

Hongyi Xin

Title: Associate Professor
Email: <hongyi.xin@sjtu.edu.cn>
Phone: +86-21-34206045 Ext. 4361
Lab Website: sites.ji.sjtu.edu.cn/Novellab

Global Institute of Future Technology
Shanghai Jiao Tong University
800 Dongchuan Road
Minhang, Shanghai, 200240

Employment

Shanghai Jiao Tong University
Global Institute of Future Technology
Shanghai Jiao Tong University
UM-SJTU Joint Institute
Shanghai Jiao Tong University
UM-SJTU Joint Institute
University of Pittsburgh
School of Medicine

Assistant Dean and Associate Professor
Nov 2022 – present
Associate Professor
Jan 2022 – Nov 2022
Assistant Professor
Aug 2020 – Dec 2021
Postdoctoral Fellow
Mar 2018 – May 2020

Education

Carnegie Mellon University
Aug 2011 – Feb 2018
University of Michigan
Aug 2009 – May 2011
Shanghai Jiao Tong University
Aug 2007 – May 2011

Ph.D. Computer Science
B.S. Computer Engineering
B.S. Electrical Engineering

Publications

- Meihong Pan, Hongyi Xin (*co-corresponding*), Chunqiu Xia, and Hongbin Shen. “Few-Shot Classification with Task-Adaptive Semantic Feature Learning” pre-print publication in *Pattern Recognition*, Volume 141, September 2023.
- Wang, Anqi, Jieli Zhou, Gang Wang, Beibei Zhang, Hongyi Xin (*co-corresponding*), and Haiyang Zhou. “Deep Learning of Endoscopic Features for Assessment of Neoadjuvant Therapy Response in Locally Advanced Rectal Cancer”, in *Asian Journal of Surgery*, online e-publication, April 2023.
- Dan Zhong, Haowei Jiang, Chengzhuo Zhou, Abrar Ahmed, Hongji Li, Xiaona Wei, Qiuyu Lian, Melodi Tastemel, Hongyi Xin, Mei Ge, Chenhong Zhang, Lili Jing “The Microbiota Regulates Hematopoietic Stem and Progenitor Cell Development by Mediating Inflammatory Signals in the Niche”, in *Cell Reports*, Volume 42, Issue 2, February 2023.
- Jieli Zhou and Hongyi Xin. “Emerging artificial intelligence methods for fighting lung cancer: A survey”, in *Clinical EHealth*, Volume 5, 19–34, December 2022.
- Dawei Yang, Jian Zhou, Rongchang Chen, Yuanlin Song, Zhenju Song, Xiaoju Zhang, Qi Wang, Kai Wang, Chengzhi Zhou, Jiayuan Sun, Lichuan Zhang, Li Bai, Yuehong Wang, Xu Wang, Yeting Lu, Hongyi Xin, Charles A. Powell, Christoph Thüemmler, Niels H. Chavannes, Wei Chen, Lian Wu, Chunxue Bai. “Expert Consensus on the Metaverse in Medicine”, in *Clinical EHealth*, Volume 5, 1–9, December 2022.
- Yiran Shan, Qian Zhang, Wenbo Guo, Yanhong Wu, Yuxin Miao, Hongyi Xin, Qiuyu Lian, and Jin Gu. “TIST: Transcriptome and Histopathological Image Integrative Analysis for Spatial Transcriptomics”, in *Genomics, Proteomics & Bioinformatics*, Volume 20, Issue 5, October 2022.
- Jieli Zhou, Baoyu Jing, Zeya Wang, Hongyi Xin (*co-corresponding*), Hanghang Tong. “SODA: Detecting COVID-19 in Chest X-rays with Semi-supervised Open Set Domain Adaptation”, in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Volume 19, Issue 5, March 2021.

