计算量。若每个任务都训练一个庞大的网络,显然会造成冗余。

(4) 应用中多模态数据的引入。随着微分方程深度学习求解方法的潜力不断被挖掘,将其应用在复杂实际系统中或探索更为复杂的科学问题自然是下一步的研究重点。不同于一般微分方程的求解所处理的数值数据,实际应用中面临的情况将更为复杂,通常需面对图像、视频、自然语言等多模态

数据。参考文献[68]尝试基于图像描述细胞-药物相互作用的形态动力学,是该类方法迈向实际应用的初步探讨。然而如何从更宏观的角度设计多模态数据的处理方法,并加速相关方法的实际应用,仍悬而未决。

(5) 实际应用中复杂社会因素的融入。在微分方程求解及其他科学问题的理论研究和实际应用中,参与、探索和解决这些科学问题的人是不可忽略的社会因素。因此,如何对人的行为进行表示,将人的影响和评价考虑到微分方程中,也是值得研究的问题。但现有分析方法和框架通常忽略社会因素,阻碍了这些方法的进一步发展。因此,亟须建立一套完备的理论框架去探索、分析复杂系统中的社会因素,并将这些因素有效地融入解决方案。

尽管研究人员针对上述问题已有初步工作,但建立新的研究范式全面覆盖上述问题,且达到令人满意的效果,仍是一个具有挑战性的难题。

利用深度学习求解微分方程(或预测动力学系统)本质上是利用神经网络的非线性函数近似能力逼近微分方程的解。在一般偏微分方程中,其解析解可由多种函数形式构成,如幂函数、指数函数、对数函数等,可将其视为伽辽金方法中的基函数。因此,不同参数下同一微分方程(或不同微分方程)的解可由相同基函数及不同参数(或不同基函数及不同参数)构成。所以在实际研究中,每次针对一个微分方程训练一个网络,无疑是耗费时间且不必要的工作。同时,在实际应用中,随着深度学习模型复杂程度的提升,模型性能和泛化性也随之提升^[34,37],此外特定应用场景下物理系统的控制方程相似,因此构建超大神经网络模型求解复杂动力系统微分方程是一条有效的途径。为此,本节提出迈向科学智能大模型——DeDAO(微分之道)¹。DeDAO的目的是通过海量的微分方程数据学习微分方程背后的数学本质,如解中的“基函数”、微分算子等,当其应用于具体微分方程时,DeDAO基于迁移学习的方式通过少量的具体微分方程数据即可迅速求得高精度的解。同时引入多任务、多模态任务和复杂社会因素,加强DeDAO解决多种科学问题的综合能力。DeDAO将客观物理事物、参与其中的人、探索相关事物的研究人员视为一个有机的整体,将构成复杂系统的

各要素有机统一地结合起来,充分考虑其中难以预测的复杂社会和人的因素,尝试基于物理信息社会系统(cyber-physical-social systems, CPSS)处理科学问题^[77]。

4.1 支撑技术

在正式引入DeDAO前,首先给出相关支撑技术的介绍,包括:平行智能和大模型。

4.1.1 平行智能

大模型的训练离不开海量的数据,DeDAO也不例外。DeDAO的实现需要一种有效生成数据并合理训练的方法。此外,训练DeDAO还需大量的物理知识提供支持。在实际系统中,基于物理定律建立其物理方程的工作有相当大的难度,需要专业人士才能完成。因此,实现DeDAO迫切需要一套能够集成物理与数据、虚拟与现实的复杂系统智能理论框架。在众多理论框架中,平行智能建立于平行系统理论,是处理该问题最为行之有效的方法。平行智能框架来源于CPSS智能管理与控制的研究^[77],该框架下物理系统与虚拟系统不断交互、相互学习,不以简单的相互模仿为目的,而是以共同进化为最终目标,非常契合物理方程大模型的训练与应用方式。平行系统理论,由王飞跃教授提出^[78-80],可以追溯到影子系统(shadow systems)^[81]。平行系统理论也被称为ACP理论,其包括^[82-86]:人工系统(artificial systems)、计算实验(computational experiments)以及平行执行(parallel execution)。人工系统被用来模拟现实世界,也被称为人工社会(artificial society)。在计算实验中,可以采用多种智能方法处理复杂的系统。通过平行执行,分析来自实际和人工系统的反馈,以进一步提高复杂系统的建模、管理和控制的性能。平行系统的基本框架如图3所示。

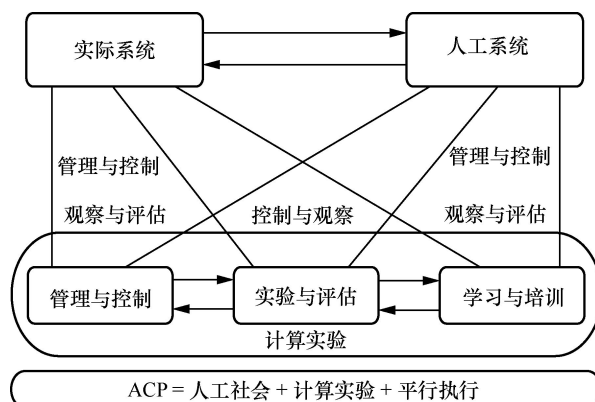


图3 平行系统的基本框架^[77]

