河南大学 2024 届本科毕业论文

基于 Fourier 神经算子的求解微分方程方法的分析与实践

论文作者姓名： 孟乾轲

作 者 学 号 ： 2024030029

所 在 学 院 ： 国际商学院

所 学 专 业 ： 计算机科学与技术

导 师 姓 名 ： 王建林

导 师 职 称 ： 副教授

2024 年 5 月 6 日

河南大学 2024 届本科生毕业论文（设计、创作）开题报告

论文（设计、创作）题目	基于 Fourier 神经算子的求解微分方程方法的分析与实践		
所在学院	国际商学院	**专　业**	计算机科学与技术
学生姓名	孟乾轲	**学　号**	2024030029

一、本课题研究意义

偏微分方程（PDEs）在工程中具有重要意义，用于解决各种复杂的物理现象。在流体动力学、热传导、结构力学和电磁场建模等领域，偏微分方程提供了强大的工具，可以用于精确建模和优化各种系统。

在 SNO（Structured Neural Operators）和 DeepONets（Deep Operator Networks）中，我们使用深度神经网络（DNN）来拟合算子。然而，传统的 DNN 在处理算子时面临一个主要限制：它们通常无法直接扩展到无限维度，因为 DNN 的结构固定在有限的维度。为了克服这一限制，我们引入了含有核函数的神经元算子，这种算子可以将 DNN 泛化到无限维度，以处理更复杂的问题，如高维偏微分方程的求解。

二、国内外有关本课题的研究动态

国内外诸多领域成功应用深度学习方法解决 PDEs，包括流体力学、材料科学、天气预测、生物医学工程等。深度学习被证明是一种灵活且高效的工具，能够处理不同类型和维度的 PDEs。

结构化神经算子（Structured Neural Operators, SNO）和 Deep Operator Networks（DeepONets）也是研究热点，它们使用深度神经网络来拟合算子，从而处理高维偏微分方程的求解问题。引入核函数的神经元算子使得这些方法能够泛化到无限维度。

三、本课题研究的基本内容

本课题旨在通过引入一种创新的神经算子，将深度学习应用于偏微分方程（PDE）的求解。采用一种新的参数化方法，在傅里叶空间中对积分核进行直接参数化，构建了一种具有强大表达能力和高效性能的神经算子。

四、本课题拟解决的主要问题及对解决复杂工程问题能力的要求

五、研究方法

傅里叶变换实现： 将傅里叶变换嵌入神经网络中，确保网络能够正确处理信息。

超参数调优： 对神经网络的超参数进行系统调优，以提高模型的性能和泛化能力。

模型训练与优化： 使用准备好的数据集对神经网络模型进行训练，并通过调整优化算法、学习率等参数进行模型的优化。

PyTorch 1.8.0

六、主要创新点（选填）

传统方法不论是有限元还是有限差分方法都是在离散空间中对方程进行求解，追求的精度越高就需要更多的计算量，而利用人工智能可能实现更快、更准或更快和更准。

七、主要参考文献

[1] 郭东亮.基于深度学习的高维非线性偏微分方程数值解法[D].南昌大学,2022.

[2] 曾壬源.基于深度神经网络的偏微分方程求解[D].中国科学技术大学,2022.

[3] 查文舒,李道伦,沈路航,等.基于神经网络的偏微分方程求解方法研究综述[J].力学学报,2022,54(03):543-556.

[4] 曹富军,郭晓斌,高飞,等.求解偏微分方程的卷积迭代方法[J].山西大学学报(自然科学版),2023,46(02):293-303.DOI:10.13451/j.sxu.ns.2022065.

[5] 汪京徽. 基于 Feynman-Kac 公式的偏微分方程的蒙特卡洛及神经网络方法[D].上海财经大学,2023.DOI:10.27296/d.cnki.gshcu.2023.000323.

[6] 李务杨. 异质偏微分方程的区域分解和深度学习算法研究[D].东北师范大学,2023.DOI:10.27011/d.cnki.gdbsu.2022.000166.

[7] Qi K ,Sun J . Gabor-Filtered Fourier Neural Operator for solving Partial Differential Equations [J]. Computers and Fluids, 2024, 274 106239-.

[8] 孟柳君. 基于多项式神经网络解偏微分方程[D].云南财经大学,2022.DOI:10.27455/d.cnki.gycmc.2022.000107.

[9] 任清华. 基于深度学习的偏微分方程求解方法[D].山东大学,2023.DOI:10.27272/d.cnki.gshdu.2022.003244.

[10] 王江元,杨耐,张英浩,等.基于深度学习方法求解偏微分方程的应用举例[J].信息与电脑(理论版),2021,33(21):48-51.

[11] Li Z, Kovachki N, Azizzadenesheli K, et al. Fourier Neural Operator for Parametric Partial Differential Equations[J]. ICRL, 2021.

[12] 高普阳,赵子桐,杨扬.基于卷积神经网络模型数值求解双曲型偏微分方程的研究[J].应用数学和力学,2021,42(09):932-947.

[13] 陈静飞.机器学习方法求解偏微分方程：从网络结构与优化算法出发[D].苏州大学,2023.

[14] 毛超利.基于深度学习的偏微分方程求解方法[J].智能物联技术,2021,4(05):18-23+30.

[15] 孙靖威.基于深度学习求解偏微分方程的研究[D].天津师范大学,2024.

[16] 韦昌,樊昱晨,周永清,等.基于Runge-Kutta的自回归物理信息神经网络求解偏微分方程[J/OL].力学学报:1-13[2024-05-01].http://kns.cnki.net/kcms/detail/11.2062.O3.20240422.1549.002.html.

[17] DAI Y ,AN Y ,LI Z , et al.Fourier neural operator with boundary conditions for efficient prediction of steady airfoil flows[J].Applied Mathematics and Mechanics(English Edition),2023,44(11):2019-2038.

具体时间及写作进度安排（表格不够可追加表格）

起止日期	主要工作内容

指导教师对开题报告的意见

选题符合本专业人才培养目标，体现本专业基本教学内容，使学生受到全面综合实践训练，有利于培养学生独立解决实际问题的工作能力。同意申报。

指导教师签名：　　　　　　2024 年 1 月 3 日

系（室）对本课题开题的意见

微分方程在各个科学领域中有广泛的应用，而 Fourier 神经算子作为一种强大的工具，其在微分方程求解中的应用具有实际意义和研究价值，研究内容合理，有一定的创新性，同意开题。

负责人签名：　　　　　　2024 年 1 月 7 日

河南大学 2024 届本科生毕业论文（设计、创作）

中期进展情况检查表

<table>
<tr><td>论文（设计、创作）题目</td><td colspan="3">基于 Fourier 神经算子的求解微分方程方法的分析与实践</td></tr>
<tr><td>所在学院</td><td>国际商学院</td><td>专　业</td><td>计算机科学与技术</td></tr>
<tr><td>学生姓名</td><td>孟乾轲</td><td>学　号</td><td>2024030029</td></tr>
<tr><td colspan="4">一、毕业论文（设计、创作）进展情况
按照学院毕业论文工作计划的相关要求，已如期提交开题报告。解决一维问题的框架已基本建成，核心功能已实现。现在需要对程序进行优化，对预测结果正确率进行提升，并按照提纲撰写毕业论文。</td></tr>
<tr><td colspan="4">二、毕业论文（设计、创作）存在问题及解决方案
存在的问题有：预测结果正确率有待提高；还未解决更高维度的偏微分方程。
解决方案：学习新的机器学习算法，提高预测正确率；继续努力解决更高维度的问题。</td></tr>
<tr><td colspan="4">三、指导教师对学生论文（设计、创作）进展等方面的评价
指导教师签字：　　　　2024 年 3 月 13 日</td></tr>
<tr><td colspan="4">系（室）对本课题中期检查的意见
负责人签名：　　　　2024 年 3 月 17 日</td></tr>
</table>

河南大学 2024 届毕业论文（设计、创作）指导教师评价表

<table>
<tr><td>论文（设计、创作）题目</td><td colspan="5">基于 Fourier 神经算子的求解微分方程方法的分析与实践</td></tr>
<tr><td>所在学院</td><td colspan="2">国际商学院</td><td>专　业</td><td colspan="2">计算机专业与技术</td></tr>
<tr><td>学生姓名</td><td colspan="2">孟乾轲</td><td>学　号</td><td colspan="2">2024030029</td></tr>
<tr><td>指导教师</td><td colspan="2">王建林</td><td>指导教师职称</td><td colspan="2">副教授</td></tr>
<tr><td rowspan="9">指导教师评分</td><td rowspan="2">序号</td><td rowspan="2">评分项目</td><td colspan="2">评分参考</td><td>得分</td></tr>
<tr><td>评阅指标（优秀标准）</td><td>满分（100 分）</td><td></td></tr>
<tr><td>1</td><td>选题质量</td><td>符合专业培养目标；有实际意义和推广价值。</td><td>20</td><td></td></tr>
<tr><td>2</td><td>文献资料 应用能力</td><td>能独立查阅文献，具有收集、加工各种信息及获取新知识的能力。</td><td>10</td><td></td></tr>
<tr><td>3</td><td>调查研究能力</td><td>能准确理解课题任务；研究方案设计合理；能独立从事调查研究；能综合运用所学知识发现与解决实际问题。</td><td>30</td><td></td></tr>
<tr><td>4</td><td>论文（设计、创作）质量</td><td>格式、图表（或图纸）规范，符合要求；结构严谨，逻辑性强；内容翔实，表达准确流畅；学术价值或实用价值高。</td><td>20</td><td></td></tr>
<tr><td>5</td><td>创新能力/解决复杂工程问题能力</td><td>观点独到，方法新颖，角度新颖/运用工程原理；涉及多个技术因素；建立合适的抽象模型；具有较高的综合性；涉及专业标准和规范之外的因素。</td><td>10</td><td></td></tr>
<tr><td>6</td><td>工作量及态度</td><td>工作量饱满；能圆满完成任务书规定的各项工作。</td><td>10</td><td></td></tr>
<tr><td colspan="3">总得分</td><td colspan="2"></td></tr>
<tr><td rowspan="2">指导教师评定意见</td><td colspan="5"></td></tr>
<tr><td colspan="5">指导教师签名：　　　　　　　　　　2024 年 4 月 25 日</td></tr>
</table>

