

## 基于神经网络的偏微分方程求解方法研究综述

查文舒, 李道伦, 沈路航, 张 雯, 刘旭亮

## REVIEW OF NEURAL NETWORK-BASED METHODS FOR SOLVING PARTIAL DIFFERENTIAL EQUATIONS

Zha Wenshu, Li Daolun, Shen Luhang, Zhang Wen, and Liu Xuliang

在线阅读 View online: <https://doi.org/10.6052/0459-1879-21-617>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 单一水平轴风电机组尾迹的模拟方法与流动机理研究综述

REVIEW OF RESEARCH ON THE SIMULATION METHOD AND FLOW MECHANISM OF A SINGLE HORIZONTAL-AXIS WIND TURBINE WAKE

力学学报. 2021, 53(12): 3169-3178

#### 线绳驱动转速提升式低频俘能器的设计与研究

DEVELOPMENT OF A LOW-FREQUENCY HARVESTER BASED ON A ROPE-DRIVEN ROTOR WITH ROTATION SPEED UP-REGULATION FUNCTION

力学学报. 2021, 53(11): 3025-3034

#### 一种压电驱动的三足爬行机器人

A PIEZOELECTRIC-DRIVEN THREE-LEGGED CRAWLING ROBOT

力学学报. 2021, 53(12): 3356-3367

#### 基于压电振动能量俘获的弯曲结构损伤监测研究

THE RESEARCH ON DAMAGE DETECTION OF CURVED BEAM BASED ON PIEZOELECTRIC VIBRATION ENERGY HARVESTER

力学学报. 2021, 53(11): 3035-3044

#### 基于S-ALE方法的圆柱体垂直出水破冰研究

RESEARCH ON VERTICAL MOVEMENT OF CYLINDRICAL STRUCTURE OUT OF WATER AND BREAKING THROUGH ICE LAYER BASED ON S-ALE METHOD

力学学报. 2021, 53(11): 3110-3123

#### 多个椭圆柱波浪力的一种解析解

AN ANALYTICAL SOLUTION FOR WAVE PRESSURE ON ARRAYS OF ELLIPTICAL BODIES

力学学报. 2021, 53(11): 3157-3167



关注微信公众号, 获得更多资讯信息

# 基于神经网络的偏微分方程求解方法研究综述<sup>1)</sup>

查文舒<sup>2)</sup> 李道伦<sup>3)</sup> 沈路航 张雯 刘旭亮

(合肥工业大学数学学院, 合肥 230009)

**摘要** 神经网络作为一种强大的信息处理工具在计算机视觉, 生物医学, 油气工程领域得到广泛应用, 引发多领域技术变革. 深度学习网络具有非常强的学习能力, 不仅能发现物理规律, 还能求解偏微分方程. 近年来基于深度学习的偏微分方程求解已是研究新热点. 遵循于传统偏微分方程解析解、偏微分方程数值解术语, 本文称用神经网络进行偏微分方程求解的方法为偏微分方程智能求解方法或偏微分方程神经网络求解方法. 本文首先简要介绍偏微分方程智能求解发展历程, 然后从反演未知偏微分方程与求解已知偏微分方程两个角度展开讨论, 重点讨论已知偏微分方程的求解方法. 根据神经网络中损失函数的构建方式, 将偏微分方程求解方法分为 3 大类: 第 1 类是数据驱动, 主要从数据中学习偏微分方程, 可以应用于恢复方程、参数反演等; 第 2 类是物理约束, 即在数据驱动的基础上, 辅以物理约束, 在损失函数中加入控制方程等物理规律, 减少网络对标签数据的依赖, 大幅提高泛化能力与应用价值; 第 3 类物理驱动 (纯物理约束), 完全不使用标签数据, 仅通过物理规律求解偏微分方程, 目前仅适用于简单偏微分方程. 本文从这 3 个方面介绍偏微分方程智能求解的研究进展, 涉及全连接神经网络、卷积神经网络、循环神经网络等多种网络结构. 最后总结偏微分方程智能求解的研究进展, 给出相应的应用场景以及未来研究展望.

**关键词** 神经网络, 偏微分方程智能求解, 数据驱动, 物理约束

中图分类号: O241 文献标识码: A doi: [10.6052/0459-1879-21-617](https://doi.org/10.6052/0459-1879-21-617)

## REVIEW OF NEURAL NETWORK-BASED METHODS FOR SOLVING PARTIAL DIFFERENTIAL EQUATIONS<sup>1)</sup>

Zha Wenshu<sup>2)</sup> Li Daolun<sup>3)</sup> Shen Luhang Zhang Wen Liu Xuliang

(School of Mathematics, Hefei University of Technology, Hefei 230009, China)

**Abstract** Neural networks are widely used as a powerful information processing tool in the fields of computer vision, biomedicine, and oil-gas engineering, triggering technological changes. Due to the powerful learning ability, deep learning networks can not only discover physical laws but also solve partial differential equations (PDEs). In recent years, PDE solving based on deep learning has been a new research hotspot. Following the terms of traditional PDE analytical solution, this paper calls the method of solving PDE by neural network as PDE intelligent solution or PDE neural-network solution. This paper briefly introduces the development history of PDE intelligent solution, and then discusses the development of recovering unknown PDEs and solving known PDEs. The main focus of this paper is on a neural network solution method for a known PDE. It is divided into three categories according to the way of constructing loss

2021-11-23 收稿, 2022-01-10 录用, 2022-01-11 网络版发表.

1) 国家自然科学基金 (1217020361) 资助项目.

2) 查文舒, 副研究员, 主要研究方向: 流动机理及数值模拟研究. E-mail: wszha@hfut.edu.cn

3) 李道伦, 教授, 主要研究方向: 流动机理及数值模拟研究. E-mail: ldaol@hfut.edu.cn

引用格式: 查文舒, 李道伦, 沈路航, 张雯, 刘旭亮. 基于神经网络的偏微分方程求解方法研究综述. 力学学报, 2022, 54(3): 543-556

Zha Wenshu, Li Daolun, Shen Luhang, Zhang Wen, Liu Xuliang. Review of neural network-based methods for solving partial differential equations. *Chinese Journal of Theoretical and Applied Mechanics*, 2022, 54(3): 543-556

functions. The first is data-driven method, which mainly learns PDEs from partially known data and can be applied to recovering physical equations, discovering unknown equations, parameter inversion, etc. The second is physical-constraint method, i.e., data-driven supplemented by physical constraints, which is manifested by adding physical laws such as governing equation to the loss function, thus reducing the network's reliance on labeled data and improving the generalization ability and application value. The third is physics-driven method (purely physical constraints), which solves PDEs by physical laws without any labeled data. However, such methods are currently only applied to solve simple PDEs and still need to be improved for complex physics. This paper introduces the research progress of intelligent solution of PDEs from these three aspects, involving various network structures such as fully-connected neural networks, convolutional neural networks, recurrent neural networks, etc. Finally, we summarize the research progress of PDE intelligent solutions, and outline the corresponding application scenarios and future research outlook.

**Key words** neural networks, PDE intelligent solution, data-driven, physical-constraint

## 引言

人工智能引发多领域技术变革,广泛应用于计算机视觉,生物医学,油气工程开发等领域.深度学习(deep learning)在工程技术,流体力学,计算力学等领域的研究具有重要的理论指导意义与工程应用价值.近年来,基于油藏动、静态数据,人工智能有望实现油藏精细描述与精准开发,提高采收率.将测井、压裂施工、生产数据等进行智能融合,大幅提升压裂改造效果,降低开发成本.大数据与智能优化方法相结合,将变革油田数据分析方法、油田开发控制与优化方法<sup>[1]</sup>.非常规油气开发难题与人工智能相结合,有望解决非常规复杂油气物理规律建立、偏微分方程求解等难题.人工智能与大数据将“实现石油勘探开发主体技术更新换代的宏伟目标,从技术层面上促进石油勘探开发行业整体转型升级”<sup>[2]</sup>.

人工智能方法因其处理高度复杂问题的突出能力,已引起油田领域的特别关注<sup>[3-6]</sup>.传统人工神经网络已在石油工程领域得到广泛应用,例如预测未知年份的测井数据<sup>[3]</sup>、预测油品压力-体积-温度属性<sup>[4]</sup>、预测注产剖面<sup>[5]</sup>、估算孔隙度<sup>[6]</sup>、井底流动压力<sup>[7]</sup>、选择页岩气藏完井方法<sup>[8]</sup>、试井解释<sup>[9-13]</sup>等.

深度学习是机器学习的一个新领域.深度学习的本质是构建含有多个隐藏层的网络模型,通过学习大规模的数据,获得更具代表性的特征,从而提高预测和分类的精度. Tian 和 Horne<sup>[14]</sup>利用递归神经网络学习永久井下压力计(PDG)数据,用于识别油藏模型及生产预测. Sudakov 等<sup>[15]</sup>将深度学习用于渗透率预测. Mosser 等<sup>[16]</sup>利用深度学习进行三维多孔介质重构. 张东晓等<sup>[17]</sup>利用循环神经网络研究测

井曲线的生成与修补.近两年,深度学习在试井参数自动反演得到了很好的应用<sup>[18-20]</sup>.

同时,在解决参数反演、数字岩心、测井曲线、试井解释等问题上,深度学习作为人工智能发展引擎有着优秀的表现<sup>[21-25]</sup>.以深度学习为核心的人工智能正在油气开发领域掀起新的研究热潮,其中最具前瞻性、颠覆性的研究当属基于深度学习的偏微分方程求解.该方法一旦突破,物理规律建立、参数反演和数值模拟方法都将发生变革,我国也将在以偏微分方程(偏微分方程)求解为核心的工业计算软件中迎来巨大机遇.自2017年来,深度学习在物理规律发现、油藏参数反演和偏微分方程求解中发挥了令人惊讶的作用<sup>[26-30]</sup>.

在实际开发现场或其他应用场景中,只能获取相关测量数据,而数据背后所潜在的物理规律需要进一步分析得到,偏微分方程模型是刻画其特性的重要工具.目前,解决基于大量数据寻找物理规律的主要思路有:在有一定的先验知识下,列出偏微分方程所描述的物理过程的备选项,利用稀疏回归技术或其他方法进行特征选择和参数估计<sup>[31-35]</sup>;其二用神经网络作为逼近器的功能,使用不同的神经网络结构表征偏微分方程系统,通过学习网络结构达到发现物理规律的目的<sup>[1,36]</sup>.

