



国防科技大学学报

Journal of National University of Defense Technology

ISSN 1001-2486, CN 43-1067/T

《国防科技大学学报》网络首发论文

题目: 面向高阶微分方程求解的预测-微分解耦式神经网络
作者: 彭博, 周仕明, 李道奎
收稿日期: 2025-03-18
网络首发日期: 2026-02-04
引用格式: 彭博, 周仕明, 李道奎. 面向高阶微分方程求解的预测-微分解耦式神经网络
[J/OL]. 国防科技大学学报.
<https://link.cnki.net/urlid/43.1067.t.20260203.1751.002>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字符、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

doi: 10.11887/j.issn.1001-2486.25030018

面向高阶微分方程求解的预测-微分解耦式神经网络

彭 博^{1,2}, 周仕明^{1,2}, 李道奎^{1,2*}

(1. 国防科技大学 空天科学学院, 湖南 长沙 410073; 2. 空天任务智能规划与仿真湖南省重点实验室, 湖南 长沙 410073)

摘要: 物理信息驱动的算子学习方法在加速求解偏微分方程 (partial differential equations, PDEs) 时, 因自动微分技术限制, 在高阶问题中面临训练成本高昂的挑战。提出一种预测-微分解耦式神经网络架构 UNet-RBF, 采用 U-Net(U-Net neural network)作为预测网络提取 PDEs 参数的空间特征, 并利用径向基函数 (radial basis function, RBF) 网络作为轻量级的微分网络施加物理约束。通过在训练中冻结 RBF 网络参数, 实现了预测任务与微分计算的解耦, 显著降低了自动微分的计算开销。数值实验结果表明, UNet-RBF 在保证高预测精度 (相对误差小于 1%) 的同时, 大幅提升了高阶 PDEs 求解的训练效率, 相较于传统物理信息驱动算子学习模型, 在四阶问题中训练效率提升超过 1500%, 且模型鲁棒性更强, 为复杂物理问题的快速、准确求解提供了有效途径。

关键词: 偏微分方程; 算子学习; 物理信息; 预测-微分解耦

中图分类号: O241.8 文献标识码: A

Decoupled prediction-differential neural networks for solving high-order differential equations

PENG Bo^{1,2}, ZHOU Shiming^{1,2}, LI Daokui^{1,2*}

(1. College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, China; 2. Hunan Key Laboratory of Intelligent Planning and Simulation for Aerospace Missions, Changsha 410073, China)

Abstract: Physics-informed operator learning methods, while advantageous for accelerating the solution of PDEs (partial differential equations), face challenges of high training costs in high-order problems due to limitations of automatic differentiation. A novel prediction-differential decoupled neural network architecture named UNet-RBF was proposed. It employed U-Net neural network as the prediction network to extract spatial features of PDEs parameters and utilized a lightweight RBF (radial basis function) network as the differential network to impose physical constraints. By freezing the RBF network parameters during training, the prediction task was decoupled from differential computation, significantly reducing the computational overhead of automatic differentiation. Numerical experiments demonstrate that UNet-RBF substantially improves the training efficiency and stability for solving high-order PDEs while maintaining high prediction accuracy (relative error less than 1%). Compared to the traditional physics-informed operator learning model, the training efficiency for fourth-order problems is increased by over 1500%, and the model exhibits stronger robustness, offering an effective pathway for the rapid and accurate solution of complex physical problems.

Keywords: partial differential equations; operator learning; physical information; prediction-differential decoupling

偏微分方程 (partial differential equations, PDEs) 作为描述复杂物理现象的核心数学工具, 其高效求解在科学研究与工程实践中至关重要。尽管传统数值方法 (如有限元法^[1]) 已取得显著成就, 但在处理大规模、高维度或实时性要求高的复杂问题时, 其高昂的计算成本往往成为瓶颈,

难以满足数字孪生^[2]等新兴领域对计算效率的严苛需求。

深度学习的快速发展为突破 PDEs 求解的效率瓶颈提供了新范式^[3]。其中, 物理信息神经网络 (physics-informed neural network, PINN)^[4]作为代表性成果, 通过将物理约束融入神经网络损失

收稿日期: 2025-03-18

基金项目: 国家自然科学基金资助项目 (12302197)

第一作者: 彭博 (1997—), 男, 湖南娄底人, 博士研究生, E-mail: pengbo@nudt.edu.cn;

*通信作者: 李道奎 (1971—), 男, 湖北随州人, 教授, 博士, 博士生导师, E-mail: lidaokui@nudt.edu.cn

引用格式: 彭博, 周仕明, 李道奎. 面向高阶微分方程求解的预测-微分解耦式神经网络[J]. 国防科技大学学报,

Citation: PENG B, ZHOU S M, LI D K. Decoupled prediction-differential neural networks for solving high-order differential equations[J]. Journal of National University of Defense Technology,

函数, 实现了对 PDEs 的无监督求解, 展现了巨大潜力。PINN 以物理方程的自变量为输入, 通过深度神经网络逼近 PDEs 的解, 并利用自动微分 (automatic differentiation, AD) 技术将 PDEs 及其边界条件编码为损失函数, 从而将求解问题转化为网络参数的优化问题。PINN 开创性地将物理信息融入神经网络的训练过程, 其简洁的架构使其能够求解多种类型 PDEs^[5-7]。在此基础上, 研究者们进一步探索了多种基于深度学习的求解策略, 如基于变分原理、适用于高维和非光滑解场景的深度 Ritz 方法 (deep Ritz method, DRM)^[8], 以及通过弱解形式结合对抗训练、有效处理间断解和非线性问题的弱对抗网络 (weak adversarial networks, WAN)^[9]。然而, 这些方法主要针对参数固定的 PDEs 问题。对于参数变化的 PDEs, 需要重新训练模型, 限制了其泛化能力和应用范围。

针对变参数 PDEs 的泛化求解需求, 算子学习 (operator learning) 应运而生, 旨在学习从输入函数 (如边界条件或方程参数) 到解函数的全局映射, 并迅速成为当前的研究热点^[10]。其中的代表性工作包括 DeepONet (deep operator network)^[11] 和傅里叶神经算子 (Fourier neural operator, FNO)^[12]。DeepONet 通过分支-主干 (branch-trunk) 网络结构学习解算子的离散表达, 而 FNO 则利用快速傅里叶变换在频域捕捉解的全局特征。此外, Koopman 神经算子^[13] 将非线性动力学映射至无限维线性空间, 实现了复杂系统长期演化的高效建模; DeepGreen Network^[14] 则结合格林函数与深度学习, 直接学习微分算子的逆, 为线性 PDEs 提供了快速反演框架。值得注意的是, Wang 等^[15] 受到 PINN 的启发, 将物理约束引入 DeepONet, 提出了 PI-DeepONet (physics-informed deep operator network)。该方法实现了网络在零样本条件下的无监督训练, 并且其预测 PDEs 解的速度相较于传统数值方法提升了数个数量级, 为 PDEs 的实时求解提供了可能。然而, PI-DeepONet 在处理高阶 PDEs 时, 其依赖自动微分计算物理残差的方式, 导致训练过程中的计算成本和内存消耗随微分阶次的增加而急剧上升, 严重制约了其在复杂高阶问题中的应用效率。

