Yaochen Xie

Department of Computer Science & Engineering
Texas A&M University
302 L.F. Peterson Building
College Station, TX 77843

Phone: (979) 676-7363
E-mail: ethanycx@tamu.edu
http://people.tamu.edu/~ethanycx/

Education

- Ph.D. Candidate in Computer Science — Aug 2018 – July 2023 (expected)
  Texas A&M University, College Station, TX, USA
- B.S. in Statistics, School of the Gifted Young — Aug 2014 – Jun 2018
  University of Science and Technology of China, Hefei, Anhui, China

Research Summary and Highlights

- I work on machine learning, deep learning, and AI4Science. My research themes include graph neural networks, self-supervised learning, and model explainability.
- I published multiple papers as the first or a co-first author at top-tier conferences and journals, including **NeruIPS**, **ICML**, **JMLR**, **TPAMI**, and **Nature Machine Intelligence**.
- Applying my work to real applications, our team achieved **first place** at the MIT AI Cures challenge for COVID-19 drug discovery and received the runner-up award (**first place of all academic teams**) at KDD Cup 2021.
- I develop an open-sourced library DIG-sslgraph to facilitate the research in self-supervised learning of graph neural networks. Together with other frontier topics in GNN, the library has received **1.3k stars and over 200 forks** from the research community on GitHub.

Research Projects and Industrial Experience

1. Research Project at **Texas A&M University** — 2020 – Now
   Theory-guided self-supervised learning approaches. We study the fundamental theory and methodology of self-supervised learning of deep models to leverage unlabeled graph data to train deep models. Multiple papers were published at NeurIPS, ICML, JMLR, TPAMI, and an open-sourced library for graph machine learning with 1.3k stars.
2. Applied Scientist Intern, **Amazon Search** — May 2022 – Aug 2022
   Causal representation learning for personality and search intent understanding. We propose a causal representation learning framework that captures and utilizes personality and search intent information from user history activities and current session data. The representation can benefit multiple downstream prediction tasks at Amazon and improves customers' search and shopping experience.
3. Research Project at **Texas A&M University** — 2020 – 2021
   Machine learning for drug discovery and quantum chemistry. We develop comprehensive machine-learning methods spanning different computational models (e.g. GNNs), molecular representations, and optimization objectives for molecular property prediction,

drug discovery, and quantum chemistry problems. We achieved 1st place at MIT AI Cures Challenge for COVID-19 drug discovery and the runner-up award at KDD-Cup 2021.

4. Applied Scientist Intern, **Amazon Search** May 2021 – Nov 2021
   Task-agnostic explanations of graph neural networks. We identify and address the task-agnostic explanation problem under the industry setting where models are usually trained in 2-stages and existing explanation methods become inapplicable to explain models trained under self-supervision without downstream tasks. The explanations help improves the reliability of fraud detection models at Amazon. We published one paper at NeurIPS 2022.

5. Research Project at **Texas A&M University** 2019 – 2021
   Advanced Deep Learning Methods for Augmented Microscopy. We develop deep learning approaches using self-supervised learning and the attention mechanism to improve the quality of 2D/3D microscopy images including denoising and super-resolution. Our work boosts biomedical and material research and potentially saves millions of dollars in microscopy instrument purchasing. We have multiple publications at Nature Machine Intelligence, NeurIPS, and TMI, which receive media coverage.

Teaching and Advising Experiences

Courses at Texas A&M University

1. Guest Lecturer, CSCE 636: Deep Learning Fall'19, Fall'20, Fall'21, Fall'22
2. Teaching Assistant, CSCE 638: Natural Language Processing Fall'22
3. Teaching Assistant, CSCE 489: Machine Learning Spring'19

Mentoring

1. **Sai Ramana Reddy**, M.S. student at TAMU 2022 – present
2. **Zhao Xu**, Ph.D. student at TAMU 2020 – present
3. **Xinyi Xu**, Visiting Ph.D. student from Xidian University 2020 – 2021
4. **Jerry Kurtin**, Undergraduate student at TAMU 2020
5. **Jingtun Zhang**, Research Engineer at Tik Tok, past M.S. student at TAMU 2020 – 2021

Grant Proposal Writing Experience

I assist in the writing of the following proposals.