现有的偏微分方程求解方法要进行网格划分、非线性方程组求解,计算成本高,技术突破难度大.基于深度学习的偏微分方程求解方法不仅能快速正演、快速反演<sup>[37-38]</sup>,而且能很好解决非线性问题<sup>[39-42]</sup>,能对更复杂、更高维的偏微分方程<sup>[27,30,43]</sup>进行求解,有望颠覆传统偏微分方程数值求解技术,引发数值模拟技术的巨大变革.

虽然技术路线存在差异,但深度学习在物理模型建立、偏微分方程求解和参数反演中的应用,其核心仍是如何用深度学习表征偏微分方程.当前的研究多集中在无源汇的偏微分方程,多是数据驱动(data driven)、物理约束(physics informed),少量为物理驱动(physical based, physical constraint, theory based or without labeled data).由于当前处于研究初始阶段,很多学者都提出了自己的术语,导致当前术语复杂,然而不同术语往往表示同一含义.为此,本文约定:数据驱动(data driven)为仅用标签数据约束的方法,物理驱动(physics driven)为不含任何标签数据约束的方法.物理约束(physics informed)介于二者之间,即标签数据约束、偏微分方程约束共存的方法.因而,若仅用偏微分方程约束,物理约束方法就是物理驱动方法.另外,遵循于传统偏微分方程解析解、偏微分方程数值解术语,这里称用深度学习进行偏微分方程求解的方法为偏微分方程智能求解方法或偏微分方程神经网络求解方法.

本文将深度学习表征偏微分方程分为两个场景:构建未知偏微分方程与求解已知偏微分方程.对于构建未知偏微分方程,本文简要介绍了网络结构与偏微分方程、微分算子或演化算子等的内在联系,概述神经网络逼近未知偏微分方程的表示方法,并给出其中有待解决的问题与难点.对于求解已知偏微分方程,本文从数据驱动、物理约束和物理驱动3个角度介绍神经网络求解已知偏微分方程方法,主要包括神经网络求解偏微分方程原理,网络框架构建,损失函数构造等,结合国内外研究现状,系统梳理该领域的研究脉络,分析神经网络求解偏微分方程中存在的关键问题和解决方案,并对可行的未来研究方向和内容进行讨论和展望.此外,虽然深度学习在近年来得到了迅猛发展,但其在求解偏微分方程等力学问题上的研究仍然有限,在实际应用中的表现仍有待考验.因此,本文主要侧重研究方法上的进展.

## 1 基于神经网络的偏微分方程求解方法探索研究

1943年 McCulloch 和 Pitts<sup>[44]</sup>建立了神经网络及其数学模型,开创了人工神经网络研究的新时代.20世纪80年代中期,首次提出的反向传播算法及其发展<sup>[45]</sup>引起了人工神经网络领域研究的第二次

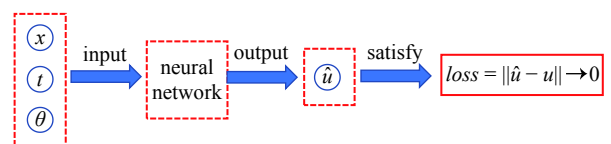
热潮.

一直以来,人们希望找到无须网格划分、无须非线性方程求解的偏微分方程数值求解新方法.其探索之一就是基于人工神经网络的求解方法.自动微分(automatic differentiation)能使用链式法则精确计算导数<sup>[46-48]</sup>,可以根据神经网络的输入坐标和网络参数对整个神经网络模型进行微分,从而代替偏微分方程中复杂的梯度计算,为基于人工神经网络的偏微分方程求解奠定了基础.

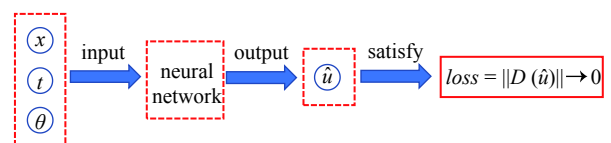
在20世纪90年代,便有学者开始研究使用神经网络求解微分方程的数学基础与方法.1990年 Wornik 等<sup>[49]</sup>证明了多层神经网络能够逼近任意函数及其导数.这为微分方程的神经网络求解奠定了理论基础.随后, Li<sup>[50]</sup>证明了一个隐藏层的神经网络可逼近多元多项式函数及其导数. Lagaris 等<sup>[51]</sup>将微分方程中的初值与边界条件独立表征,提出颇为新颖的偏微分方程求解方法.随后,不少学者进行了探索研究,例如, Aarts 和 Van<sup>[52]</sup>将表征不同阶微分算子的单隐层前馈网络联合起来,共同训练来求解偏微分方程;又如, Ramuhalli 等<sup>[53]</sup>将有限元模型嵌入到神经网络中,提出了有限元神经网络.由于早期多层前馈神经网络模型的局限性,早期方法只能求解简单的偏微分方程,基于神经网络的偏微分方程求解方法没引起足够的重视.

早期方法主要基于数据驱动,即事先获得偏微分方程的输入及精确解(常称为“标签数据”),然后用神经网络逼近标签数据,从而获得能够表征偏微分方程的神经网络模型,如图1(a)所示.网络的输入可以是参数或空间、时间等,可根据需要选择.

随着深度学习算法在多个领域的成功应用<sup>[54-55]</sup>,



(a) 数据驱动的 PDE 求解方法,精确解是标签数据  
(a) Data-driven PDE solution method, the exact solution is label data



(b) 物理驱动的 PDE 求解方法,无须任何标签数据  
(b) Physics-driven PDE solution without any label data

图 1 基于 PDE 智能求解方法的两种技术路线

Fig. 1 Two technical routes based on PDE intelligent solution method

国内外学者重新开启了基于神经网络的偏微分方程求解方法研究,取得了系列突破,提出了新方法,如纯物理驱动的偏微分方程求解方法.该方法用控制方程进行约束,无需标签数据,如图 1(b)所示.

根据不同的应用场景,本文将从深度学习反演构建未知偏微分方程和求解已知偏微分方程两个方面展开介绍.下节主要介绍如何通过神经网络逼近线性或非线性算子,从数据中找出隐藏的偏微分方程模型.

## 2 基于深度学习反演未知偏微分方程

利用深度学习从数据中反演未知的偏微分方程是当前的研究热点之一.对于未知的偏微分方程,主要的研究目标是通过深度学习找出数据背后蕴藏的偏微分方程模型,从数据中反演未知的偏微分方程(例如方程的右端项、方程的积分形式或方程的演化算子等),进一步构建模型用于求解.

恢复方程的传统思路是构建简单函数和偏导数的备选字典.这些函数和偏导数很可能出现在未知的控制方程中.根据已知偏微分方程的非线性响应的形式构建模型,然后利用稀疏回归方法来学习这些未知参数,选择最准确代表数据的项.这种传统的恢复方式要求假设非线性响应形式已知或确定微分算子的有限差分逼近的方法,而深度学习大大降低了对偏微分方程先验知识的要求,只需要简单的先验知识,如方程最大可能的阶.此外稀疏回归方法需要事先确定字典中空间差分的数值近似,限制了字典的表达能力和预测能力且需要建立一个足够大的字典,这可能会导致高的内存负载和计算成本,特别是当模型变量的数量很大的时候.深度学习采用可学习的卷积近似微分算子或近似演化算子,从根本上提高从噪声数据中识别动力学的能力,从而使模型具有更强的表达能力和预测准确性.如果没有足够的先验知识,也有可能通过调整多项式的微分来获得更好的表征效果,神经网络在偏微分方程求解和恢复问题中都大有可为.

近几年,国内外学者致力于探究网络结构与偏微分方程、各项微分算子或方程的演化算子等的内在联系,从理论上支撑用深度学习来表征偏微分方程.2018年 Long 等<sup>[56-57]</sup>提出一种基于数据驱动的前馈神经网络(PDE-Net),其核心思想是:时间导数项做欧拉离散,受约束卷积核近似微分算子,进而使

用神经网络或其他机器学习方法近似方程右端项,构建网络来逼近偏微分方程系统,并对其解进行长期预测.此外结合 Symnet (symbolic neural network)<sup>[58-59]</sup>,使用可学习滤波器 PDE-Net2.0<sup>[56-57]</sup>更加灵活,能够在少量先验知识的情况下揭示方程的解析形式,尤其对于非线性问题有更好的结果,并能够更强大地逼近未知动态和更长的时间预测.

变分和偏微分方程框架下卷积与微分之间存在内在关联<sup>[60-61]</sup>.据此关联,Long 等<sup>[56]</sup>提出受约束的卷积核,即在数学上能证明该卷积核能表征微分算子,因而卷积核是受限的.深度学习在受限的基础上学习卷积核,从而有更好的偏微分算子表征能力.例如,对于卷积核  $\mathbf{a}_{0,1} = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ ,将该卷积核作用在二维空间变量  $\mathbf{u}$  上,得到

$$\begin{aligned} \frac{2}{h}(\mathbf{a}_{0,1}[-\cdot] \otimes \mathbf{u})[i, j] &= \frac{1}{2h}(u(x_i, y_j) - u(x_i, y_j - h)) + \\ &\frac{1}{2h}(u(x_i - h, y_j) - u(x_i - h, y_j - h)) \\ &\rightarrow u_y(x_i, y_j) \text{ as } h \rightarrow 0. \end{aligned}$$

类似于小波理论中的消失矩的阶,对于卷积核  $\mathbf{q}$ ,定义  $\mathbf{q}$  的消失矩的阶  $\alpha = (\alpha_1, \alpha_2) \in \mathbb{Z}_+^2$  满足

$$\sum_{\mathbf{k} \in \mathbb{Z}^2} \mathbf{k}^\beta \mathbf{q}[\mathbf{k}] = 0 \quad (1)$$

其中  $\beta = (\beta_1, \beta_2) \in \mathbb{Z}_+^2$ , 对于满足  $|\beta| < |\alpha|$  或  $|\beta| = |\alpha|$ , 但  $\beta \neq \alpha$  的所有  $\beta$  都有式 (1) 成立. 满足以上条件的  $\alpha = (\alpha_1, \alpha_2)$  称为  $\mathbf{q}$  的消失矩的阶. 卷积核  $\mathbf{q}$  有  $\alpha \in \mathbb{Z}_+^2$  阶消失矩, 其作用于一个连续函数  $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^2$  有, 应用泰勒展开进行推导得

$$\frac{1}{\varepsilon^{|\alpha|}} \sum_{\mathbf{k} \in \mathbb{Z}^2} \mathbf{q}[\mathbf{k}] \mathbf{F}(\mathbf{x} + \mathbf{k}\varepsilon) = C_\alpha \frac{\partial^\alpha \mathbf{F}(\mathbf{x})}{\partial \mathbf{x}^\alpha} + o(\varepsilon), \quad \varepsilon \rightarrow 0 \quad (2)$$

其中  $C_\alpha = \frac{1}{\alpha!} \sum_{\mathbf{k} \in \mathbb{Z}^2} \mathbf{k}^\alpha \mathbf{q}[\mathbf{k}]$ , 此时卷积核  $\mathbf{q}$  相当于逼近  $\frac{\partial^\alpha}{\partial \mathbf{x}^\alpha}$ , 从而卷积网络可逼近偏微分方程.

