

Xinyi Yu

Ph.D. Candidate, Chinese University of Hong Kong, Shenzhen
Email: 223040233@link.cuhk.edu.cn

EDUCATION

Ph.D., Data Science

2023/09 – Present

School of Data Science, Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen)
Advisor: Prof. Tianwei Yu

M.Phil., Mathematics

2020/09 – 2022/06

Department of Mathematics, Hong Kong University of Science and Technology (HKUST)
Advisor: Prof. Can Yang

B.Sc., Mathematics

2016/09 – 2020/06

Cuiying Honors College (Top-Notch Undergraduate Training Program), Lanzhou University

RESEARCH INTERESTS

My research focuses on statistical genetics and computational genomics, with an emphasis on developing scalable statistical and machine learning frameworks for multi-omics integration and genetic analysis of complex diseases. I have worked on cell-type-specific eQTL mapping, spatially resolved multicellular niche detection, and cross-population association analysis, leveraging my interdisciplinary background in statistics, computation, and AI-driven modeling.

PUBLICATIONS

Statistical Genetics and Genomics

- [1] **Xinyi Yu**, Xiaomeng Wan, Leqi Tian, Yuheng Chen, Yuyao Liu, Tianwei Yu, Can Yang, Jiashun Xiao (2025). NicheScope: Identifying Multicellular Niches and Niche-Regulated Cell States in Spatial Transcriptomics. *Submitted*. [[preprint](#)][[software](#)]
- [2] **Xinyi Yu**, Xianghong Hu, Xiaomeng Wan, Zhiyong Zhang, Xiang Wan, Mingxuan Cai, Tianwei Yu, Jiashun Xiao (2024). A unified framework for cell-type-specific eQTLs prioritization by integrating bulk and scRNA-seq data. *American Journal of Human Genetics*. [[paper](#)][[software](#)]
- [3] **Xinyi Yu**, Jiashun Xiao, Mingxuan Cai, Yuling Jiao, Xiang Wan, Jin Liu, Can Yang (2023). PALM: A powerful and adaptive latent model for prioritizing risk variants with functional annotations. *Bioinformatics*. [[paper](#)][[software](#)]
- [4] Jiashun Xiao, Mingxuan Cai, **Xinyi Yu**, Xianghong Hu, Gang Chen, Xiang Wan, Can Yang (2022). Leveraging the local genetic structure for trans-ancestry association mapping. *The American Journal of Human Genetic* [[paper](#)][[software](#)]

Computational and AI Methods in Biomedical Image Analysis

- [1] **Xinyi Yu**, Guanbin Li, Wei Lou, Siqi Liu, Xiang Wan, Yan Chen, Haofeng Li (2023). Diffusion-based data augmentation for nuclei image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. [[paper](#)][[software](#)]
- [2] Wei Lou, **Xinyi Yu**, Chenyu Liu, Xiang Wan, Guanbin Li, Siqi Liu, Haofeng Li (2023). Multi-stream Cell Segmentation with Low-level Cues for Multi-modality Images. *Competitions in Neural Information Processing Systems*. [[paper](#)][[software](#)]

RESEARCH AND ACADEMIC EXPERIENCES

Algorithm Engineer, Shenzhen Research Institute of Big Data

- Developed a diffusion-based augmentation method for histology image segmentation.
- Published a first-author paper at *MICCAI 2023* and obtained a national patent.
- Achieved 2nd place in the *NeurIPS 2022* multimodal cell segmentation challenge.

2022/06 – 2023/06

Teaching Assistant
STA4100: Statistical Inference II, CUHK-Shenzhen
STA2001: Probability and Statistics II, CUHK-Shenzhen
STA4100: Statistical Inference II, CUHK-Shenzhen
STA2001: Probability and Statistics I, CUHK-Shenzhen
MATH2011: Introduction to Multivariable Calculus, HKUST
MATH1013: Calculus I, HKUST
MATH2011: Introduction to Multivariable Calculus, HKUST

2025 Spring
2024 Fall
2023 Spring
2023 Fall
2022 Spring
2021 Fall
2021 Spring

Research Intern
Department of Statistics, Pennsylvania State University
Advisor: Prof. Qunhua Li

2019/06 – 2020/01

Visiting Student
University of Wisconsin–Madison
University of California, Berkeley

2019/06 – 2020/01
2018/07 – 2018/08

HONORS AND AWARDS

Dean’s Fellowship, CUHK-Shenzhen
NeurIPS 2022 Medical Image Segmentation Challenge – Winner Finalist Award, SRIBD
RedBird PhD Scholarship, HKUST
National Scholarship (top 1%), Lanzhou University
First-Class Scholarship (top 2%), Lanzhou University
Merit Student of Gansu Province, Lanzhou University

2023
2023
2020
2018
2017
2017

TECHNICAL SKILLS

Programming & Computing: Python (PyTorch for deep learning), R, Linux
Statistical & Machine Learning: Probabilistic modeling, Bayesian inference, high-dimensional data analysis
Genomics & Bioinformatics: Integrative multi-omics analysis, GWAS, eQTL, spatial transcriptomics