针对上述挑战, 在 DeepONet 的基础上提出一种新颖的预测-微分解耦式神经网络架构——UNet-RBF, 采用 U-Net 网络和径向基函数 (radial basis function, RBF) 网络分别作为 DeepONet 的分支网络和主干网络, U-Net 用以预测控制点, 并利用控制点求解 RBF 网络结果。通过 RBF 网络代替原始 DeepONet 中的全连接主干网络, 有效降低主干网络规模, 从而减少模型训练成本, 旨在实现训练效率与精度的双重突破。

1 物理信息驱动的算子学习方法

1.1 微分问题表述

本研究致力于构建一种神经网络模型, 旨在求解参数化 PDEs^[16], 其通用形式可表述为:

$$\mathcal{N}[u, s(x)] = 0, x \in \Omega \quad (1)$$

其中: $\Omega \subset \mathbb{R}^d$ 代表 d 维物理域; x 表示 d 维空间坐标; \mathcal{N} 是一个广义微分算子; u 代表 PDEs 的参数, 这些参数可涵盖控制方程的系数、源项以及边界条件; $s(x)$ 表示在给定参数 u 和相应边界条件下, 该 PDEs 的解。

1.2 DeepONet 网络

算子通用逼近定理^[17] 揭示了, 存在某个函数集合的线性组合, 能够以任意精度逼近定义在函数空间上的连续算子。受此启发, Lu 等^[11] 设计了 DeepONet 的网络架构 G_θ , 用以近似求解算子 $G: u \rightarrow s$:

$$G_\theta(u)(x) = \sum_{k=1}^q \underbrace{b_k(u(y_1), u(y_2), \dots, u(y_m))}_{\text{branch}} \cdot \underbrace{t_k(x)}_{\text{trunk}} \quad (2)$$

其中: $\{y_1, y_2, \dots, y_m\}$ 是计算域内预先选定的一组位置点, 称为传感器 (sensors) 位置; θ 代表网络的全部可训练参数。

DeepONet 的核心思想是将参数化 PDE 的输入函数 u 在预设的传感器位置 $\{y_j\}_{j=1}^m$ 进行离散化采样, 得到向量 $[u(y_1), u(y_2), \dots, u(y_m)]$, 并将其作为分支网络 (branch net) 的输入。同时, 将计算域内待求解位置的坐标 x 作为主干网络 (trunk net) 的输入。最后, 通过计算分支网络输出向量 $b = [b_1, \dots, b_q]$ 与主干网络输出向量 $t = [t_1, \dots, t_q]$ 的点积, 得到 DeepONet 的最终输出, 即解在位置 x 的近似值, 如图 1 所示。

DeepONet 提供了一个模块化的算子学习框架, 其分支网络和主干网络的具体架构可以灵活选择^[18]。鉴于坐标 x 的输入维度通常较低, 主干网络常采用前馈神经网络 (feedforward neural network, FNN)。分支网络的选择则依赖于输入函数 u 的特性, 例如当 u 表现为结构化的离散数据时 (如图像), 卷积神经网络 (convolutional neural networks, CNN) 是理想的选择; 而对于非结构化的离散数据 (如点云或图结构数据), 图神经网络 (graph neural network, GNN) 则更为适用。

1.3 物理信息驱动的 DeepONet 网络

Wang 等^[15] 注意到 DeepONet 网络的输出对其输入坐标 x 是连续且可微的。基于这一特性, 从 PINN 中汲取灵感, 他们提出将自动微分技术应用于 DeepONet, 通过计算网络输出关于坐标 x 的偏导数, 并将这些导数带入原始 PDEs 以构造

残差损失项，从而使模型输出在训练过程中满足 PDEs 约束。这种方法被称为物理信息驱动的 DeepONet (PI-DeepONet)，其损失函数具体定义如下：

$$L_{\text{PINO}} = L_{\text{BC}} + L_{\text{PDE}} \quad (3)$$

此处， L_{BC} 代表边界条件产生的损失， L_{PDE} 代表 PDEs 残差产生的物理信息损失。

$$\begin{cases} L_{\text{BC}} = \frac{1}{N} \sum_{i=1}^N \|G_{\theta}(u)(x_i) - g(x_i)\|^2, & x_i \in \partial\Omega \\ L_{\text{PDE}} = \frac{1}{M} \sum_{j=1}^M \|\mathcal{N}[u, G_{\theta}(u)(x_j)]\|^2, & x_j \in \Omega \end{cases} \quad (4)$$

其中： N 和 M 分别表示在边界 $\partial\Omega$ 和计算域内部 Ω 随机采样的配点总数； $\{x_i\}_{i=1}^N$ 是边界上的采样点， $\{x_j\}_{j=1}^M$ 是域内的采样点； $g(x_i)$ 表示在边界点 x_i 处应满足的边界条件值，该信息包含在参数 u 中。

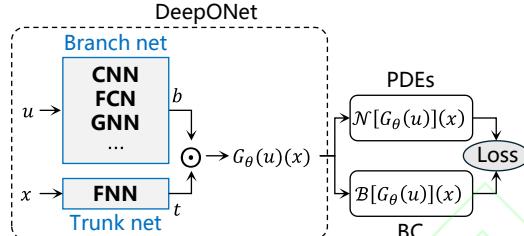


图 1 物理信息驱动的 DeepONet 架构示意图

Fig. 1 Schematic of the PI-DeepONet architecture

2 预测-微分解耦式神经网络

尽管物理信息驱动的 DeepONet (PI-DeepONet) 在算子学习中表现出显著优势，但其在求解高阶 PDEs 时，对自动微分技术的依赖导致了计算成本和内存消耗随微分阶次的升高而急剧增长的问题，严重制约了其应用效能。为应对这一挑战，本研究提出一种新颖的预测-微分解耦式神经网络架构，命名为 UNet-RBF。该架构的核心思想在于将 PDEs 解的预测任务与物理约束施加所需的微分运算进行解耦，从而在保证模型表达能力的同时，显著优化高阶 PDEs 求解的训练效率与稳定性。

UNet-RBF 架构设计的核心原则是策略性地实现“预测-微分解耦”。具体而言，该模型在结构上将复杂的算子映射学习(即从输入函数到 PDEs 解的预测任务)与计算密集型的微分运算(即通过物理残差施加约束的任务)分离开来。通过为微分计算模块选用轻量化设计，UNet-RBF 旨在不牺牲模型表达能力的前提下，大幅提升训练效率，特别是针对高阶微分问题。