最终可通过式 (2) 构造卷积核来近似微分算子, 详见文献<sup>[56]</sup>. 基于受约束的卷积核, 构造卷积神经网络来逼近偏微分方程, 并使用神经网络或其他机器学习方法确定非线性响应项. 当偏微分方程对应以下格式时

$$u_t(t, x, y) = F(x, y, u_x, u_y, u_{xx}, u_{yy}, u_{xy}, \dots) \quad (3)$$

对时间导数项做欧拉离散, 空间类导数项做约束卷积近似, 方程 (3) 可表示为

$$\hat{u}(t_{i+1}, \cdot) = D_0 u(t_i, \cdot) + \delta t \times (c_{00} D_{00} u + c_{10} D_{10} u + c_{01} D_{01} u + c_{01} D_{01} u, \dots) \quad (4)$$

其中  $D_0$  和  $D_{ij}$  表示卷积核,  $c_{ij}$  是二维矩阵, 相当于偏导项的系数, 可以插值得到, 也可以直接在训练过程中学习得到, 再根据矩阵是否为零来判断对应的项是否存在, 从而反演出未知偏微分方程。

González-García 等<sup>[62]</sup> 基于人工神经网络体系结构提出物理模型建立方法, 其本质是在一定的先验知识下, 列出描述物理过程的偏微分方程所有的备选项, 利用人工神经网络进行自动选择和参数估计, 从而发现数据后隐藏的物理规律。

Wu 等<sup>[63-64]</sup> 首次基于残差网络 (ResNet) 构建了从数据中学习未知微分方程的新框架。该框架以微分方程内在的积分形式为基础, 以逼近方程的流谱 (flow map, 针对常微分方程) 和演化算子 (evolution operator, 针对偏微分方程) 为目标, 从根本上避免了传统框架 (以逼近方程的右端项为目标) 所依赖的数值微分。文献 [63-64] 提出了两种多步的 ResNet 神经网络结构, 从精确演化算子的角度, 首次在理论上建立了 ResNet 与精确演化算子的内在数学关系, 由此给出了该深度学习方法的数学解释。不同于 Wu 和 Xiu<sup>[64]</sup> 在模态/傅里叶空间学习方程, Chen 等<sup>[65]</sup> 在物理空间进行学习和建模, 利用 DNN 学习测量数据, 从而学习未知偏微分方程。Chen 等<sup>[66]</sup> 提出一种无梯度的符号遗传算法 (SGA-PDE), 使用符号数学灵活表示任意给定偏微分方程, 优化其表示形式, 从数据中发现开放形式的偏微分方程。Xu 和 Zhang<sup>[67]</sup> 在 PINN 的基础上提出一种更具鲁棒性的深度学习遗传算法 (R-DLGA), 将深度学习-遗传算法提供的潜在项的初步结果作为物理约束加入损失函数, 提升了在高阶导数等影响下导数的计算精度, 从而在高噪声稀疏数据中获得偏微分方程。

迄今为止, 该领域提出的许多方法都存在一些局限性。特别是, 目前的方法通常研究  $u_t = N(u, x, t)$  形式的方程, 但许多物理方程不在此类。此外, 如果测量一个具有参数依赖性的系统, 如何消除演化动力学及其参数依赖性之间的歧义是有待解决的问题。尽管神经网络表现出强大的数据学习能力, 但对于有噪声数据的学习, 尤其在非线性、多耦合的复

杂物理系统中, 网络模型的精确性以及稳定性有待提升。

而对于已知的给定的偏微分方程, 神经网络可用于逼近偏微分方程的解或表征方程, 本文下节将从数据驱动、物理约束和物理驱动 3 个方面对神经网络求解偏微分方程展开介绍, 并简述所用的神经网络, 如全连接神经网络, 卷积神经网络 (CNN)、残差网络 (ResNet)、DenseNet、自编码网络 (autoencoder)、长短期记忆 (LSTM) 网络等, 总结现有研究的重要进展, 并探讨下一步的发展趋势, 对未来偏微分方程智能求解的研究提出建议。

### 3 基于深度学习的偏微分方程求解方法

#### 3.1 偏微分方程神经网络求解方法概述

神经网络的基本结构是前馈全连接深度神经网络<sup>[68]</sup>, 以此为例介绍已知偏微分方程的神经网络求解方法。以  $d$  维行向量  $x \in R^d$  为网络输入, 一个单隐层神经网络的  $k$  维输出形式为

$$y = \sigma(xW_1 + b_1)W_2 + b_2 \quad (5)$$

式中,  $W_1$  和  $W_2$  分别为  $d \times q$  和  $q \times k$  的权重矩阵,  $b_1$  和  $b_2$  分别为  $1 \times q$  和  $1 \times k$  的偏置向量;  $\sigma(\cdot)$  是一个非线性模型, 称为激活函数。

对于多层神经网络, 模型参数可以表示为

$$(W^*, b^*) = \arg \min_{(W, b)} J(\theta; X, Y) \quad (6)$$

式中,  $\theta$  表示网络参数  $\{W, b\}$  的集合。参数的优化采用随机梯度下降 (SGD) 或其变体方法<sup>[69-71]</sup>。以 SGD 为例, 第  $i$  次迭代过程如下

$$\theta^{(i+1)} = \theta^{(i)} - \eta \nabla_{\theta} J(\theta^{(i)}; X, Y) \quad (7)$$

式中,  $\eta$  为第  $i$  次迭代的步长。损失函数相对于模型参数的梯度  $\nabla_{\theta} J$  通常使用反向传播<sup>[72]</sup> 计算, 这是反向模式自动微分<sup>[48]</sup> 技术的特殊情况。关于神经网络的优化过程不再详细描述, 具体可参考文献 [48, 68, 72]。

对于给定的一般偏微分方程, 在初始条件  $I(\cdot)$  (IC) 以及边界条件  $\mathcal{B}(\cdot)$  (BC) 的约束下可表述为<sup>[73]</sup>

$$\left. \begin{aligned} N(t, x; u(t, x; \theta)) &= 0, \quad t \in [0, T], x \in \mathcal{D} \\ I(x; u(0, x; \theta)) &= 0, \quad x \in \mathcal{D} \\ \mathcal{B}(t, x; u(t, x; \theta)) &= 0, \quad t \in [0, T], x \in \partial \mathcal{D} \end{aligned} \right\} \quad (8)$$

式中,  $u(t, x; \theta)$  是方程的近似解,  $\theta$  是近似解在方程中对应的参数;  $\mathcal{N}(\cdot)$  是一个微分算子包含时间微分, 空间微分等组成的线性或非线项;  $x$  为定义在有界连续空间域  $\mathcal{D} \subseteq \mathbb{R}^D$  中的位置向量,  $\partial\mathcal{D}$  为边界.

一般来说, 根据训练方法的不同, 基于深度学习的偏微分方程求解方法可以分为数据驱动和物理驱动两种方法. 数据驱动方法所需的标签数据形如  $u(t, x)$ , 通过寻找一组最优的网络参数  $(W, b)$ , 以局部最小化训练数据  $u(t, x)$  和神经网络预测  $\hat{u}^*(t, x; W, b)$  之间的差. 也就是说, 可以将数据驱动的优化问题表示为

$$\left. \begin{aligned} loss_{data}(W, b) &= \|u(t, x) - \hat{u}^*(t, x; W, b)\| \\ W^*, b^* &= \arg \min_{W, b} loss_{data}(W, b) \end{aligned} \right\} \quad (9)$$

式中,  $W^*$  和  $b^*$  为神经网络的优化目标.

对于物理驱动方法, 通常使用控制方程和 IC, BC 构造残差, 然后将残差添加到损失函数中, 从而优化网络参数. 其中, 网络的输出  $\hat{u}^*(t, x; W, b)$  被代入控制方程构造残差, 再通过最小化残差来优化参数. 基于物理约束的优化问题如下所示

$$\left. \begin{aligned} loss_{psy}(W, b) &= \|N(t, x, \hat{u}^*)\| + \|B(t, x, \hat{u}^*)\| + \\ &\quad \|\mathcal{I}(0, x, \hat{u}^*)\| \\ W^*, b^* &= \arg \min_{W, b} loss_{psy}(W, b) \end{aligned} \right\} \quad (10)$$

### 3.2 数据驱动下的偏微分方程神经网络求解方法

偏微分方程已知情况下, 基于数据驱动求解偏微分方程, 其核心问题是探究方程及其中各项微分算子的表征方法, 进而得到方程的解.

#### 3.2.1 基于 CNN 的求解方法

由于受约束的卷积核数学上具有偏微分算子特征, 用其来表征与求解偏微分方程会具有很好的效果. 为此, 基于受限卷积核来表征微分算子的思想<sup>[57]</sup>, Zha 等<sup>[73]</sup> 将二维受限卷积核推广到三维受限卷积核, 构建新的三维偏微分方程智能求解方法.