1 道(DAO)取自道德经“道生一,一生二,二生三,三生万物”。

学系统，弥补实际方程或数据不足的缺点，同时为仿真和计算实验提供基本条件。

计算实验：计算实验提供计算实验平台，为 DeDAO 的训练和长期优化打下坚实基础，其中数据处理、模型训练均在计算实验中完成。

平行执行：通过平行执行获得人工和实际系统的运行结果，并进一步反馈优化计算实验。

同时，基于 CPSS 框架，DeDAO 可有效处理科学理论研究和实际应用中的社会要素，包括参与具体活动的人及探索和改进这些活动的人。这些人又可划分为：探索科学问题和从事实际应用的自然人（生物意义上的自然人，biological human）、协助探索科学问题和实际应用的机器人（robot）以及协助探索科学问题和实际应用的数字人（digital human）。自然人根据实际情况和需求提出问题并寻求答案；机器人主要协助自然人在物理世界完成繁杂的任务，减轻自然人的体力劳动，其主要活动范围为物理社会；数字人的主要任务为降低实际物理实验的昂贵成本、突破物理世界严苛的约束、为自然人的探索提供更为广阔的想象空间，为获得更为适合的解决方案打下基础，其主要活动范围为人工社会。3 类人基于 DAO（全中心化自主组织及全中心化自主运行）框架进行通信、组织和协调，其基本框架如图 5 所示。对于一个具体的科学问题，该框架主要包括：问题和需求提出层、问题求解层、求解验证层及应用层。自然人主要活动于问题和需求提出层，也出现在验证层和应用层，其借助 Web 3.0、区块链、边缘计算等智能技术与其他自然人、机器人和数字人实时加密通信，再协同完成工作；各层的自然人、机器人和数字人根据各层具体任务分工合作，并根据下一层反馈信息优化其工作。因此，该框架不仅包含层内的组织和协调，同时覆盖了层与层之间的组织和协调，各层的“人”完成所分配的工作即达成该层任务，每层工作均达成即实现自然人提出的总体目标，“人”与“人”之间的工作独立而又互补。故该工作框架具有典型分布式、去中心化、自主性、自动化、组织化与有序性的特征。以求解微分方程为例，在问题和需求提出层，自然人根据物理社会的现象和需求提出求解微分方程的问题和客观约束，问题和约束传递至数字人和机器人。在自然人的指导和协助下，数字人和机器人借助大模型提供的智能源泉，完成求解、验证和应用等任务。其中

层与层之间紧密耦合，每层的工作结果都是下层的起点，同时每层都为上层提供反馈信息，构成了一个包含社会因素的复杂大闭环系统。

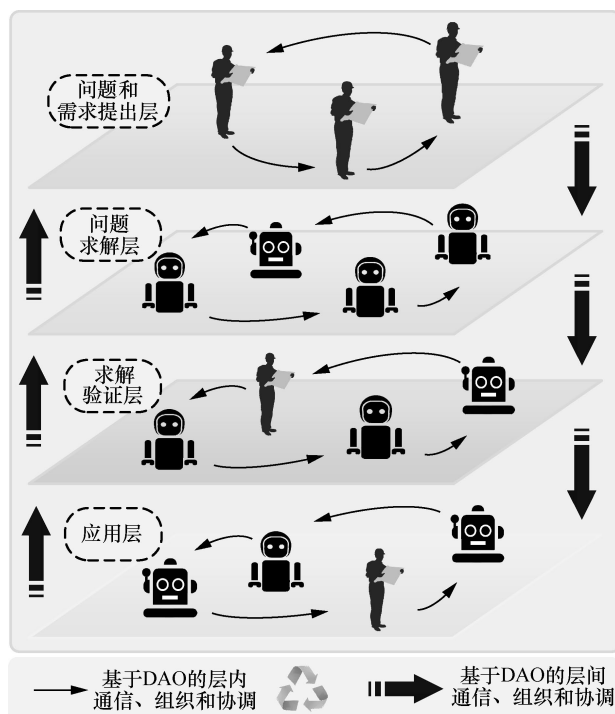


图5 自然人、机器人和数字人通信、组织和协调框架

DeDAO 的大平台基于 DAO 框架^[93]，采用云边协同架构，分为硬件平台和软件平台两个部分。硬件平台中的数据库连接各种传感器，存储多模态数据，包括环境数据、设备数据、人员数据等，为构建知识图谱和定义人工系统等高级功能提供可靠的数据服务。高性能计算环境提供云计算服务，采用分布式计算和并行计算等方式，快速实现大规模数据的处理和分析。移动端相对灵活，可以是专业设备、手机等具备一定计算能力的便携式设备，提供最低限度的通信、计算等功能，能够支撑起基于区块链的去中心化互联网 Web 3.0。软件平台包含了网络库、知识库和方法库等基本库，用于构建和训练大模型，提供两个重要的服务，即人工系统设计和任务导向的计算实验。子任务模块面向下游任务，支持场景建模、科学问题计算、可视化等功能。软件平台的操作系统应既能满足真人的使用需要，也能为数字人等提供管理操作的接口，提供实时的虚实交互和可视化功能，能够以场景工程的方式展示人工系统和实际系统的变化趋势。