UNet-RBF 架构如图 2 所示，其分支网络采

用 U-Net 模型。鉴于 U-Net 独特的编码器-解码器结构及跳跃连接机制，其在处理具有显著空间结构特征的输入数据(例如，离散于结构化网格上的 PDEs 参数场)时，能够有效地捕获多尺度特征，从而为后续的解算子近似提供高质量的系数表征。U-Net 的输入维度是 $C \times H \times W$ 的 PDEs 参数在结构化网络 $\{y_k\}_{k=1}^{HW}$ 上的离散张量，其中 C 表示参数 u 中变量的总量， $H \times W$ 是该结构化网格的大小。U-Net 的输出是一个维度为 $O \times H \times W$ 的控制张量，其中 O 表示 PDEs 解的参数化分量数量。该控制张量随后被展平，形成一个维度为 $O \times HW$ 的控制向量 b ，实质上构成了 DeepONet 框架中分支网络输出的系数。

本文的关键创新在于主干网络的改进：将传统 DeepONet 中的前馈神经网络 (FNN) 主干替换为径向基函数 (RBF) 网络，此设计借鉴了基于 RBF 的偏微分方程数值解法^[19]。RBF 网络不仅具有良好的函数逼近能力，其通常采用的浅层结构以及解析可微的基函数特性，使其非常适合作为本架构中的轻量级微分模块。这一选择是实现预测与微分计算有效分离进而降低自动微分计算复杂度的关键。

UNet-RBF 的独特之处在于其主干 RBF 网络参数的处理方式(如图 2 所示)。RBF 基函数的中心 $\{y_k\}_{k=1}^{HW}$ 被设定为与分支 U-Net 输入对应的传感器(结构化网格点)位置，而基函数的宽度 $\{r_k\}_{k=1}^{HW}$ 则通过预训练或经验设定后固定。分支 U-Net 的输出 $\{b_k(u)\}_{k=1}^{HW}$ 直接作为这些 RBF 基函数的权系数。因此，对于给定的输入函数 u 和待求解坐标 x ，UNet-RBF 的预测值 $G_{\theta}(u)(x)$ 可表示为：

$$\begin{aligned} G_{\theta}(u)(x) &= b(u) \cdot t(x) \\ &= \sum_{k=1}^{HW} b_k(u) \cdot \varphi_k(x) \\ &= \sum_{k=1}^{HW} b_k(u) \cdot \exp\left(-\frac{\|x - y_k\|^2}{r_k^2}\right) \end{aligned} \quad (5)$$

其中： $\varphi_k(x)$ 为高斯基函数。

至关重要的是，在这种设计下，RBF 网络自身不包含在主训练阶段需要优化的参数(或仅宽度参数进行少量预训练后即固定)，其主要功能是提供一组固定的、易于求导的基函数。这使得微分运算主要作用于结构固定的 RBF 网络，有效避免了微分模块深度参与到复杂的预测网络参数优化过程中，从而实现了预测任务与微分任务的实质性解耦。

UNet-RBF 架构在训练效率上的一个核心突破在于，当用于评估 PDEs 残差的配点集合 $X = \{x_i\}$ 固定时，RBF 网络的基函数数值及其各

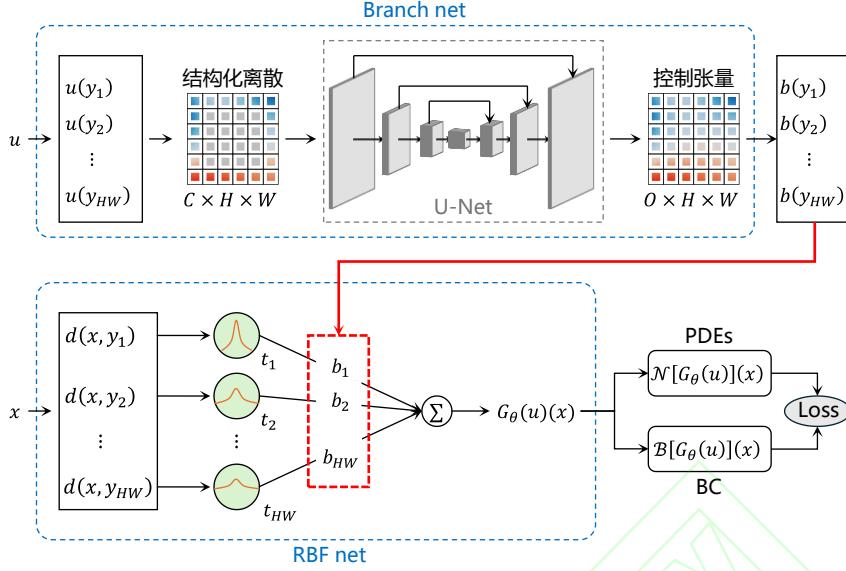


图 2 预测-微分解耦式 UNet-RBF 网络架构示意图

Fig.2 Schematic of the UNet-RBF prediction-differential decoupled network architecture

阶导数值 $\mathbf{T}^{(n)}(X)$ (对于 $n=0,1,\dots$) 均可预先计算并存储。因此, 在训练 U-Net 分支网络以优化系数 $b(u)$ 的过程中, PDEs 残差的计算可以直接利用这些预存的常数项, 而无需在每个训练迭代中对主干网络重复执行高成本的自动微分。此时, 解的预测 (如式(6)所示) 及其导数的计算均不涉及对 RBF 网络参数的实时微分, 这从根本上降低了与物理信息约束相关的计算负担, 尤其在高阶微分问题中效果更为显著。

$$G_\theta(u) = \mathbf{T}^{(0)}(X) \cdot b(u) \quad (6)$$

3 数值实验与分析

为全面评估所提出的 UNet-RBF 架构在求解偏微分方程 (PDEs) 方面的性能, 本章设计并实施了一系列数值实验。选取了涵盖一阶、二阶及四阶微分算子的三个典型 PDEs 求解问题作为基准测试。通过将 UNet-RBF 的预测精度、训练效率及稳定性与不同规模的传统物理信息驱动 DeepONet (PI-DeepONet) 模型进行对比分析, 旨在定量地验证 UNet-RBF 架构的有效性和优越性。本章将首先详细介绍实验设定与模型配置, 随后对实验结果进行深入分析与讨论。

3.1 实验设定

3.1.1 一阶微分算例

首先, 从一维微分问题入手, 该问题亦是 PI-DeepONet 原始文献^[15]使用的基准测试之一:

$$\begin{cases} \frac{ds(x)}{dx} = u(x), x \in [0,1] \\ s(0) = 0 \end{cases} \quad (7)$$