对卷积核  $q$ , 其矩阵  $M(q) = (m_{i,j,t})_{N \times N \times N}$ ,

$$m_{i,j,t} = \frac{1}{i!j!t!} \sum_{k_1, k_2, k_3 = -\frac{N-1}{2}}^{\frac{N-1}{2}} k_1^i k_2^j k_3^t q[k_1, k_2, k_3], \quad i, j, t = 0, 1, \dots, N-1.$$

当三维微分算子最大可能的阶为 1 时, 在卷积核表征微分算子过程中, 约束  $q$  最大阶  $i + j + t = 1$  时,  $m_{i,j,t} \neq 0$ . 例如, 对  $2^3$  受限卷积核近似微分算子有

$$q_{100} = \frac{1}{8} \begin{pmatrix} A_2 \\ A_2 \end{pmatrix}, \quad q_{010} = \frac{1}{8} \begin{pmatrix} A_3 \\ A_3 \end{pmatrix}, \quad q_{001} = \frac{1}{8} \begin{pmatrix} A_1 \\ A_4 \end{pmatrix},$$

其中  $A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ ,  $A_2 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ ,  $A_3 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ ,  $A_4 = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$ , 其约束条件  $i + j + t = 1$  时,  $m_{i,j,t} \neq 0$ .  $q$  作用于三维空间变量  $u$ , 有

$$\begin{aligned} \frac{2}{\varepsilon} (q_{100}[-\cdot] \otimes u)[x, y, z] &= \frac{1}{4\varepsilon} (u(x, y, z) - u(x - \varepsilon, y, z)) + \\ &\quad \frac{1}{4\varepsilon} (u(x, y - \varepsilon, z) - u(x - \varepsilon, y - \varepsilon, z)) + \\ &\quad \frac{1}{4\varepsilon} (u(x, y, z - \varepsilon) - u(x - \varepsilon, y, z - \varepsilon)) + \\ &\quad \frac{1}{4\varepsilon} (u(x, y - \varepsilon, z - \varepsilon) - u(x - \varepsilon, y - \varepsilon, z - \varepsilon)) \approx \\ &u_x(x, y, z), \quad \text{as } \varepsilon \rightarrow 0 \end{aligned}$$

类似地,  $\frac{2}{\varepsilon} (q_{010}[-\cdot] \otimes u) \approx u_y(x, y, z)$ ,  $\frac{2}{\varepsilon} (q_{001}[-\cdot] \otimes u) \approx u_z(x, y, z)$ .

显然, 可以用约束下的卷积核近似对应阶数的微分算子, 对偏微分方程  $\mathcal{N}(t, x, y, z; u) := u_t - F(x, y, z, u_x, u_y, u_z, u_{xx}, \dots)$  进行卷积近似时, 有

$$\hat{u}(t_{i+1}, \cdot) = D_0 u(t_i, \cdot) + \delta t \times F(x, y, z, D_{000} u, D_{100} u, \dots)$$

在此基础上, 引入分层自适应激活函数构建 3D-PDE-Net, 如图 2 所示, 其中激活函数可以表示为  $\sigma(na^i D_{ijl} u)$ ,  $n \geq 1$  为预定义的比例因子, 参数  $a^i$  是激活函数的斜率, 可学习参数,  $\sigma$  采用 Tanh 激活函数.

数值实验表明, 3D-PDE-Net 求解精度  $L_\infty$  误差比求解比数值格式 Douglas-Gunn ADI 降低 20 倍<sup>[73]</sup>; 所加入分层的自适应激活函数可 10 倍提高训练速度, 且局部误差得到改善. 但此时 3D-PDE-Net 不是显式可解释的.

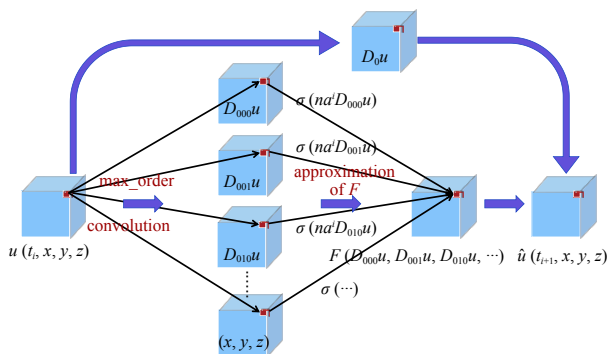


图 2 3D-PDE-Net 网络结构示意图<sup>[73]</sup>  
Fig. 2 The schematic diagram of a  $\delta t$ -block<sup>[73]</sup>

### 3.2.2 基于其他网络的偏微分方程求解方法

Liu 等<sup>[26]</sup>探讨了全连接神经网络在函数逼近中的应用, 并提出了一个通用的基础微分方程求解器, 主要利用自动微分对方程的初值问题和边值问题进行求解. E 等<sup>[30]</sup>与 Han 等<sup>[27]</sup>用深度学习逼近梯度算子, 基于偏微分方程的离散格式, 对高维偏微分方程给出深度学习求解新方法. 对有  $H$  个隐藏层、 $N$  个时间间隔的半线性抛物型偏微分方程的网络结构如图 3 所示, 整个网络共有  $(H+1)(N-1)$  层, 通过损失函数共同优化所有的网络参数.  $t = t_1, t_2, \dots, t_{N-1}$  中的每一列对应一个  $t$  时间步的子网络,  $h_n^1, h_n^2, \dots, h_n^H$  是每一个子网络中的中间神经元. 基于标签数据, 用多层前馈神经网络逼近梯度算子, 从而可得到高于 100 维的偏微分方程解, 并给出了多种类型高维偏微分方程的求解结果.

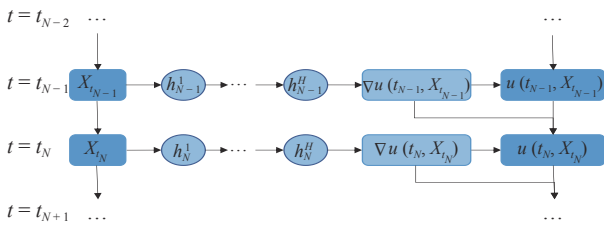


图 3 有  $H$  个隐藏层、 $N$  个时间间隔的半线性抛物型偏微分方程的网络结构 (修改自文献 [30])

Fig. 3 Illustration of the network architecture for solving semilinear parabolic PDEs with  $H$  hidden layers for each sub-network and  $N$  time intervals (modified from Ref. [30])

### 3.3 物理约束下的偏微分方程神经网络求解方法

由于数据驱动存在泛化能力弱等缺点, 物理驱动可提高泛化能力, 减少标签数据. 物理驱动与数据驱动相融合, 即物理约束的方法, 受到广泛的关注.

在近几年的研究中<sup>[74-76]</sup>, 已经看到利用结构化先验信息构建基于数据和物理信息的机器学习算法的研究前景. Sirignano 等<sup>[77]</sup>给出了类似于 LSTM 人工神经网络的 deep Galerkin method (DGM) 网络, 提出了基于 Galerkin 方法的二阶微分算子计算方法, 同时给出了物理约束下的神经网络逼近定理.

#### 3.3.1 PINN

Raissi 等<sup>[39-40]</sup>利用偏微分方程的控制方程以及边界条件等恒等式构造残差, 利用各项残差之和构造损失函数, 并将此方法拓展到解决非线性问题, 提出了物理约束下的神经网络 (physics informed neural network, PINN). PINN 将数据驱动与物理约束相结

合, 从而提出了偏微分方程建立与求解的新思路, 即, 对偏微分方程

$$\left. \begin{aligned} D(u(x)) &= 0, \quad x \in \Omega \\ B(u(x)) &= f(x), \quad x \in \partial\Omega \end{aligned} \right\} \quad (11)$$

PINN 中的损失函数主要由 3 部分组成, 分别为偏微分方程的控制方程, 网络输出与初始条件、边界条件的标签数据的残差.

同时, Raissi 等<sup>[41]</sup>还研究了连续时间模型与离散时间模型在方程求解和方程恢复两种场景下的应用, 以及标签数据噪声对求解精度的影响及误差传播.

以求解 Dirichlet 边界下的 Burgers 方程为例, 其方程为

$$\left. \begin{aligned} u_t + uu_x - (0.01/\pi)u_{xx} &= 0, \quad x \in [-1, 1], t \in [0, 1] \\ u(0, x) &= -\sin(\pi x) \\ u(t, -1) = u(t, 1) &= 0 \end{aligned} \right\} \quad (12)$$

则有函数  $D(u)$  为

$$D(u) := u_t + uu_x - (0.01/\pi)u_{xx} \quad (13)$$

可定义损失函数为

$$L = \frac{1}{N_D} \sum_{n=1}^{N_D} \|D(\hat{u}(x_n^D; \mathbf{W}, \mathbf{b}))\|^2 + \frac{1}{N_u} \sum_{n=1}^{N_u} \|\hat{u}(x_n) - u\|^2$$

通过最小化损失函数优化神经网络参数, 使得网络输出逼近 Burgers 方程的解.

#### 3.3.2 基于 PINN 的改进方法

基于 PINN 算法, Toshiyuki 等<sup>[78]</sup>使用由 3 个 DNNs 组成的 PINN 框架对 Richardson-Richards 方程进行参数反演, 并估算保水曲线和水力传导函数. Han 等<sup>[79]</sup>介绍了一种基于深度学习的一般高维抛物型偏微分方程的求解方法. 先对偏微分方程进行重新构造, 再利用神经网络逼近未知解的梯度, 在非线形方程计算中得到了满意的数值结果. Meng 等<sup>[80]</sup>提出了一种改进的 PINN 方法, 称为 PPINN, 将一个长时间的问题分解为多个独立的短时间问题, 以加速偏微分方程的求解. Michoski 等<sup>[81]</sup>研究了激波偏微分方程神经网络求解方法, 神经网络方法与传统方法结果对比表明, 基于神经网络的求解方法有优势, 标签数据可有效提升求解精度. Kani 和 Elsheikh<sup>[82]</sup>将物理约束求解偏微分方程方法与正交分解 (POD) 和离散经验插值方法 (DEIM) 相结合, 提供了一个高精度的非线性动力系统降阶模型, 降低了高保真数



值模拟的计算复杂度.

