

基于梯度增强物理信息神经网络的一维弱噪声随机Burgers方程求解

王 蓉, 刘宏业

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2024年4月21日; 录用日期: 2024年5月23日; 发布日期: 2024年6月30日

摘 要

物理信息神经网络PINNs (Physics-Informed Neural Networks, PINNs)因其良好的可解释性和泛化能力, 被当作求解偏微分方程的一类通用逼近器。传统PINNs通过在损失函数中引入方程残差, 将方程所携带的物理先验信息编译至网络中, 但在求解实际问题时往往需要数量足够大的训练数据集, 其输出才具有较好的网络表现力和求解精度。针对一维弱噪声随机Burgers方程, 提出梯度增强的物理信息神经网络(Gradient-enhanced Physics-Informed Neural Networks, G-PINNs), 采用方程导数的残差作为一种优化的软约束指导网络迭代更新, 理论分析PINNs与G-PINNs两种网络架构寻找方程计算域内最优解的能力。改变数据集的噪声比例和配置点数量, 评估两种网络架构拟合一维Burgers方程解结构的性能差异。数值算例表明, G-PINNs在无噪声、1%、3%和5%弱噪声下求解一维Burgers方程的精度分别提高了约80%、44%、31%、和20%。相较于传统PINNs, 当迭代次数相同时G-PINNs在无噪声与弱噪声的情况下均能在更少的数据集上拟合一维Burgers方程的解结构, 获得精度更高的预测解。

关键词

PINNs, 梯度增强, Burgers方程, 弱噪声, 精度

Solution of One-Dimensional Stochastic Burgers Equation by Gradient-Enhanced Physics-Informed Neural Networks with Weak Noise

Rong Wang, Hongye Liu

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Abstract

Physics-Informed Neural Networks (PINNs) are regarded as a general approximator for solving partial differential equations because of their good interpretability and generalization ability. Traditional PINNs compile the physical prior information carried by the equation into the network by introducing the residual of the equation into the loss function. However, when solving practical problems, it often requires a large number of training datasets to achieve better network performance and solution accuracy. For one-dimensional weak noise random Burgers equation, Gradient-enhanced Physics-Informed Neural Networks (G-PINNs) are proposed. The residual of the derivative of the equation is used as an optimization soft constraint to guide the iterative updating of the network. The ability of PINNs and G-PINNs network architectures to find the optimal solution in the computational domain of the equation is analyzed theoretically. The noise ratio and the number of configuration points of the dataset were changed to evaluate the performance difference of the two network architectures in fitting the solution structure of the one-dimensional Burgers equation. Numerical examples show that the accuracy of G-PINNs in solving one-dimensional Burgers equation with no noise, 1%, 3% and 5% weak noise is improved by about 80%, 44%, 31% and 20%, respectively. Compared with traditional PINNs, G-PINNs can fit the solution structure of one-dimensional Burgers equation on fewer datasets with no noise and weak noise, and obtain a more accurate predictive solution with the same number of iterations.

Keywords

PINNS, Gradient-Enhanced, Burgers Equation, Weak-Noise, Accuracy

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

Burgers 方程是流体力学中描述高速流体运动的一类非线性扩散类型的重要偏微分方程(Partial Differential Equation, PDE), 已广泛应用于流体力学, 空气动力学, 交通流动力学等领域[1]。Burgers 方程传统的数值解法在迭代求解大型方程组时计算成本高昂, 网格的划分精度需要频繁的人机交互进行不断优化[2]。在实际工程应用中, 由于测量设备的限制、环境因素和其他外部干扰, 数据集不可避免的包含一定程度的弱噪声, 导致数值解法在逼近微分方程的导数值时产生不可忽视的偏差, 影响最终的求解精度。

深度学习在工程技术, 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[3]。随着计算机数据分析能力的提高, 神经网络开始应用于 PDE 的研究。作为传统数值求解器的替代方法, 神经网络能够从高维非线性空间中学习到潜在的规则和特征, 拟合得到输入与输出之间的函数映射关系[4]。

韩忠华[5]提出的基于 Kriging 模型的代理优化算法, 可以有效应用于具有光滑、连续设计空间的任意优化问题。孙斌等[6]提出高斯过程回归方法对风速时间序列进行预测, 模型达到较好的稳定性。Roska 等[7]提出卷积神经网络(Convolutional Neural Network, CNN), 将非线性偏微分方程的离散表达式抽象为特定的 CNN 网络模型。Long 等[8] [9]提出偏微分方程网络(Partial Differential Equation-Net, PDE-Net)和偏

微分方程网络 2.0 (Partial Differential Equation-Net2.0, PDE-Net2.0), 从观测的动态数据中发现偏微分方程的先验知识。Regazzoni 等[10]使用人工神经网络(Artificial Neural Network, ANN)对可用输入输出对的误差进行最小化, 以此选择最优模型。

然而上述神经网络大多基于纯数据驱动的深度神经网络模型, 在处理偏微分方程时依赖整个系统的输入输出, 需要大量的数据集用于训练。若数据集携带噪声, 神经网络往往会将训练集中的噪声标签一起学习, 大大降低模型训练过程中的准确性和泛化能力。

为了减轻数据集携带噪声带来的影响, Schaeffer 与 Mccalla [11]采用微分方程的积分形式处理有噪声的数据, 但只能在一定程度上减轻噪声数据的干扰。陈庆强等[12]以数据分布为中心, 根据数据集自身的分布特点对标签噪声进行特殊处理。陈倩等[13]提出一类基于重要性重加权的半监督分类算法, 但仅能用于解决数据集标签含有噪声的二分类问题。2019 年 Raissi 等[14]提出的物理信息神经网络(Physics-Informed Neural Networks, PINNs)在噪声数据下对偏微分方程的正问题和反问题进行求解。然而 PINNs 在解决实际问题时的预测精度难以达到任务要求, 即使增加用于训练的配置点数量, 最终网络的输出精度也有限, 误差值难以收敛到较小量级[15]。唐明健与唐和生[16]利用 PINNs 框架求解矩形薄板中的力学方程, 结果表明 PINNs 模型可以显著减小所需数据的范围, 但是模型最终的精度有限。Cai [17]等结合物理学习、无缝集成数据和数学模型, 验证了 PINNs 在超音速流和生物医学流等相关流体力学逆问题中的有效性。Majid 等[18]运用 PINNs 求解声波波动方程, 将 PINNs 作为地球物理反演求解器, 探索 PINNs 处理不同程度地理结构复杂性的能力, 结果表明 PINNs 在全波形地震反演具备优秀的反演结果。尽管上述研究表明 PINNs 在相关偏微分方程的反演或逆问题的应用中表现出了较高的性能, 但目前尚未有研究对 PINNs 正向求解相关偏微分方程的高精度数值结果进行探讨分析。

本文提出基于 PINNs 的梯度增强的物理信息神经网络(Gradient-enhanced Physics-Informed Neural Networks, G-PINNs), 引入方程导数的残差并在训练过程中不断学习方程及其导数形式, 同时探究 G-PINNs 各项损失的权重配比, 增强对偏微分方程梯度的约束以提高预测结果的精度并优化网络架构。采用 PINNs 与 G-PINNs 模型对一维 Burgers 方程的数值结果进行正向预测求解, 对比 PINNs 和 G-PINNs 在无噪声与弱噪声下求解一维 Burgers 方程的相对误差, 理论分析其网络架构在噪声环境下稳健性和求解性能的提升, 同时为数据集稀缺情况下偏微分方程的求解提供了理论参考和实验依据。