DeDAO 的模型设计和实现主要包括如下 5 个

模块：数据预处理、单模态特征提取、多模态特征和专家知识融合、预训练任务设计及下游任务适配。各模块介绍如下。

- 数据预处理。DeDAO 旨在处理微分方程数值求解及其应用的相关问题，所以处理的数据不仅包含一般的数值数据，还包括实际应用中的自然语言、图像、视频等数据。为此，需要对数据进行预处理，包括填补数据、修复数据、结构化数据以及标准化（Tokenization）等，将这些数据统一处理成计算机可处理的形式。

- 单模态特征提取、多模态特征和专家知识融合。首先是单模态特征提取模块，该模块采用自注意机制从单模态数据中初步提取特征。但是不同微分方程背后的数学本质是一样的，可采用相同或类似的“基函数”进行描述，因此单模态特征的初步提取是不够的。为了进一步提取多模态数据中的特征并获得混合多模态特征，多模态特征和专家知识融合模块采用混合注意机制将单模态特征和专家知识（物理定律）进一步融合。同时，值得指出的是，单模态数据中提取出的特征并非都被使用，而是根据具体需求选择。因此多模态特征和专家知识融合模块采用了混合专家（mixture-of-experts）机制^[94]，即哪个特征参与最后的多模态和知识混合由该模块决定，这既减小了计算量，也有助于提升性能。

- 预训练任务。预训练任务设计是训练大模型的关键步骤。通常来说，预训练任务需满足两个条件：预训练任务与下游任务相关；预训练任务需以自监督的形式实现。第一个条件主要确保预训练任务提取的特征是适用于下游任务的；第二个条件则是因为大模型需要海量训练数据，如果每个数据都需手工标注，则很难实现，同时也不利于长期学习优化。为此，DeDAO 的预训练任务可设计为如下3类：单步/多步状态预测、微分方程求解及虚拟空间任务。由于 DeDAO 处理动力学系统，其状态预测是一个关键问题，且具有广泛的应用。对于单步/多步状态预测任务，可以先挖空一段数据再对其预测，做到自监督学习。对于微分方程求解任务，基于自然语言描述的微分方程知识、微分方程定律和数据等，可做到自监督，如在式（5）中令 $f(\mathbf{x}, t) = 0$ 。而对于虚拟空间任务，由计算机技术产生的虚拟任务，其数据和标签通常是成对出现的，无须额外的人工标注，这也是采用人工系统的好处。

- 下游任务适配。下游任务适配模块主要针对实际应用。由于预训练任务很难覆盖所有实际的下游任务，所以预训练任务的主要功能是从数据中提取尽可能全面而精确的特征。在下游任务适配模块，基于迁移学习等方式，通过少量数据的微调 DeDAO，使其可高效处理下游任务，并通过模型压缩、知识蒸馏等方式缩小模型，使其符合部署条件。同时，微调后的模型将同时应用于人工系统和实际系统，云端系统将持续跟踪和分析 DeDAO 在实际系统和人工系统中的性能。当性能不符合预期时，DeDAO 和人工系统会被优化以达到期望的效果。DeDAO 的设计、训练、运行和优化遵循平行系统“边缘简单，云端复杂”的基本原则。

4.3 DeDAO 之基于最佳格点集的最小二乘法

早在 20 世纪 80 年代，深度学习未兴起之前，研究人员已探索过数据与物理相结合的方法求解微分方程^[23-24]。与现如今方案不同的是，该方案将试函数作为近似分析的基本单元，残差中只包含控制方程的计算值，通过最小二乘法优化其中的待定参数。类似于式（4），可取如式（11）所示的试函数：

$$\hat{u}(\mathbf{x}, t, c) = u(\mathbf{x}, t) \quad (11)$$

其中， c 为待定参数，将式（11）带入式（3）可得残差如式（12）所示：

$$R^i(c, \mathbf{x}^i, t^i) = \hat{u}_i + F(\mathbf{x}, \hat{u}, \nabla \hat{u}, \nabla^2 \hat{u}, \lambda) \quad (12)$$

采用最小二乘法优化该残差的平方和，即可得到物理方程的近似解。从图 1 来看，将深度神经网络替换为试函数，物理方程中只包含微分方程，损失函数采用其平方和，即最小二乘法求解微分方程的经典结构之一。