Jagtap 等<sup>[83-84]</sup> 提出自适应激活函数, 有效地提高了 PINN 逼近非线性函数和偏微分方程的效率、鲁棒性和准确性, 自适应激活函数如下, 图 4 为各自适应激活函数的图像

$$\left. \begin{aligned} \text{sigmoid} &: \frac{1}{1 + e^{-ax}} \\ \text{tanh} &: \frac{e^{ax} - e^{-ax}}{e^{ax} + e^{-ax}} \\ \text{ReLU} &: \max(0, ax) \\ \text{leaky ReLU} &: \max(0, ax) - \nu \max(0, -ax) \end{aligned} \right\} \quad (14)$$

然而, 含标签数据的偏微分方程神经网络求解

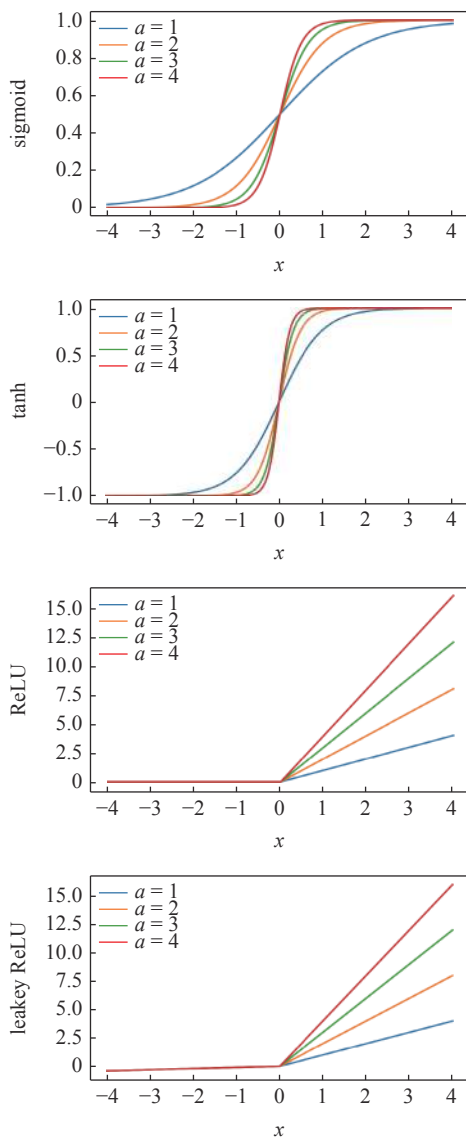


图 4 Sigmoid, tanh, ReLU 和 leaky-ReLU 的对应变量  $a$  的激活函数<sup>[83]</sup>

Fig. 4 Sigmoid, tanh, ReLU and leaky-ReLU activation functions for various values of  $a$ <sup>[83]</sup>

方法, 存在很大的局限性. 对很多问题, 其精确解是未知的. 若需要偏微分方程的精确解才能构造损失函数, 这大大限制了其应用范围. 例如, 在油田开发过程中, 仪器只能测量井底的压力、井口的产量, 而不能获得其他地区的压力. 这意味着基于标签数据的偏微分方程求解方法无效. 从而, 基于纯物理约束 (即物理驱动) 的求解方法具有更广阔的应用前景, 有着与传统求解方法一样的便利性 (无须任何标签数据). 这一旦突破, 将引发偏微分方程求解技术的真正变革.

### 3.3.3 可测量标签数据下的偏微分方程神经网络求解方法

上述数据驱动下的偏微分方程求解方法往往需要未知量的分布数据, 例如需要知道压力空间分布数据. 这往往在实验条件下才能获得. 例如, 可在实验中布置多个压力传感器, 才能获得压力的空间时变数据. 但对实际工程问题, 这部分数据是不可测量的. 例如, 在油藏开发中, 只能测量井中的压力, 其他的压力数据无法获知. 因而, 上述数据驱动下的偏微分方程智能求解方法难以有真正应用.

实际工程中, 存在一部分可测量的数据, 若能使用少量的可测量数据作为标签, 就能对偏微分方程进行求解, 将具有重要的理论意义与应用价值. 部分学者对此进行了积极的探索研究.

Wang 等<sup>[85-86]</sup> 将实际工程中的专家经验、物理规律和稀疏观测数据等整合为理论指导神经网络 (theory-guided neural network, TgNN), 如图 5 所示, 利用 TgNN 解决地下流动建模, 不确定性量化等问题.

Li 等<sup>[87]</sup> 使用深度神经网络解决单相渗流问题, 加入部分可测量的井底流压数据作为标签, 有效提高非稳态、具有源汇的偏微分方程问题的求解精度. 该方法的重大特点是, 除将可观测的井底压力数

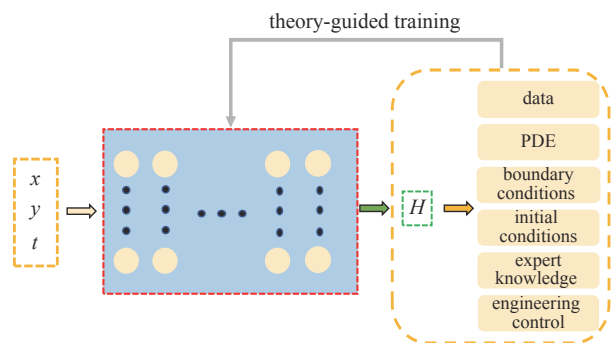


图 5 TgNN 模型的网络结构<sup>[85]</sup>

Fig. 5 Structure of the TgNN model<sup>[85]</sup>

据作为标签外, 不再需要任何其他的标签数据, 而是用偏微分方程约束代替压力分布的标签数据, 从而大大提高该方法实际应用可行性. 此外, 利用源汇项引起的梯度特征构造梯度模型, 作为“路标”加入神经网络, 通过添加固定神经元的方式帮助网络提高优化能力, 同时提出了预训练获得“路标”的解决思路. 图 6 为智能求解所获得的压力分布、井底压力 (BHP) 图.

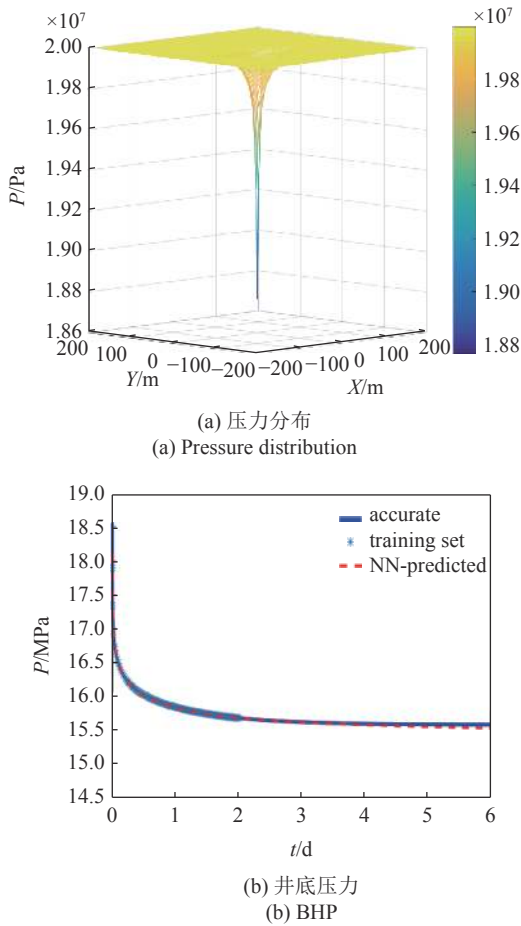


图 6 智能求解得到的压力分布和井底压力图<sup>[87]</sup>  
Fig. 6 Pressure distribution and BHP obtained by intelligent solution<sup>[87]</sup>

Chen 等<sup>[88]</sup> 提出一种基于协方差矩阵优化的无梯度神经网络, 有效提升学习小数据样本的鲁棒性, 适合实际工程应用. 在后续研究中, Chen 等<sup>[89]</sup> 提出一种硬约束投影 (hard constraint projection, HCP) 的方法提升机器学习方法对小样本数据的学习能力.

### 3.4 物理驱动下的偏微分方程神经网络求解方法

相较于传统数值求解, 偏微分方程智能求解仍受标签数据的约束, 在实际应用中, 往往会面临数据获取困难的情况. 对此, 无需标签数据的物理驱动方法成为重要的研究方向, 是最终的解决方案.

#### 3.4.1 全连接神经网络 (FC-NN) 求解偏微分方程

Nabian 等<sup>[90]</sup> 使用无监督的前馈深度残差神经网络近似高维偏微分方程, 利用随机梯度下降 (SGD) 算法优化神经网络参数. Cai 等<sup>[91]</sup> 研究了基于无监督深度学习的一维二阶椭圆偏微分方程数值求解方法, 并利用一阶系统最小二乘 (FOSLS) 作为损失函数来优化神经网络的参数. Sun 等<sup>[92]</sup> 提出了基于“硬边界约束”的神经网络求解方法, 将偏微分方程的控制方程和边界条件作为损失函数来约束神经网络, 通过构造“硬编码”的结构化深度神经网络来加强初始条件和 Dirichlet 边界条件, 有效增强物理约束下的偏微分方程智能求解, 如图 7 所示. 对于下述不可压缩 Navier-Stokes 方程<sup>[92]</sup>

$$\left. \begin{aligned} \nabla \cdot \mathbf{u} &= 0, \quad x, t \in \Omega_{f,t}, \theta \in R^d \\ \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \frac{1}{\rho} \nabla p - \nu \nabla^2 \mathbf{u} + b_f &= 0, \quad x, t \in \Omega_{f,t}, \theta \in R^d \end{aligned} \right\} \quad (15)$$

其中  $t$  和  $x$  分别代表时间和空间,  $\Omega_{f,t} \triangleq \Omega_f \times [0, T]$ ;  $\theta$  是  $d$  维向量, 包括流体性质、域的几何形状等参数; 速度  $u(t, x, \theta)$  和压力  $p(t, x, \theta)$  都是时间、空间和可变参数的函数;  $\rho$  和  $\nu$  分别表示流体的密度和黏度;  $b_f$  是身体力参数;  $\Omega_f \subset R^3$  为流体域.