2. Burgers 方程

Burgers 方程又称粘性 Burgers 方程, 展示了流体运动中对流项和扩散项的耦合, 是流体力学领域中的基础非线性偏微分方程之一。选取单一空间维度的 Burgers 方程, 其一般形式如下[19]:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} - \mu \frac{\partial^2 u}{\partial x^2} = 0 \quad (1)$$

式中, u 表示速度, μ 表示介质的粘度, t 表示时间, x 表示空间坐标, $\frac{\partial u}{\partial t}$ 为非稳态项, $u \frac{\partial u}{\partial x}$ 为对流项,

$\mu \frac{\partial^2 u}{\partial x^2}$ 为扩散项。

本文中选定计算域 $\Omega: -1 \leq x \leq 1, 0 \leq t \leq 1$, 使用的 Burgers 方程为:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} - \left(\frac{0.01}{\pi}\right) \frac{\partial^2 u}{\partial x^2} = 0 \quad (2)$$

初始条件为:

$$u(x, 0) = -\sin(\pi x), -1 \leq x \leq 1 \quad (3)$$

选择 Dirichlet 边界条件

$$u(-1,t) = u(1,t) = 0, 0 \leq t \leq 1 \quad (4)$$

20 世纪 50 年代, Hopf [20] 提出了一种精确求解 Burgers 方程的方法, 即 Hopf-Cole 变换, 该方法允许在任意初始条件下获得 Burgers 方程的解。在 $x \in [-1,1]$ 内等间隔取 256 个点, $t \in [0,1]$ 内等间隔取 100 个点, 构成 25,600 个时空坐标点作为空间离散化的基准。利用 Hopf-Cole 变换和有限元法得到高精度的数值解, 该数值解被视作本次实验的参考真值。为保证训练集的多样性与代表性, 从坐标点随机选取配置点构成网络训练所用的数据集, 确保神经网络模型能够对广泛的输入条件进行准确预测。

3. 算法比较

本小节将 PINNs 与 G-PINNs 进行详细阐述并比较, 对 PINNs 和 G-PINNs 的原理从数学角度进行理论性的推导分析, 侧重探讨两种网络架构的损失函数以及改进前后对方程形式的约束, 研究 G-PINNs 引入的改进机制的有效性, 更全面的证明 G-PINNs 架构相较于原 PINNs 性能提升这一优势。

3.1. PINNs

PINNs 的基本思想是将经过自动微分计算后的 PDE 残余项作为一种软约束嵌入到特定问题的损失函数中, 利用神经网络来近似训练出方程的解, 并在此过程中保持对物理方程的满足[21]。PINNs 引入方程的约束可以极大降低假设空间的复杂度, 以便模型在学习数据的同时也保持对物理定律的遵循, 因此 PINNs 的网络输出能够较好的近似表示 PDE 的解结构[22]。

以 Burgers 方程为例, 定义如下函数:

$$f = u_t + uu_x - \frac{0.01}{\pi} u_{xx} \quad (5)$$

式中: $u_t = \frac{\partial u}{\partial t}$, $u_x = \frac{\partial u}{\partial x}$, $u_{xx} = \frac{\partial^2 u}{\partial x^2}$, u 表示速度, t 表示时间, x 表示空间坐标。

PINNs 训练过程中通过最小化均方误差(Mean Squared Error, MSE)损失函数来逐步调整参数, 使预测值与真实值之间的差异最小化。在 PINNs 中, 损失函数是一个多项式组合, 包括数据拟合项和物理约束项。数据拟合项衡量了网络预测值与真实标签值之间的差异, 而物理约束项则强制要求网络输出满足给定的物理方程。定义 PINNs 的损失函数如下:

$$MSE = MSE_u + MSE_f \quad (6)$$

式中: MSE_u 为数据拟合损失, 衡量了模型对预测解和真值之间的拟合程度; MSE_f 为方程的残差, 用来确保模型满足所描述的物理方程。数据拟合损失 MSE_u 是预测值与真值间的均方误差, 定义如下:

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

式中: N_u 为用来训练预测值与真值之间均方误差的配置点个数, $u^{DNN}(x_u^i)$ 为 PINNs 在对应坐标点深度神经网络的预测值, u^i 为该坐标点对应的真值。

PINNs 计算出预测值对应的导数, 当其满足方程时, 式(5)应等于 0, 由此计算方程的残差来量化模型是否满足方程所描述的物理定律。方程的残差:

$$MSE_f = \frac{1}{N_f} \sum_{i=1}^{N_f} |f^{PINNs}(x_u^i) - 0|^2 \quad (8)$$

式中: N_f 为用于计算方程残差的配置点个数, $f^{PINNs}(x_u^i)$ 为 PINNs 将对应预测值及其导数表达式代入式(5)之后的函数值。

PINNs 在训练过程中计算两种损失, 共同指导神经网络参数的迭代更新, 但是并非所有的训练点均需计算两种损失函数。训练集中的数据点分为两类, 第一类是初边值点计算 MSE_u 和 MSE_f 两种损失; 第二类是内部的点, 这类点只受方程约束, 只需计算方程的残差 MSE_f 。方程的残差 MSE_f 旨在衡量神经网络的预测解是否满足给定的偏微分方程, 确保求解结果符合方程的物理规律, 增强预测解的物理一致性。即使在缺乏真实解的情况下, PINNs 在训练过程中仍可以最小化方程的残差来调整网络参数, 尽可能的让 f 趋近于 0, 实现非监督学习的策略。

3.2. G-PINNs

PINNs 算法对方程的求解形式与解析解非常相似, 但是在实际训练过程中最终的逼近精度往往难以达到较高水准。此外, 从数据集中抽样选取的配置点数量较为庞大, 真实的流体力学工程涉及到众多变量和条件, 需要通过精密的测量和实时监测才能获得准确的数据, 从中筛选得到满足训练条件的可靠数据集更是具有挑战性。针对此局限性, 需要在现有 PINNs 模型的基础上进行改进, 使得改进后的 PINNs 可以在更少配置点训练的基础上得到更加精确的解, 以确保模型在实际复杂场景中表现出色。

PINNs 算法引入 PDE 的残差作为训练过程中的软约束, 在此基础上可以推论出既然 PDE 的残差指导网络训练的效果良好, 那么 PDE 导数的残差同样可以对训练过程进行有效约束。如果 PDE 的残差为零, 显然 PDE 残差的梯度也应为零。改进后的网络 G-PINNs 强制 PDE 导数的残差为零, 在网络收敛的过程中加强梯度的约束, 增强网络对 PDE 导数信息的学习能力, 从而在训练过程中实现对方程形式的更强约束。增强梯度的限制能够有效提升网络架构在训练过程中对方程形式的约束能力, 使其能够更充分地学习梯度信息和物理先验知识, 提升模型在预测方程解时的逼近能力以及对复杂物理系统的适应性, 降低流体力学工程环境下获取精确数据集的难度, 并提高数值模拟的效率和可靠性。