此问题的目标是利用神经网络学习从输入函数 $u(x)$ 到其反导数 $s(x)$ 的映射算子:

$$G: u(x) \rightarrow s(x) = s(0) + \int_0^x u(t) dt \quad (8)$$

输入函数 $u(x)$ 通过零均值高斯随机过程 (gaussian process, GP)^[20]生成, 选择 GP 是因为其能够便捷地生成具有不同平滑度和复杂度地函数样本:

$$u(x) \sim \mathcal{GP}(0, k_l(x_1, x_2)) \quad (9)$$

其中: $k_l(x_1, x_2) = \exp(-\|x_1 - x_2\|^2 / 2l^2)$ 为二次协方差核函数 (squared exponential kernel); $l > 0$ 是长度尺度, 用于控制输入函数的复杂度, l 越大, 生成的 u 越平滑。本算例中, 训练样本在训练时动态生成, 长度尺度设定为 $l = 0.2$ 。

由于物理信息驱动的算子学习属于无监督或自监督学习的范畴, 因此无需真实的标签数据即可构造损失函数。总损失函数 L 由边界条件损失 L_{BC} 和 PDEs 残差损失 L_{PDE} 构成, 且有:

$$\begin{aligned} L_{BC} &= \|G_\theta(u(Y))(0) - s(0)\|^2 \\ L_{PDE} &= \frac{1}{N} \sum_{x \in X} \left\| \frac{dG_\theta(u(Y))(x)}{dx} - u(x) \right\|^2 \end{aligned} \quad (10)$$

其中: $Y = [y_1, y_2, \dots, y_{16}]$ 表示在计算域 $[0,1]$ 上均匀选取的 16 个传感器位置的集合, 用于对输入函数 $u(y)$ 进行采样; $X = [x_1, x_2, \dots, x_N]$ 表示在计算域 $[0,1]$ 上均匀选取的 $N = 256$ 个配点的集合, 用于评估 PDEs 残差。

3.1.2 二阶稳态热传导算例

接下来, 考虑一个二维稳态热传导问题。根据傅里叶定律, 当导热系数为常数且无内热源时,

温度场 $T(x_1, x_2)$ 满足拉普拉斯方程:

$$\frac{\partial^2 T}{\partial x_1^2} + \frac{\partial^2 T}{\partial x_2^2} = 0, (x_1, x_2) \in \Omega \quad (11)$$

其中: 计算域 Ω 为一个大小为 2×2 的矩形区域。该问题的边界条件设定为: 计算域的三边 ($x_1 = -1$ 和 $x_2 = \pm 1$) 满足第一类边界条件, 其温度由参数函数 $u(x_1, x_2)$ 给定; 第四边 ($x_1 = 1$) 满足绝热边界条件, 即该边界上的法向热通量为零。具体表示为:

$$\begin{cases} T(x_1, x_2) = u(x_1, x_2), & (x_1, x_2) \in \partial\Omega_1 \\ \partial T / \partial x_1 = 0, & (x_1, x_2) \in \partial\Omega_2 \end{cases} \quad (12)$$

其中: $\partial\Omega_1 = (x_1, x_2) \in \partial\Omega |_{x_1=-1 \text{ or } x_2=\pm 1}$ 为第一类边界条件施加的边界部分; $\partial\Omega_2 = (x_1, x_2) \in \partial\Omega |_{x_1=1}$ 为第二类边界条件施加的边界部分; 参数函数 $u(x_1, x_2)$ 仍通过式(9)的高斯随机过程生成, 作为该参数化 PDEs 问题的输入, 长度尺度为 $l = 0.5$ 。具体如图 3 (a) 所示。

该问题的损失函数由三部分组成: L_{BC1} 对应第一类边界条件, L_{BC2} 对应第二类边界条件, 以及 L_{PDE} 对应 PDE 残差:

$$\begin{aligned} L_{\text{BC1}} &= \frac{1}{|X_{\text{BC1}}|} \sum_{x \in X_{\text{BC1}}} \|G_\theta(u(Y))(x) - u(x)\|^2 \\ L_{\text{BC2}} &= \frac{1}{|X_{\text{BC2}}|} \sum_{x \in X_{\text{BC2}}} \left\| \frac{\partial G_\theta(u(Y))(x)}{\partial n} \right\|^2 \\ L_{\text{PDE}} &= \frac{1}{|X_{\text{PDE}}|} \sum_{x \in X_{\text{PDE}}} \left\| \frac{\partial^2 G_\theta(u)(x)}{\partial x_1^2} + \frac{\partial^2 G_\theta(u)(x)}{\partial x_2^2} \right\|^2 \end{aligned} \quad (13)$$

其中: Y 是在计算域 Ω 内均匀离散的 $16 \times 16 = 256$ 个传感器位置的集合; X_{PDE} 是在 Ω 内部均匀离散的 $N \times N$ 个配点的集合, 此处选取 $N = 64$; X_{BC1} 和 X_{BC2} 分别是 X_{PDE} 在相应边界 $\partial\Omega_1$ 和 $\partial\Omega_2$ 上的子集; $|\cdot|$ 表示集合中点的数量; $\partial / \partial n$ 表示沿边界法线方向的导数, 对于 $\partial\Omega_2$ 边界即为 $\partial / \partial x_1$ 。

3.1.3 四阶薄板弯曲算例

为进一步评估模型处理高阶 PDEs 的能力, 考虑一个四边固支的矩形薄板在均布载荷作用下的弯曲问题, 如图 3 (b) 所示。薄板长宽均为 1000mm, 基于小挠度理论和 Kirchhoff 假设建立的矩形薄板弯曲基本方程 (Sophie Germain 方程) 为:

$$\frac{\partial^4 w}{\partial x_1^4} + 2 \frac{\partial^4 w}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 w}{\partial x_2^4} = \frac{q}{D} \quad (14)$$

其中: $w(x_1, x_2)$ 为薄板的挠度;

$D = E\delta_t^3 / (12(1 - \mu^2))$ 为薄板弯曲刚度; E 为薄板弹性模量, 取 70 GPa; δ_t 为薄板厚度, 取 5 mm; μ 为薄板泊松比, 取 0.33; q 为均布载荷大小, 是该 PDEs 的可变参数。

模型采用四边固支边界条件, 即在边界上挠度为零, 且挠度沿边界法线方向的导数也为零:

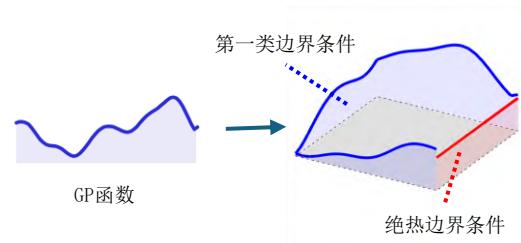
$$\begin{cases} w = 0, \partial w / \partial x_1 = 0, & x_1 = \pm 500 \\ w = 0, \partial w / \partial x_2 = 0, & x_2 = \pm 500 \end{cases} \quad (15)$$