1. C3CLEAR: Collaborative CurriCular Conceptual LEarning, Analysis & Reasoning. DARPA (pending), Assisted Co-PI Shuiwang Ji, Texas A&M University
2. Advanced Deep Learning Methods and Tools for High-Throughput Microconnectivity Analysis. NIH (declined), Assisted PI Shuiwang Ji, Texas A&M University
3. Integrative 2D-3D Graph Representation Learning for Molecular Property Prediction. NSF (declined), Assisted PI Shuiwang Ji, Texas A&M University

4. Administrative Supplement Request for "Genetics of Deep-Learning-Derived Neuroimaging Endophenotypes for Alzheimer's Disease." NIH (awarded), Assisted PI Shuiwang Ji, Texas A&M University

List of Publications

* The authors contributed equally

1. **Yaochen Xie**, Sumeet Katariya, Xianfeng Tang, Edward Huang, Nikhil Rao, Karthik Subbian, Shuiwang Ji
   Task-Agnostic Graph Explanations
   The 36th Annual Conference on Neural Information Processing Systems (**NeurIPS**), 2022

2. **Yaochen Xie***, Zhao Xu*, Shuiwang Ji
   Self-Supervised Representation Learning via Latent Graph Prediction
   International Conference on Machine Learning (**ICML**), 24460–24477, 2022

3. **Yaochen Xie**, Yu Ding, Shuiwang Ji
   Augmented Equivariant Attention Networks for Microscopy Image Reconstruction
   IEEE Transactions on Medical Imaging (**TMI**), 41(11): 3194–3206, 2022

4. Xinyi Xu, Cheng Deng, **Yaochen Xie**, Shuiwang Ji
   Group Contrastive Self-Supervised Learning on Graphs
   IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**), 2022

5. **Yaochen Xie**, Zhao Xu, Zhengyang Wang, Shuiwang Ji
   Self-supervised Learning of Graph Neural Networks: A Unified Review
   IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**), 2022

6. Zhengyang Wang*, Meng Liu*, Youzhi Luo*, Zhao Xu*, **Yaochen Xie***, Limei Wang*, Lei Cai*, Qi Qi, Zhuoning Yuan, Tianbao Yang, Shuiwang Ji
   Advanced Graph and Sequence Neural Networks for Molecular Property Prediction and Drug Discovery
   Bioinformatics, 38(9): 2579–2586, 2022

7. Meng Liu*, Youzhi Luo*, Limei Wang*, **Yaochen Xie***, Hao Yuan*, Shurui Gui, Zhao Xu, Haiyang Yu, Jingtun Zhang, Yi Liu, Keqiang Yan, Bora Oztekin, Haoran Liu, Xuan Zhang, Cong Fu, Shuiwang Ji
   DIG: A Turnkey Library for Diving into Graph Deep Learning Research
   Journal of Machine Learning Research (**JMLR**), 22(240): 1–9, 2021

8. Zhengyang Wang*, **Yaochen Xie***, Shuiwang Ji
   Global Voxel Transformer Networks for Augmented Microscopy
   Nature Machine Intelligence, 3: 161–171, 2021

9. **Yaochen Xie**, Zhengyang Wang, Shuiwang Ji
   Noise2Same: Optimizing A Self-Supervised Bound for Image Denoising
   The 34th Annual Conference on Neural Information Processing Systems (**NeurIPS**), 2020

10. Xiang-Yang Li, Huiqi Liu, Lan Zhang, Zhenan Wu, **Yaochen Xie**, Ge Chen, Chunxiao Wan, Zhongwei Liang
Finding the stars in the fireworks: Deep understanding of motion sensor fingerprint
IEEE/ACM Transactions on Networking 27(5), 1945–1958, 2020

11. Jianjun He, Ye Yu, **Yaochen Xie**, Hongjun Mao, Lin Wu, Na Liu, Suping Zhao
Numerical model-based artificial neural network model and its application for quantifying impact factors of urban air quality
Water, Air, & Soil Pollution 227(7), 1–16, 2016

Awards & Honors

- **Scholar Award**, Annual Conference on Neural Information Processing Systems 2022
- **1st Place**, MIT AI Cures Challenge for COVID-19 Drug Discovery 2021
- **Runner-up Award**, Open Graph Benchmark PCQM4M-LSC Track, KDD Cup 2021 2021
- **Outstanding Award**, The 4th National College Cloud Computing Contest, China 2018
- **Outstanding Student Scholarship**, University of Science and Technology of China 2015

Professional Services

Program Committee Member

- International Conference on Machine Learning (**ICML**) 2022
- International Conference on Learning Representations (**ICLR**) 2022, 2023
- Annual Conference on Neural Information Processing System (**NeurIPS**) 2021, 2022
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**) 2021, 2022

Journal Reviewer

- IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**)
- Transactions on Machine Learning Research (**TMLR**)
- IEEE Transactions on Neural Networks and Learning Systems (**TNNLS**)
- IEEE Transactions on Image Processing (**TIP**)

References

- **Shuiwang Ji**, Professor and Presidential Impact Fellow, Department of Computer Science & Engineering, Texas A&M University
- **Yu Ding**, Mike and Sugar Barnes Professor, Department of Industrial & Systems Engineering, Texas A&M University
- **Karthik Subbian**, Director and Sr. Principal Scientist, Amazon