以 Burgers 方程为例, 对 f 进行求导后得到如下函数:

$$f_t = u_{tt} + u_t * u_x + uu_{xt} - \frac{0.01}{\pi} u_{xxt} \tag{9}$$

$$f_x = u_{tx} + u_x^2 + uu_{xx} - \frac{0.01}{\pi} u_{xxx} \tag{10}$$

式中 $f_t = \frac{\partial f}{\partial t}$ 为 f 对 t 求导后的导数表达式, $f_x = \frac{\partial f}{\partial x}$ 为 f 对 x 求导后的导数表达式。

G-PINNs 在训练过程中不仅约束式(5)为 0, 同时约束式(5)的导数也为 0。迭代优化过程中持续进行梯度更新, 寻求最接近真值且同时满足方程及其导数的预测解, 提高模型对物理系统行为的建模能力。G-PINNs 方程导数的残差计算沿用 PINNs 中的均方误差函数, 引入方程导数的残差后, G-PINNs 的损失函数也随之做出相应的调整:

$$MSE = w_u MSE_u + w_f MSE_f + w_{ft} MSE_{ft} + w_{fx} MSE_{fx} \tag{11}$$

式中: w 为不同损失的权重配比, MSE_u 是预测值与真值间的均方误差, MSE_f 为方程的残差, MSE_{ft} 为方程对 t 求导后导数的残差, MSE_{fx} 为方程对 x 求导后导数的残差。方程导数残差的表达式如下:

$$\begin{cases} MSE_{ft} = \frac{1}{N_{ft}} \sum_{i=1}^{N_{ft}} |f_t(x_{ft}^i, t_{ft}^i) - 0|^2 \\ MSE_{fx} = \frac{1}{N_{fx}} \sum_{i=1}^{N_{fx}} |f_x(x_{fx}^i, t_{fx}^i) - 0|^2 \end{cases} \tag{12}$$

式中: $f(x_{f_t}^i, t_{f_t}^i)$ 为 G-PINNs 在坐标点 (x, t) 将对应预测值及其导数表达式代入式(9)后的函数值, $f(x_{f_x}^i, t_{f_x}^i)$ 为 G-PINNs 在坐标点 (x, t) 将对应预测值及其导数表达式代入式(10)后的函数值, N_{f_t} 为用于计算 $f(x_{f_t}^i, t_{f_t}^i)$ 对 t 求导后导数残差的配置点个数, N_{f_x} 为用于计算 $f(x_{f_x}^i, t_{f_x}^i)$ 对 x 求导后导数残差的配置点个数。

图 1 所示为求解 Burgers 方程时 G-PINNs 模型示意图, G-PINNs 包含神经网络和梯度增强两部分, 训练过程中不断迭代优化以更好的满足各项损失, 从而使模型收敛到更为精确的预测解。

数据预处理阶段将位移 x 和时间 t 构成的 25600 个时空坐标点 (x, t) 和每个点对应的真值构成数据集, 采用拉丁超立方抽样(Latin Hypercube Sampling, LHS)对坐标点进行随机选取, LHS 对参数空间进行等概率划分, 能够很好的覆盖到每个参数空间, 进而可以通过较少次迭代产生更加准确的概率估计。LHS 选取的坐标点包含初边值点和内部点, 构成 G-PINNs 的训练样本集。

DNN 网络输入是坐标点 (x, t) , 经由全连接层组成的隐藏层, 对输入数据进行非线性变换, 学习复杂的特征和表示。隐藏层应用双曲正切(Tanh)激活函数, 引入非线性和稀疏激活性, 学习更加高级的非线性映射关系及有效的抽象特征, 提高网络的表达能力。Tanh 函数对输入的微小变化更为敏感, 提升了模型的抗干扰能力, 能够更有效地传播梯度并促进权重的更新。

G-PINNs 将 DNN 输出的预测解 u_{DNN} 作为网络输入, 利用自动微分技术对其进行求导, 精确计算输入变量的导数, 使得方程及其导数项能够在区域内的配置点上进行准确计算, 由此得到方程的残差和方程导数的残差[23]。综合考虑预测解与真值间的误差, 为各项损失分配相应的权重, 构成最终的损失函数。G-PINNs 模型强调对方程导数的准确建模, 因此训练过程中优化算法的全局收敛性至关重要。

本文选取机器学习算法中的“L-BFGS”优化算法对损失函数进行优化。L-BFGS 优化算法在每一步迭代中, 计算目标函数的梯度, 利用近似 Hessian 矩阵的逆对梯度进行修正, 而不需要显式地存储整个矩阵, 可以极大的提升内存效率, 在处理高维参数空间和大规模数据时具有良好的收敛性和计算效率, 能够更有效地在搜索空间内找到物理约束下的全局最优解。结合 G-PINNs 模型求解 Burgers 方程过程中数据集规模及非线性目标函数, 综合考虑后选择该算法。图 1 中 ϵ 是允许的最小误差, $n.Iter$ 为迭代的最大步骤, 两者作为循环的条件, 神经网络不断迭代更新, 直至满足 ϵ 或达到 $n.Iter$ 的设定条件之一, 跳出循环结束训练, 输出最终的预测解 $u(x^i, t^i)$ 。

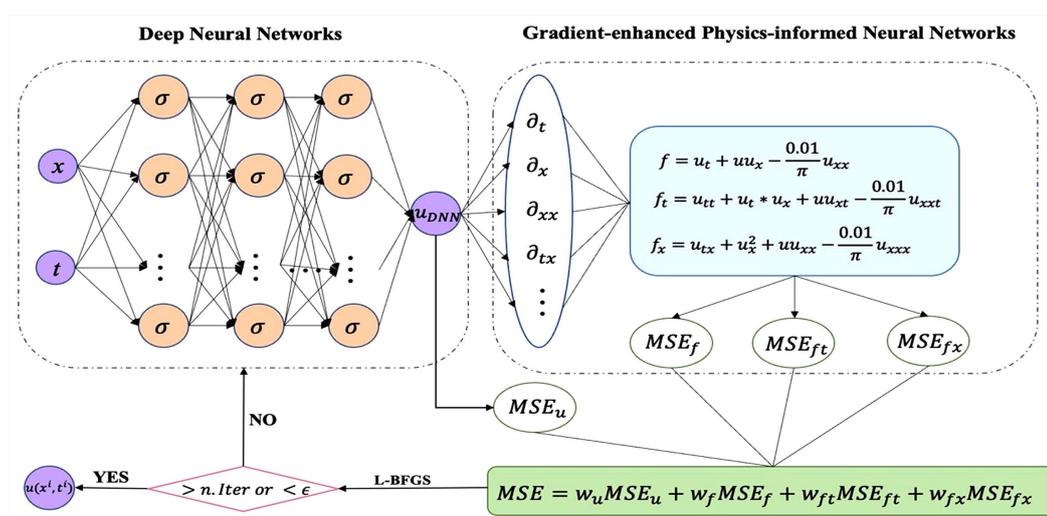


Figure 1. Schematic diagram of G-PINNs model

图 1. G-PINNs 模型示意图

G-PINNs 不仅将物理信息和数学方程嵌入到损失函数, 更在优化过程中强调对方程导数的约束, 不断的训练求解, 获得神经网络的最佳参数和近似描述该力学系统的通用逼近算法。数据集的内部点不被真值约束, 即没有标签, 只需计算方程及其导数的残差, 降低了对数据集的需求, 有效节约数据资源和实际应用时采集数据的成本。引入梯度增强后的网络能提高最终预测解的精度和可解释性, 在少量训练配置点的基础上也能取得更为显著的训练效果。结合实际流体力学场景中数据资源有限且存在噪声的情况, G-PINNs 模型提供了应对复杂系统行为更优秀的建模能力。