该问题的损失函数包含三部分: L_{BC1} 对应挠度为零的边界条件, L_{BC2} 对应法向导数为零的边界条件, 以及 L_{PDE} 对应 PDEs 残差:

$$\begin{aligned} L_{\text{BC1}} &= \frac{1}{|X_{\text{BC}}|} \sum_{x \in X_{\text{BC}}} \|G_\theta(q)(x)\|^2 \\ L_{\text{BC2}} &= \frac{1}{|X_{\text{BC}}|} \sum_{x \in X_{\text{BC}}} \left\| \frac{\partial G_\theta(q)(x)}{\partial n} \right\|^2 \\ L_{\text{PDE}} &= \frac{1}{|X_{\text{PDE}}|} \sum_{x \in X_{\text{PDE}}} \left\| \frac{\partial^4 G_\theta(q)(x)}{\partial x_1^4} + 2 \frac{\partial^4 G_\theta(q)(x)}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 G_\theta(q)(x)}{\partial x_2^4} - \frac{q}{D} \right\|^2 \end{aligned} \quad (16)$$

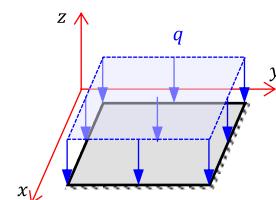
其中: X_{PDE} 为在计算域内部均匀离散的 $N \times N$ 个配点的集合, 此处 $N = 64$; X_{BC} 为 X_{PDE} 在 $x_1 = \pm 500$ mm 和 $x_2 = \pm 500$ mm 上的子集; $\partial / \partial n$ 为沿 x_1 或 x_2 方向边界的法向导数。

该问题可变参数 q (均布载荷) 的取值范围为 200~1 000 Pa, 在训练时从该范围内随机均匀采样生成。



(a) 二阶稳态热传导问题

(a) The second-order steady-state heat conduction problem



(b) 四阶薄板弯曲问题

(b) The fourth-order thin plate bending problem

图 3 数值算例的边界条件示意图

Fig.3 Boundary condition schematics for numerical examples

3.2 模型配置

在本文的数值实验中, 所提出的 UNet-RBF 架构具体采用了 Huang 等^[21]提出的 UNet3+模型作为分支网络。为适应本研究的需求, 我们对原始 UNet3+模型进行如下调整: 移除了原模型中的批量归一化层 (Batch Normalization Layer), 并将激活函数统一替换为 \tanh 函数。此外, 针对一维微分算例, UNet3+中的二维卷积模块相应地修改为一维卷积模块。UNet-RBF 的主干网络采用了 RBF 网络, 其基函数的中心位置与分支网络输入对应的传感器(即结构化网格点)位置保持一致。

为进行全面的性能对比, 引入两种不同规模的传统 PI-DeepONet 模型作为基准。这两种对比模型的分支网络均采用与 UNet-RBF 相同的、经过上述调整的 UNet3+结构, 其主要区别在于主干网络的设计: (1) DeepONet-S (Small): 其主干网络采用了一个与 RBF 网络结构相似的三层前馈神经网络 (FNN), 即输入层-单隐藏层-输出层;

(2) DeepONet-L (Large): 参考文献^[13]的经验, 其主干网络采用层数更多、神经元数量也更多的 FNN。DeepONet-S 与 DeepONet-L 的主干网络均采用 \tanh 作为激活函数。所有模型的具体网络结构与训练参数配置详见表 1。

所有模型均采用 Adam 优化器进行训练, 初始学习率设置为 0.001, 学习率调整策略为等间隔衰减: $lr = lr_0 \times 0.99^{\lfloor iteration/500 \rfloor}$ 。对于 UNet-RBF 模型, 训练过程中的前 50 个迭代步包含 RBF 网络的基函数宽度优化, 此后冻结 RBF 网络参数, 仅训练分支 U-Net 网络的参数。所有实验均在 Python 3.12 环境下, 于 NVIDIA GeForce 4090 GPU 上完成。

为定量评估模型的预测精度, 针对三个测试问题分别生成了包含 500 个独立样本的测试集。这些测试集样本的精确解通过成熟的数值方法获

得。模型性能的评价指标采用相对 L^2 范数误差, 其定义如下:

$$L^2 = \frac{1}{N} \sum_{i=1}^N \left(\frac{\sum_{j=1}^P \left\| \mathcal{G}_\theta(u^{(i)})(x_j) - s^{(i)}(x_j) \right\|_2^2}{\sum_{j=1}^P \left\| s^{(i)}(x_j) \right\|_2^2} \right)^{1/2} \quad (17)$$

其中: N 为测试集样本数量; P 为每个样本中用于评估误差的配点数量, s 为测试集精确数值解。

3.3 结果分析与讨论

本节旨在对 UNet-RBF、DeepONet-S 及 DeepONet-L 三种模型在不同偏微分方程 (PDEs) 求解算例中的性能表现进行系统的量化评估与深入比较。核心评价维度包括预测精度、模型稳定性(通过相对误差的标准差衡量)以及训练效率。

表 2 系统地总结了各模型在三个代表性测试问题上所获得的平均相对 L^2 范数误差、相对误差的标准差以及完成模型训练所需的总计算时间。在预测精度与稳定性方面, 分析表 2 数据可以发现, UNet-RBF 模型在所有测试问题中均展现出具有竞争力的预测精度。与具有相近计算复杂度的 DeepONet-S 模型(其主干网络层数与 RBF 网络相似)相比, UNet-RBF 模型在三个算例中的预测精度均观测到一个数量级的提升, 这初步证实了 RBF 网络作为主干网络在特定任务中的有效性。在与具有更深主干网络、理论上具有更高表达能力的 DeepONet-L 模型的比较中, 对于二阶稳态热传导问题, 两种模型的平均相对误差基本持平, 均达到了 0.9% 左右的水平。然而, 在一阶微分问题和四阶薄板弯曲问题中, UNet-RBF 模型的预测精度相较于 DeepONet-L 模型分别提升了 29.4% (从 0.34% 降至 0.24%) 和 30.0% (从 0.13% 降至 0.091%)。同时, 在模型稳定性方面,

表 1 各模型网络结构与训练参数配置

Tab.1 Configuration of network architectures and training parameters for all models

算例	模型	分支网络结构			主干网络结构	Batch size
		输入/输出尺寸	编码器通道数	解码器通道数		
一阶微分问题	UNet-RBF	16	[8,16,32,64]	[32,32,32,32]	[1,16]	256
	DeepONet-S	16	[8,16,32,64]	[32,32,32,32]	[1,16]	256
	DeepONet-L	16	[8,16,32,64]	[32,32,32,32]	[1,50,50,16]	256
二阶稳态热传导问题	UNet-RBF	16×16	[16,32,64,128,256]	[80,80,80,80,80]	[2,256]	16
	DeepONet-S	16×16	[16,32,64,128,256]	[80,80,80,80,80]	[2,256]	16
	DeepONet-L	16×16	[16,32,64,128,256]	[80,80,80,80,80]	[2,256,256,256]	16
四阶薄板弯曲问题	UNet-RBF	16×16	[4,8,16,32,64]	[20,20,20,20]	[2,256]	16
	DeepONet-S	16×16	[4,8,16,32,64]	[20,20,20,20]	[2,256]	16
	DeepONet-L	16×16	[4,8,16,32,64]	[20,20,20,20]	[2,256,256,256]	16