对初始条件和 Dirichlet 边界条件的“硬编码”过程如下所示

$$\left. \begin{aligned} \hat{u}_p(t, x, \theta; \mathbf{W}, \mathbf{b}) &= u_{\text{particular}}(t, x, \theta) + \\ &\quad D(t, x, \theta) u^p(t, x, \theta; \mathbf{W}, \mathbf{b}) \\ \hat{p}_p(t, x, \theta; \mathbf{W}, \mathbf{b}) &= p_{\text{particular}}(t, x, \theta) + \\ &\quad D(t, x, \theta) p^p(t, x, \theta; \mathbf{W}, \mathbf{b}) \end{aligned} \right\} \quad (16)$$

其中,  $u_{\text{particular}}, p_{\text{particular}}$  是满足初始和边界条件的特解,  $D(t, x, \theta)$  是定义从内部点到“边界”的平滑函数. 即,  $D$  在边界  $\partial\Omega_f \times [0, T]$  和  $\Omega_f \times 0$  处为零且在远离边界处增大. 对于 IC/BC 和简单几何域  $\Omega_{f,t}$  的问题, 可以用解析的方法写出函数  $D$  和特解.

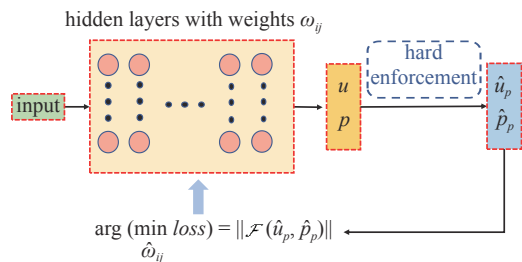


图 7 基于“硬边界约束”的 FC-NN 框架 (修改自文献<sup>[92]</sup>)  
Fig. 7 Schematic diagram of FC-NN framework based on “hard boundary constraint” (modified from Ref. <sup>[92]</sup>)

然而, 对于具有 Neumann 边界条件的偏微分方程问题, 仍将 Neumann 边界条件以惩罚项的形式加入损失函数, 在求解精度上仍有所欠缺. 综上所述, Sun 等<sup>[92]</sup> 为一部分具有 Dirichlet 边界的稳态问题的物理约束求解提供了很好的解决方法, 但由于需要解析构造边界平滑函数, 对非稳态, 具有 Neumann 边界的问题存在一定的局限性.

### 3.4.2 CNN 求解偏微分方程

在人脸识别、AlphaGo 等大显身手的 CNN 在偏微分方程求解中也受到广泛关注. 偏微分方程的属性空间 (如渗流方程中的渗透率) 与解空间的对应关系, 非常适合用卷积算子进行表征. 基于这一特性, 卷积神经网络求解偏微分方程会有很大的优势. 然而, 多数偏微分方程求解网络都基于 FC-NN 的“点态”(基于时空域中离散分布点) 方式进行训练, 这意味着 FC-NN 训练样本分布自由度很大. 而 CNN 则需要输入相对完整的样本分布, 是以“点阵”图像的形式进行训练. 这就带来了诸多新挑战, 如不规则域、卷积网络损失函数构造等. 对于不规则域, Gao 等<sup>[93]</sup> 通过对 CNN 网络输入的物理量约束, 得到偏微分方程求解神经网络模型, 再通过保形变换, 实现对不规则区域下的偏微分方程求解.

对于卷积网络的损失函数, Zhu 和 Zabaras<sup>[94]</sup> 提出了贝叶斯损失函数约束下的新偏微分方程求解网络. 该网络模型将 CNN 与编解码器网络相融合, 吸取了 DenseNet 特征重用的优点, 使得在参数和计算成本更少的情形下实现比 ResNet 更优的性能<sup>[95]</sup>. 如图 8 所示.

随后, Zhu 等<sup>[96]</sup> 研究了数据驱动、物理驱动等损失约束下的偏微分方程求解方法, 并提出物理约束的稠密卷积编解码器网络 (如图 9 所示), 提高求解精度以及泛化能力, 并使用 Sobel 算子计算 CNN 框架下的导数. 与 FC-NN 不同, CNN 以卷积的形式提取图像特征, 其本质就是对图像的像素求导数, Sobel 算子是以滤波算子的形式计算一阶导数, 从而可利用卷积函数快速计算.

然而, 目前研究的物理驱动方法 (无标签数据) 尚不能有效解决非稳态与源汇的问题. 物理驱动的研究仍在起步阶段, 离实际需求还很远.

总体上, 基于深度学习的偏微分方程建立、求解与参数反演研究还处于初始阶段, 主要存在以下问题: (1) 很多方法只能应用于简单的偏微分方程,

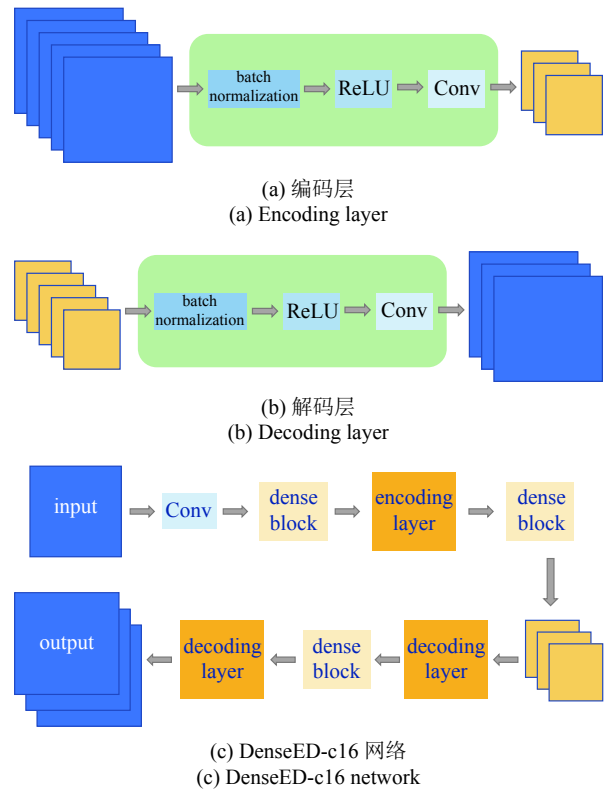


图 8 贝叶斯损失函数约束下的 DenseED-c16 网络 (修改自文献 [94])  
Fig. 8 DenseED-c16 network with Bayesian loss function constraints (modified from Ref. [94])

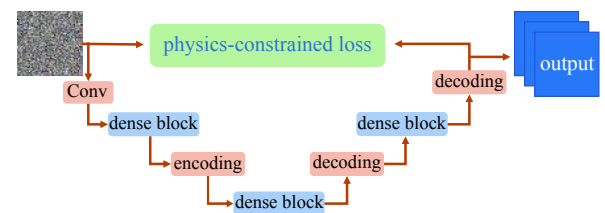


图 9 物理约束下的稠密卷积编解码器网络 (修改自文献 [96])  
Fig. 9 Dense convolutional encoder-decoder network as the deterministic surrogate (modified from Ref. [96])

例如, 偏微分方程中没有时间导数项, 或没有源汇项. 这说明求解方法还有待深入研究. (2) 多数方法都是基于数据驱动的: 数据驱动下, 神经网络受到了强有力的约束, 从而更容易收敛, 但其场景有限. (3) 物理驱动方法亟待突破: 一旦物理驱动下的偏微分方程求解方法获得突破, 偏微分方程求解方法将获得颠覆性的变革, 基于深度学习的偏微分方程重建、参数反演方法都将随之变革. 然而, 少了标签数据的约束, 深度学习收敛将十分困难, 大量的科学技术问题亟待解决.

## 4 研究前景与展望

国内外研究进展表明, 当前的应用场景主要有

3 种.

(1) 基于神经网络的偏微分方程建立: 在大量实测数据的基础上, 可以利用深度学习网络在大量备选的偏微分算子中, 寻找合适的算子来“拟合”实测数据, 从而“人工智能”地建立偏微分方程. 相关研究很多, Raissi 和 Karniadakis<sup>[42]</sup> 认为, 深度学习有能力在少量的数据中就可发现其背后的物理规律; 对高噪声稀疏数据, Xu 和 Zhang<sup>[67]</sup> 将深度学习和遗传算法相融合来提算法高鲁棒性.

(2) 参数反演: 参数反演是众多工程问题的难点之一. 利用神经网络的逼近能力, 可以进行快速反演. 该方法一般是基于标签数据的. 数据驱动下的参数反演可无须考虑物理过程, 例如试井参数自动反演就可从数据到数据<sup>[18-19]</sup>. 然而, 将数据驱动与物理信息相结合, 可能会有更好的反演效果<sup>[86,97]</sup>. 从另一个角度, 将物理信息(如控制方程)加入预测模型, 能够根据容易获得的数据预测更难获得的数据<sup>[89,98-99]</sup>. 为此, 将物理驱动与数据驱动相结合, 将大幅降低标签数据的数据量, 相关研究将给工业应用带来很大的便利.

(3) 偏微分方程智能求解: 基于深度网络的偏微分方程求解, 无须网格划分、线性方程组求解、没有维度灾难. 然而, 由于约束少, 纯物理约束下的求解方法挑战仍很大. 这也是偏微分方程深度学习求解必须攻克的难题.

另一求解偏微分方程的思路是利用神经网络对非线性方程组的逼近能力, 进行偏微分方程求解. 神经网络的权值修正方法可以用来逼近与求解非线性方程组, 且能证明, 通过对权值的限制, 该方法一定是收敛的<sup>[100]</sup>. 由于偏微分方程离散后仍是非线性的, 从而可用此方法进行求解. 按此思路, Ebadi 等<sup>[101]</sup> 用自适应神经网络代替传统的牛顿迭代法, 提出了一种技术路线与上完全不同的偏微分方程求解方法, 并利用该方法对一维的单项与两相流动进行了求解.