备注：

1. 总得分可保留两位小数。
2. 评阅教师要在学生参与答辩前对论文整体情况做出客观评价，并说明是否同意该生参与毕业论文答辩。

河南大学 2024 届毕业论文（设计、创作）评阅教师评价表

<table>
<tr><td>论文（设计、创作）题目</td><td colspan="5">基于 Fourier 神经算子的求解微分方程方法的分析与实践</td></tr>
<tr><td>所在学院</td><td colspan="2">国际商学院</td><td>专　业</td><td colspan="2">计算机科学与技术</td></tr>
<tr><td>学生姓名</td><td colspan="2">孟乾轲</td><td>学　号</td><td colspan="2">2024030029</td></tr>
<tr><td>评阅教师</td><td colspan="2"></td><td>评阅教师职称</td><td colspan="2"></td></tr>
<tr><td rowspan="9">评阅教师评分</td><td rowspan="2">序号</td><td rowspan="2">评分项目</td><td colspan="2">评分参考</td><td>得分</td></tr>
<tr><td>评阅指标（优秀标准）</td><td>满分（100 分）</td><td></td></tr>
<tr><td>1</td><td>选题质量</td><td>符合专业培养目标：有实际意义和推广价值。</td><td>20</td><td></td></tr>
<tr><td>2</td><td>文献资料应用能力</td><td>能独立查看文献，具有收集，加工各种信息及获取新知识的能力。</td><td>10</td><td></td></tr>
<tr><td>3</td><td>调查研究能力</td><td>能准确理解课题任务：研究方案设计合理:能独立从事调查研究，能综合运用所学知识发现与解决实际问题。</td><td>30</td><td></td></tr>
<tr><td>4</td><td>论文（设计、创作）质量</td><td>格式，图表（或图纸）规范符合要求：结构严谨，逻辑性强 内容翔实，表达准确流畅：学术价值或使用价值高。</td><td>30</td><td></td></tr>
<tr><td>5</td><td>创新能力/解决复杂工程问题能力</td><td>观点独到，方法新颖，角度新颖/运用工程原理；涉及多个技术因素；建立合适的抽象模型；具有较高的综合性；涉及专业标准和规范之外的因素。</td><td>10</td><td></td></tr>
<tr><td colspan="3">总得分</td><td colspan="2"></td></tr>
<tr><td colspan="5"></td></tr>
<tr><td rowspan="2">评阅教师评定意见</td><td colspan="5"></td></tr>
<tr><td colspan="5">评阅教师签名：　　　　　　　　　　2024 年 4 月 26 日</td></tr>
</table>

备注：

1. 总得分可保留两位小数。

2. 评阅教师要在学生参与答辩前对论文整体情况做出客观评价，并说明是否同意该生参与毕业论文答辩。

河南大学2024届本科生毕业论文（设计、创作）

答辩成绩表

<table>
<tr><td colspan="2">论文（设计、创作）题目</td><td colspan="5">基于Fourier神经算子的求解微分方程方法的分析与实践</td></tr>
<tr><td colspan="2">所在学院</td><td colspan="2">国际商学院</td><td>专　业</td><td colspan="2">计算机科学与技术</td></tr>
<tr><td colspan="2">学生姓名</td><td colspan="2">孟乾轲</td><td>学　号</td><td colspan="2">2024030029</td></tr>
<tr><td rowspan="5">答辩委员会（组）评分及评定结论</td><td rowspan="3">评分项目及分值</td><td colspan="5">答辩委员会（组）专家评分</td></tr>
<tr><td colspan="2">答　辩　情　况</td><td colspan="2">论　文　质　量</td><td rowspan="2">合计
（100分）</td></tr>
<tr><td>内容表达情况
（15分）</td><td>回答问题情况
（25分）</td><td>规范要求与文字表达
（20分）</td><td>论文（设计、创作）质量和创新意识
（40分）</td></tr>
<tr><td>得分</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>答辩委员会（组）评定结论</td><td colspan="5">答辩委员会（组）签字：

2024年4月29日</td></tr>
<tr><td colspan="7">毕业论文（设计、创作）答辩成绩：　　　　　　分

答辩委员会（组）负责人签字：

2024年4月29日</td></tr>
</table>

备注：

1. 论文的质量评定，应包括对论文的语言表达、结构层次、逻辑性理论分析、设计计算、数据处理、分析和概括能力及在论文中是否有新的见解或创造性成果等做出评价。从论文来看学生掌握本专业基础理论和基本技能的程度。
2. 评分由答辩委员会（组）专家（原则上5或7人一组）操作，根据学生答辩情况，答辩委员会专家及时给出每个学生的成绩。内部对评分如有争议时，应当场进行表决。
3. 答辩委员会（组）评定结论应说明是否通过答辩。
4. 各院亦可根据本专业的不同情况，制定相应的具有自己特色的评价表格。
5. 以上表格所有签字部分必须手签。

河南大学 2024 届本科生毕业论文（设计、创作）
答辩纪要

<table>
<tr><td>论文（设计、创作）题目</td><td colspan="3">基于 Fourier 神经算子的求解微分方程方法的分析与实践</td></tr>
<tr><td>答辩人姓名</td><td>孟乾轲</td><td>指导教师</td><td>王建林</td></tr>
<tr><td>答辩时间</td><td></td><td>答辩地点</td><td></td></tr>
<tr><td>答辩委员会(组)负责人</td><td></td><td>答辩委员会(组)成员</td><td></td></tr>
<tr><td colspan="4">答辩中提出的问题及回答的简要情况记录：

记录人签名：　　　　　　　　　　　　　　　　　　2024 年 4 月 28 日</td></tr>
</table>

河南大学2024届本科生毕业论文（设计、创作）

综合成绩表

<table>
<tr><td>论文（设计、创作）题目</td><td colspan="5"></td></tr>
<tr><td>所在学院</td><td colspan="2"></td><td>专　业</td><td colspan="2"></td></tr>
<tr><td>学生姓名</td><td colspan="2"></td><td>学　号</td><td colspan="2"></td></tr>
<tr><td>指导教师评分</td><td></td><td>评阅教师评分</td><td></td><td>答辩评分</td><td></td></tr>
<tr><td>综合成绩</td><td></td><td>成绩等级</td><td></td><td>是否推优</td><td></td></tr>
<tr><td>所在系（室）意见</td><td colspan="5">负责人签名：　　　　　　　　2024年4月30日</td></tr>
<tr><td>学院意见（仅推优论文填写）</td><td colspan="5">负责人签名（盖章）：　　　　　　　　2024年4月30日</td></tr>
</table>

备注：

毕业论文（设计、创作）综合成绩以百分制计，构成方式为：指导教师评分（30%）+评阅教师评分（20%）+答辩成绩（50%），各部分评分可保留两位小数，最终综合成绩需要四舍五入保留整数。

毕业论文（设计、创作）综合成绩对应四个等级：优秀：100～90；良好：89～76；中等：75～60；不及格：59～0。

河南大学本科生毕业论文（设计、创作）承诺书

论文（设计、创作）题目	基于 Fourier 神经算子的求解微分方程方法的分析与实践				
姓　名	孟乾轲	**学　号**	2024030029	**专　业**	计算机科学与技术
指导教师姓名	王建林	**职　　称**	副教授		
完成时间	**年　　月　　日　—　　　年　　月　　日**				

承诺内容：

1、本毕业论文（设计、创作）是学生 孟乾轲 在导师 王建林 的指导下独立完成的，没有抄袭、剽窃他人成果，没有请人代做，若在毕业论文（设计、创作）的各种检查、评比中被发现有以上行为，愿按学校有关规定接受处理，并承担相应的法律责任。

2、学校有权保留并向上级有关部门送交本毕业论文（设计、创作）的复印件和电子稿件。学校可以将毕业论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编本论文。

备注：

学生签名：　　　　　　　　　　**指导教师签名：**

2024 年 4 月 21 日　　　　　　　　**2024 年 4 月 22 日**

说明：学生毕业论文（设计、创作）如有保密等要求，请在备注中明确，承诺内容第 2 条即以备注为准

摘 要

偏微分方程（PDEs）的高维数值求解在许多科学和工程问题中至关重要，但传统数值方法在处理这些问题时往往受限于昂贵的计算成本[1]。深度学习技术，尤其是 Fourier 神经算子（FNO），为降低这些问题的求解难度提供了前所未有的途径。本文通过优化深度学习模型参数，提升了 FNO 在求解 PDEs 时的性能和准确度。

论文首先分析了 FNO 在解决 PDE 问题中的应用原理，接着围绕如何通过调整和优化网络参数来提升模型性能展开深入探讨。通过综合性的实验设计来测试如何调整参数，这些设计包括了两种类型 PDE 和多组参数设置，通过这些实验，我们不仅验证了参数优化前后的性能提升，还展示了 FNO 模型在特定测试集上的具体改进数据。

论文结果表明，通过精确调整深度学习模型的参数，可以优化 PDE 求解过程。未来的工作将聚焦于优化算法的进一步改进，以及 FNO 在更广泛类型的 PDEs 中的应用。

关键词：偏微分方程；深度学习；傅里叶神经算子；参数优化

ABSTRACT

The high-dimensional numerical resolution of partial differential equations (PDEs) is crucial for many scientific and engineering problems. However, traditional numerical methods are often limited by expensive computational costs when addressing these issues. Deep learning technologies, especially Fourier Neural Operators (FNO), offer unprecedented pathways to reduce the difficulty of solving these problems. This paper enhances the performance and accuracy of FNO in solving PDEs through the optimization of deep learning model parameters.

The paper first analyzes the application principles of FNO in solving PDE problems, and then delves into how to enhance model performance by adjusting and optimizing network parameters. Comprehensive experimental designs were implemented to test how parameter adjustments could be made, including two types of PDEs and multiple parameter settings. Through these experiments, not only was the improvement in performance before and after parameter optimization verified, but specific improvements in the FNO model on particular test sets were also demonstrated.

The results indicate that precise adjustments of deep learning model parameters can optimize the PDE solving process. Future work will focus on further improvements to optimization algorithms and the application of FNO in a broader range of PDE types.

