

忻成 (Cheng Xin)

邮箱: xin.job2025@gmail.com | 个人主页: jackal092927.github.io | Google Scholar: jackal092927.github.io/scholar

个人简介

忻成, 计算机科学博士, 现为美国罗格斯大学 (Rutgers University) 计算机科学系博士后研究员, 合作导师为 Jie Gao 教授。2023 年获美国普渡大学 (Purdue University) 计算机科学博士学位, 博士导师为 Tamal K. Dey 教授; 此前在美国理海大学 (Lehigh University) 获计算机科学硕士学位, 本科毕业于同济大学软件工程专业。研究方向包括拓扑数据分析、拓扑机器学习、非欧几何表示学习、可解释人工智能、AI for Science 与 3D/视频生成基准。已在 NeurIPS、ICML、CVPR、SoCG、TMLR、Pattern Recognition、Journal of Applied and Computational Topology 等国际会议和期刊发表论文, 并担任 TAG-DS Workshop 2026 Area Chair 及 ICML、ICLR、NeurIPS、SoCG 等国际会议审稿人。

海外学习与科研经历

- 罗格斯大学 (Rutgers University) 计算机科学系, 博士后研究员, 2023.10 至今。导师: Prof. Jie Gao。围绕可解释图学习、拓扑机器学习、非欧几何表示学习、超曲空间算法、AI for Science 与 3D/视频生成基准开展研究。
- 普渡大学 (Purdue University) 计算机科学系, 博士研究生/研究助理, 2020.08-2023.08。导师: Prof. Tamal K. Dey。博士论文: Decomposition and Stability of Multiparameter Persistence Modules。
- 俄亥俄州立大学 (The Ohio State University) 计算机科学与工程系, 博士研究生/研究助理, 2017.01-2020.08。导师: Prof. Tamal K. Dey。研究多参数持久同调分解算法与稳定性理论。
- 理海大学 (Lehigh University) 计算机科学系, 硕士研究生, 2014.01-2016.05。导师: Prof. Xiaolei Huang。硕士论文: Machine Learning Techniques for Cervigram Image Analysis。

教育背景

- Ph.D. in Computer Science, Purdue University, 美国, 2020.08-2023.08
- Ph.D. student in Computer Science and Engineering, The Ohio State University, 美国, 2017.01-2020.08
- M.S. in Computer Science, Lehigh University, 美国, 2014.01-2016.05
- 软件工程学士, 同济大学, 中国上海, 2009.09-2013.07

主要科研项目与本人贡献

1. 拓扑可解释图学习 TopInG (ICML 2025)

- 项目内容: 提出基于 persistent rationale filtration 的图神经网络可解释框架, 在分子图等任务中同时提升预测性能与解释质量。
- 本人角色: 第一作者, 负责核心问题建模、方法设计、算法实现、实验设计与论文写作。

2. 端到端拓扑学习 D-GRIL (SoCG 2026, to appear)

- 项目内容: 将二维持久同调特征以可微方式融入机器学习模型, 推进多参数拓扑特征在端到端学习中的应用。
- 本人角色: 核心参与者, 参与方法设计、理论分析、算法实现与论文撰写。

3. 非欧几何与超曲空间表示学习系列工作 (NeurIPS 2024、NeurIPS 2025、SoCG 2026)

- 项目内容: 研究非欧多维尺度分析、欧氏空间之外的 Johnson-Lindenstrauss 引理, 以及超曲空间 locality sensitive hashing 等问题, 为非欧表示学习提供理论与算法基础。
- 本人角色: 核心参与者/字母序作者, 参与理论推导、算法设计、实验验证与论文写作。

4. 多参数持久同调分解与稳定性理论 (SoCG 2018; Journal of Applied and Computational Topology 2022; 博士论文 2023)

- 项目内容: 研究多参数 persistence modules 的分解算法、瓶颈距离计算、矩形近似与稳定性, 为拓扑数据分析提供基础算法。
- 本人角色: 主要研究者/字母序作者, 负责算法设计、理论证明、复杂度分析与论文写作。

