# Fangchen Yu

*Ph.D. Candidate*

*2001 Longxiang Boulevard*
*Shenzhen, China, 518172*
*+86-18810100813*
*fangchenyu@link.cuhk.edu.cn*
*https://sciyu.github.io*


## Education

**The Chinese University of Hong Kong, Shenzhen (CUHK-SZ)** Sep. 2020 - Present
*Ph.D. Candidate in Computer and Information Engineering* (GPA: 3.82/4.00)
Supervisors: **Prof. Wenye Li**, **Prof. Jianfeng Mao**

**Mohamed Bin Zayed Univerisity of Artificial Intelligence (MBZUAI)** Oct. 2024 - Apr. 2025
*Visiting Student*; Supervisor: **Prof. Qiang Sun**

**University of Chinese Academy of Sciences (UCAS)** Sep. 2016 - Jul. 2020
*Bachelor Degree in Physics* (GPA: 3.61/4.00)

**University of California, Davis (UCD)** Aug. 2019 - Dec. 2019
*Exchange Student* (GPA: 3.77/4.00)

## Professional Experience

**Shenzhen Research Institute of Big Data**, Shenzhen, China Sep. 2020 - Aug. 2024
*Graduate Research Assistant*, Supervisors:**Prof. Wenye Li**, **Prof. Yicheng Zeng**

**Vivo AI Lab**, Shenzhen, China Mar. 2023 - Jul. 2023
*Research Intern*, Area: Visual Understanding and Generation

## Research Interests

Statistical Machine Learning, Optimization, Optimal Transport, Generative Model.

## Research Topics

- **Efficient Similarity and Distance Learning for Incomplete Data (Previous Work)**
  Optimize similarity matrices for offline and online incomplete data (Published in NeurIPS, WWW, and UAI)
  Estimate distance matrices for incomplete data in similarity search tasks (Published in AAAI, ECAI, and ECML)
- **Optimization and Generalization of Wasserstein Distance (Ongoing Work)**
  Develop an accurate tree-Wasserstein distance for approximating the 1-Wasserstein distance (Accepted in ICML)
  Design a novel Wasserstein distance for unbalanced point clouds
- **Optimal Transport for Visual Generative Models (Future Work)**
  Apply optimal transport to flow matching and diffusion models
  Investigate optimal transport techniques for video generation and multi-modal learning

## Selected Publications

- **Efficient Similarity Learning for Incomplete Data**

**1. A Theory-Driven Approach to Inner Product Matrix Estimation for Incomplete Data: An Eigenvalue Perspective** **WWW-2025**
**Fangchen Yu**, Yicheng Zeng, Jianfeng Mao, Wenye Li
International World Wide Web Conference (WWW), 2025. [Github]

**2. Boosting Spectral Clustering on Incomplete Data via Kernel Correction and Affinity Learning** **NeurIPS-2023**
**Fangchen Yu**, Runze Zhao, Zhan Shi, Yiwen Lu, Jicong Fan, Yicheng Zeng, Jianfeng Mao, Wenye Li
37th Conference on Neural Information Processing Systems (NeurIPS), 2023. [Github]

**3. Online Estimation of Similarity Matrices with Incomplete Data** **UAI-2023**
**Fangchen Yu**, Yicheng Zeng, Jianfeng Mao, Wenye Li
39th Conference on Uncertainty in Artificial Intelligence (UAI), 2023. [Github]

- **Robust Distance Learning for Incomplete Data**

**4. Highly-Efficient Robinson-Foulds Distance Estimation with Matrix Correction** **ECAI-2023**
**Fangchen Yu**, Rui Bao, Jianfeng Mao, Wenye Li
26th European Conference on Artificial Intelligence (ECAI), 2023. [Github]

**5. Metric Nearness Made Practical** **AAAI-2023**
Wenye Li, **Fangchen Yu**, Zichen Ma
37th AAAI Conference on Artificial Intelligence (AAAI), 2023. [Github]

**6. Calibrating Distance Metrics Under Uncertainty** **ECML-2022**
Wenye Li, **Fangchen Yu**
Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML), 2022.

- **Optimization in Computer Vision and Natural Language Processing**

**7. From Incompleteness to Unity: A Framework for Multi-view Clustering with Missing Values** **ICONIP-2023**
**Fangchen Yu**, Zhan Shi, Yuqi Ma, Jianfeng Mao, Wenye Li.
International Conference on Neural Information Processing (ICONIP), 2023.

**8. DocReal: Robust Document Dewarping of Real-Life Images via Attention-Enhanced Control Point Prediction**
**Fangchen Yu**, Yina Xie, Lei Wu, Yafei Wen, Guozhi Wang, Shuai Ren, Xiaoxin Chen, Jianfeng Mao, Wenye Li
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024. [Github] **WACV-2024**

**9. UltraTWD: Optimizing Ultrametric Trees for Tree-Wasserstein Distance** **ICML-2025**
**Fangchen Yu**, Yanzhen Chen, Jiaxing Wei, Jianfeng Mao, Wenye Li, Qiang Sun.
42nd International Conference on Machine Learning (ICML), 2025.

## Honors and Awards

**PhD Fellowship** at Shenzhen Research Institute of Big Data 2020 - 2024
**Class I Outstanding Teaching Assistant Award** at The Chinese University of Hong Kong, Shenzhen Oct. 2022
**Class II Outstanding Teaching Assistant Award** at The Chinese University of Hong Kong, Shenzhen Jul. 2021
**Class III Scholarship** at University of Chinese Academy of Sciences Nov. 2019
**Merit Student** at University of Chinese Academy of Sciences Dec. 2017

## Academic Service

**Conference Reviewer:** ICML 2025, ICLR 2025, NeurIPS 2025/2024, WWW 2025, AAAI 2025/2024, IJCAI 2025/2024.

## Teaching Assistant (in English)

**MAT3007 Optimization** Summer, 2024
**DDA4210 Advanced Machine Learning** Spring & Fall, 2023
**MAT3300 Mathematical Modeling** Fall, 2022
**STA3010 Regression Analysis** Spring, 2022
**MAT4003 Number Theory** Fall, 2021
**MAT4004 Graph Theory** Spring, 2021
**MAT3280 Probability Theory** Fall, 2020

## Skills

**Programming** Python (PyTorch, NumPy, Pandas, Sklearn), Linux, Git, MATLAB, LaTeX
**Language** Fluent in English (TOEFL, CET-6), Mandarin

## References


**Prof. Wenye Li**
Chinese University of Hong Kong, Shenzhen
wyli@cuhk.edu.cn

**Prof. Qiang Sun**
University of Toronto
qiang.sun@utoronto.ca

**Prof. Yicheng Zeng**
Sun Yat-sen University
zengyicheng@mail.sysu.edu.cn