Keywords: Partial Differential Equations; Deep Learning; Fourier Neural Operators; Parameter Optimization

目　录

第 1 章　绪　论

在科研和工程实践中,经常面临的各种复杂的物理问题。这些问题往往可以被建模转化为偏微分方程。它们在众多领域中扮演着关键角色。大多数情况下,这些方程的解不能用公式来表示,所以要寻找数值解法。常见的数值解法分为两类：传统方法和机器学习方法。

求解 PDEs 往往伴随着复杂的计算过程，这对传统数值解法来说是一个高成本且低效的挑战。随着深度学习技术的发展，其在数据处理和模式识别方面的强大能力开始被应用于解决这一挑战，尤其是结合 FNO 原理的方法展现了解决复杂 PDEs 问题的高效潜力。本文致力于通过优化深度学习模型的参数来提高 Fourier 神经算子求解 PDEs 的性能和精度，目的是减少计算时间并提高解的准确性。

对于 FNO 在解决 PDE 问题上的应用，虽然近年来已有研究取得进展，但现有文献主要集中在其基本应用，对如何优化模型参数以进一步提升性能和精度的探索还相对有限。因此，本文的创新之处在于参数优化方法，充分发挥 FNO 在解决 PDEs 中的潜力。

研究方法将基于 FNO 原理，并结合深度学习技术进行模型的参数优化。通过详细的实验设计和结果验证，本文将展示参数优化前后在求解特定 PDEs 问题上性能的提升。

总的来说，本文的主要贡献在于提出 FNO 模型参数优化方法，这提高了利用 FNO 原理解 PDEs 的效率和精度，本论文的结构为逐步揭示研究内容，从绪论到理论背景与相关工作、方法论与实验设计，最后是分析与讨论。

1.1　课题来源

本课题由老师指派。

1.2　课题背景

在深度学习技术迅猛发展的背景下，其在物理科学领域，尤其是偏微分方程（PDEs）求解方面展现出显著潜力[2]。偏微分方程是描述多种自然和工程现象的核心数学模型，但传统求解方法面临高计算成本和维数灾难的挑战。Fourier Neural Operators（FNO）作为一种创新的深度学习方法，通过傅里叶变换的性质，有效优化 PDEs 的建模和求解过程，提高了处理高维问题的效率和精度。本文聚焦于 FNO 在 PDEs 求解中的应用，通过

优化 FNO 模型，进一步探索其在科学研究和工程实践中的潜力，为传统数学问题求解提供新的视角。

1.3 国内外在该方向的研究现状及分析

在深度学习领域，利用 Fourier Neural Operators（FNO）解偏微分方程已成为一个研究热点，表现了深度学习技术在解决传统数学和物理问题上的广泛应用潜力。整体上，无论是国内还是国外的研究，都集中于提高模型的求解精度、计算效率以及处理高维数据的能力。研究者们通过不断优化 FNO 的结构和参数，探索更为高效的训练方法和损失函数，以适应不同类型和复杂度的 PDE 问题。同时，跨学科的合作也促进了理论方法与实际应用的结合，拓宽了 FNO 技术的应用范围[3]。

1.3.1 国内研究现状

国内在 FNO 应用于偏微分方程求解的研究中，呈现出积极的探索态势。多个高校和研究机构的团队[4]，侧重于理论与方法的创新，例如通过结合中国特有的数学理论[5]，优化 FNO 的网络结构，或是在特定领域，如流体力学和材料科学中，应用 FNO 解决实际问题。借助深度学习的强大功能，国内学者开始探索传统数值方法与机器学习技术的结合。例如，使用神经网络来近似解的表示或改进区域分解算法的边界处理技术[6]，这些研究显示了良好的初步成果，特别是在处理高维和非线性复杂问题方面表现出了潜力。国内研究不仅注重算法的提升，也在探索 FNO 技术在工业、环境科学等领域的应用，以期实现科学研究与经济发展的双赢。

1.3.2 国外研究现状

国外在使用 FNO 解偏微分方程的研究方面较早开始，拥有广泛的应用实例。海外研究机构和大学通常集中在技术的原理探索、算法优化以及跨学科应用等方面。例如，他们在提高 FNO 模型泛化能力、处理复杂边界条件的 PDEs 上取得了显著成果。在 Qi K 和 Sun J 的研究中，通过将 Gabor 滤波与 FNO 结合，展示了这一方法在处理具有复杂模式和动态变化的系统中的有效性[7]。

第 2 章　理论背景与相关工作

第二章“理论背景与相关工作”提供关于偏微分方程（PDEs）解法的必要理论基础及其在深度学习领域应用的最新进展。本章将简要介绍偏微分方程的基本概念以及其在多个科学和工程领域中的重要性。探讨传统的 PDE 求解方法，包括解析解法及数值解法如有限差分法、有限元法等，每种方法的优缺点也将被讨论。深度学习如何被用于 PDE 的求解，特别是傅立叶神经算子（FNO）的原理和应用。

2.1　偏微分方程（PDEs）简介

在本文中，通过深度学习方法求解两类具有代表性的偏微分方程——Burgers 方程和 Darcy 方程。

Burgers 方程是一个非线性偏微分方程，广泛应用于流体力学、交通流等领域，用于描述一维粘性流体中的波动和冲击波形成等现象。它简洁而富有挑战性的特点，使之成为测试和展示新求解方法效果的理想选择。

Darcy 方程是描述地下水流动的线性偏微分方程，主要应用于土木工程和石油工程等领域。它可以模拟二维或三维空间中的流体流动，特别是在多孔介质中的流动。Darcy 方程的线性特性与 Burgers 方程的非线性特性使得本文中覆盖宽广的 PDE 类型。

2.2　传统的 PDE 求解方法

偏微分方程在描述自然界和工程领域中的各种现象时发挥着重要作用，它们广泛应用于物理、工程、金融和生命科学等领域。为了解决这些方程，人们创造了解析方法、有限差分法（FDM）、有限元法（FEM）、有限体积法（FVM）和谱方法等诸多方法。每种方法都有其特定的适用场景和局限性，而选择最合适的方法需要综合考虑问题的性质、所需解的精确度，以及可用的计算资源。本论文总结了传统偏微分方程求解方法的优缺点比较，如表 1.1 所示。

表 1.1　传统偏微分方程求解方法的优缺点比较

方法	优点	缺点
解析方法	可以提供精确解 对于简单问题，解法直观且易于理解	仅限于简单几何和边界条件 对复杂或非线性方程难以找到解
有限差分法 (FDM)	实现简单 对于线性问题和常规几何形状适用性好	对复杂几何边界处理困难 精度受到网格划分的限制
有限元法 (FEM)	灵活处理复杂几何和边界条件 可以用于非线性问题 高精度求解	计算量大 网格生成和元素选择复杂
有限体积法 (FVM)	保守性好，适合流体动力学问题 可以直接处理复杂的边界条件	网格划分和重构可能复杂 难以模拟对流体流动中的高速和湍流
谱方法	对于平滑问题，提供高精度 效率高于有限差分法和有限元法	对非平滑问题或有尖锐边界的问题效果不佳

总的来说，传统方法在求解一些高维偏微分方程的过程中会遭遇“维数灾难”， 传统求解方法中使用的网格随着维度增加，计算量呈指数级增长，因此划分网格的方法在高维 PDE 的数值计算上不太有效[8]。

而对于低维的偏微分方程，Burgers 方程和 Darcy 方程分别代表了非线性和线性偏微分方程（PDE）的典型案例。针对这两类方程，传统求解方法已经相当成熟，能够有效地处理各种复杂情况。

对于 Burgers 方程，它通过模拟各类流体动力学现象，如湍流和波的传播，展现了非线性 PDE 的特性。虽然在一些特定条件下，如无粘情况，可以通过解析方法找到精确解，但更普遍的情况是通过数值方法，如有限差分法（FDM）和有限体积法（FVM），来求解。这些方法将连续方程离散化，并在网格上近似求解，这些传统方法已经发展得相当成熟[8]，能够高效处理包括边界条件和初始条件在内的复杂问题。

Darcy 方程描述了流体通过多孔介质的流动，其线性特性使得解析方法在特定简单条件下依然有效。但在处理复杂边界和异质介质时，数值方法如有限元法（FEM）和有限差分法（FDM）提供了更大的灵活性和适应性。这些方法允许对求解域进行灵活的网格划分，从而精确模拟实际物理过程。

在现代工程和科学研究中，随着计算能力的提升和数值方法的发展，数值求解已成为这两类方程常用的解决方案。

2.3 深度学习在 PDE 求解中的应用

2.3.1 深度学习基础知识简介

深度学习，作为机器学习的一个分支，主要依赖于神经网络的架构来处理数据。神经网络由相互连接的节点（或称为神经元）组成，这些节点在不同的层次中组织起来。输入数据在网络中流动，通过各层的加权、偏置调整和激活函数处理，最终产生输出。这个过程可以通过反向传播算法和梯度下降等优化技术进行学习，使得网络能够逼近复杂函数[9]。

2.3.2 深度学习方法解 PDE 的早期尝试和发展历程

早期尝试将深度学习应用于 PDE 求解主要集中在使用全连接网络来直接逼近 PDE 的解。这些方法通常将 PDE 的边界条件和初始条件整合进损失函数中，通过训练神经网络来最小化这个损失函数，寻求逼近真实解的网络参数。随着研究的深入，更多专门针对 PDE 求解的网络结构被提出，卷积神经网络（CNN）在处理空间信息方面的优势使其成为一个有力的工具。

2.3.3 现代深度学习方法与 PDE

现代深度学习法在解决偏微分方程问题方面显示出了较强的潜能。如物理信息神经网络（PINN）和 Fourier 神经算子两种方式。PINN 在神经网络的训练过程中使用物理定律的限制，使得训练出的模型不仅能够拟合数据，而且能满足物理定律。即使在数据稀疏的情况下，PINN 也能利用物理定律的先验知识，得到较好的预测结果。然而，PINNs 的训练过程比较复杂，需要大量的计算资源。FNO 是一种基于傅里叶变换的神经算子，它可以有效地处理周期性的偏微分方程问题。FNO 利用快速傅立叶变换将偏微分方程从空间域转换到频率域在频率域中进行计算,最后再将结果通过逆傅立叶变换转换回空间域。这种方法的好处是运算效率高。

2.3.4 Fourier 神经算子求解 PDE

与机器学习的其他方法相比，Fourier 神经算子的优势在于其结合了傅里叶变换的能力，能在频域进行高效运算，从而避免了在时域进行复杂的卷积计算。这一特性使其在

解决偏微分方程时表现出高效性。在处理周期性或近似周期性的问题时，由于经傅里叶变换后更易于提取周期性问题的特征，Fourier 神经算子能够提供更稳定的解。相对应的，对于非周期性的问题，或者在数据噪声较大的情况下，Fourier 神经算子可能无法展现优势。

与传统方法相比，FNO 在处理高维 PDE 问题时展现出了显著优势。FNO 能够通过学习 PDE 的频域特征来高效地进行近似，这不仅大幅度减少了计算资源和时间的需求，而且一旦训练完成，同一个 FNO 模型就能够应对多种初始和边界条件下的 PDE 求解，显示出良好的泛化能力。

在本文中，通过求解流体动力学中的 Burgers 方程和多孔介质流动的 Darcy 方程，FNO 展示了其在从简单到复杂的 PDE 问题上的强大潜力[10]。

第 3 章　方法论与实验设计

本章详细描述了 FNO 的架构，介绍模型如何通过增加输入数据维度、卷积和傅里叶变换以及进行频域内的线性和非线性操作来增强处理偏微分方程的能力。实验部分则侧重于如何通过调整关键参数来优化模型性能。介绍使用的数据集、训练过程和评估指标。