4. 数值实验

本节研究针对有噪声与无噪声两种情景, 对比分析 PINNs 和 G-PINNs 两种网络架构在相同网络配置下求解 Burgers 方程的效果。实验使用 Inter(R) Core(TM) i9-10980XE(CPU)和 NVIDIA T600(GPU)作为测试平台, 基于 Python3.10 和 Pytorch1.12.1 搭建实验环境。PINNs 与 G-PINNs 架构均构建 6 层隐藏层, 每层 20 个神经元的深度神经网络, 网络的初始学习率设置为 1.0, 选用 L-BFGS 优化器进行迭代优化。

G-PINNs 模型的训练是一个多任务学习, 其损失函数为网络训练过程中多项损失的复合函数[24]。如果将不同学习任务的损失项直接相加可能会导致训练过程被其中某个任务所主导, 因此需要对 G-PINNs 损失之间的权重配比进行讨论分析。此处主要讨论方程导数残差的权重配比, 即 w_{f_t} 和 w_{f_x} (本次实验中 $w_{f_t} = w_{f_x} = w$)。评估各项损失权重的过程中使用神经网络输出的预测解与真值之间的 L2 范数相对误差 δ 来衡量模型性能。定义相对误差 δ 如下:

$$\delta = \frac{\sqrt{\sum_{i=1}^N (u(x^i, t^i) - u^i)^2}}{\sqrt{\sum_{i=1}^N u^i^2}} \quad (13)$$

式中: $u(x^i, t^i)$ 为网络模型在坐标点 (x, t) 对应解的预测值, u^i 为坐标点 (x, t) 对应的真值。

图 2 展示了不同 w 取值下, G-PINNs 经过 3000、6000、9000 和 12,000 次迭代后获得的预测解的相对误差对比情况。由图 2 可以观察到在 $w = 1.0$ 与 $w = 0.01$ 时, G-PINNs 相对误差较大; $w = 0.1$ 与 $w = 0.001$ 时, 相对误差远小于前两者, 且可以明显观察到 $w = 0.001$ 时整个训练过程相较于 $w = 0.1$ 时的误差降低速度更快, 最终稳定在一个较小范围附近。故实验将 G-PINNs 模型 PDE 导数残差的静态权重设置为 0.001, 此权重下 G-PINNs 模型在训练过程中能够合理考虑方程导数残差的影响, 达到最佳的预测效果。

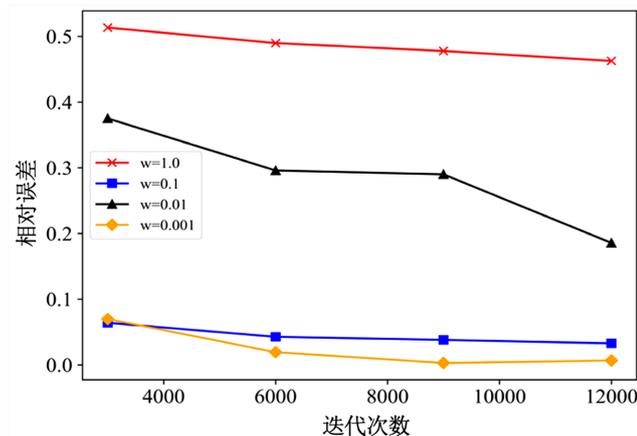


Figure 2. Relative error of G-PINNs when PDE derivative residuals weight ratio is different

图 2. PDE 导数残差的权重配比不同时 G-PINNs 的相对误差

4.1. 无噪声下 PINNs 与 G-PINNs 对 Burgers 方程的求解结果

PINNs 和 G-PINNs 均在计算域内采用 LHS 选取 100 个边界点与 10,000 个内部点进行训练, 将两种网络架构的预测解以及最终误差在整个计算域内进行对比分析。为了细化结果, 实验选取 $t = 0.5$ 时刻 PINNs 和 G-PINNs 的预测解及误差进行单独的分析, 由此涵盖整个计算域和特定时刻两方面的综合比较。这一分析方法旨在全面评估两种网络架构的性能表现以及在时间维度上的稳定性和精确性, 为后续实验结果的解释提供有力支持。实验选取绝对误差作为衡量模型性能的评估指标。绝对误差直接反映模型的预测值与真实值之间的距离, 对于数据集中可能存在的极端值具有较强的稳健性, 在真实的复杂环境中具有较强的适用性, 可以提供清晰、客观的评估结果。绝对误差 $\Delta_{\text{error of } u}$ 计算如下:

$$\Delta_{\text{error of } u} = |u(x^i, t^i) - u^i| \quad (14)$$

式中: $u(x^i, t^i)$ 为网络模型在坐标点 (x, t) 对应解的预测值, u^i 为坐标点 (x, t) 对应的真值。

图 3 为 PINNs 训练 10,000 次后对 Burgers 方程计算域内各坐标点的真值、预测解及绝对误差值, 图 4 为 G-PINNs 训练 10,000 次后对 Burgers 方程计算域内各坐标点的真值、预测解及绝对误差值。如图 3 所示, PINNs 求解 Burgers 方程最终预测解与真值间的绝对误差的峰值约在 6×10^{-3} 左右, 而图 4 中 G-PINNs 最终绝对误差的峰值约 1.4×10^{-3} 左右。将全域内预测值与真实值之间进行误差归一化后, G-PINNs 对方程解的预测精度相较于 PINNs 提升了近 76.7%, 在处理 Burgers 方程时显现出更强的精度和预测能力。