另一研究思路是, 将偏微分方程与深度学习混合使用, 相互学习、相互支撑. 存在这样的情形, 数据中隐藏的物理规律部分已知, 部分未知. 未知部分用深度学习表征, 已知部分用偏微分方程表征, 从而整个损失函数将包含这两部分内容. 这样, 在神经网络训练时, 二者能相互约束、相互修正, 既能防止过拟合, 也能弥补数据丢失等引起的误差<sup>[102]</sup>, 使精度

更高, 效果更好.

除传统数值计算具备的应用前景外, 基于深度神经网络的偏微分方程求解还将大幅提升强非线性偏微分方程求解能力, 借助深度学习的并行能力, 提升数值模拟速度. 再借助物理约束与深度学习本身认知能力, 在历史拟合中, 有望自我完善偏微分方程, 从而极大丰富与增强建模途径与能力, 流动模型建立与数值模拟能力都将发生巨大变化.

## 5 结语

从反演未知偏微分方程与求解已知偏微分方程两个角度, 本文总结了偏微分方程智能求解方法的发展历程, 从数据驱动、物理约束和物理驱动 3 个方面, 重点介绍了已知偏微分方程的智能求解方法, 简要介绍了应用场景和未来研究方向. 数学界更注重一般偏微分方程智能求解方法研究, 结合具体领域的应用研究正受到重视. 若能突破物理驱动求解方法瓶颈, 有望颠覆传统偏微分方程数值求解技术, 引发数值模拟技术的巨大变革.

偏微分方程深度学习求解具有深厚的科学内涵, 需要将深度学习理论、数值模拟技术、偏微分方程数学本质、偏微分方程物理意义和工程背景等有机融合, 深度交叉, 方能得到物理意义明确、数学基础坚实、能解决工程问题的偏微分方程求解新方法, 将推动数学、力学、人工智能和油藏工程等学科融合与学科发展.

## 参 考 文 献

- 1 Zhang K, Zhao XG, Zhang LM, et al. Current status and prospect for the research and application of big data and intelligent optimization methods in oilfield development. *Journal of China University of Petroleum (Edition of Natural Science)*, 2020, 44(4): 28-38
- 2 韩大匡. 加快推进智慧油气田建设, 助力企业降本增效. 中国智慧油气田技术交流会, 2019 (Han Dakuang. Accelerate the construction of intelligent oil and gas fields to help enterprises reduce cost and increase efficiency//China Intelligent Oil and Gas Field Technology Exchange Conference, 2019 (in Chinese))
- 3 李道伦, 卢德唐, 孔祥言等. BP 神经网络隐式法在测井数据处理中的应用. *石油学报*, 2007, 28(3): 105-108 (Li Daolun, Lu Detang, Kong Xiangyan, et al. Application of BP neural network implicit method in logging data processing. *Acta Petrolei Sinica*, 2007, 28(3): 105-108 (in Chinese))
- 4 Asadisaghani J, Tahmasebi P. Comparative evaluation of back-propagation neural network learning algorithms and empirical correlations for prediction of oil PVT properties in Iran oilfields. *Journal of Petroleum Science & Engineering*, 2011, 78(2): 464-475

- 5 Enab K, Ertekin T. Artificial neural network based design for dual lateral well applications. *Journal of Petroleum Science and Engineering*, 2014, 123: 84-95
- 6 Singh S, Kanli A, Sevgen S. A general approach for porosity estimation using artificial neural network method: a case study from Kansas gas field. *Studia Geophysica et Geodaetica*, 2016, 60(1): 130-140
- 7 Memon P, Yong S, Pao W, et al. Dynamic well bottom-hole flowing pressure prediction based on radial basis neural network//Science and Information Conference. Springer, Cham, 2014: 279-292
- 8 Kim C, Kim Y, Shin C, et al. A comprehensive approach to select completion and fracturing fluid in shale gas reservoirs using the artificial neural network. *Environmental Earth Sciences*, 2017, 76(20): 1-18
- 9 Athichanagorn S, Horne R. Automatic parameter estimation from well test data using artificial neural network//Software - Practice and Experience, 1995
- 10 邓远忠, 陈钦雷. 试井解释图版拟合分析的神经网络方法. *石油勘探与开发*, 2000, 27(1): 2-5 (Deng Yuanzhong, Chen Qinlei. Neural network method for well test interpretation plate fitting analysis. *Petroleum Exploration and Development*, 2000, 27(1): 2-5 (in Chinese))
- 11 Jeirani Z, Mohebbi A. Estimating the initial pressure, permeability and skin factor of oil reservoirs using artificial neural networks. *Journal of petroleum science and engineering*, 2006, 50(1): 11-20
- 12 Adibifard M, Tabatabaei-Nejad S, Khodapanah E. Artificial Neural Network (ANN) to estimate reservoir parameters in Naturally Fractured Reservoirs using well test data. *Journal of Petroleum Science and Engineering*, 2014, 122: 585-594
- 13 Ghaffarian N, Eslamloueyan R, Vaferi B. Model identification for gas condensate reservoirs by using ANN method based on well test data. *Journal of Petroleum Science and Engineering*, 2014, 123: 20-29
- 14 Tian C, Horne R. Recurrent neural networks for permanent down-hole gauge data analysis//SPE Annual Technical Conference and Exhibition, 2017, SPE-187181-MS
- 15 Sudakov O, Burnaev E, Koroteev D. Driving digital rock towards machine learning: Predicting permeability with gradient boosting and deep neural networks. *Computers & Geosciences*, 2019, 127: 91-98
- 16 Mosser L, Blévec T, Dubrule O. Reconstruction of three-dimensional porous media: Statistical or deep learning approach? *Statistical Data Science*, 2018: 125-139
- 17 张东晓, 陈云天, 孟晋. 基于循环神经网络的测井曲线生成方法. *石油勘探与开发*, 2018, 45(4): 598-607 (Zhang Dongxiao, Chen Yuntian, Meng Jin. Logging curve generation method based on cyclic neural network. *Petroleum Exploration and Development*, 2018, 45(4): 598-607 (in Chinese))
- 18 李道伦, 刘旭亮, 查文舒等. 基于卷积神经网络的径向复合油藏自动试井解释方法. *石油勘探与开发*, 2020, 47(3): 583-591 (Li Daolun, Liu Xuliang, Zha Wenshu, et al. Automatic well test interpretation method for radial composite reservoir based on convolutional neural network. *Petroleum Exploration and Development*, 2020, 47(3): 583-591 (in Chinese))
- 19 Liu XL, Li DL, Yang JH, et al. Automatic well test interpretation based on convolutional neural network for infinite reservoir. *Journal of Petroleum Science and Engineering*, 2020, 195: 107618
- 20 Liu XL, Zha WS, Qi ZK, et al. Automatic reservoir model identification method based on convolutional neural network. *Journal of Energy Resources Technology*, 2021, 144(4): 043002
- 21 Esmaili S, Mohaghegh S. Full field reservoir modeling of shale assets using advanced data-driven analytics. *Geoscience Frontiers*, 2016, 7(1): 11-20
- 22 Zhang DX, Chen YT, Jin M. Synthetic well logs generation via Recurrent Neural Networks. *Petroleum Exploration and Development*, 2018, 45(4): 629-639
- 23 Wang NZ, Chang HB, Zhang DX. Efficient uncertainty quantification for dynamic subsurface flow with surrogate by Theory-guided Neural Network. *Computer Methods in Applied Mechanics and Engineering*, 2021, 373: 113492
- 24 Zha WS, Li XB, Xing Y, et al. Reconstruction of shale image based on Wasserstein Generative Adversarial Networks with gradient penalty. *Advances in Geo-Energy Research*, 2020, 4(1): 107-114
- 25 Zha WS, Li XB, Li DL, et al. Shale digital core image generation based on generative adversarial networks. *Journal of Energy Resources Technology*, 2021, 143(3): 033003
- 26 Liu ZY, Yang YT, Cai QD. Neural network as a function approximator and its application in solving differential equations. *Applied Mathematics and Mechanics*, 2019, 40(2): 237-248
- 27 Han JQ, Jentzen A, E WN. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 2018, 115(34): 8505-8510
- 28 Mo S, Zabarar N, Shi X, et al. Deep autoregressive neural networks for high - dimensional inverse problems in groundwater contaminant source identification. *Water Resources Research*, 2019, 55(5): 3856-3881
- 29 He XP, Santoso R, Hoteit H. Application of machine-learning to construct equivalent continuum models from high-resolution discrete-fracture models//International Petroleum Technology Conference, 2020
- 30 E WN, Han JQ, Jentzen A. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics*, 2017, 5(4): 349-380
- 31 Brunton S, Proctor J, Kutz J. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 2016, 113(15): 3932-3937
- 32 Rudy S, Brunton S, Proctor J, et al. Data-driven discovery of partial differential equations. *Science Advances*, 2017, 3(4)
- 33 Schaeffer H. Learning partial differential equations via data discovery and sparse optimization. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2017, 473(2197): 20160446
- 34 Chang HB, Zhang DX. Identification of physical processes via combined data-driven and data-assimilation methods. *Journal of Computational Physics*, 2019, 393: 337-350
- 35 Reinbold P, Grigoriev R. Data-driven discovery of partial differential equation models with latent variables. *Physical Review E*, 2019, 100(2): 022219
- 36 Xu H, Chang HB, Zhang DX. DLGA-PDE: Discovery of PDEs with incomplete candidate library via combination of deep learning and genetic algorithm. *Journal of Computational Physics*, 2020, 418: 109584