3.1　问题定义

本文的核心内容是 Fourier 神经算子（FNO）在解决特定偏微分方程（PDE），即 Burgers 方程和 Darcy 流方程的分析与实践。这两种方程分别代表非线性和线性的 PDE 问题，为深入评估 FNO 的效能提供了理想的测试场景。核心目标是通过优化网络参数来提高 FNO 在解这些方程中的性能和准确度。

实验的主要内容围绕如何调整 FNO 模型的关键参数，如傅里叶模式的数量、网络宽度、学习率等参数，以达到最优的模型性能。这包括系统地测试不同参数配置对模型预测精度的影响，并通过大量实验确定最佳参数组合。通过这种方式，不仅证明 FNO 在解决 Burgers 和 Darcy 方程中的有效性，而且通过参数优化提升了深度学习模型的性能。

3.2　FNO 架构的详细描述

在神经网络中。神经网络的输入和输出是有限维的。而神经算子将输入视为定义在一段连续域上的函数，将输入和输出提升到无限维。在 FNO 中，使用傅立叶变换将其转换为有限维的频率表示，这种神经算子可以保留尽可能多的信息，同时又使得算子需要的计算资源能够在现实中满足。

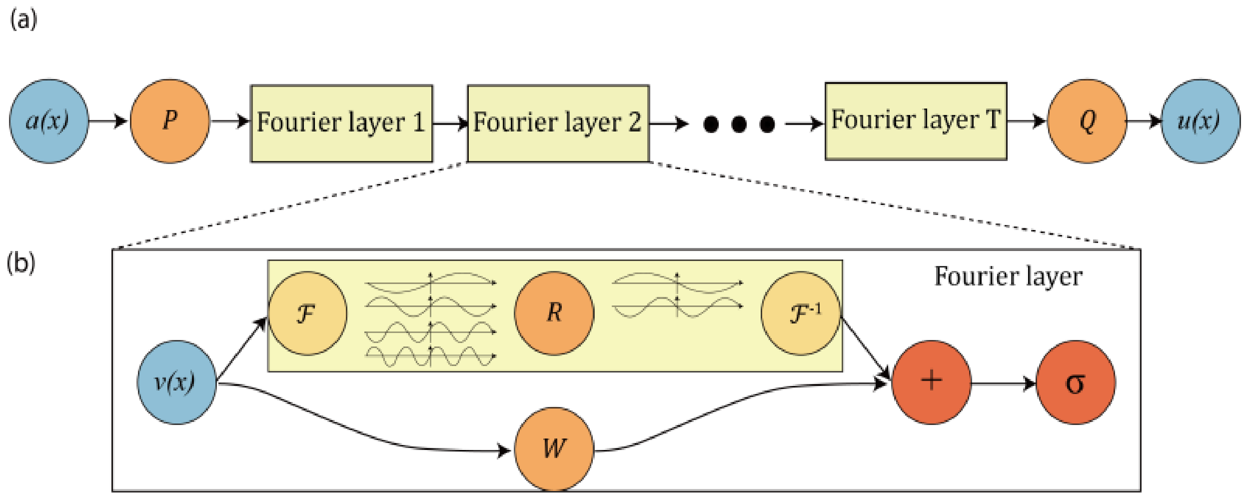

图 3. 1 FNO 架构图[11]

FNO 的输入是系数场，表示为a。输入a首先通过一个神经网络 P 处理，该网络将输入数据“提升”到一个更高维度的特征空间，有助于增加网络的表示能力。之后应用一系列的积分算子层。积分算子层包括一个积分算子和一个激活函数。最后，将高维特征空间通过 Q 转换回原来的维度。最终输出解空间u。

Fourier 层的输入是来自前一层的特征表示，表示为 v。对v进行傅里叶变换F，将其转换到频域上。对变换后输入的低频傅里叶模式应用线性算子 R。可以在频域中学习重要特征。然后对结果进行逆傅里叶变换F^{-1}，将其映射回空间域。与 Fourier 层操作并行的是局部线性变换，直接对输入v应用局部线性变换W。

3.2.1 输入和输出表示

Burgers 方程是一个非线性偏微分方程，常用于流体力学中描述流体中的粘性效应。在一维空间中，Burgers 方程可表示为:

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}=\nu\frac{\partial^2 u}{\partial x^2} \tag{1}$$

其中，u是未知函数，代表流体速度；ν是粘性系数；t是时间变量；x是空间变量

对于 Burger’s 方程，输入张量包含初始速度分布（初始条件）和边界条件。例如，如果考虑一个固定边界的问题，边界条件可以是在区域的两端固定速度值。输出张量表示随时间变化的速度分布。对于时间依赖的问题，输出将包括多个时间步骤，每个步骤都显示整个空间域内的速度预测[12]。

Darcy 流方程用于描述流体通过多孔介质的流动。在简化形式下，稳态 Darcy 流方程可以表示为：

$$\nabla\cdot\left(\frac{k}{\mu}\nabla P\right)=0 \tag{2}$$

其中$\nabla\cdot$表示散度运算符，k是介质的渗透率，μ是流体的动力粘性系数，P是压力。这个方程表示在稳态条件下，流体通过孔隙介质的流动不随时间变化。

对于 Darcy 流方程，输入通常包括渗透性分布$\frac{k}{\mu}$是介质的渗透性（可能是空间变量），以及边界条件（如固定边界上的压力值）。在一些情况下，源项f也可以作为输入的一部分。输出张量是空间域内的压力分布p。由于 Darcy 流方程常在稳态条件下求解，输出可能不涉及时间维度，而是集中在空间分布上。

3.2.2　频域转换

在 Fourier 神经算子处理偏微分方程时，核心优势之一是利用频域转换。频域转换部分的主要内容为将输入信号从时域转换到频域，执行频域内的线性操作，最后将处理过的频域信号转换回时域。首先，对时域内的输入信号应用傅里叶变换，这步操作将信号从时域转换到频域。在频域内，信号以其频率成分表示，这使得频域内的线性操作成为可能。在信号被转换到频域后，对其进行操作。经过频域线性操作处理的信号，通过逆傅里叶变换再次被转换回时域。这一步骤将处理后的频域数据转换为可以解释的时域信号，以完成模型的预测过程。此外，FNO 在处理输入信号时，并不需要使用所有的傅里叶模式，而是可以根据问题的特性选择重要的模式。这种选择机制减少了模型的复杂度和计算成本，同时保持了对 PDE 动态的有效捕捉[13]。

快速傅里叶变换（FFT）算法可以将一个函数或一组数据转换到频域上。FFT 利用数学方法提高计算效率，能够在 O(N log N)的时间复杂度内计算离散傅立叶变换。在 Fourier 神经算子中，用 FFT 算法将输入数据从时域转换到频域。逆快速傅立叶变换（IFFT）是 FFT 的逆操作，可以将频域的数据转换回时域。

离散傅里叶变换（DFT）公式：

$$X[k]=\sum_{n=0}^{N-1}x[n]e^{-\frac{2\pi i}{N}nk} \tag{3}$$

其中$X[k]$是频率域的序列，i是虚数单位，N是序列长度。

3.2.3　频域处理层

经过傅里叶变换后，时空数据被转换成频率空间中的表示。在这个表示中，数据被

表现为不同频率成分的组合，这使得模型能多聚焦于解的不同频率特征，尤其是那些在时空域不易捕捉的特性。在频域处理层，主要进行的是线性变换。这些变换通过与可学习的复数权重矩阵相乘实现。这些权重矩阵代表不同频率成分的重要性，以及它们之间的关系。这种方式使得模型能够根据训练数据调整对不同频率成分的关注度和处理方式。与传统的基于局部操作的网络结构不同，频域处理层能够捕捉到数据的全局特征。因为频率成分代表了整个数据集的统计特性，而不是局部区域的特征[14]。这使得模型在处理 PDEs 时能够更全面地考虑整个时空域的信息。为了提高模型的表达能力，在线性变换后会引入非线性激活函数。这种结合线性和非线性操作的策略可以进一步提升模型处理复杂 PDEs 的能力。总的来说，频域处理层是神经算子层的核心，通过在频域中进行有效的线性和非线性处理，使模型能够高效并准确地处理 PDEs。这一层的设计充分利用了频域表示的优势，提高了对复杂动态的捕捉能力和模型的整体性能。

3.2.4 神经网络结构图

神经网络的结构首先是一个全连接层，接着多个 Fourier 层，最后两个全连接层。如图 3.2 神经网络结构图。

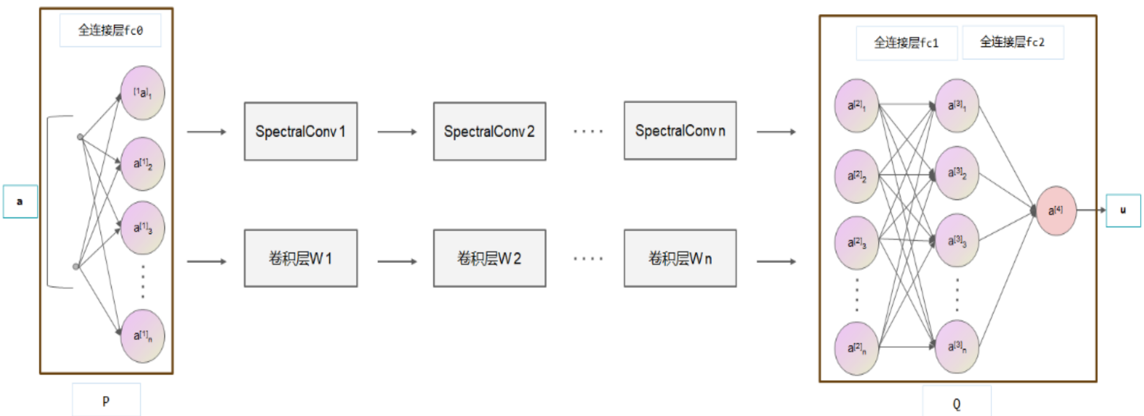

图 3. 2 神经网络结构图

图中的全连接层 fc0，这层的目的是将原始的输入数据转换到一个更高维度的空间，以便后续的神经网络层可以从中提取更复杂的特征。SpectralConv 层是 Fourier 卷积层，这些层会将数据利用快速傅里叶转换的方法转换到频域上再进行操作，这有利于对于周期性或振荡性的函数提取特征。卷积层 W 对每个特征进行独立的线性变换，而不改变