- Hongyi Xin, Qiuyu Lian, Yale Jiang, Jiadi Luo, Xinjun Wang, Carla Erb, Xiaoyi Zhang, Elisa Heidrich-O’Hare, Qi Yan, Richard Duerr, Kong Chen, Wei Chen. “GMM-Demux: sample demultiplexing, multiplet detection, experiment planning, and novel cell-type verification in single cell sequencing”, in *Genome Biology*, Volume 21, Issue 1, July 2020.
- Qiuyu Lian, Hongyi Xin (*co-first*), Jianzhu Ma, Lina Konnikova, Wei Chen, Jin Gu, Kong Chen. “Artificial-Cell-Type Aware Cell Type Classification in CITE-seq”, in *Oxford Bioinformatics*, Volume 33 Issue Supplement_1, November 2020
- Hongyi Xin, Mingfu Shao, Carl Kingsford. “Context-Aware Seeds for Read Mapping”, in *Algorithms in Molecular Biology*, Volume 15, Issue 10, May 2020.
- Xinjun Wang, Zhe Sun, Yanfu Zhang, Zhongli Xu, Hongyi Xin, Heng Huang, Richard H Duerr, Kong Chen, Ying Ding, Wei Chen. “BREM-SC: a bayesian random effects mixture model for joint clustering single cell multi-omics data”, in *Nucleic Acids Research*, Volume 48, Issue 11, May 2020.
- Qi Yan, Daniel Weeks, Hongyi Xin, Anand Swaroop, Emily Chew, Heng Huang, Ying Ding, Wei Chen. “Deep-Learning-Based Prediction of Late Age-Related Macular Degeneration Progression”, in *Nature Machine Intelligence*, Volume 2, Issue 2, 141-150, February 2020.
- KR Jayaram, Anshul Gandhi, Hongyi Xin, Shu Tao. “Adaptively Accelerating Map-Reduce/Spark with GPUs: A Case Study”, in Proceedings of the 16th The International Conference on Autonomic Computing (*ICAC 2019*), Umea, Sweden, June 2019.
- Zhe Sun, Li Chen, Hongyi Xin, Qianhui Huang, Anthony Cillo, Tracy Tabib, Ying Ding, Jay Kolls, Tullia Bruno, Robert Lafyatis, Dario Vignali, Kong Chen, Ming Hu, and Wei Chen. “BAMM-SC: A Bayesian Mixture Model for Clustering Droplet-Based Single Cell Transcriptomic Data from Population Studies”, in *Nature Communications*, Volume 10, Issue 1, April 2019.
- Jeremie Kim, Damla Cali, Hongyi Xin, Donghyuk Lee, Saugata Ghose, Mohammed Alser, Hasan Hassan, Oğuz Ergin, Can Alkan, Onur Mutlu. “GRIM-Filter: Fast Seed Location Filtering in DNA Read Mapping Using Processing-in-Memory Technologies”, in *BMC Genomics*, Volume 19, Issue 2, May 2018.
- Hongyi Xin, Jeremie Kim, Sunny Nahar, Can Alkan, Onur Mutlu. “LEAP: A Generalization of the Landau-Vishkin Algorithm with Custom Gap Penalties”, in Proceedings of the 7th RECOMB Satellite Workshop on Massively Parallel Sequencing (*RECOMB-seq 2017*), Hong Kong, China, May 2017.
- Mohammed Alser; Hasan Hassan; Hongyi Xin; Oğuz Ergin, Onur Mutlu, Can Alkan. “GateKeeper: A New Hardware Architecture for Accelerating Pre-Alignment in DNA Short Read Mapping”, in *Oxford Bioinformatics*, Volume 33 Issue 21, November 2017.
- Hongyi Xin, Sunny Nahar, Richard Zhu, John Emmons, Gennady Pekhimenko, Carl Kingsford, Can Alkan, Onur Mutlu. “Optimal Seed Solver: Optimizing Seed Selection in Read Mapping”, in *Oxford Bioinformatics*, Volume 32, Issue 11, May 2016.
- Hongyi Xin, John Greth, John Emmons, Gennady Pekhimenko, Carl Kingsford, Can Alkan, Onur Mutlu. “Shifted Hamming Distance: A Fast and Accurate SIMD-Friendly Filter for Local Alignment in Read Mapping”, in *Oxford Bioinformatics*, Volume 31, Issue 10, June 2015.
- Donghyuk Lee, Farhad Hormozdiari, Hongyi Xin, Faraz Hach, Can Alkan, Onur Mutlu. “Fast and Accurate Mapping of Complete Genomics Reads”, in *Methods (Elsevier)*, Volume 79-80, Issue 1, June 2015.
- Vivek Seshadri, Samihan Yedkar, Hongyi Xin, Onur Mutlu, Phillip B. Gibbons, Michael A. Kozuch, Todd Mowry. “Mitigating Prefetcher-Caused Pollution using Informed Caching Policies for Prefetched Blocks”, in *ACM Transactions on Architecture and Code Optimization (TACO)*, Volume 11, Issue 4, January 2015.
- Gennady Pekhimenko, Vivek Seshadri, Yoongu Kim, Hongyi Xin, Onur Mutlu, Todd Mowry. “Linearly Compressed Pages: A Low-complexity, low-latency main memory compression framework” in Proceedings of the 46th International Symposium on Microarchitecture (*MICRO 2013*), Davis, CA, USA, December 2013.