由于微分方程的解为连续函数，因此定义域中点列 $\{\mathbf{x}^i\}_i^N$ 的选取与近似解的精度相关，也就是说在离散空间中表示原函数，需要采取一定的标准。对此，参考文献[23]采用最佳格点集来实现最小二乘法求解微分方程的配点，式（13）是常见的一种佳点集：

$$(\{kw_1\}, \{kw_2\}, \dots, \{kw_s\}), 1 \leq k \leq n \quad (13)$$

其中， $w_i = 2 \cos \frac{2\pi i}{p}$ ($1 \leq i \leq s$)， p 为素数，具有偏差 $\varphi(n) = c(\varepsilon)n^{-1+\varepsilon}$ 。采用一致分布使配点分布均匀，离散残差能较好地逼近连续型残差，这一点比

高斯配点法优越。

在计算机技术快速发展的当下, 采用人工神经元取代试函数作为基底能够形成通用的近似结构, 方便工程实践。采用物理知识辅助进行深度学习, 可实现无网格的求解, 借助人工智能解决物理中的数学问题。基于神经网络的通用逼近定律如式(10), 深度学习求解微分方程的方法, 有望取得与有限元法与有限体积法等常用数值方法相似甚至更好的精度。因此, 在 DeDAO 中, 本文采用深度神经网络描述方程解的特征, 以取代试函数的组合, 给出使用最小二乘法优化网络参数的一种求解方法框架, 其残差加权损失函数设计如式(14)所示:

$$\text{Loss} = \frac{1}{2} \begin{bmatrix} R_{\text{ph}} \\ R_{\text{ob}} \end{bmatrix}^T \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} R_{\text{ph}} \\ R_{\text{ob}} \end{bmatrix} \quad (14)$$

其中, R_{ph} 及 R_{ob} 如式(15)所示:

$$R_{\text{ph}} = \begin{bmatrix} R_{\text{ph}}^i(\theta, x^i, t^i) \\ \vdots \\ R_{\text{ph}}^{I_1}(\theta, x^{I_1}, t^{I_1}) \end{bmatrix}, R_{\text{ob}} = \begin{bmatrix} R_{\text{ob}}^i(\theta, x^i, t^i) \\ \vdots \\ R_{\text{ob}}^{I_2}(\theta, x^{I_2}, t^{I_2}) \end{bmatrix} \quad (15)$$

R_{ph} 和 R_{ob} 表示根据物理方程与已知观测观察数据计算的残差, 分别是物理和数据驱动的部分, 并乘以权重 A 和 B , 物理方程和观测数据由 DeDAO 数据预处理模块提供; θ 表示神经网络的权重参数。考虑到高斯法的局限性, 选择一致分布点集构造最小二乘法最佳配点, 对各类近似解的“基函数”进行编码^[95]。通过最小化损失函数确定一组权重参数, 满足对微分方程的近似求解。针对非线性优化问题, 采用迭代法进行网络参数优化。

针对不同应用场景, 存在数据有限或有噪声等问题, 可以在预训练任务过程中, 调节权重参数以缓解过拟合。此外, 基于最小二乘法优化损失函数的同时, 考虑将网络的结构和参数与有限元法、有限体积法等近似结果比较, 再借助数论分析待定参数的含义与变化, 探索物理知情深度神经网络的可解释性。

5 结束语

为促进微分方程深度学习求解方法研究的进一步深入, 本文对常用微分方程深度学习求解方法的研究现状进行简要综述, 将相关深度学习求解方法分为数据驱动和物理知情两类。同时介绍了相关方法在实际工程中的典型应用。在总结和调研的基础上, 本文

进一步给出了科学智能大模型——DeDAO。DeDAO 将深度学习和微分方程求解深度融合, 利用大模型强大的函数近似和泛化能力充分挖掘微分方程的数学本质, 提升微分方程深度学习求解方法的效率和精度, 同时 DeDAO 基于平行系统理论, 可处理包含复杂社会因素的科学问题及相关应用, 极大地促进了科学智能的发展和落地应用。

参考文献:

- [1] ZILL D. A first course in differential equations with modeling applications[M]. Stamford: Cengage Learning, 2012.
- [2] HART J K, MARTINEZ K. Environmental sensor networks: a revolution in the earth system science?[J]. Earth-Science Reviews, 2006, 78(3/4): 177-191.
- [3] THOMAS J. Numerical partial differential equations: finite difference methods[M]. New York: Springer Science & Business Media, 2013.
- [4] DONEA J, HUERTA A. Finite element methods for flow problems[M]. Hoboken: John Wiley & Sons, 2003.
- [5] FEIT M, FLECK J, STEIGER A. Solution of the Schrödinger equation by a spectral method[J]. Journal of Computational Physics, 1982, 47(3): 412-433.
- [6] LIU G R, GUY T. An introduction to meshfree methods and their programming[M]. New York: Springer Science & Business Media, 2005.
- [7] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [8] SUTTON R S, BARTO A G. Reinforcement learning: an introduction[M]. Cambridge: MIT Press, 2018.
- [9] 李涛, 魏庆来. 基于深度强化学习的智能暖气温度控制系统[J]. 智能科学与技术学报, 2020, 2(4): 348-353.
- [10] LI T, WEI Q L. Intelligent heating temperature control system based on deep reinforcement learning[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(4): 348-353.
- [11] 覃缘琪, 季青原, 葛俊, 等. 城市交通路网动态短时推理与精准预测研究[J]. 智能科学与技术学报, 2022, 4(3): 380-395.
- [12] QIN Y Q, JI Q Y, GE J, et al. Short-term traffic state reasoning and precise prediction in urban networks[J]. Chinese Journal of Intelligent Science and Technology, 2022, 4(3): 380-395.
- [13] 曾淦雄, 柯道. 基于 3D 卷积的图像序列特征提取与自注意力的车牌识别方法[J]. 智能科学与技术学报, 2021, 3(3): 268-279.
- [14] ZENG G X, KE X. 3D convolution-based image sequence feature extraction and self-attention for license plate recognition method[J]. Chinese Journal of Intelligent Science and Technology, 2021, 3(3): 268-279.
- [15] 杜倩龙, 宗成庆, 苏克毅. 利用上下文相似度增强词对齐效果的自然语言推理方法[J]. 智能科学与技术学报, 2020, 2(1): 26-35.
- [16] DU Q L, ZONG C Q, SU K Y. Enhancing alignment with context similarity for natural language inference[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(1): 26-35.
- [17] 洪依, 孙成立, 冷严. 基于超轻量通道注意力的端对端语音增强方法[J]. 智能科学与技术学报, 2021, 3(3): 351-358.
- [18] HONG Y, SUN C L, LENG Y. End-to-end speech enhancement based on ultra-lightweight channel attention[J]. Chinese Journal of Intelligent Science and Technology, 2021, 3(3): 351-358.