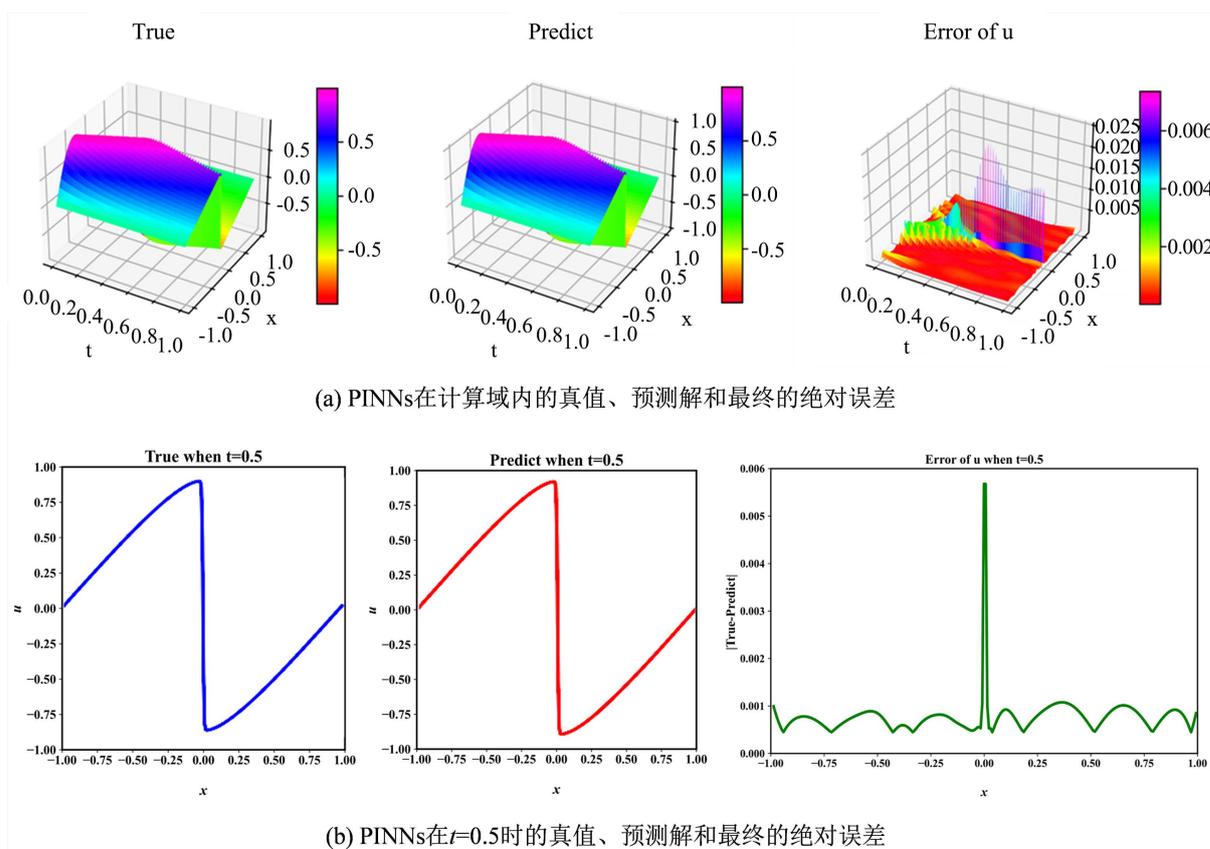


Figure 3. Training results of PINNs

图 3. PINNs 的训练结果图

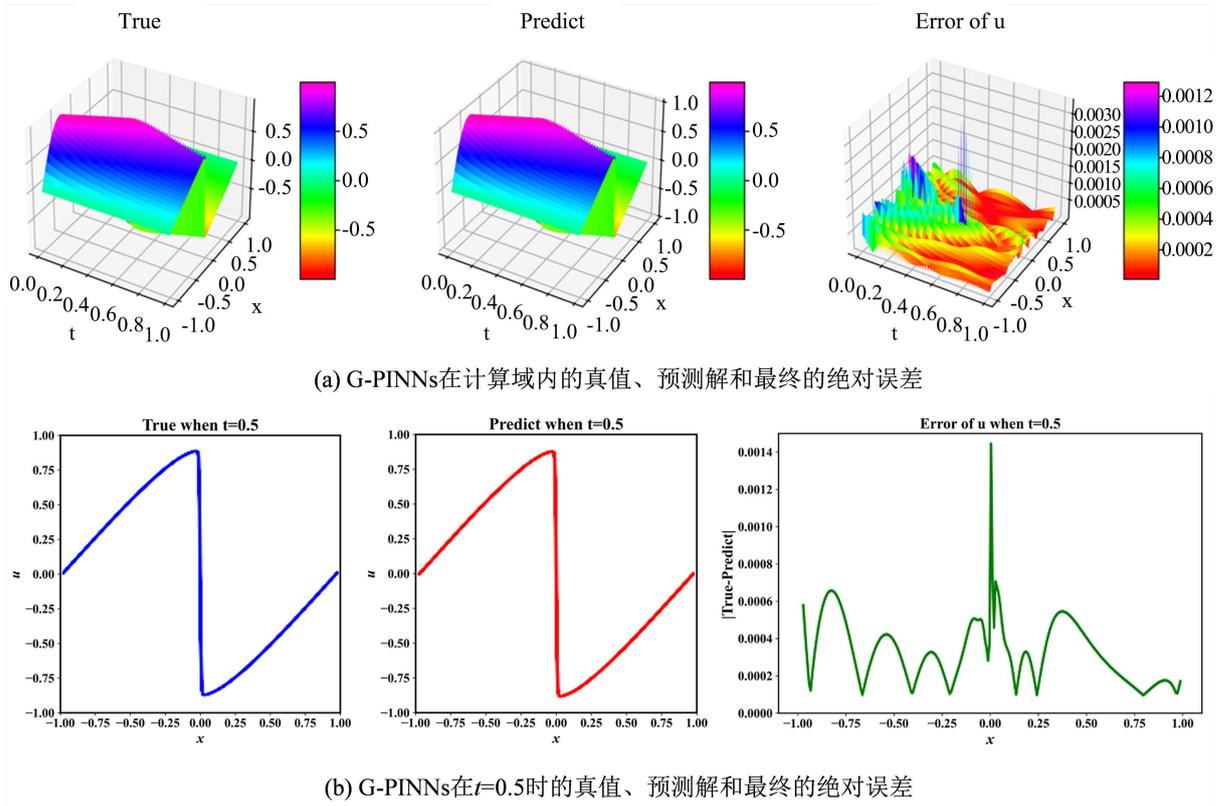


Figure 4. Training results of G-PINNs
图 4. G-PINNs 的训练结果图

为全面评估 PINNs 与 G-PINNs 在不同训练集规模下的表现, 实验选择在不同数据配置条件下比较二者的性能。表 1 列出了不同数据配置条件下 PINNs 与 G-PINNs 迭代 3000、6000、9000 次后的绝对误差。通过考察模型在不同数据情况下的性能表现, 可以评估模型对于数据量变化的鲁棒性及泛化能力。

Table 1. Absolute error of PINNs and G-PINNs when network configuration points are different
表 1. 网络配置点不同时 PINNs 与 G-PINNs 训练的绝对误差值

	3000 (Iters)	6000 (Iters)	9000 (Iters)
Error of PINNs ($N_u = 100, N_f = 5000$)	8.917022×10^{-2}	5.132573×10^{-2}	1.855752×10^{-2}
Error of G-PINNs ($N_u = 100, N_f = 5000$)	3.272040×10^{-2}	2.645900×10^{-2}	7.259635×10^{-3}
Error of PINNs ($N_u = 100, N_f = 10,000$)	7.268971×10^{-2}	3.792699×10^{-2}	6.582435×10^{-3}
Error of G-PINNs ($N_u = 100, N_f = 10,000$)	6.970254×10^{-2}	1.925936×10^{-2}	1.265518×10^{-3}

表 1 的实验结果表明 G-PINNs 在相同数据配置点情况下的绝对误差相较于 PINNs 降低了约 36%~63%。进一步减少一半配置点后, PINNs 与 G-PINNs 的训练误差均呈现出一定程度的增加。然而 G-PINNs 在使用一半配置点训练时的误差比 PINNs 不减少配置点时的误差更小, 该结果表明 PINNs 模型在处理 PDE 时其性能在很大程度上依赖于物理信息的准确性和完整性, 在数据集数量减少的情况下