表 2 不同模型在各测试算例上的性能对比

Tab.2 Performance comparison of different models on test examples

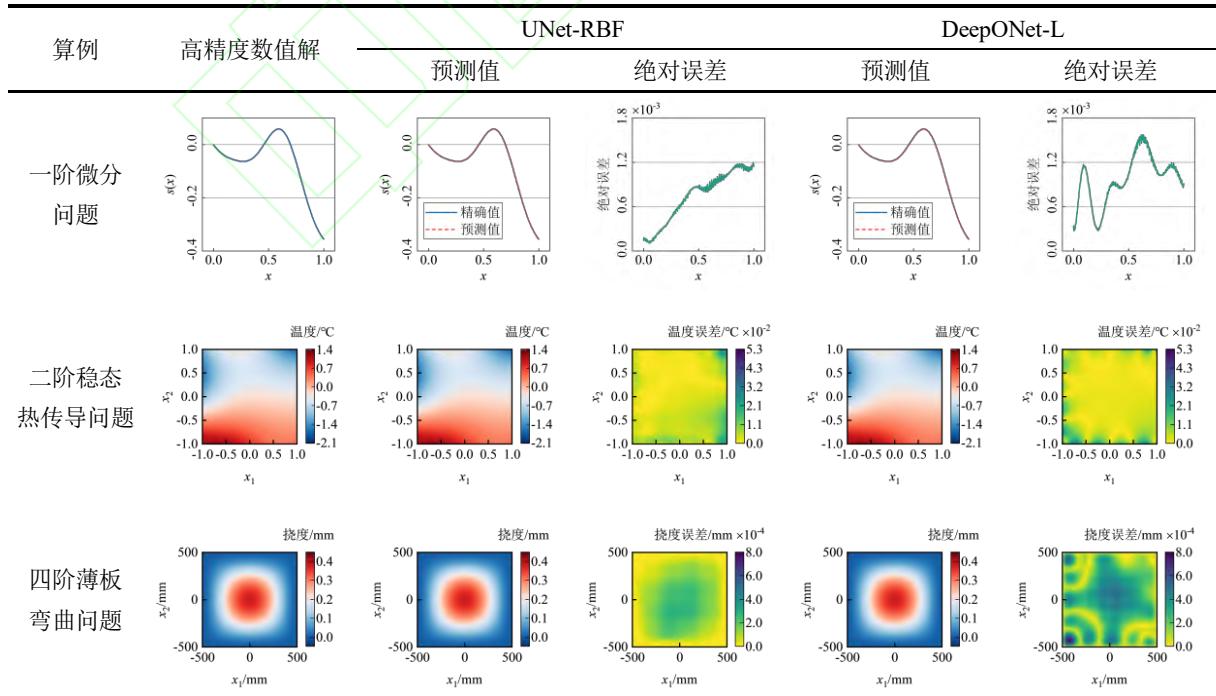
评价指标	算例	UNet-RBF	DeepONet-S	DeepONet-L
相对误差/%	微分问题	0.24%±0.23%	4.93%±4.80%	0.34%±0.27%
	稳态热传导问题	0.93%±0.31%	8.27%±6.42%	0.92%±0.33%
	薄板弯曲问题	0.091%±0.052%	0.72%±0.39%	0.13%±0.09%
训练时间/min	微分问题	12.41	6.14	17.10
	稳态热传导问题	38.18	13.71	67.75
	薄板弯曲问题	122.46	22.43	2003.32

UNet-RBF 模型在三个测试问题中的相对误差标准差相较于 DeepONet-L 模型分别降低了 14.8%、6.1% 和 42.2%，这一结果揭示了 UNet-RBF 模型结构在不同算例和不同输入函数下，其预测性能的波动更小，具有更优的稳定性。UNet-RBF 模型的核心优势则体现在其训练效率上。凭借其冻结参数的 RBF 网络（在少量预训练后）以及预测-微分解耦机制，UNet-RBF 有效避免了在训练过程中对复杂主干网络进行重复的高成本梯度计算。因此，如表 2 所示，在三个测试问题中，UNet-RBF 的训练时间相较于 DeepONet-L 模型分别显著减少了 37.79%、74.44% 和高达 1535.80%。尤其是在四阶薄板弯曲问题中，DeepONet-L 的训练时间超过 2000 分钟，而 UNet-RBF 仅需约 122 分钟，显示出在处理高阶微分问题时，UNet-RBF 在训练效率上的压倒性优势。

表 3 进一步详细展示了 UNet-RBF 模型与性能更优的 DeepONet-L 模型在不同算例中特定样本的预测结果及其对应的逐点绝对误差分布。从表 3 的预测值云图可以看出，两种模型均能以较高精度捕捉 PDEs 数值解的主要特征和分布趋势。对比两种模型的绝对误差云图则可以发现，UNet-RBF 模型的误差分布通常表现出更优的平滑性和稳定性，具体体现在其误差波动幅度相对较小，误差场的非线性程度较低，并且误差在边界与计算域内部之间的过渡也更为自然。相较之下，DeepONet-L 模型的误差分布在某些区域呈现出更为剧烈和集中的变化。这种现象可能归因于其深层 FNN 主干网络在物理信息驱动的训练范式下，更易陷入局部最优解或未能完全收敛，尤其是在高阶微分算子引入复杂梯度信息时。例如，在四阶薄板弯曲问题中，高阶微分运算可能加剧

表 3 UNet-RBF 与 DeepONet-L 模型在各算例中的预测结果及绝对误差分布

Tab.3 Prediction results and absolute error distributions for UNet-RBF and DeepONet-L models on test examples