- [14] SILVER D, SCHRITTWIESER J, SIMONYAN K, et al. Mastering the game of Go without human knowledge[J]. *Nature*, 2017, 550(7676): 354-359.
- [15] JUMPER J, EVANS R, PRITZEL A, et al. Highly accurate protein structure prediction with AlphaFold[J]. *Nature*, 2021, 596(7873): 583-589.
- [16] CIODARO T, DEVA D, DE SEIXAS J M, et al. Online particle detection with neural networks based on topological calorimetry information[J]. *Journal of Physics: Conference Series*, 2012, 368.
- [17] HELMSTAEDTER M, BRIGGMAN K L, TURAGA S C, et al. Connectomic reconstruction of the inner plexiform layer in the mouse retina[J]. *Nature*, 2013, 500(7461): 168-174.
- [18] BOMMASANI R, HUDSON D A, ADELI E, et al. On the opportunities and risks of foundation models[J]. *arXiv preprint*, 2021, arXiv: 2108.07258.
- [19] LI X, WANG Y T, YAN L, et al. ParallelEye-CS: a new dataset of synthetic images for testing the visual intelligence of intelligent vehicles[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(10): 9619-9631.
- [20] EVERINGHAM M, GOOL L, WILLIAMS C K I, et al. The pascal visual object classes (VOC) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303-338.
- [21] KARNIADAKIS G E, KEVREKIDIS I G, LU L, et al. Physics-informed machine learning[J]. *Nature Reviews Physics*, 2021, 3(6): 422-440.
- [22] RAISSI M, PERDIKARIS P, KARNIADAKIS G E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. *Journal of Computational Physics*, 2019, 378: 686-707.
- [23] 王飞跃, 王景美, 谢贻权. 最佳格点集在最小二乘法中的应用[J]. *浙江大学学报(自然科学版)*, 1988, 22(3): 33-43.
WANG F Y, WANG J M, XIE Y Q. An application of good lattice points in least square method[J]. *Journal of Zhejiang University (Natural Science)*, 1988, 22(3): 33-43.
- [24] 王飞跃, 谢贻权. 最小二乘加权残值方法的递进算法[C]//第三届全国加权残值法会议论文集. 成都: 西南交通大学出版社, 1989: 312-316.
WANG F Y, XIE Y Q. Sequential algorithms for least square weighted residual methods[C]//*Proceedings of the 3rd National Conference on Weighted Residual Methods*. Chengdu: Southwest Jiaotong University Press, 1989: 312-316.
- [25] LAGARIS I E, LIKAS A, FOTIADIS D I. Artificial neural networks for solving ordinary and partial differential equations[J]. *IEEE Transactions on Neural Networks*, 1998, 9(5): 987-1000.
- [26] LU L, MENG X H, MAO Z P, et al. DeepXDE: a deep learning library for solving differential equations[J]. *SIAM Review*, 2021, 63(1): 208-228.
- [27] WINOVICH N, RAMANI K, LIN G. ConvPDE-UQ: convolutional neural networks with quantified uncertainty for heterogeneous elliptic partial differential equations on varied domains[J]. *Journal of Computational Physics*, 2019, 394: 263-279.
- [28] LI Z J, MEIDANI K, FARIMANI A. Transformer for partial differential equations' operator learning[J]. *arXiv preprint*, 2022, arXiv: 2205.13671.
- [29] DONG B, JIANG Q T, SHEN Z W. Image restoration: wavelet frame shrinkage, nonlinear evolution PDEs, and beyond[J]. *Multiscale Modeling & Simulation*, 2017, 15(1): 606-660.
- [30] LONG Z C, LU Y P, DONG B. PDE-Net 2.0: learning PDEs from data with a numeric-symbolic hybrid deep network[J]. *Journal of Computational Physics*, 2019, 399.
- [31] LU L, JIN P Z, PANG G F, et al. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators[J]. *Nature Machine Intelligence*, 2021, 3(3): 218-229.
- [32] WANG S F, WANG H W, PERDIKARIS P. Improved architectures and training algorithms for deep operator networks[J]. *Journal of Scientific Computing*, 2022, 92(2): 1-42.
- [33] HAN J, JENTZEN A, E W. Solving high-dimensional partial differential equations using deep learning[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2018, 115(34): 8505-8510.
- [34] SIRIGNANO J, SPILIOPOULOS K. DGM: a deep learning algorithm for solving partial differential equations[J]. *Journal of Computational Physics*, 2018, 375: 1339-1364.
- [35] CHEN Z, CHURCHILL V, WU K L, et al. Deep neural network modeling of unknown partial differential equations in nodal space[J]. *Journal of Computational Physics*, 2022, 449.
- [36] ZHA W S, ZHANG W, LI D L, et al. Convolution-based model-solving method for three-dimensional, unsteady, partial differential equations[J]. *Neural Computation*, 2022, 34(2): 518-540.
- [37] MEADE A J, FERNANDEZ A A. Solution of nonlinear ordinary differential equations by feedforward neural networks[J]. *Mathematical and Computer Modelling*, 1994, 20(9): 19-44.
- [38] VAN MILLIGEN B P, TRIBALDOS V, JIMÉNEZ J A. Neural network differential equation and plasma equilibrium solver[J]. *Physical Review Letters*, 1995, 75(20): 3594-3597.
- [39] MAI-DUY N, TRAN-CONG T. Mesh-free radial basis function network methods with domain decomposition for approximation of functions and numerical solution of Poisson's equations[J]. *Engineering Analysis with Boundary Elements*, 2002, 26(2): 133-156.
- [40] REYNALDI A, LUKAS S, MARGARETHA H. Backpropagation and levenberg-marquardt algorithm for training finite element neural network[C]//*Proceedings of 2012 6th UKSim/AMSS European Symposium on Computer Modeling and Simulation*. Piscataway: IEEE Press, 2012: 89-94.
- [41] JUMPER J, EVANS R, PRITZEL A, et al. Highly accurate protein structure prediction with AlphaFold[J]. *Nature*, 2021, 596(7873): 583-589.
- [42] RAISSI M, KARNIADAKIS G E. Hidden physics models: machine learning of nonlinear partial differential equations[J]. *Journal of Computational Physics*, 2018, 357: 125-141.
- [43] BAR-SINAI Y, HOYER S, HICKEY J, et al. Learning data-driven discretizations for partial differential equations[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2019, 116(31): 15344-15349.
- [44] PANG G F, LU L, KARNIADAKIS G E. fPINNs: fractional physics-informed neural networks[J]. *SIAM Journal on Scientific Computing*, 2019, 41(4): A2603-A2626.
- [45] GAO H, SUN L N, WANG J X. PhyGeoNet: physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain[J]. *Journal of Computational Physics*, 2021, 428.