PINNs 的模型性能也随之降低并且求解精度存在局限性。G-PINNs 通过增强对方程梯度的约束提升了模型物理拟合的能力以及方程建模的准确性, 其表现力显著优于 PINNs 模型。对于数据集数量有限或者方程的解难以求得的复杂流体扩散场景, G-PINNs 能够以相对较少的配置点实现更为精准的预测, 表现出更佳的网络逼近能力。在资源有限的情况下, G-PINNs 权衡时间效益与预测精度, 为工程领域的问题求解提供了一种高效、精确的选择。

通过考察 PINNs 与 G-PINNs 模型在不同数据情况下的性能表现, 可以评估二者对于数据量变化的鲁棒性及泛化能力。此外, 不同的数据配置条件涉及到实际应用中不同的数据采集成本、时间成本和人力资源投入, 此比较研究有助于做出合理的资源分配和成本考量, 为实际应用提供了更可靠的求解策略以及更全面的评估框架。

4.2. 弱噪声下 PINNs 与 G-PINNs 对 Burgers 方程的求解结果

实验对 PINNs 与 G-PINNs 添加 5% 服从标准正态分布的噪声, 模拟实际应用中各种随机因素对数据产生的噪声干扰, 如测量误差、环境干扰等, 使实验结果更贴近真实情况[25]。在 $[-1,1] \times [0,1]$ 的计算域内采用 LHS 选取 100 个边界点与 10,000 个内部点训练至收敛, 在整个计算域内对 PINNs 与 G-PINNs 的预测结果和绝对误差进行比较分析。通过对比两种网络架构在实际弱噪声干扰情况下的表现, 评估噪声引入后模型的稳健性及其在噪声环境下的性能[26]。为了更详细的了解结果, 仍选取 $t = 0.5$ 时刻 PINNs 和 G-PINNs 的预测结果和误差值进行单独的展示分析, 得到的 PINNs 和 G-PINNs 的训练结果如图 5 和图 6 所示。

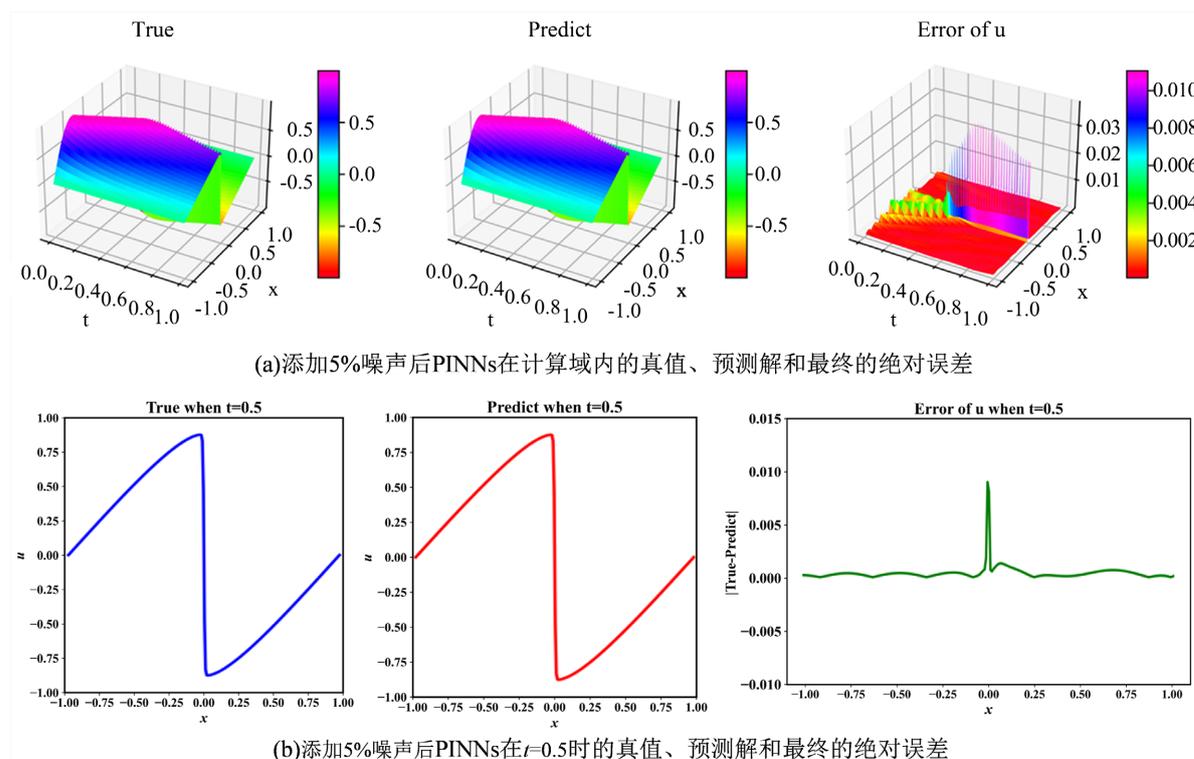


Figure 5. Training results of PINNs with 5% noise

图 5. 添加 5% 弱噪声后 PINNs 的训练结果图

结果显示, 与无噪声训练相比, PINNs 和 G-PINNs 模型在包含噪声的训练条件下最终预测解误差均

呈现出一定程度的增加, 该现象反映了噪声对模型预测精度的不利影响。但是 G-PINNs 模型在相同条件下的误差相对于 PINNs 模型在全域范围内以及 $t = 0.5$ 时刻均显著降低了近 20%。这一结果表明, 在 5% 噪声率的情况下 G-PINNs 模型在处理噪声数据方面具有更强的鲁棒性, 能够更有效地抑制噪声带来的不利影响, 从而在噪声存在的条件下提供更为准确的预测结果。