梯度不均衡问题，增加了深层网络精确寻优的难度，从而可能在一定程度上导致过拟合或欠拟合现象，表现为误差分布的不稳定。

图4展示了各模型在一阶微分问题训练过程中的迭代损失曲线。UNet-RBF模型与DeepONet-L模型均表现出良好的收敛特性，其损失函数值能够持续下降并最终达到较低水平。相比之下，DeepONet-S模型由于其主干网络深度和表达能力受限，难以有效学习复杂的算子映射，其损失函数在早期便停滞在较高水平，未能有效收敛至真实解附近，因此其训练过程基于预设的收敛停滞标准被提前终止。若不考虑提前终止的因素，在相同的迭代次数内，UNet-RBF模型的训练时间消耗通常低于DeepONet-S模型（因其主干网络更简单且部分参数冻结），这进一步从侧面证实了“预测-微分解耦”框架在提升单次迭代效率方面的潜力。通过对UNet-RBF模型与DeepONet-L模型的收敛曲线可以观察到，UNet-RBF模型的损失函数呈现出更为平滑和稳定的下降趋势，而DeepONet-L模型的迭代曲线则表现出相对明显的阶梯式下降特征，即损失函数在某些阶段可能长时间停滞，然后突然下降。此现象可能揭示了采用深层网络作为主干网络时的一个潜在挑战：在物理信息驱动的训练范式中，深层网络可能在复杂的损失曲面上进行优化时，更容易在平坦区域或鞍点附近探索，导致收敛速度阶段性减缓。尽管理论上深层全连接网络具备更强的函数拟合能力，但其实际应用效果可能受到优化算法、初始化、损失函数形态等多重因素的影响。RBF网络因其固有的局部响应和良好插值特性，在某些场求解问题中可能表现出更强的鲁棒性和更易于优化的特性。为缓解深层网络可能面临的收敛问题，通常需要更精细的超参数调整、损失函数权重设计或配点策略优化。

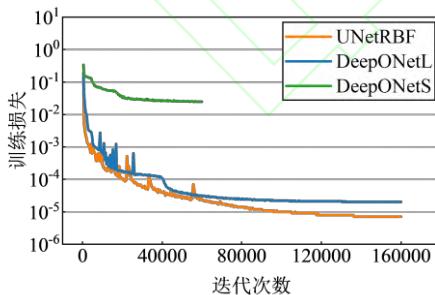


图4 一阶微分问题中不同模型的训练迭代损失曲线

Fig.4 Training loss curves for different models on the first-order differential problem

图5详细对比了不同模型在三个算例的训练过程中，各项主要计算任务（前向传播、偏微分项计算、反向传播与参数更新）所消耗的时间。UNet-RBF模型与DeepONet-L模型在前向传播阶段（即从输入到预测输出）所需时间较为接近，

这符合预期，因为它们的分支网络结构相同。UNet-RBF的核心效率优势体现在偏微分项计算以及反向传播更新参数阶段。由于RBF网络参数的冻结和其结构的简洁性，基于自动微分计算物理残差时，UNet-RBF的计算图远小于DeepONet-L，这使得UNet-RBF在计算偏导数和后续梯度反向传播时所需时间大幅减少。尤其是在高阶微分问题（如四阶薄板弯曲）中，偏微分计算和反向传播的计算复杂度随微分阶次的增加而显著上升（对于传统方法而言接近指数级增长）。在此类问题中，UNet-RBF通过解耦和预算（如果配点固定）有效规避了这一瓶颈，其在这两个环节节省的时间构成了其整体训练效率提升的主要来源。值得注意的是，由于RBF网络参数的冻结，UNet-RBF模型在偏微分项计算时所消耗的时间甚至低于主干网络更简单但仍需全程训练的DeepONet-S模型（在其未提前终止的训练阶段内），再次凸显了解耦策略的有效性。

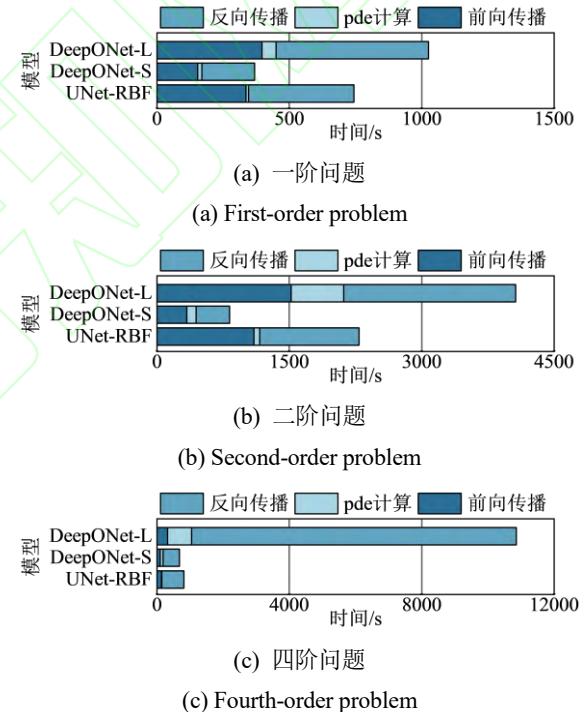


图5 不同模型在各计算任务上的训练时间分解对比

Fig.5 Breakdown of training time for different models across various computational tasks

图6展示了通过调整分支U-Net网络的宽度与深度后，三种模型在一阶微分问题中的预测精度与训练时间之间的关系。结果显示，DeepONet-S模型由于其主干网络表达能力严重受限，即使增加分支网络的复杂度，其预测精度也远逊于其他两种模型，且改善幅度有限。DeepONet-L模型在一定程度上展现出模型复杂度提升带来的精度改善；然而，当模型复杂度增加到一定程度后，其精度提升可能趋于饱和，甚至在某些情况下可

能因陷入局部最优或优化困难而出现精度略微下降或训练时间急剧增加的现象。相较而言, UNet-RBF 模型展现出更优的性能权衡, 其模型精度能够随着分支网络复杂度的提升而稳定提高, 并且较少出现因模型过大导致的训练困难或过拟合现象。在相似的预测精度水平下, UNet-RBF 通常需要更少的训练时间。在整体性能上, UNet-RBF 在所测试的模型复杂度范围内普遍优于 DeepONet-L 模型。结合前述分析, 可以合理推断, 在高阶微分问题中, 由于训练成本的差异被进一步放大, UNet-RBF 模型在训练效率与稳定性方面的优势将更为显著和关键。

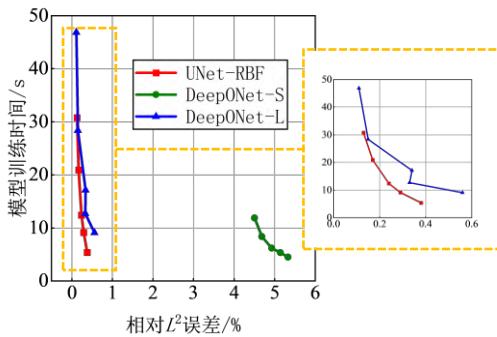


图 6 模型复杂度对预测精度与训练时间影响关系
Fig.6 Influence of model complexity on prediction accuracy and training time

4 结论与展望

本研究针对基于机器学习方法求解高阶偏微分方程 (PDEs) 的需求, 构建了一种新颖的预测-微分解耦式神经网络架构——UNet-RBF。该架构创新性地采用 U-Net 网络作为预测模块, 径向基函数 (RBF) 网络作为微分模块, 并通过在训练过程中冻结 RBF 网络参数的策略, 显著提升了面向高阶 PDEs 求解的神经网络模型的训练效率。通过对两个具有代表性的工程算例进行深入分析, 可以得出以下主要结论:

(1) 模型有效性与精度: 所提出的预测-微分解耦式神经网络架构, 凭借其强大的预测网络与精确的微分网络, 能够较为准确地计算各类偏微分方程的数值解。在所测试的算例中, 模型预测结果与高精度数值解的相对误差达到了 10^{-3} 量级。尤为重要的是, 该模型的训练过程无需额外的标签数据, 充分展现了此方法在工程应用中的实用性与合理性;

(2) 训练效率提升: 通过实现预测网络与微分网络的分离, 并在模型训练过程中冻结微分网络参数, 成功避免了因反复进行微分计算而产生的高昂计算成本。这一设计显著提升了模型在处

理高阶问题时的训练效率。与传统的 DeepONet 方法相比, 本研究提出的 UNet-RBF 模型在一阶、二阶及四阶微分问题中的训练效率分别提升了 37.79%、74.44% 和 1535.80%, 显示出卓越的训练加速效果;

(3) 训练稳定性与鲁棒性: UNet-RBF 模型中的 RBF 网络因其固有的强鲁棒性, 使得模型在处理多种类型的 PDEs 时均表现出更优的训练稳定性, 有助于模型稳定收敛至更优解。具体而言, 在一阶和四阶微分问题中, UNet-RBF 模型的预测精度相较于 DeepONet 模型分别提升了 29.4% 和 30.0%。同时, 其预测结果的相对误差标准差在三个测试问题中分别降低了 14.8%、6.1% 和 42.2%。这些结果充分体现了 UNet-RBF 模型在训练过程中的强大鲁棒性, 使其更适用于求解复杂的物理问题。

当前 UNet-RBF 网络展现了优越的性能, 未来的工作还可包括更全面的消融实验 (ablation studies), 以进一步分析 UNet-RBF 架构中不同组件 (例如采用其他类型的浅层网络代替 RBF 网络等) 对模型整体性能的具体贡献和影响。此外, 当前 UNet-RBF 网络中 RBF 样本点的选取策略为在计算域内进行均匀选取。未来的研究方向可以探索基于 U-Net 网络离散预测结果自适应地分配 RBF 样本点位置的策略。这种自适应采样方法有望进一步提升模型对解的关键区域 (如高梯度区或奇异点附近) 的预测能力与分辨率, 从而期望在更少的训练时间下实现模型预测精度的进一步提升。

参考文献 (References)

- [1] HAFEEZ M B, KRAWCZUK M. A review: applications of the spectral finite element method[J]. Archives of Computational Methods in Engineering, 2023, 30(5): 3453-3465.
- [2] SHI G L, LIU Z S, LU D C, et al. Digital twin - based model updating method for mechanical behaviors analysis of cable truss structure[J]. Journal of Constructional Steel Research, 2024, 221: 108917.
- [3] 任清华. 基于深度学习的偏微分方程求解方法[D]. 山东: 山东大学, 2022.
- [4] RAISSI M, PERDIKARIS P, KARNIADAKIS G E. Physics informed deep learning (part I): data-driven solutions of nonlinear partial

- differential equations[EB/OL]. (2017-11-28)[2025-03-18].
<https://arxiv.org/abs/1711.10561>.
- [5] BELTRÁN-PULIDO A, BILIONIS I, ALIPRANTIS D. Physics-informed neural networks for solving parametric magnetostatic problems[J]. *IEEE Transactions on Energy Conversion*, 2022, 37(4): 2678-2689.
- [6] HANSEN D, MADDIX D C, ALIZADEH S, et al. Learning physical models that can respect conservation laws[J]. *Physica D: Nonlinear Phenomena*, 2024, 457: 133952.
- [7] ZHU Q M, ZHAO Z, YAN J H. Physics-informed machine learning for surrogate modeling of wind pressure and optimization of pressure sensor placement[J]. *Computational Mechanics*, 2023, 71(3): 481-491.
- [8] E W N, YU B. The deep ritz method: a deep learning-based numerical algorithm for solving variational problems[J]. *Communications in Mathematics and Statistics*, 2018, 6(1): 1-12.
- [9] ZANG Y H, BAO G, YE X J, et al. Weak adversarial networks for high-dimensional partial differential equations[J]. *Journal of Computational Physics*, 2020, 411: 109409.
- [10] 李道伦, 沈路航, 查文舒, 等. 基于神经算子与类物理信息神经网络智能求解新进展[J]. *力学学报*: 2024, 56(4): 875-889.
 LI D L, Shen L H, CHA W S, et al. New progress in intelligent solution of neural operators and physics-informed-based methods [J]. *Chinese Jornal of Theoretical and Applied Mechanics*, 2024, 56(4): 875-889. (in Chinese)
- [11] 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.
- [12] LI Z Y, KOVACHKI N, AZIZZADENESHELI K, et al. Fourier neural operator for parametric partial differential equations[C]//Proceedings of the International Conference on Learning Representations, 2021.
- [13] XIONG W, HUANG X M, ZHANG Z Y, et al. Koopman neural operator as a mesh-free solver of non-linear partial differential equations[J]. *Journal of Computational Physics*, 2024, 513: 113194.
- [14] GIN C R, SHEA D E, BRUNTON S L, et al. DeepGreen: deep learning of Green's functions for nonlinear boundary value problems[J]. *Scientific Reports*, 2021, 11(1): 21614.
- [15] WANG S F, WANG H W, PERDIKARIS P. Learning the solution operator of parametric partial differential equations with physics-informed DeepONets[J]. *Science Advances*, 2021, 7(40): eabi8605.
- [16] KHOO Y, LU J F, YING L X. Solving parametric PDE problems with artificial neural networks[J]. *European Journal of Applied Mathematics*, 2021, 32(3): 421-435.
- [17] CHEN T P, CHEN H. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems[J]. *IEEE Transactions on Neural Networks*, 1995, 6(4): 911-917.
- [18] LU L, MENG X H, CAI S Z, et al. A comprehensive and fair comparison of two neural operators (with practical extensions) based on FAIR data[J]. *Computer Methods in Applied Mechanics and Engineering*, 2022, 393: 114778.
- [19] KOUSHKI M, JABBARI E, AHMADINIA M. Evaluating RBF methods for solving PDEs using Padua points distribution[J]. *Alexandria Engineering Journal*, 2020, 59(5): 2999-3018.
- [20] RASMUSSEN C E, WILLIAMS C K I. Gaussian processes for machine learning[M]. 2nd ed. Cambridge, Mass.: MIT Press, 2006.
- [21] HUANG H M, LIN L F, TONG R F, et al. UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation[C]//Proceedings of the 2020 IEEE International Conference on Acoustics, 2020.