- Chuanjun Zhang, Glenn G. Ko, Jungwook Choi, Shang-nien Tsai, Minje Kim, Abner Guzmán-Rivera, Rob A. Rutenbar, Paris Smaragdis, Mi Sun Park, Narayanan Vijaykrishnan, Hongyi Xin, Onur Mutlu, Bin Li, Li Zhao, and Mei Chen “EMERALD: Characterization of emerging applications and algorithms for low-power devices”, in Proceedings of the *2013 IEEE International Symposium on Performance Analysis of Systems Software (ISPASS 2013)*, Austin, TX, USA, April 2013
- Hongyi Xin, Donghyuk Lee, Farhad Hormozdiari, Samihan Yedkar, Onur Mutlu, Can Alkan. “Accelerating Read Mapping with FastHASH”, in *BMC Genomics*, Volume 14, January 2013.
- Joseph L. Greathouse, Hongyi Xin, Yixin Luo, and Todd Austin. “A Case for Unlimited Watchpoints”, in Proceedings of the *17th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2012)*, London, United Kingdom, March 2012.

Conference Talks

- “Context-Aware Seeds for Read Mapping”, in *WABI 2019*, Niagara Falls, USA, September 2019.
- “LEAP: A Generalization of the Landau-Vishkin Algorithm with Custom Gap Penalties”, in (*Recomb-seq 2017*), Hong Kong, China, May 2017.
- “Accelerating Read Mapping with FastHASH”, in *11th Asia Pacific Bioinformatics Conference (APBC 2013)*, Vancouver, Canada, January 2013.

Software Projects

- TOMAS. TOMAS is a total-mRNA ratio estimation tool for scRNA-seq analytics. It adopts a Bayesian probabilistic model to synthetically deconvolute heterotypic doublets. It contains an in-house marker-gene co-expression based heterotypic doublet identification mechanism that is robust against inter-cell-type mRNA recovery efficiency discrepancies.
- GMM-Demux. GMM-Demux is a sample demultiplexing and cell type authentication method designed for processing single-cell sample barcoding datasets, including cell hashing and MULTI-seq. GMM-Demux also packs an online experiment planner that helps estimate single cell library preparation outcomes.
- CITE-sort. CITE-sort is a artificial-cell-type-aware clustering method for processing CITE-seq data. It produces a sort tree which guides accurate cell type annotations.
- BAMM-SC. BAMM-SC is a Bayesian Mixture Model for Single Cell sequencing (BAMM-SC) method to cluster scRNA-seq data from multiple individuals simultaneously.
- SNAP. SNAP is a new sequence aligner that is 3-20x faster and just as accurate as existing tools like BWA-mem, Bowtie2 and Novoalign. It supports a rich error model that lets it cheaply match reads with more differences from the reference than other tools.
- Optimal Seed Solver (OSS). OSS is a dynamic-programming based, linear average-case complexity algorithm that enhances the performance of seed-and-extend based read mappers. OSS finds the least frequent set of $e+1$ non-overlapping seeds from a read to tolerate a maximum of e errors.
- Shifted Hamming Distance (SHD). SHD is a SIMD-friendly edit-distance filter that quickly estimates the edit-distances between string pairs and filters out string pairs that have large edit-distances.
- Micro-Read Fast Alignment Search Tool (mrFAST). mrFAST is a DNA read mapper which map short DNA strings generated with the Illumina platform to reference genome assemblies; in a fast and memory-efficient manner.

Teaching

Shanghai Jiao Tong University, Associate Professor, UM-SJTU Joint Institute.

- Lecturer, Unsupervised Learning for Science, ECE6706J, in Summer 2023.
- Lecturer, Algorithms and Data Structures, ECE281, in Fall 2020 - Fall 2022.

Carnegie Mellon University, Teaching Assistant, CS Dept.

- TA for Computer Architecture, 15-740, in Fall 2013.
- TA for Parallel Programming and Parallel Architecture, 15-418, in Spring 2013.