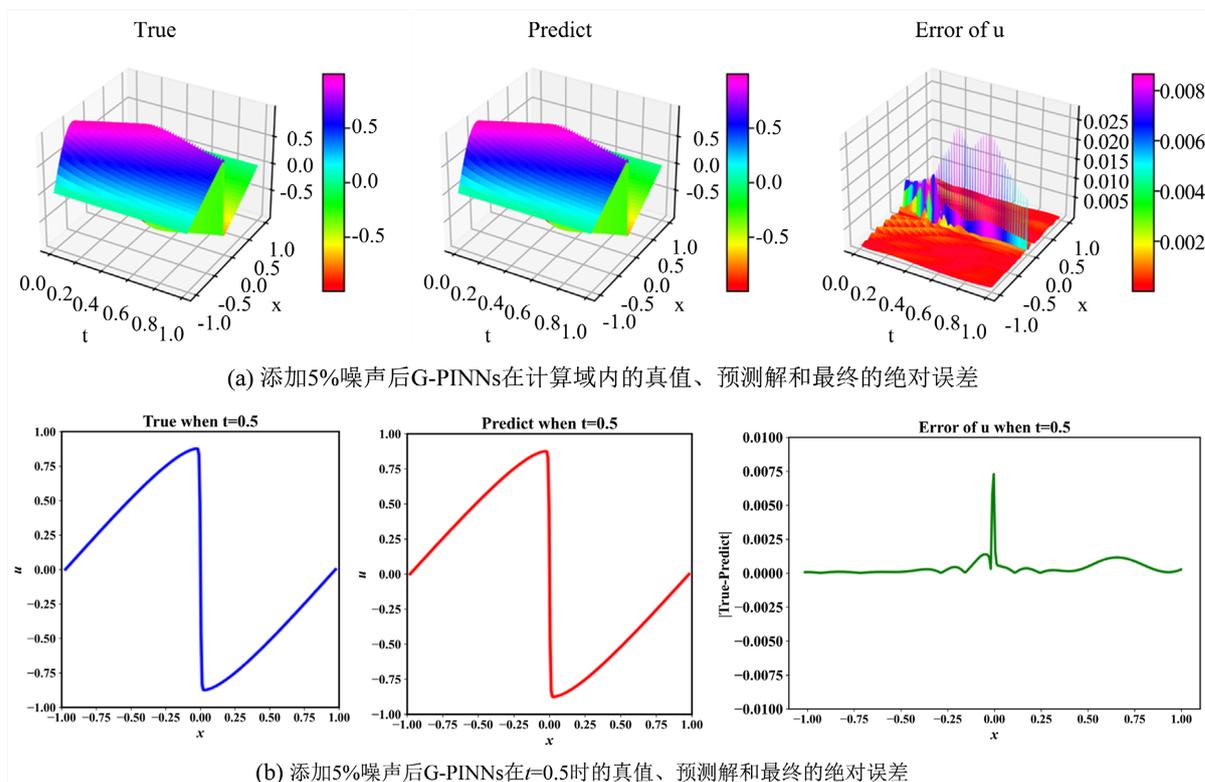


Figure 6. Training results of G-PINNs with 5% noise

图 6. 添加 5% 弱噪声后 G-PINNs 的训练结果图

表 2 列出了添加 5% 弱噪声后, PINNs 与 G-PINNs 在不同数据配置条件下迭代 3000、6000、9000 次后绝对误差值。表 2 中的误差结果表明, 在受到弱噪声干扰的情况下, G-PINNs 模型寻找最优解的能力仍然优于 PINNs 模型。由表 2 数据可看出 PINNs 处理噪声数据时表现出较低的稳定性和准确度, 而 G-PINNs 引入方程导数的残差增强了模型对物理过程的拟合精度, 提高了模型在面对噪声数据时的鲁棒性。

Table 2. Absolute error of PINNs and G-PINNs when network configuration points are different (with 5% noise)

表 2. 网络配置点不同时 PINNs 与 G-PINNs 训练的绝对误差值(5%噪声)

	3000 (Iters)	6000 (Iters)	9000 (Iters)
Error of PINNs ($N_u = 100, N_f = 5000$)	1.869648×10^{-1}	5.516061×10^{-2}	4.288181×10^{-2}
Error of G-PINNs ($N_u = 100, N_f = 5000$)	5.468906×10^{-2}	4.320308×10^{-2}	3.192474×10^{-2}
Error of PINNs ($N_u = 100, N_f = 10,000$)	1.015703×10^{-1}	4.789887×10^{-2}	3.911240×10^{-2}
Error of G-PINNs ($N_u = 100, N_f = 10,000$)	3.728751×10^{-2}	2.218031×10^{-2}	1.591892×10^{-2}

实际流体力学场景复杂, 难以获取完备且充分的数据集, 数据集的准确度及质量也较难达到高水平。综合考虑数据集引入噪声后 G-PINNs 的表现可以得出, G-PINNs 投入更多的计算成本换取模型在不确定性噪声环境下的稳健性和卓越性能, 其模型对数据噪声的容忍度高于 PINNs。G-PINNs 模型的优良表现可大幅减少对数据集的依赖, 降低实际数据采集过程中的成本及难度, 为复杂流体力学场景的应用提供强有力的理论支持, 而且在真实应用环境中也具有显著的求解潜力。

本实验中对数据集分别引入 1%、3% 和 5% 服从标准正态分布的噪声, 模拟 PINNs 与 G-PINNs 在不同噪声情况下的逼近能力。计算域内选取 100 个边界点与 10,000 个内部点进行弱噪声下的实验, 并记录网络输出的预测解的误差值, 结果如图 7 所示。

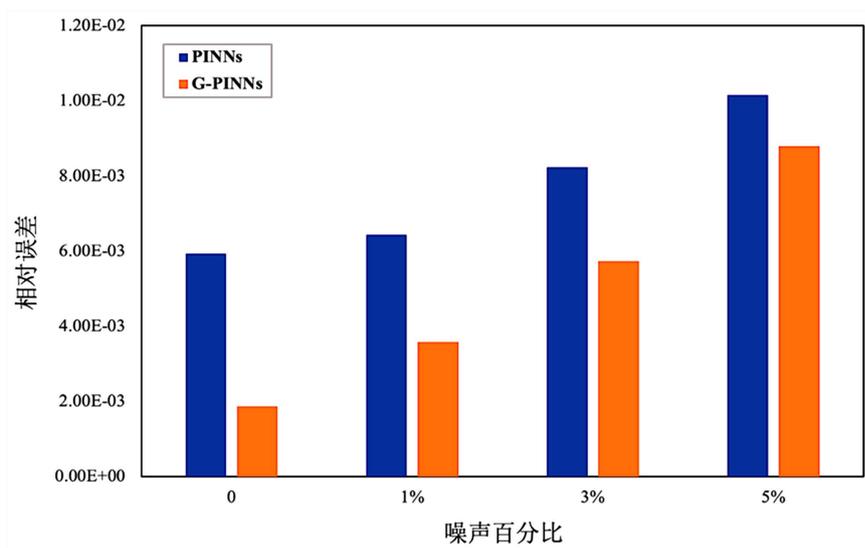


Figure 7. Error of PINNs and G-PINNs with different noises
图 7. 不同噪声下 PINNs 与 G-PINNs 的误差

针对 Burgers 方程数值求解的过程, 以跳出循环后最终的相对误差值作为评估 PINNs 与 G-PINNs 两种网络架构表现力的依据。由图 7 可观察到在无噪声情况下 PINNs 与 G-PINNs 最终的相对误差存在显著差异, 表明两者在求解精度上具有较大的性能差距。当引入数据噪声后, PINNs 和 G-PINNs 预测解的误差均随着噪声强度的增加而增加, 但在弱噪声环境下 G-PINNs 的误差始终保持在较低水平且明显优于 PINNs。G-PINNs 通过引入导数的残差加强对梯度的约束, 使得网络在训练过程中不仅关注函数值的准确性, 还考虑到导数信息的准确性, 可以有效减少数值解在导数上的震荡, 从而提高整体解的平滑度和精确度。传统 PINNs 缺乏对梯度的强约束从而导致求解精度的局限性, 而 G-PINNs 则能够更好地捕捉解的物理特性, 为数值求解偏微分方程提供了一种更为有效和鲁棒的方法。

然而对于神经网络来说, 在算力和算法都基本已经确定的前提下, 数据的数量与质量是决定深度学习任务最终效果的关键因素。数据集噪声过多直接降低了数据的质量, 对网络训练不利, 最终的效果也随之不理想。当数据集中的噪声增大到一定程度时, 两种网络结构的训练效果均变得不理想, 最终误差都显著增加, 性能差距也逐渐减小。因此, 在数据集中噪声过多的情况下, 网络能力的对比不再具有明显的实质性意义。