5. DL3DV-10K 大规模 3D 视觉数据集与基准 (CVPR 2024)

- 项目内容: 构建面向深度学习 3D 视觉、novel view synthesis 与视频生成的大规模场景数据集。
- 本人角色: 参与数据集与基准相关工作, 支持 3D/视频生成方向的实验与论文产出。

6. 医学图像机器学习与宫颈病变分类基准 (MLMI 2015; Pattern Recognition 2017)

- 项目内容: 构建宫颈图像疾病分类数据集与多特征机器学习基准, 比较多种经典机器学习模型在医学图像分类中的表现。
- 本人角色: MLMI 2015 共同第一作者, 参与特征工程、模型评估、实验分析与论文写作。

论文发表情况

注: † 表示作者按姓氏字母顺序排序; * 表示共同第一作者。

1. Chengyuan Deng, Jie Gao, Kevin Lu, Feng Luo, and **Cheng Xin**[†]. "Locality Sensitive Hashing in Hyperbolic Space." 42nd International Symposium on Computational Geometry (SoCG), to appear, 2026. arXiv:2603.19724. 类别: 国际理论计算会议; 本人角色: 字母序作者。

2. Soham Mukherjee, Shreyas N. Samaga, **Cheng Xin**, Steve Oudot, and Tamal K. Dey. "D-GRIL: End-to-End Topological Learning with 2-parameter Persistence." 42nd International Symposium on Computational Geometry (**SoCG**), to appear, 2026. arXiv:2406.07100. 类别: 国际理论计算会议; 本人角色: 核心参与。
3. **Cheng Xin**, Fan Xu, Xin Ding, Jie Gao, and Jiaxin Ding. "TopInG: Topologically Interpretable Graph Learning via Persistent Rationale Filtration." The 42nd International Conference on Machine Learning (**ICML**), 2025. 类别: 机器学习顶级国际会议; 本人角色: 第一作者。
4. Chengyuan Deng, Jie Gao, Kevin Lu, Feng Luo, and **Cheng Xin**[†]. "Johnson-Lindenstrauss Lemma Beyond Euclidean Geometry." The 39th Advances in Neural Information Processing Systems (**NeurIPS**), 2025. 类别: 机器学习顶级国际会议; 本人角色: 字母序作者。
5. Chengyuan Deng, Jie Gao, Kevin Lu, Feng Luo, Hongbin Sun, and **Cheng Xin**[†]. "Neuc-MDS: Non-Euclidean Multidimensional Scaling Through Bilinear Forms." Advances in Neural Information Processing Systems (**NeurIPS**), Vol. 37, 2024, pp. 121539-121569. 类别: 机器学习顶级国际会议; 本人角色: 字母序作者。
6. Shahrzad Haddadan, **Cheng Xin**, and Jie Gao. "Optimally Improving Cooperative Learning in a Social Setting." Proceedings of the 41st International Conference on Machine Learning (**ICML**), PMLR 235, 2024, pp. 17148-17188. 类别: 机器学习顶级国际会议; 本人角色: 核心参与。
7. Lu Ling, Yichen Sheng, Zhi Tu, Wentian Zhao, **Cheng Xin**, Kun Wan, Lantao Yu, Qianyu Guo, Zixun Yu, Yawen Lu, et al. "DL3DV-10K: A Large-Scale Scene Dataset for Deep Learning-Based 3D Vision." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2024, pp. 22160-22169. 类别: 计算机视觉顶级国际会议; 本人角色: 参与作者。
8. Simon Zhang, **Cheng Xin**, and Tamal K. Dey. "Expressive Higher-Order Link Prediction through Hypergraph Symmetry Breaking." Transactions on Machine Learning Research (**TMLR**), 2024. 类别: 机器学习国际期刊; 本人角色: 核心参与。
9. **Cheng Xin**. "Decomposition and Stability of Multiparameter Persistence Modules." Ph.D. Thesis, Purdue University Graduate School, 2023. DOI:10.25394/PGS.23848995.v1.
10. **Cheng Xin**, Soham Mukherjee, Shreyas N. Samaga, and Tamal K. Dey. "GRIL: A 2-parameter Persistence Based Vectorization for Machine Learning." Proceedings of Machine Learning Research, Vol. 221, 2023, pp. 313-333. 类别: 机器学习/拓扑学习会议论文; 本人角色: 第一作者。
11. Tamal K. Dey and **Cheng Xin**[†]. "Generalized Persistence Algorithm for Decomposing Multiparameter Persistence Modules." Journal of Applied and Computational Topology 6(3), 2022, pp. 271-322. 类别: 国际期刊; 本人角色: 字母序作者。
12. Tamal K. Dey and **Cheng Xin**[†]. "Rectangular Approximation and Stability of 2-parameter Persistence Modules." arXiv:2108.07429, 2021. 类别: 预印本; 本人角色: 字母序作者。
13. Tamal K. Dey and **Cheng Xin**[†]. "Computing Bottleneck Distance for 2-D Interval Decomposable Modules." 34th International Symposium on Computational Geometry (**SoCG**), LIPIcs 99, 2018, 32:1-32:15. 类别: 国际理论计算会议; 本人角色: 字母序作者。
14. Tao Xu, Han Zhang, **Cheng Xin**, Edward Kim, L. Rodney Long, Zhiyun Xue, Sameer Antani, and Xiaolei Huang. "Multi-feature Based Benchmark for Cervical Dysplasia Classification Evaluation." Pattern Recognition 63, 2017, pp. 468-475. 类别: SCI/SCIE 期刊; 本人角色: 参与作者。
15. Tao Xu, **Cheng Xin**^{*}, L. Rodney Long, Sameer Antani, Zhiyun Xue, Edward Kim, and Xiaolei Huang. "A New Image Data Set and Benchmark for Cervical Dysplasia Classification Evaluation." Machine Learning in Medical Imaging (**MLMI**), Springer, 2015, pp. 26-35. 类别: 医学图像机器学习会议; 本人角色: 共同第一作者。