特征的数量。全连接层 fc1 将卷积层的输出转换为一个更高维度的空间，以便后续的神经网络层可以从中提取更复杂的特征。全连接层 c2 是网络的最后一层，它将高维特征映射到一个单一的输出。

Fourier 卷积层和卷积层 W 在处理数据时各有其特点和优势，将它们结合在一起使用可以有效地捕捉和处理周期性或振荡性的数据，同时可以捕捉和处理局部的数据信息。这种结合使用的方式使得网络能够更好地适应各种类型的偏微分方程问题，提高了模型的泛化能力和解决问题的精度。

在 FNO 中，网络的第一个全连接层（fc0）具有 Width 个神经元，每个神经元接收 n 个输入。全连接层（fc1）拥有 128 个神经元，每个神经元接收 Width 个输入，网络的最后一个全连接层（fc2）只包含一个神经元，该神经元接收来自前一层的 128 个输入。

在每个傅立叶层和非线性变换层之间以及全连接层 fc1 之后应用 Gaussian Error Linear Unit（GELU）激活函数。GELU 函数表达式为

$$\mathrm{GELU}(x) = 0.5x\left(1 + \tanh\left[\sqrt{\frac{2}{\pi}}\left(x + 0.044715x^3\right)\right]\right) \tag{4}$$

GELU 激活函数图像如图 3.3 所示：

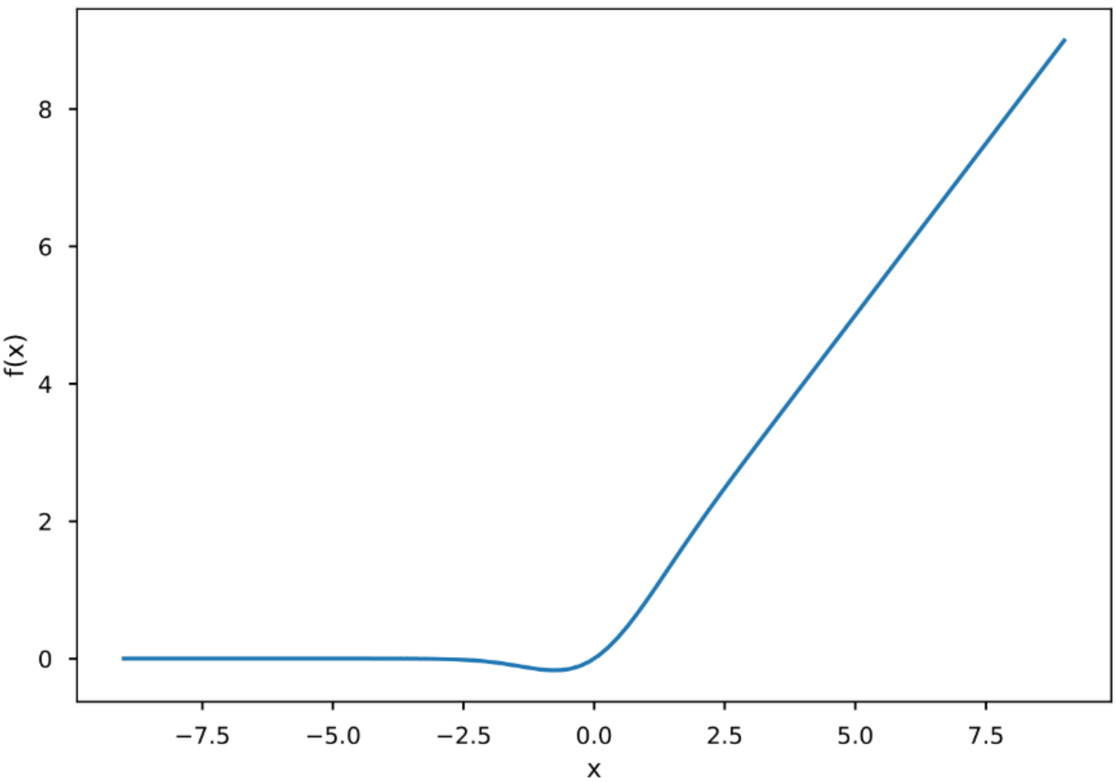

图 3. 3 GELU 激活函数图像

FNO 架构中的 Fourier 层和全连接层都是线性操作，而一些偏微分方程的解通常具有非线性的性质，GULE 激活函数可以在神经网络中引入非线性，有利于提高模型的泛化能力。GELU 函数在输入值接近 0 的区域近似于线性函数，可以在一定程度上防止训练过程中出现的梯度爆炸的问题。GELU 函数本身已被证明可以提供更好的性能和更快的训练速度，在每个傅立叶层和非线性变换层之间引人 GELU 函数有助于 FNO 模型的性能提升。

3.3 数据集

3.3.1 Burgers 方程数据集

Burgers 方程数据集来源于公开资源，可以在以下链接中获取：Burgers_R10.mat。该数据集专门用于研究 Burgers 方程的数值解。数据集的形状为[1000, 8192]，包含了 1000 个样本。每个样本都在一个 8192 的网格上描述了 Burgers 方程的行为。

这些样本数据是通过高斯随机场（GRF）生成的，使用 Burgers 方程进行演化。'a'代表了每个样本的初始状态，'u'代表了对应的 Burgers 方程的解，这些解是在给定初始条件'a'下，通过数值方法得出的。如图 3.3 是 Burgers 方程一个样本的初始状态和解。

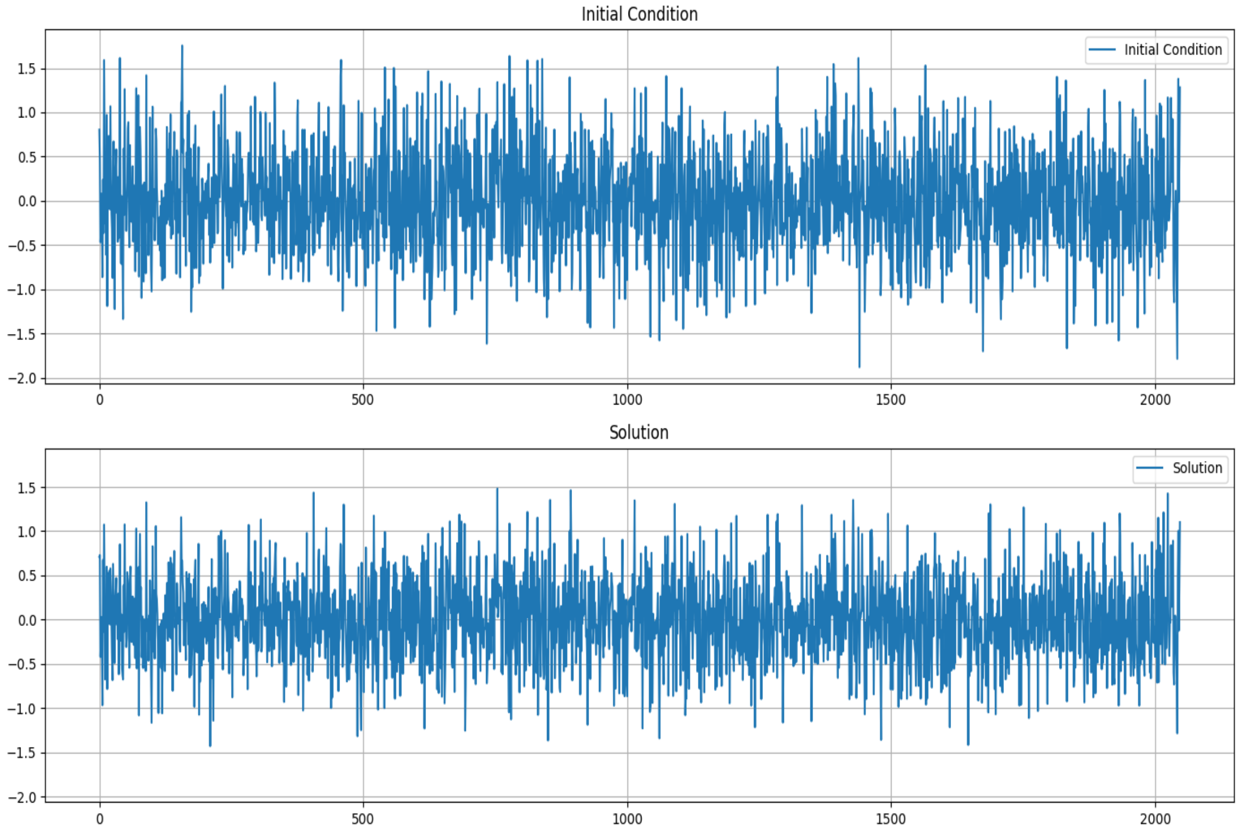

图 3．4 burgers 方程一个样本的初始状态和解

上方的图显示的是 Burgers 方程数据集的一个样本的初始条件，这个初始状态是由高斯随机场生成的，具有随机性。

下方的图显示在给定的初始条件下，通过数值法求解 Burgers 方程的结果。解的形状相比于初始条件更加平滑。这是因为 Burgers 方程具有粘性。解的范围相比初始条件有所减小。因为 Burgers 方程中的非线性项会导致能量的耗散，使得解的范围变小。

3.3.2 Darcy 方程数据集

Darcy 方程数据集也来源于公开资源，可以在以下链接中获取：Darcy_241。该数据集专门用于研究 Darcy 方程的数值解。数据集的形状为[1024, 241, 241]，包含了 1024 个样本。每个样本都在一个 241x241 的网格上描述了 Darcy 方程的行为。

其中'Kcoeff'代表了每个样本的初始状态，而'sol'则代表了对应的 Darcy 方程的解，这些解是在给定初始条件'Kcoeff'下，通过数值方法得出的。如图 3.4

