Yicheng Gao

Department of bioinformatics, Tongji University, China
gao.yicheng.98@gmail.com https://gaoyichengtj.github.io

Education

Tongji University (985, 211), Shanghai, China **Sep. 2020 – Present**
Bioinformatics PhD student supervised by Prof. Qi Liu (AI for life science)

Huazhong Agricultural University (211), Wuhan, China **Sep. 2016 – Jun. 2020**
Bachelor of Bioinformatics (with Highest Honor), supervised by Prof. Wen Zhang (AI for life science)
GPA: 3.71/4, Rank: 3/65

Relevant Coursework

- Data Structures and Algorithms
- Discrete Mathematics
- Probability Theory and Statistics
- Calculus
- Database Systems
- Biostatistics
- Linear Algebra
- Bioinformatics Principles
- Computer Network

Main Research Experience

Algorithm design for genetic perturbation prediction, PhD candidate **2024**
Advisor: Prof. Qi Liu *Tongji University*

- Developed a new subtask decomposition-based genetic perturbation prediction model, called STAMP.
- Proposed a three-subtasks-based benchmark evaluation strategy.
- STAMP was implemented based on PyTorch and achieved superior or comparable performance in this task, deployed on github https://github.com/bm2-lab/STAMP.
- Paper was published on Nature Computational Science.

Algorithm design for multi-modal analysis in T-cells, PhD candidate **2022**
Advisor: Prof. Qi Liu *Tongji University*

- Developed a new low-resource-aware multi-modal representation learning for cross-modality integration and analysis of T-cell receptor and T-cell transcriptomes in a unified way, called UniTCR.
- UniTCR was used for an array of downstream tasks, including single modality analysis, modality gap analysis, epitope-TCR binding prediction and cross-modality generation task.
- UniTCR was implemented based on PyTorch and achieved superior or comparable performance in these tasks, deployed on github https://github.com/bm2-lab/UniTCR.
- Paper was published on Cell Genomics.

Algorithm design for predicting peptide-TCR binding, PhD candidate **2021**
Advisor: Prof. Qi Liu *Tongji University*

- Developed a new meta-learning framework combined with the ideas of meta learning and neural turning machine (NTM) for tackling the peptide-TCR binding prediction task, called PanPep.
- Based on the NTM, we proposed a disentanglement distillation module for generalizing few-shot learning to the zero-shot learning.
- PanPep was implemented with PyTorch and achieved SOTA performance in this task, deployed on github https://github.com/bm2-lab/PanPep.
- Paper was published on Nature Machine Intelligence.

Algorithm design for neo-antigen identification, PhD candidate **2020**
Advisor: Prof. Qi Liu *Tongji University*

- Developed a weakly-supervised learning framework for identifying neo-antigens, called TCRBagger.
- TCRBagger was implemented with TensorFlow, and deployed on github https://github.com/bm2-lab/TCRBagger
- Paper was preprinted on bioRxiv

Publications

- **Yicheng Gao***, Yuli Gao*, Yuxiao Fan, Chengyu Zhu, Zhiting Wei, Chi Zhou, Guohui Chuai, Qinchang Chen, He Zhang, Qi Liu, **Pan-Peptide Meta Learning for T-cell receptor–antigen binding recognition**, **Nature Machine Intelligence, (Research Highlight, ESI highly cited)**, 2023
- **Yicheng Gao***, Zhiting Wei*, Kejing Dong, Jingya Yang, Guohui Chuai, Qi Liu, **Toward subtask decomposition-based learning and benchmarking for genetic perturbation outcome prediction and beyond**, **Nature Computational Science**, 2024
- **Yicheng Gao***, Kejing Dong*, Yuli Gao, Xuan Jin, Qi Liu, **Unified cross-modality integration and analysis of T-cell receptors and T-cell transcriptomes**, **Cell Genomics, (Featured Article)**, 2024
- **Yicheng Gao**, Qi Liu, **Delineating the cell types with transcriptional kinetics**, **Nature Computational Science**, 2024
- **Yicheng Gao***, Yuli Gao*, Kejing Dong, Siqi Wu, Qi Liu, **Reply to: The pitfalls of negative data bias for the T-cell epitope specificity challenge**, **Nature Machine Intelligence**, 2023
- Zhiting Wei*, Duanmiao Si*, Bin Duan*, **Yicheng Gao***, Qian Yu, Ling Guo, Qi Liu, **PerturBase: a comprehensive database for single-cell perturbation data analysis and visualization**, **Nucleic Acid Research**, 2024
- Yuli Gao*, **Yicheng Gao***, Wannian Li, Siqi Wu, Feiyang Xing, Chi Zhou, Shaliu Fu, Guohui Chuai, Qinchang Chen, He Zhang, Qi Liu, **Neo-epitope identification by weakly-supervised peptide-TCR binding prediction**, **Preprint on bioRxiv**, 2023
- Shaoqi Chen, Bin Duan, Chenyu Zhu, Chen Tang, Shuguang Wang, **Yicheng Gao**, Shaliu Fu, Lixin Fan, Qiang Yang, Qi Liu, **Privacy-preserving integration of multiple institutional data for single-cell type identification with scPrivacy**, **Science China Life Sciences**, 2023
- Bin Duan, Shaoqi Chen, Xiaohan Chen, Chenyu Zhu, Chen Tang, Shuguang Wang, **Yicheng Gao**, Shaliu Fu, Qi Liu, **Integrating multiple references for single-cell assignment**, **Nucleic acids research**, 2021
- **Yi-Cheng Gao**, Xiong-Hui Zhou, Wen Zhang, **An ensemble strategy to predict prognosis in ovarian cancer based on gene modules**, **Frontiers in Genetics**, 2019

Selected Honors and Awards

- **National scholarship** for PhD students in China (Top 0.2%), 2023
- Tongji University **Scholarship for Outstanding Ph.D. Freshman** (Top 5%), 2020
- The 16th **'Top Ten Students'** of Huazhong Agricultural University (Top 0.1%), **ranked first**, 2020
- Huazhong Agricultural University 2020 **Outstanding Graduates** (Top 5%), 2020
- The 9th MathorCup University Mathematical Modeling Challenge, **First Prize**, 2019
- The 12th Central China Mathematical Modeling Invitational Competition, **Second Prize**, 2019
- China SAS Data Analysis Competition, the twelfth place in Hubei Province, 2018

Intern Experience

Microsoft | *AI/ML group intern* **Dec. 2023 – Present**

- Disentanglement representation method design and proof
- Ongoing project

BGI Genomics | *Algorithm Engineer* **Jul. 2020 – Sep. 2020**

- Algorithm optimization for detection of structural variation

BGI Genomics | *Algorithm Engineer* **Jul. 2019 – Sep. 2019**

- Algorithm optimization for detection of SNP variation.

Technical Skills

Languages: Python, R, Matlab, C++, Objective-C
Developer Tools/APIs: VS Code, R Studio, PyTorch, TensorFlow, Adobe Illustrator
Technologies/Frameworks: Linux, GitHub

English Communication Skills

IELTS: 6.5 (Listening 6.0, Reading 7.0, Writing 6.0, Speaking 6.0)
CET-6: 596/710 (Listening 209, Reading 241, Writing 146)
CET-4: 556/710 (Listening 160, Reading 214, Writing 183)

Leadership / Extracurricular

iOSClub **Jun. 2018 – Sep. 2019**
President *HZAU*

- Lead a team to participate in a summer camp organized by Apple.
- Completed the iOS front-end of "Little Farmer-Farm Management APP" (https://github.com/greysonyc/xiaonongfu)