教学经历

- 研究生课程 Lecturer: Design and Analysis of Algorithms, Rutgers University, 2025 年秋季, 约 45 名学生。
- 本科生课程助教: Data Structures and Algorithms, Purdue University, 2023 年春季, 约 200 名学生。
- 研究生课程助教: Computational Geometry, Purdue University, 2020 年秋季, 约 30 名学生。

学术服务

- Area Chair, TAG-DS Workshop, 2026。
- 审稿人: ICML、ICLR、NeurIPS、SoCG。

邀请报告

- "TopInG: Topologically Interpretable Graph Learning via Persistent Rationale Filtration," Conference on Topological Data Analysis: Recent Developments and Applications, University of Missouri, 2025.11.
- "Understanding through Shape of Data: Topological Data Analysis for Interpretable AI," Management Science and Information Systems Department Colloquium, Rutgers University, 2024.10.
- "Exploring Representations Beyond Euclidean Geometry," John Hopcroft Center Seminar, Shanghai Jiao Tong University, 2024.06.
- "Multiparameter Persistence and Its Applications," Theory Seminar, Department of Computer Science, Rutgers University, 2023.11.
- "Generalized Persistence Algorithm for Decomposing Multi-parameter Persistence Modules," Applied Algebraic Topology Network Seminar, 2020.07.

业界与工程经历

- Machine Learning Scientist Intern, Electronic Arts, Big Data Group, 2018.05-2018.08。参与基于 Spark 的大规模机器学习模型、关系型数据库上的图学习、属性评估与选择、数据集压缩。

- Software Development Engineer Intern, Amazon AWS Infrastructure Group, 2015.05-2015.08。开发支持网络消息接收、解析、存储和检索的数据管理系统。
- 开发工程师, 格尔软件, 上海, 2013.07-2013.12。参与后端数据库、业务逻辑、接口和前端 UI 开发。
- Developer Support Intern, Microsoft 数据智能组, 上海, 2012.07-2012.11。参与 SQL Server 相关技术支持案例处理。

专业技能

Python、PyTorch、Spark、Keras、Java、C/C++、MATLAB、R；拓扑数据分析、图机器学习、非欧几何表示学习、机器学习实验设计与大规模数据处理。