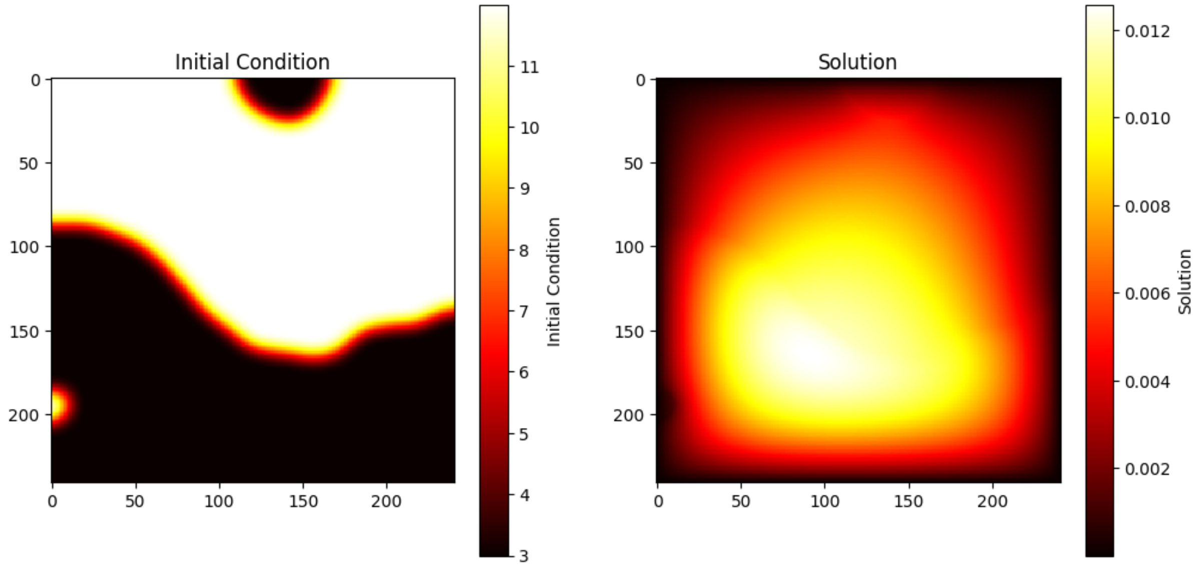

图 3. 5 Darcy 方程的一个样本的初始状态和数值解

左侧的图像显示了样本的初始状态，这是通过'Kcoeff'参数表示的。'Kcoeff'参数表示物质的导热系数。在这个热力图中，颜色的深浅代表了导热系数的大小。

右侧的图像则表示了在给定初始条件'Kcoeff'下，通过数值方法求解 Darcy 方程得到的解，我们用'sol'来表示。同样，颜色的深浅代表了解的大小。

3.4 优化器的选择

是一种常用的优化算法，它结合了两种扩展的随机梯度下降方法：Adaptive Gradient Algorithm (AdaGrad) 和 Root Mean Square Propagation (RMSProp)。Adam 通过计算梯度的一阶矩估计和二阶矩估计，为每个参数提供了自适应的学习率。这使得它在处理包含大规模数据和高维参数空间的复杂问题时，能表现出优越的性能。

在 FNO 解偏微分方程的实践中，需要一个能够自适应调整学习率并且能够有效处理大规模数据和高维参数空间的优化器。相比于梯度下降，Adam 不需要在每一步都使用全部数据，因此在处理大规模数据集时，内存需求更低，同时 Adam 的收敛速度更快，因此更适用于 FNO 解偏微方程的实践。

3.5 训练过程

在每个 epoch 开始时，模型被设置为训练模式，并记录训练开始时间。设置训练均方误差为 0。将训练数据集中的每一批数据，将输入 x 和输出 y 移至 GPU，同时清零优化器梯度，然后，将 x 传递给模型，进行向前传播，生成预测输出 out。计算 out 与 y 的

均方误差。得到 MSE 后，执行反向传播算法，计算出每个参数的梯度，使用 Adam 优化器更新参数。在更新参数后，将当前的 MSE 累加到训练误差中。在每个 epoch 结束时，使用学习率调度器调整学习率。

反向传播算法是一种在神经网络中计算梯度的方法，反向传播算法的主要目标是计算损失函数关于网络权重的梯度。从输出层开始，向输入层方向逐层计算并传播误差。对于每一层，计算损失函数关于该层权重的梯度。这个梯度表示如果改变这一层的权重，损失函数会如何变化。计算出梯度后，用这个梯度来更新网络的权重。在实验中使用了 PyTorch 的框架只需要调用.backward()方法就可以自动计算梯度，然后使用 Adan 优化器就可以进行权重更新。

学习率调度器（scheduler）是一种工具，实验中，它每个 epoch 结束动态地调整学习率。在训练深度学习模型的过程中，在开始时使用较大的学习率，以快速地达到一个相对优的解，然后随着训练的进行，我们逐渐减小学习率，以更精细地优化模型。同时可以在一定程度上防止过拟合的发生。

3.6 评估指标

MSE 用于评估模型对偏微分方程解的预测准确性。在每个训练周期后，通过计算训练集和验证集上的 MSE，监控模型的学习进度和过拟合情况。MSE 计算公式为

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{4}$$

其中y_i是实际值，$\hat{y}_i$是模型预测值，n是样本数量。

第 4 章　结果分析与讨论

本章详细介绍了实验设置，并分析了 Fourier 神经算子（FNO）在求解一维 Burgers 方程和二维 Darcy 方程中的应用效果。实验部分重点探讨了通过调整模型关键参数如学习率、傅里叶模式的数量、网络宽度等来提升模型的预测准确度。具体内容包括数据集的准备、模型架构与参数配置、实验流程描述以及性能评估。此外，本章还通过图表形式直观展示了不同参数配置下模型在训练和测试集上的表现，最终确定了最优参数组合。最后分析了当前实验存在的限制。

4.1　实验设置

本实验的目的是探索和分析 Fourier 神经算子（FNO）在求解一维 Burgers 方程和二维 Darcy 方程中的应用效果。通过实验比较不同参数配置下 FNO 的性能，最终确定最佳参数组合以提高模型的预测准确度和泛化能力。主要从以下四个方面介绍这一实验：数据集准备、模型架构与参数配置、实验流程、以及性能评估。

4.1.1　数据集的准备

对于一维 Burgers 方程，数据集包含 2048 个样本，其中 1000 个用于训练，100 个用于测试。通过对数据进行子采样，通过每隔 8 个数据点取一个数据点来减少每个样本的特征数量，并将其重塑成模型所需的格式。其中使用 PyTorch 的 DataLoader，因为 DataLoader 提供的 shuffle 功能可以确保了每次迭代时数据的随机性。

对于二维 Darcy 方程的，数据集包含 1024 个样本，其中 1000 个用于训练，1000 个用于测试。首先，使用 MatReader 从指定文件中读取数据，并通过每隔 5 个数据点取一个数据点将数据维度从[241, 241] 缩减到[49, 49]。之后，对输入和输出数据进行标准化。将数据转换为均值为 0，标准差为 1 的分布，这可以帮助我们更好地训练模型。，与生成的二维网格坐标结合，形成模型的输入。同样，使用 PyTorch 的 DataLoader 进行数据加载，优化内存使用并支持批量处理。

4.1.2　模型架构与参数配置

在解 Burgers 方程时，我们使用基于 Fourier 神经算子的 1 维模型，在解 Darcy 方程

时，我们使用基于 Fourier 神经算子的 2 维模型。该模型在解决偏微分方程方面展现出了优异的性能。选择了学习率、学习率衰减步长、衰减率、傅里叶模式数以及网络宽度 5 个参数作为优化的重点。设计 run_experiment 函数，该函数接受这些参数作为输入，并执行模型的训练与评估流程，该函数返回在训练集和测试集上的 l2 损失。

4.1.3 实验流程

模型训练在 Kaggle 平台上进行，使用 Adam 优化器和 L2 损失函数。训练过程中，批量大小设置为 20，总共进行 500 个训练周期。学习率调度器根据预设的步长和衰减率调整学习率，以期在训练过程中实现更好的收敛性能。通过一系列循环自动化地运行实验，分别测试各个预定义的学习率、学习率衰减步长、衰减率、傅里叶模式数和网络宽度。每次循环中，run_experiment 函数被调用，以当前循环的参数值作为输入。实验结果，包括训练损失和测试集上的 L2 损失，被记录下来以评估模型性能。

4.1.4 性能评估

模型的性能通过计算测试集上的归一化 L2 损失来评估。这一指标能够量化模型输出与真实解之间的差异，是评价模型预测精度的重要指标。实验的最终目标是确定一组最优的超参数配置，以最小化测试集上的 L2 损失。

4.2 模型性能对比

在本节中，展示了使用 Fourier 神经算子原理解偏微分方程的深度学习模型的实验结果。图表直观地显示了不同模型参数下的训练集和测试集的均方误差（MSE）。图 4.1.1-4.1.5 展示了随着参数的变化，一维 Fourier 神经算子在训练集上和测试集上的 MSE 如何变化。每个图表都精心设计了双 y 轴。左侧 y 轴以蓝色表示训练集 MSE，右侧 y 轴用红色表示测试集 MSE。双 y 轴的设计有助于发现潜在的过拟合问题。

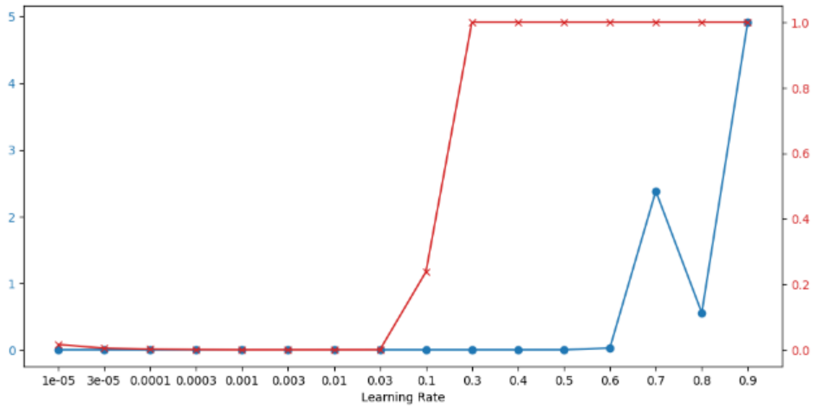

图 4. 1. 1 一维 FNO 模型对学习率调优效果图

学习率决定模型参数每次更新的步长。这里的学习率是指在 Adam 优化器中的全局学习率。如图 4.1.1 所示，当学习率超过 0.1 时，模型出现了明显的过拟合现象。这是由于过高的学习率使模型在训练集上快速找到了一个局部最优解。这个局部最优解可能并不是全局最优解，也可能并不适用于测试集。因此，模型在测试集上的准确度较差