- 37 Wang NZ, Chang HB, Zhang DX. Deep - Learning - Based Inverse Modeling Approaches: A Subsurface Flow Example. *Journal of Geophysical Research: Solid Earth*, 2021, 126(2)
- 38 Xu H, Chang HB, Zhang DX. DL-PDE: Deep-learning based data-driven discovery of partial differential equations from discrete and noisy data. *Communications in Computational Physics*, 2021, 29(3): 698-728
- 39 Raissi M, Perdikaris P, Karniadakis G. Physics informed deep learning (part I): Data-driven solutions of nonlinear partial differential equations. *ArXiv*(2017): 1711.10561
- 40 Raissi M, Perdikaris P, Karniadakis G. Physics Informed Deep Learning (Part II): Data-driven solutions of nonlinear partial differential equations. *ArXiv*(2017): 1711.10566
- 41 Raissi M, Perdikaris P, Karniadakis G. Numerical Gaussian processes for time-dependent and nonlinear partial differential equations. *SIAM Journal on Scientific Computing*, 2018, 40(1): A172-A198
- 42 Raissi M, Karniadakis G. Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics*, 2018, 357: 125-141
- 43 Nabian M, Meidani H. A deep learning solution approach for high-dimensional random differential equations. *Probabilistic Engineering Mechanics*, 2019, 57: 14-25
- 44 McCulloch W, Pitts W. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 1943, 5(4): 115-133
- 45 Rumelhart D, Hinton G, Williams R. Learning representations by back-propagating errors. *Nature*, 1986, 323(6088): 533-536
- 46 Rall L. Automatic Differentiation: Techniques and Applications. *Springer Berlin Heidelberg*, 1981
- 47 Rall L, Corliss G. An introduction to automatic differentiation. *Computational Differentiation: Techniques, Applications, and Tools*, 1996: 89
- 48 Baydin A, Pearlmutter B, Radul A, et al. Automatic differentiation in machine learning: a survey. *Journal of Machine Learning Research*, 2018, 18(153): 1-43
- 49 Wornik K, Stinchcombe M, White H. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 1990, 3(5): 551-560
- 50 Li X. Simultaneous approximations of multivariate functions and their derivatives by neural networks with one hidden layer. *Neurocomputing*, 1996, 12(4): 327-343
- 51 Lagaris I, Likas A, Fotiadis D. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Transactions on Neural Networks*, 1998, 9(5): 987-1000
- 52 Aarts L, Van D. Neural network method for solving partial differential equations. *Neural Processing Letters*, 2001, 14(3): 261-271
- 53 Ramuhalli P, Udpa L, Udpa S. Finite-element neural networks for solving differential equations. *IEEE Transactions on Neural Networks*, 2005, 16(6): 1381-1392
- 54 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *ArXiv*(2014): 1409.1556
- 55 Sutskever I, Vinyals O, Le Q. Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*, 2014: 3104-3112
- 56 Long ZC, Lu YP, Ma XZ, et al. Pde-net: Learning pdes from data. *International Conference on Machine Learning. PMLR*, 2018: 3208-3216
- 57 Long ZC, Lu YP, Dong B. PDE-Net 2.0: Learning PDEs from data with a numeric-symbolic hybrid deep network. *Journal of Computational Physics*, 2019, 399: 108925
- 58 Martius G, Lampert C. Extrapolation and learning equations. *ArXiv*(2016): 1610.02995
- 59 Sahoo S, Lampert C, Martius G. Learning equations for extrapolation and control//International Conference on Machine Learning, 2018, 4439-4447
- 60 Cai JF, Dong B, Osher S, et al. Image restoration: total variation, wavelet frames, and beyond. *Journal of the American Mathematical Society*, 2012, 25(4): 1033-1089
- 61 Dong B, Jiang QT, Shen ZW. Image restoration: Wavelet frame shrinkage, nonlinear evolution pdes, and beyond. *Multiscale Modeling & Simulation*, 2017, 15(1): 606-660
- 62 González-García R, Rico-Martinez R, Kevrekidis I. Identification of distributed parameter systems: A neural net based approach. *Computers & Chemical Engineering*, 1998, 22: S965-S968
- 63 Qin T, Wu KL, Xiu DB. Data driven governing equations approximation using deep neural networks. *Journal of Computational Physics*, 2019, 395: 620-635
- 64 Wu KL, Xiu DB. Data-driven deep learning of partial differential equations in modal space. *Journal of Computational Physics*, 2020, 408: 109307
- 65 Chen Z, Churchill V, Wu KL, et al. Deep neural network modeling of unknown partial differential equations in nodal space. *Journal of Computational Physics*, 2022, 449: 110782
- 66 Chen YT, Luo YT, Liu Q, et al. Any equation is a forest: Symbolic genetic algorithm for discovering open-form partial differential equations (SGA-PDE). *ArXiv: Neural and Evolutionary Computing*, 2021, 2106.11927
- 67 Xu H, Zhang DX. Robust discovery of partial differential equations in complex situations. *Phys Rev Research*, 3, 2021: 033270
- 68 Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press, 2016
- 69 Bottou L. Stochastic gradient descent tricks. Springer, Berlin, Heidelberg, 2012: 421-436
- 70 Duchi J, Hazan E, Singer Y. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 2011, 12(7): 2121-2159
- 71 Sutskever I, Martens J, Dahl G, et al. On the importance of initialization and momentum in deep learning. *International Conference on Machine Learning*, 2013: 1139-1147
- 72 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015, 521(7553): 436-444
- 73 Zha WS, Zhang W, Li DL, et al. Convolution-based model-solving method for three-dimensional, unsteady, partial differential equations. *Neural Comput*, 2021, 30: 1-23
- 74 Raissi M, Perdikaris P, Karniadakis G. Inferring solutions of differential equations using noisy multi-fidelity data. *Journal of Computational Physics*, 2017, 335: 736-746
- 75 Raissi M, Perdikaris P, Karniadakis G. Machine learning of linear differential equations using Gaussian processes. *Journal of Computational Physics*, 2017, 348: 683-693
- 76 Owahdi H. Bayesian numerical homogenization. *Multiscale Modeling & Simulation*, 2015, 13(3): 812-828
- 77 Sirignano J, Spiliopoulos K. DGM: A deep learning algorithm for

- solving partial differential equations. *Journal of Computational Physics*, 2018, 375: 1339-1364
- 78 Bandai T, Ghezzehei T. Physics - informed neural networks with monotonicity constraints for Richardson - Richards equation: Estimation of constitutive relationships and soil water flux density from volumetric water content measurements. *Water Resources Research*, 2021, 57(2)
- 79 Han J, Nica M, Stinchcombe A. A derivative-free method for solving elliptic partial differential equations with deep neural networks. *Journal of Computational Physics*, 2020, 419: 109672: 1-18
- 80 Meng XH, Li Z, Zhang DK, et al. PPINN: Parareal physics-informed neural network for time-dependent PDEs. *Computer Methods in Applied Mechanics and Engineering*, 2020, 370: 113250
- 81 Michoski C, Milosavljević M, Oliver T, et al. Solving differential equations using deep neural networks. *Neurocomputing*, 2020, 399: 193-212
- 82 Kani J, Elsheikh A. Reduced-order modeling of subsurface multi-phase flow models using deep residual recurrent neural networks. *Transport in Porous Media*, 2019, 126(3): 713-741
- 83 Jagtap A, Kawaguchi K, Karniadakis G. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *Journal of Computational Physics*, 2020, 404: 109136
- 84 Jagtap A, Kawaguchi K, Em Karniadakis G. Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks. *Proceedings of the Royal Society A*, 2020, 476(2239): 20200334
- 85 Wang NZ, Zhang DX, Chang HB, et al. Deep learning of subsurface flow via theory-guided neural network. *Journal of Hydrology*, 2020, 584: 124700
- 86 Wang NZ, Chang HB, Zhang DX. Theory-guided auto-encoder for surrogate construction and inverse modeling. *Computer Methods in Applied Mechanics and Engineering*, 2021, 385: 114037
- 87 Li DL, Shen LH, Zha WS, et al. Physics-constrained deep learning for solving seepage equation. *Journal of Petroleum Science and Engineering*, 2021: 109046
- 88 Chen YT, Chang HB, Jin M, et al. Ensemble neural networks (ENN): A gradient-free stochastic method. *Neural Networks*, 2019, 110: 170-185
- 89 Chen YT, Huang D, Zhang DX, et al. Theory-guided hard constraint projection (HCP): A knowledge-based data-driven scientific machine learning method. *Journal of Computational Physics*, 2021, 445: 110624
- 90 Nabian M, Gladstone R, Meidani H. Efficient training of physics - informed neural networks via importance sampling. *Computer-Aided Civil and Infrastructure Engineering*, 2021, 36(8): 962-977
- 91 Cai ZQ, Chen JS, Liu M, et al. Deep least-squares methods: An unsupervised learning-based numerical method for solving elliptic PDEs. *Journal of Computational Physics*, 2020, 420: 109707
- 92 Sun LN, Gao H, Pan SW, et al. Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data. *Computer Methods in Applied Mechanics and Engineering*, 2020, 361: 112732
- 93 Gao H, Sun LN, Wang JX. PhyGeoNet: Physics-informed geometry-adaptive convolutional neural networks for solving parametric PDEs on irregular domain. ArXiv (2020): 2004.13145
- 94 Zhu YH, Zabaras N. Bayesian deep convolutional encoder - decoder networks for surrogate modeling and uncertainty quantification. *Journal of Computational Physics*, 2018, 366: 415-447
- 95 Iandola F, Moskewicz M, Karayev S, et al. DenseNet: implementing efficient convnet descriptor pyramids. *ArXiv(2014)*: 1404.1869
- 96 Zhu YH, Zabaras N, Koutsourelakis P, et al. Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data. *Journal of Computational Physics*, 2019, 394: 56-81
- 97 Tartakovsky A, Marrero C, Perdikaris P, et al. Physics - informed deep neural networks for learning parameters and constitutive relationships in subsurface flow problems. *Water Resources Research*, 2020, 56(5)
- 98 Chen YT, Zhang DX. Theory-guided deep-learning for electrical load forecasting (TgDLF) via ensemble long short-term memory. *Advances in Applied Energy*, 2021, 1: 100004
- 99 Chen YT, Zhang DX. Physics-constrained deep learning of geomechanical logs. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(8): 5932-5943
- 100 Li GM, Zeng ZZ. A neural-network algorithm for solving nonlinear equation systems//International Conference on Computational Intelligence and Security, 2008, 1: 20-23
- 101 Ebadi M, Zabihiyar S, Bezyan Y, et al. A nonlinear solver based on an adaptive neural network, introduction and application to porous media flow. *Journal of Natural Gas Science and Engineering*, 2021, 87: 103749
- 102 Sirignano J, MacArt J, Freund J. DPM: A deep learning PDE augmentation method with application to large-eddy simulation. *Journal of Computational Physics*, 2020, 423: 109811