- [46] ZHU Y H, ZABARAS N, KOUTSOURELAKIS P-S, et al. Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data[J]. *Journal of Computational Physics*, 2019, 394: 56-81.
- [47] SUN L N, GAO H, PAN S W, et al. Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data[J]. *Computer Methods in Applied Mechanics and Engineering*, 2020, 361.
- [48] MENG X H, LI Z, ZHANG D K, et al. PPINN: parareal physics-informed neural network for time-dependent PDEs[J]. *Computer Methods in Applied Mechanics and Engineering*, 2020, 370.
- [49] YANG L, MENG X H, KARNIADAKIS G E. B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data[J]. *Journal of Computational Physics*, 2021, 425.
- [50] WANG Y Z, SUN H, SUN G Z. DSP-PIGAN: a precision-consistency machine learning algorithm for solving partial differential equations[C]//*Proceedings of 2021 13th International Conference on Machine Learning and Computing*. [S.l.:s.n.], 2021: 21-26.
- [51] WANG S, WANG H, PERDIKARIS P. Learning the solution operator of parametric partial differential equations with physics-informed DeepONets[J]. *Science Advances*, 2021, 7(40).
- [52] PASZKE A, GROSS S, MASSA F, et al. Pytorch: an imperative style, high-performance deep learning library[C]//*Proceedings of Advances in Neural Information Processing Systems*. [S.l.:s.n.], 2019.
- [53] ABADI M, AGARWAL A, BARHAM P, et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems[J]. *arXiv preprint*, 2015, arXiv: 1603.04467.
- [54] 王昀卓. 求解偏微分方程的神经网络方法[D]. 合肥: 中国科学技术大学, 2021.
WANG Y Z. Neural network for solving partial differential equations[D]. Hefei: University of Science and Technology of China, 2021.
- [55] HAN J H, NICA M, STINCHCOMBE A R. A derivative-free method for solving elliptic partial differential equations with deep neural networks[J]. *Journal of Computational Physics*, 2020, 419.
- [56] OSORIO J D, WANG Z, KARNIADAKIS G, et al. Forecasting solar-thermal systems performance under transient operation using a data-driven machine learning approach based on the deep operator network architecture[J]. *Energy Conversion and Management*, 2022, 252.
- [57] FRANKLIN T S, SOUZA L S, FONTES R M, et al. A physics-informed neural networks (PINN) oriented approach to flow metering in oil wells: an ESP lifted oil well system as a case study[J]. *Digital Chemical Engineering*, 2022, 5.
- [58] SHI H J, MENG M Q H. Deep koopman operator with control for nonlinear systems[J]. *IEEE Robotics and Automation Letters*, 2022, 7(3): 7700-7707.
- [59] 张子龙, 潘秋景, 仇文岗, 等. 基于物理信息神经网络的盾构隧道诱发地表沉降预测[J]. *工程力学*, 2022, 39: 1-14.
ZHANG Z L, PAN Q J, ZHANG W G, et al. Prediction of surface settlements induced by shield tunnelling using physics-informed neural networks[J]. *Engineering Mechanics*, 2022, 39: 1-14.
- [60] 万鹏, 李迎光, 华家圻, 等. 基于元学习和 PINN 的变工况刀具磨损精确预测方法[J]. *南京航空航天大学学报*, 2022, 54(3): 387-396.
WAN P, LI Y G, HUA J Q, et al. Accurate prediction method of tool wear under varying cutting conditions based on meta learning and PINN[J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2022, 54(3): 387-396.
- [61] 郑素佩, 林云云, 封建湖, 等. 浅水波方程的粘性正则化 PINN 算法[J]. *计算物理*, 2022: 1-15.
ZHENG S P, LIN Y Y, FENG J H, et al. Viscous regularization PINN algorithm for shallow water equations[J]. *Chinese Journal of Computational Physics*, 2022: 1-15.
- [62] GOSWAMI S, ANITESCU C, CHAKRABORTY S, et al. Transfer learning enhanced physics informed neural network for phase-field modeling of fracture[J]. *Theoretical and Applied Fracture Mechanics*, 2020, 106.
- [63] ZHANG E R, YIN M L, KARNIADAKIS G E. Physics-informed neural networks for nonhomogeneous material identification in elasticity imaging[J]. *arXiv preprint*, 2020, arXiv:2009.04525.
- [64] BECK C, BECKER S, GROHS P, et al. Solving the Kolmogorov PDE by means of deep learning[J]. *Journal of Scientific Computing*, 2021, 88(3): 1-28.
- [65] GLAU K, WUNDERLICH L. The deep parametric PDE method and applications to option pricing[J]. *Applied Mathematics and Computation*, 2022, 432.
- [66] BARMPARIS G D, TSIRONIS G P. Physics-informed machine learning for the COVID-19 pandemic: adherence to social distancing and short-term predictions for eight countries[J]. *Quantitative Biology*, 2022, 10(2): 139.
- [67] GRIMM V, HEINLEIN A, KLAUWONN A, et al. Estimating the time-dependent contact rate of SIR and SEIR models in mathematical epidemiology using physics-informed neural networks[J]. *Electronic Transactions on Numerical Analysis*, 2021, 56: 1-27.
- [68] CAVANAGH H, MOSBACH A, SCALLIET G, et al. Physics-informed deep learning characterizes morphodynamics of Asian soybean rust disease[J]. *Nature Communications*, 2021, 12: 1-9.
- [69] ZHANG S T, WU J R, WANG Y G, et al. A physics-informed statistical learning framework for forecasting local suspended sediment concentrations in marine environment[J]. *Water Research*, 2022, 218.
- [70] GROSS M R, HYMAN J D, SRINIVASAN S, et al. A physics-informed machine learning workflow to forecast production in a fractured Marcellus shale reservoir[C]//*Proceedings of the 9th Unconventional Resources Technology Conference*. Tulsa: American Association of Petroleum Geologists, 2021.
- [71] KASHINATH K, MUSTAFA M, ALBERT A, et al. Physics-informed machine learning: case studies for weather and climate modelling[J]. *Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences*, 2021, 379(2194).
- [72] POMBO D V, RINCON M J, BACHER P, et al. Assessing stacked physics-informed machine learning models for co-located wind-solar power forecasting[J]. *Sustainable Energy, Grids and Networks*, 2022, 32.
- [73] JHONG B C, LIN C Y, JHONG Y D, et al. Assessing the effective spatial characteristics of input features through physics-informed machine learning models in inundation forecasting during typhoons[J]. *Hydrological Sciences Journal*, 2022, 67(10): 1527-1545.
- [74] BUKHARI A H, RAJA M A Z, SHOAIB M, et al. Fractional order Lorenz based physics informed SARFIMA-NARX model to monitor and mitigate megacities air pollution[J]. *Chaos, Solitons & Fractals*, 2022, 161.
- [75] RICE J L, XU W W, AUGUST A. Analyzing Koopman approaches to physics-informed machine learning for long-term sea-surface temper-