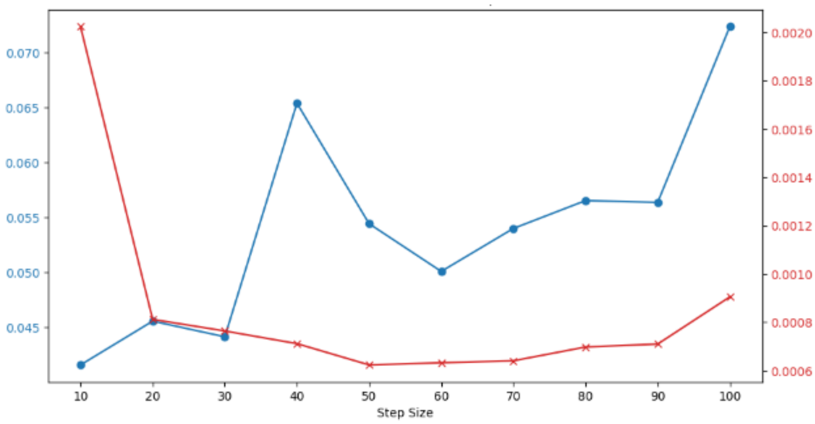

图 4. 1. 2 一维 FNO 模型对 stepsize 调优效果图

图 4.1.2 展示 stepsize 对模型精度的影响。当 stepsize 设置过大时，模型的离散程度降低，导致模型无法精确地捕捉到微分方程解的细节，模型在训练集和测试集上的准确度都较低。当 stepsize<20 时，训练集和测试集的准确度之间的差距开始显著增大，此时发生了过拟合。

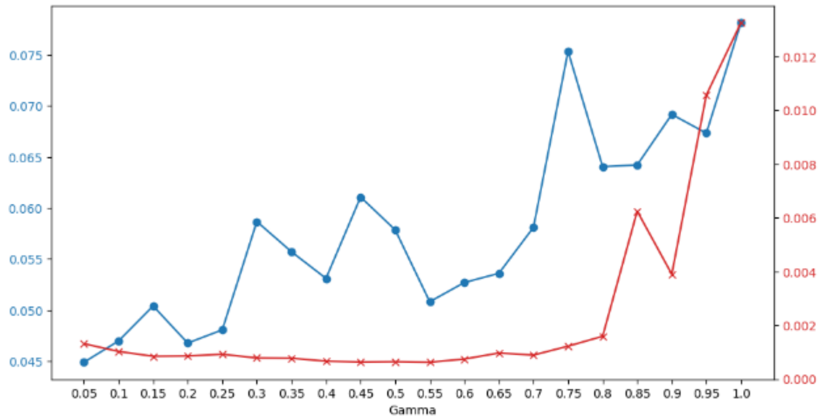

图 4. 1. 3 一维 FNO 模型对 Gamma 调优效果图

Gamma 用于计算梯度的指数加权平均。当 Gamma 值过大，模型对新的梯度信息反应慢，导致训练过程很长。图 4.1.3 显示出当 Gamma 值处于 0.7 到 1 之间时，模型出现了过拟合的现象。

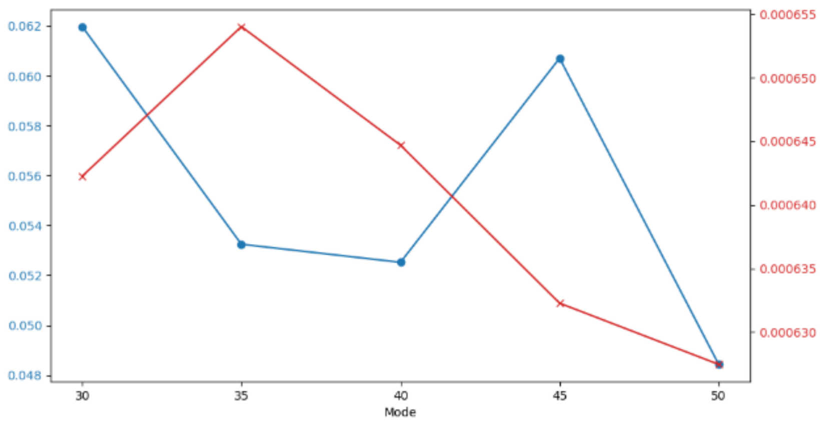

图 4. 1. 4 一维 FNO 模型对 mode 调优效果图

mode 指的是频率组件的数量，也就是在傅立叶神经算子中使用的正弦和余弦函数的数量。这个数量决定了模型能够捕获的信号复杂度。过多的 mode 会引发梯度消失或梯度爆炸的问题。如果 mode 数量较少，模型无法捕获到复杂的信号。

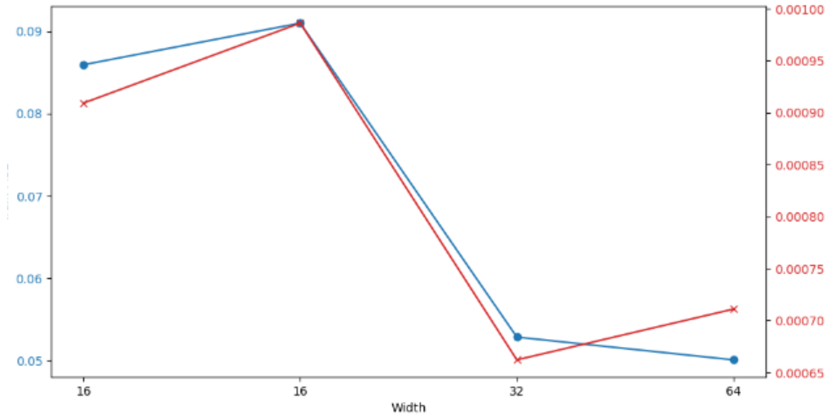

图 4. 1. 5 一维 FNO 模型对 width 调优效果图

width 指的是神经网络中隐藏层的宽度，也就是每一层的神经元数量。当 width 的值增大时，模型的参数数量也会相应增大。每次前向传播和反向传播时，都需要进行更多的计算资源。Width 越大，模型的训练和预测速度就会越慢。Width 越小，模型的预测值就越不精确。

图 4.2.1-4.2.5 展示了随着参数的变化，Darcy 方程在训练集上和测试集上的 MSE 如何变化。同样使用了双 y 轴的设计。

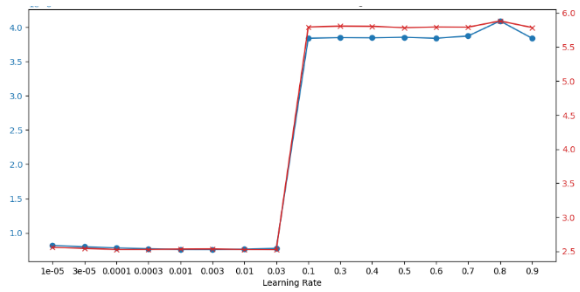

图 4. 2. 1 二维 FNO 模型对学习率调优效果图

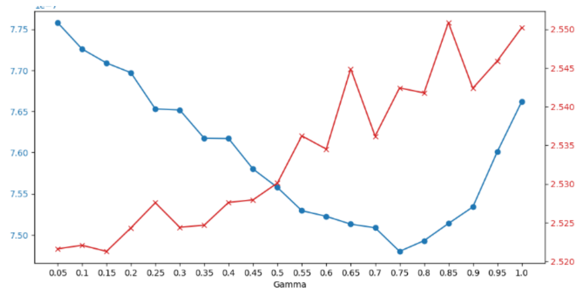

图 4. 2. 7 二维 FNO 模型对 Gamma 调优效果图

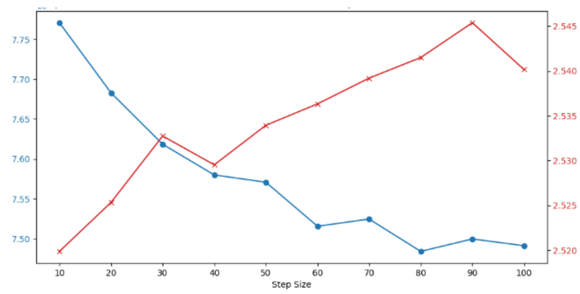

图 4. 2. 6 二维 FNO 模型对 stepsize 调优效果图

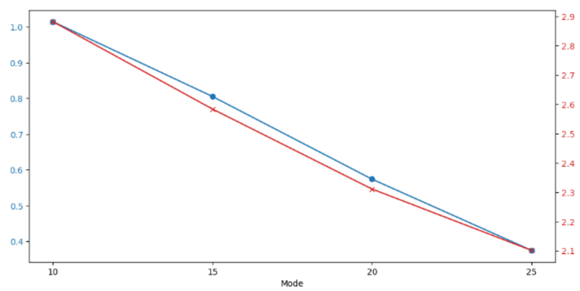

图 4. 2. 8 二维 FNO 模型对 mode 调优效果图

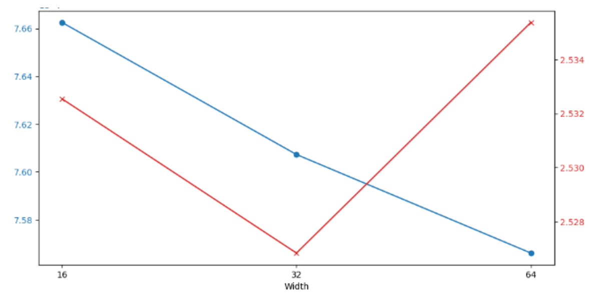

图 4. 2. 9 一维 FNO 模型对 width 调优效果图

通过多轮实验，调整神经网络的关键参数，经过一系列的实验对比，最终找到最佳的一组参数组合: learning_rate=0.003，step_size=60，Gamma=0.45，mode=20，width=32。这一参数组合有效提高了模型在求解微分方程上的准确性，同时在计算效率上的表现也较为合适。

优化后的模型在一维和二维的 FNO 网络中均表现出出色的性能，这种一维和二维性能的一致性也验证了这一参数组合对 FNO 网络性能提升的鲁棒性，即使在不同的问题和数据集下，这种参数组合仍然能够有效地提高模型的性能，而不会出现明显的性能下降或不稳定现象。

