Jian Luo

University of Science and Technology of China, Hefei
jianluo@mail.ustc.edu.cn | https://smart-jluo.github.io/

RESEARCH INTERESTS

LLM inference, Knowledge Editing, Prompt Compression, LLM agent, Natural Language Processing, AI4PDE

EDUCATION

University of Science and Technology, China September 2022 — Now
M.E. in Electronic Engineering GPA: 3.76 / 4.30

Central China Normal University September 2017 — June 2021
B.E. in Electronic Engineering GPA: 85 / 100

PUBLICATION

- Neural Krylov Iteration for Accelerating Linear System Solving
  Luo, J., Wang, J., Wang, H., Dong, H., & Wang, J. In Thirty-Eighth Annual Conference on Neural Information Processing Systems, NeurIPS **(Accepted as Spotlight)**.
- Accelerating PDE Data Generation via Differential Operator Action in Solution Space
  Dong, H., Wang, H., Liu, H., **Luo, J.**, & Wang, J. In Forty-first International Conference on Machine Learning, ICML.

RESEARCH EXPERIENCE

University of Science and Technology of China September 2022 — Now
Master's Student

- Explored AI for Science problems, especially how to incorporate AI models into traditional solvers and improve data-driven methods. Developed an AI-aided solver for solving linear systems and contributed to a data-driven eigenvalue problem solution. Contributed to the implementation of operator action code.
- Contributed to the knowledge graph completion project. Identified and resolved the label imbalance problem. Designed and implemented several sampling strategies.

Stony Brook University October 2023 — May 2024
Research Intern

- Identified the translation symmetry problem in FNO. Developed a dynamic kernel to enhance the expressiveness of the Fourier Neural Operator.
- Participated in the Geometric Deep Learning seminar. Presented several times on group theory concepts and classic papers in Geometric Deep Learning.

RESEARCH PROJECT

Accelerating Linear System Solving via Neural Krylov Iteration
Accepted to NIPS'24, Spotlight

- Traditional numerical solvers based on the Krylov subspace iteration algorithm suffer from the low-efficiency problem.
- We addressed this challenge by training a neural operator to predict the invariant subspace of the given linear system, using QR decomposition for stability and a projection loss function as our training objective.
- Experimental results show that our algorithm achieves a 5.5x speedup compared with GMRES.

Accelerating PDE Data Generation via Operator Action
Published in ICML'24

- One significant challenge for Neural Operators is their reliance on vast amounts of high-precision training data, which require substantial computational resources to generate.
- To address this challenge, we propose a novel PDE dataset generation algorithm, which generates basic PDE solutions, combines them, and applies differential operators to efficiently produce accurate PDE data points.
- Our algorithm boosts large-scale dataset generation by 300 times, with resulting model performance nearly equivalent to that trained on the original dataset.

Dynamic Schwartz-Fourier Neural Operator for Enhanced Expressive Power
Submitted to CVPR'25

- One significant limitation of Fourier Neural Operators (FNOs) is their reliance on translation-invariant kernels, which restrict their expressiveness in systems lacking translational symmetries.
- To address this limitation, we propose a novel approach by introducing Dynamic Kernel Fourier Neural Operators (DSFNOs), which enhance FNOs with dynamic Schwartz operators to learn non-translational symmetric kernels.
- Our method significantly improves the expressive power of FNOs, achieving superior performance across multiple tasks and enabling better representation of complex physical phenomena.

Neural Projector for Large Scale Eigenvalue Problems

- Eigenvalues problem of large scale matrix is crucial in many fields, but traditional solvers, with their cubic time complexity, are inefficient for practical large-scale problems.
- We introduce a neural operator-based module to predict eigenvalues, modeling the problem as an operator learning task.
- We developed a Krylov-based module that projects matrices, reducing sapce complexity, which enables large matrix eigenvalue prediction.
- Experiments shows that our algorithm achieves up to 28.3-fold acceleration compared to traditional algorithms.

Multi-Prompt Triple Classification for Knowledge Graph Completion

- Triple classification is key for knowledge graph prediction, yet current text-based models are costly (requiring manual templates) or uncontrollable.
- We designed an auto-complete method based on the perplexity evaluation metric, effectively addressing the high costs of manual template design and the uncontrolled generation problem.
- Our algorithm achieved state-of-the-art results across multiple datasets. Moreover, our method converges faster during training with the same pre-trained model and outperforms the baseline in few-shot scenarios.

AWARDS

USTC Second-Class Academic Scholarship September 2023
USTC Second-Class Academic Scholarship September 2022
National Scholarship October 2020
CCNU Silver Academic Scholarship September 2019
CCNU Bronze Academic Scholarship September 2018

SKILLS

- **Programming:** Python, MATLAB, C/C++
- **Tools:** PyTorch, Transformers, Numpy, SciPy, PETSc
- **Lnaguage:** Mandarin(Native), Engalish(TOEFL 95)