- ature forecasting[J]. arXiv preprint, 2020, arXiv:2010.00399.
- [76] HORNIK K, STINCHCOMBE M, WHITE H. Multilayer feedforward networks are universal approximators[J]. Neural Networks, 1989, 2(5): 359-366.
- [77] WANG F Y. The emergence of intelligent enterprises: from CPS to CPSS[J]. IEEE Intelligent Systems, 2010, 25(4): 85-88.
- [78] 王飞跃. 平行系统方法与复杂系统的管理和控制[J]. 控制与决策, 2004, 19(5): 485-489, 514.
- WANG F Y. Parallel system methods for management and control of complex systems[J]. Control and Decision, 2004, 19(5): 485-489, 514.
- [79] 王飞跃. 平行控制与数字孪生: 经典控制理论的回顾与重铸[J]. 智能科学与技术学报, 2020, 2(3): 293-300.
- WANG F Y. Parallel control and digital twins: control theory revisited and reshaped[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(3): 293-300.
- [80] WANG F Y. The DAO to MetaControl for MetaSystems in metaverses: the system of parallel control systems for knowledge automation and control intelligence in CPSS[J]. IEEE/CAA Journal of Automatica Sinica, 2022, 9(11): 1899-1908.
- [81] WANG F Y. Shadow systems: a new concept for nested and embedded co-simulation for intelligent systems[R]. 1994.
- [82] WEI Q L, LI H Y, WANG F Y. Parallel control for continuous-time linear systems: a case study[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(4): 919-928.
- [83] LU J W, WEI Q L, WANG F Y. Parallel control for optimal tracking via adaptive dynamic programming[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(6): 1662-1674.
- [84] LU J W, WEI Q L, WANG Z Y, et al. Event-triggered optimal control for discrete-time multi-player non-zero-sum games using parallel control[J]. Information Sciences, 2022, 584: 519-535.
- [85] LU J W, WEI Q L, LIU Y J, et al. Event-triggered optimal parallel tracking control for discrete-time nonlinear systems[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(6): 3772-3784.
- [86] LU J W, WEI Q L, ZHOU T M, et al. Event-triggered near-optimal control for unknown discrete-time nonlinear systems using parallel control[J]. IEEE Transactions on Cybernetics, 2022.
- [87] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. [S.l.:s.n.], 2019: 4171-4186.
- [88] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners[C]//Proceedings of the Conference on Neural Information Processing Systems. [S.l.:s.n.], 2020: 1877-1901.
- [89] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision[C]//Proceedings of the International Conference on Machine Learning. [S.l.:s.n.], 2021: 8748-8763.
- [90] ALAYRAC J B, DONAHUE J, LUC P, et al. Flamingo: a visual language model for few-shot learning[J]. arXiv preprint, 2022, arXiv:2204.14198.
- [91] SU B, DU D, YANG Z, et al. A molecular multimodal foundation model associating molecule graphs with natural language[J]. arXiv preprint, 2022, arXiv:1810.05481.
- [92] XIE Z D, ZHANG Z, CAO Y, et al. SimMIM: a simple framework for masked image modeling[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 9643-9653.
- [93] 王飞跃. 平行管理: 复杂性管理智能的生态科技与智慧管理之DAO[J]. 自动化学报, 2022, 48(11): 2655-2669.
- WANG F Y. Parallel management: the DAO to smart ecological technology for complexity management intelligence[J]. Acta Automatica Sinica, 2022, 48(11): 2655-2669.
- [94] SHAZEER N, MIRHOSEINI A, MAZIARZ K, et al. Outrageously large neural networks: the sparsely-gated mixture-of-experts layer[C]//Proceedings of the International Conference on Learning Representations. [S.l.:s.n.], 2017.
- [95] CHENG X, LU J, WANG F Y. Deep learning meets optimal lattices for continuous traffic flow prediction[J]. International Journal of Intelligent Control and Systems, 2022, 2(1): 7-12.

[作者简介]



卢经纬(1990—), 男, 青岛智能产业技术研究院助理研究员, 主要研究方向为最优控制、平行控制、自适应动态规划和深度强化学习。



程相(1994—), 男, 中国科学院自动化研究所博士生, 主要研究方向为智慧油田、深度学习和平行控制。



王飞跃(1961—), 男, 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主任, 主要研究方向为平行系统的方法与应用、社会计算、平行智能以及知识自动化。