4.3　优化后求解的具体例子

使用优化后的参数进行训练后，取数据集中最后一个方程作为例子，对比模型的预测解与方程的实际解。Burgers 方程的对比结果如图 4.3 所示：

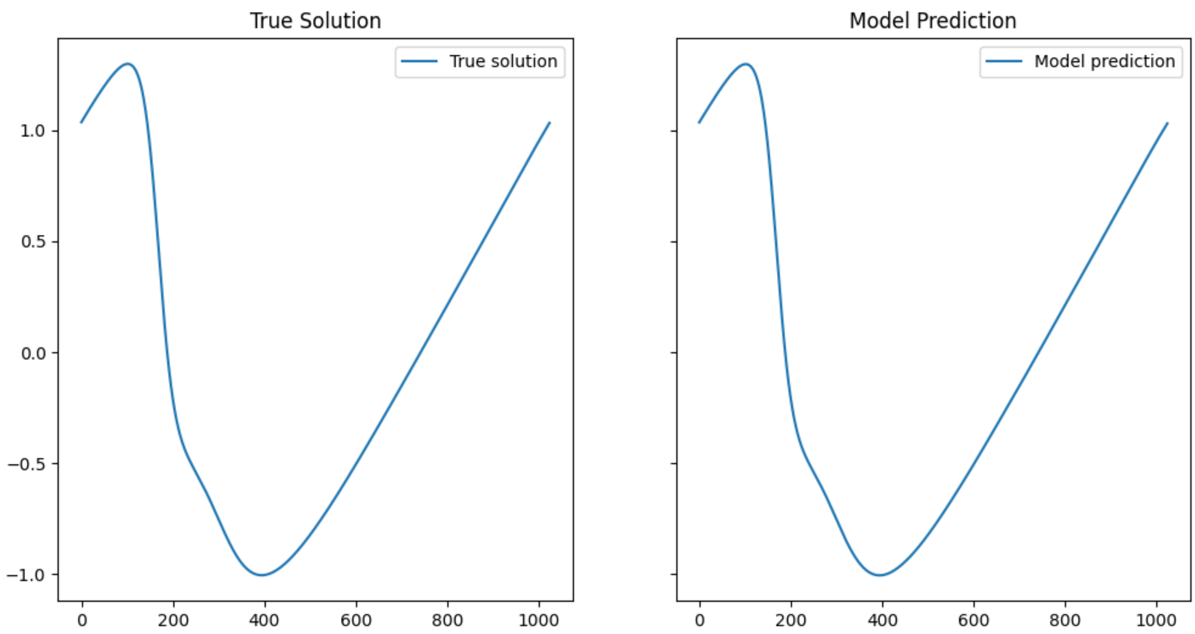

图 4.　3 Burgers 方程预测解与真实解的对比

Darcy 方程的对比结果如图 4.4 所示：

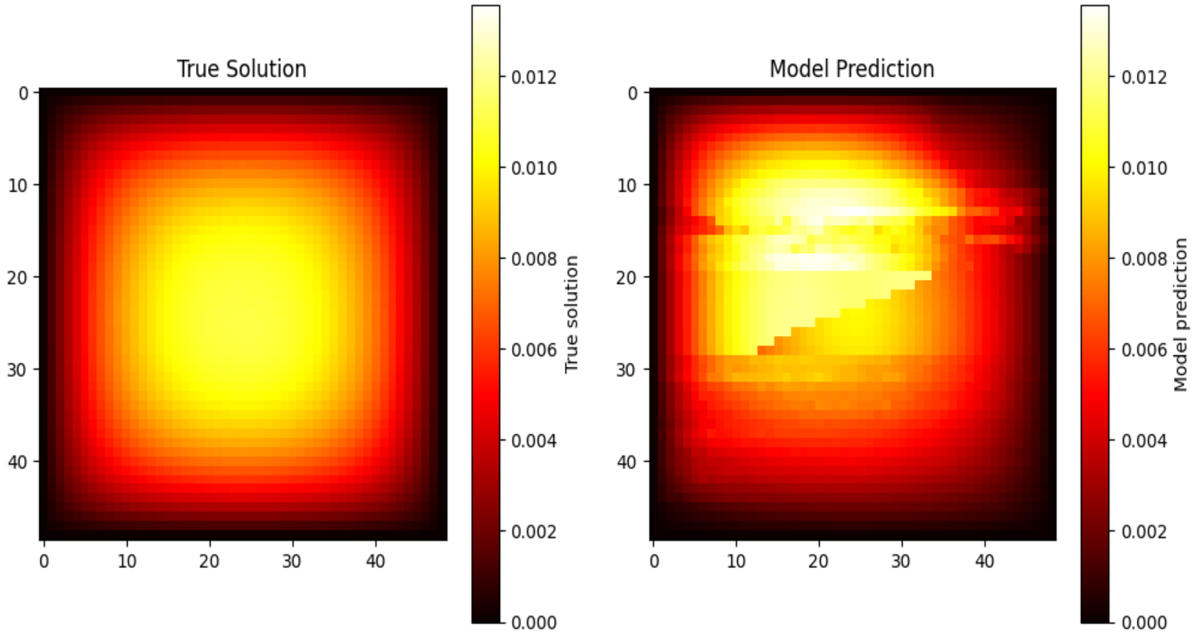

图 4. 4 Darcy 方程预测解与真实解的对比

在实验中，对优化后的 Fourier 神经算子的模型对微分方程的求解结果与真实解进行比较。如图 4. 3 和 4. 4 所示，模型预测的解在结构上与真实解相似，这说明模型能够有效地捕捉到微分方程的主要特性。然而在细节上，模型预测的解与真实解之间存在一些差异。

4.4 实验的限制

在本次实验中，使用 Fourier 神经算子（FNO）的模型对微分方程的求解，研究结果受限于若干关键因素。首先，模型依赖于一个较小的数据集进行模型训练，这可能导致了对问题全貌的捕捉不足，进而限制了模型在更广泛场景中的泛化能力。其次，实验在 Kaggle 平台上进行，由于是使用平台的计算资源，对更大模型架构的探索以及超参数的调整都受到了诸多限制，这些都是提升模型性能的潜在因素。此外，虽然观察到特定参数配置能够提升模型性能，但对于这些参数为何有效的深入理解尚不充分，这限制了对模型的调整，同时可能影响到这些优化在其他种类的偏微分方程的适用性。

模型预测的解在结构上与真实解相似，但在细节上仍然存在一些偏差。这是由于模型容量、训练数据的数量、优化算法的选择等因素的影响。在未来工作中需要进一步改进和优化这些因素。

尽管存在这些限制，研究成果为使用深度学习技术解决复杂偏微分方程提供了有益

的见解，并指出了未来研究可能的改进方向，包括扩充数据集、增强计算资源以及深化对模型参数与性能关系的分析。

结　论

利用傅里叶神经算子（FNO）原理，求解一维和二维的偏微分方程，实验结果表明，通过对 FNO 的结构和训练过程中的参数进行优化，神经网络的准确性和泛化能力得到了显著增强。这些改进不仅提高了在低维 PDEs 上的求解效率，而且预示了在处理更高维度 PDEs 时深度学习方法的巨大潜力[16]。

通过细致的参数调整，深度学习模型能够更好地捕捉和学习 PDEs 的复杂动态[17]，从而提供更精确的解决方案。尽管当前的实验是在较低维度的 PDEs 上进行的，但已经可以看出，这种性能提升对于更高维度问题的求解将是至关重要的。在高维空间，解偏微分方程的复杂性迅速增加，而传统方法在求解这些高维偏微分方程的过程中会遭遇“维数灾难”[10]。

但在未来的研究中，仍需探索如何在保持模型性能的同时，进一步提升模型的计算效率和解释性。此外，我们还需在更大规模的数据集上验证模型的泛化能力，并探讨模型在更高维度 PDEs 求解中的应用。

参考文献

[18] 郭东亮.基于深度学习的高维非线性偏微分方程数值解法[D].南昌大学,2022.

[19] 曾壬源.基于深度神经网络的偏微分方程求解[D].中国科学技术大学,2022.

[20] 查文舒,李道伦,沈路航,等.基于神经网络的偏微分方程求解方法研究综述[J].力学学报,2022,54(03):543-556.

[21] 曹富军,郭晓斌,高飞,等.求解偏微分方程的卷积迭代方法[J].山西大学学报(自然科学版),2023,46(02):293-303.DOI:10.13451/j.sxu.ns.2022065.

[22] 汪京徽. 基于 Feynman-Kac 公式的偏微分方程的蒙特卡洛及神经网络方法[D].上海财经大学,2023.DOI:10.27296/d.cnki.gshcu.2023.000323.

[23] 李务杨. 异质偏微分方程的区域分解和深度学习算法研究[D].东北师范大学,2023.DOI:10.27011/d.cnki.gdbsu.2022.000166.

[24] Qi K ,Sun J . Gabor-Filtered Fourier Neural Operator for solving Partial Differential Equations [J]. Computers and Fluids, 2024, 274 106239-.

[25] 王江元,杨耐,张英浩,等.基于深度学习方法求解偏微分方程的应用举例[J].信息与电脑(理论版),2021,33(21):48-51.

[26] 任清华. 基于深度学习的偏微分方程求解方法[D].山东大学,2023.DOI:10.27272/d.cnki.gshdu.2022.003244.

[27] 孟柳君. 基于多项式神经网络解偏微分方程[D].云南财经大学,2022.DOI:10.27455/d.cnki.gycmc.2022.000107.

[28] Li Z, Kovachki N, Azizzadenesheli K, et al. Fourier Neural Operator for Parametric Partial Differential Equations[J]. ICRL, 2021.

[29] 高普阳,赵子桐,杨扬.基于卷积神经网络模型数值求解双曲型偏微分方程的研究[J].应用数学和力学,2021,42(09):932-947.

[30] 陈静飞.机器学习方法求解偏微分方程：从网络结构与优化算法出发[D].苏州大学,2023.

[31] 毛超利.基于深度学习的偏微分方程求解方法[J].智能物联技术,2021,4(05):18-23+30.

[32] 孙靖威.基于深度学习求解偏微分方程的研究[D].天津师范大学,2024.

[33] 韦昌,樊昱晨,周永清,等.基于 Runge-Kutta 的自回归物理信息神经网络求解偏微分方程[J/OL].力学学报:1-13[2024-05-01].http://kns.cnki.net/kcms/detail/11.2062.O3.20240422.1549.002.html.

[34] DAI Y ,AN Y ,LI Z , et al.Fourier neural operator with boundary conditions for efficient prediction of steady airfoil flows[J].Applied Mathematics and Mechanics(English Edition),2023,44(11):2019-2038.

致 谢

在本论文的写作过程中，我收获颇丰，不仅积累了新的知识，对于问题的分析和解决能力也有了显著的提升。这一过程中，我深切地感受到了来自许多人的帮助和支持。

首先，我要衷心感谢我的指导教师王建林教授，他严谨的学术态度对我影响深远。在论文的选题、研究方法的确定以及实验设计等各个环节，王建林教授都给予了我细致的指导和宝贵的建议。

同时，特别感谢 Kaggle 平台提供的 GPU 资源支持，这为我进行模型训练和实验提供了极大的便利和加速。在研究过程中，GPU 的高效计算能力帮助我快速验证了各种假设和方法，为论文的完成提供了有力的支持。

此外，我还要感谢我的同学和朋友们，在整个学习和研究过程中给予我无私的帮助和支持。他们的鼓励和陪伴使我能够在学术探索的道路上不断前行。

最后，我要感谢我的家人，他们对我的学业和生活给予了无条件的支持和理解，是我坚持下去的最大动力。

感谢各位评阅本论文以及答辩组的各位老师们，感谢你们能够抽出宝贵的时间对我的论文进行评阅并提出宝贵的意见。