5. 结论

本文基于 PINNs 架构做出了优化, 将梯度增强后的网络用于正向求解流体力学中的一维 Burgers 方

程, 阐述了 G-PINNs 的损失函数和神经网络的训练参数, 并对比分析两种模型寻找最优解的逼近能力。G-PINNs 引入导数残差的概念, 融合物理系统和神经网络的计算模型, 充分利用方程的物理特性实现高精度的计算。考虑到神经网络求解偏微分方程的效果受到数据集数量和质量的影响, 实验中将配置点减半模拟数据量的减小, 在数据集中加入噪声模拟数据质量的降低。基于数据集的数量与质量两方面说明 G-PINNs 的求解精度相较于传统 PINNs 得到了显著提升, 且 G-PINNs 在面对噪声和干扰时表现出良好的抗干扰性, 整体性能明显优于传统 PINNs。G-PINNs 引入导数残差并结合数据拟合和物理拟合共同指导模型训练, 由此实现半监督学习的策略。未来考虑将 G-PINNs 与自监督学习、迁移学习等相结合, 探索混合方法在偏微分方程求解中的优势和应用场景, 进一步提高模型的泛化能力和适应性。

参考文献

- [1] Alwehebi, F., Hobiny, A. and Maturi, D. (2023) Variational Iteration Method for Solving Time Fractional Burgers Equation Using Maple. *Applied Mathematics*, **14**, 336-348. <https://doi.org/10.4236/am.2023.145021>
- [2] Ohkitani, K. and Dowker, M. (2012) Numerical Study on Comparison of Navier-Stokes and Burgers Equations. *Physics of Fluids*, **24**, Article ID: 055113. <https://doi.org/10.1063/1.4719787>
- [3] Sirignano, J. and Spiliopoulos, K. (2018) DGM: A Deep Learning Algorithm for Solving Partial Differential Equations. *Journal of Computational Physics*, **375**, 1339-1364. <https://doi.org/10.1016/j.jcp.2018.08.029>
- [4] Xu, H., Zhang, D. and Zeng, J. (2021) Deep-Learning of Parametric Partial Differential Equations from Sparse and Noisy Data. *Physics of Fluids*, **33**, Article ID: 037132. <https://doi.org/10.1063/5.0042868>
- [5] 韩忠华. Kriging 模型及代理优化算法研究进展[J]. 航空学报, 2016, 37(11): 3197-3225.
- [6] 孙斌, 姚海涛, 刘婷. 基于高斯过程回归的短期风速预测[J]. 中国电机工程学报, 2012, 32(29): 104-109.
- [7] Roska, T., Chua, L.O., Wolf, D., Kozek, T., Tetzlaff, R. and Puffer, F. (1995) Simulating Nonlinear Waves and Partial Differential Equations via CNN. I. Basic Techniques. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, **42**, 807-815. <https://doi.org/10.1109/81.473590>
- [8] Long, Z., Lu, Y., Ma, X. and Dong, B. (2018) PDE-Net: Learning PDEs from Data. *International Conference on Machine Learning (PMLR)*, Stockholm, 10-15 July 2018, 3208-3216.
- [9] Long, Z., Lu, Y. and Dong, B. (2019) PDE-Net 2.0: Learning PDEs from Data with a numeric-Symbolic Hybrid Deep Network. *Journal of Computational Physics*, **399**, Article ID: 108925. <https://doi.org/10.1016/j.jcp.2019.108925>
- [10] Regazzoni, F., Dede, L. and Quarteroni, A. (2019) Machine Learning for Fast and Reliable Solution of Time-Dependent Differential Equations. *Journal of Computational Physics*, **397**, Article ID: 108852. <https://doi.org/10.1016/j.jcp.2019.07.050>
- [11] Schaeffer, H. and McCalla, S.G. (2017) Sparse Model Selection via Integral Terms. *Physical Review E*, **96**, Article ID: 023302. <https://doi.org/10.1103/PhysRevE.96.023302>
- [12] 陈庆强, 王文剑, 姜高霞. 基于数据分布的标签噪声过滤[J]. 清华大学学报(自然科学版), 2019, 59(4): 262-269.
- [13] 陈倩, 杨旻, 魏鹏飞. 标签带噪声数据的重加权半监督分类方法[J]. 烟台大学学报(自然科学与工程版), 2019, 32(3): 205-209.
- [14] Raissi, M., Perdikaris, P. and Karniadakis, G.E. (2019) Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations. *Journal of Computational Physics*, **378**, 686-707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- [15] Bai, G., Koley, U., Mishra, S. and Molinaro, R. (2021) Physics Informed Neural Networks (PINNs) for Approximating Nonlinear Dispersive PDEs. <https://doi.org/10.4208/jcm.2101-m2020-0342>
- [16] 唐明健, 唐和生. 基于物理信息的深度学习求解矩形薄板力学正反问题[J]. 计算力学学报, 2022, 39(1): 120-128.
- [17] Cai, S., Mao, Z., Wang, Z., Yin, M. and Karniadakis, G.E. (2021) Physics-Informed Neural Networks (PINNs) for Fluid Mechanics: A Review. *Acta Mechanica Sinica*, **37**, 1727-1738. <https://doi.org/10.1007/s10409-021-01148-1>
- [18] Rasht-Behesht, M., Huber, C., Shukla, K. and Karniadakis, G.E. (2022) Physics-Informed Neural Networks (PINNs) for Wave Propagation and Full Waveform Inversions. *Journal of Geophysical Research: Solid Earth*, **127**, e2021JB023120. <https://doi.org/10.1029/2021JB023120>
- [19] Burgers, J.M. (1948) A Mathematical Model Illustrating the Theory of Turbulence. *Advances in Applied Mechanics*, **1**, 171-199. [https://doi.org/10.1016/S0065-2156\(08\)70100-5](https://doi.org/10.1016/S0065-2156(08)70100-5)

-
- [20] Hopf, E. (1950) The Partial Differential Equation $ut + uux = \mu xx$. *Communications on Pure and Applied mathematics*, **3**, 201-230. <https://doi.org/10.1002/cpa.3160030302>
- [21] Meng, X., Yang, L., Mao, Z., Águila Ferrandis, J. and Karniadakis, G.E. (2022) Learning Functional Priors and Posteriors from Data and Physics. *Journal of Computational Physics*, **457**, Article ID: 111073. <https://doi.org/10.1016/j.jcp.2022.111073>
- [22] Vadyala, S.R., Betgeri, S.N. and Betgeri, N.P. (2022) Physics-Informed Neural Network Method for Solving One-Dimensional Advection Equation Using PyTorch. *Array*, **13**, Article ID: 100110. <https://doi.org/10.1016/j.array.2021.100110>
- [23] 刘金国, 许开来. 自动微分及其在物理模拟中的应用[J]. 物理学报, 2021, 70(14): 97-107.
- [24] 张钰, 刘建伟, 左信. 多任务学习[J]. 计算机学报, 2020, 43(7): 1340-1378.
- [25] Knowles, I. and Renka, R.J. (2014) Methods for Numerical Differentiation of Noisy Data. *Electronic Journal of Differential Equations*, **21**, 235-246.
- [26] 蒲削鋸. 标签噪声下的故障诊断算法研究[D]: [硕士学位论文]. 杭州: 浙江大学, 2023: 27-39.