

# Yujun He


+86-177-7310-7739 | [heyj2022@mail.sustech.edu.cn](mailto:heyj2022@mail.sustech.edu.cn)/[he443@wisc.edu](mailto:he443@wisc.edu)

Shenzhen, Guangdong Province - 518055, China

## EDUCATION

- **Turing Honor Class, Southern University of Science and Technology (SUSTech)** 2022 – 2026  
*B.E. in Computer Science and Technology* Shenzhen, China
  - Advisor: [Prof. Bo Tang](#)
  - GPA: 3.94/4.00; Rank: 3/166 (first 3 years)
- **University of Wisconsin–Madison** Feb 2025 – Aug 2025  
*Visiting Undergraduate Researcher* Madison, WI, U.S.
  - Advisor: [Prof. Xiangyao Yu](#). I also work very closely with [Prof. Paris Koutris](#).

## RESEARCH PROJECTS

- **A Novel Partition-based Approach Applied in Database Query Optimization** Feb 2025 – Present  
*Project Leader*
  - Explored the design space of partition-based query optimization. This novel approach, inspired by techniques from worst-case optimal join (WCOJ), systematically investigates how partitioning can improve join performance, especially for cyclic queries where traditional optimizers struggle.
  - Proposed and implemented a set of heuristic optimizations, which enable state-of-the-art practical performance while keeping the optimization overhead minimal.
  - Built a system layer that can be put in front of any databases, to enable the partition operator.
  - We plan to submit the work to VLDB 2026 soon.
- **A System for Supporting Graph-based Vector Index Updates** June 2024 – Dec 2024  
*Project Co-Leader*
  - Identified limitations and performance bottlenecks in the on-disk state-of-the-art out-of-place graph-based vector update system (i.e., FreshDiskANN).
  - Proposed a novel algorithm for efficient deletion handling in graph-based indexes and designed a VectorDB system architecture tailored for dynamic vector updates.
  - Developed a graph-based VectorDB from scratch, achieving state-of-the-art performance in vector update scenarios and outperforming SOTA competitor systems such as SPFresh, FreshDiskANN, and Greator.
  - The work is currently under submission to SIGMOD 2026.
- **Vector Search Hybrid with Scalar Attributes** Mar 2024 – May 2024  
*Core Contributor* 
  - Proposed a novel indexing strategy to support hybrid vector search with interval-type scalar filters.
  - Implemented data quantization techniques (e.g., SQ8 quantization) on graph-based vector indexes (i.e., HNSW), and conducted parameter tuning for optimal indexing and search performance.
  - We achieved 1st place among all teams in the 2024 SIGMOD Programming Contest.

## PATENTS AND PUBLICATIONS (\* INDICATES EQUAL CONTRIBUTION)

C=CONFERENCE, J=JOURNAL, P=PATENT, S=IN SUBMISSION, T=THESIS

- [S.1] Haotian Liu\*, Yujun He\*, Bo Tang. [Title fuzzed] A system to support graph-based vector index in-place updates. Manuscript is under submission to SIGMOD2026.
- [S.2] Yujun He, et al. [Title fuzzed] A novel partition-based approach in query optimization. Manuscript plans to submit for publication in VLDB2026.

## HONORS AND AWARDS

---

- **World Champion** May 2024  
2024 SIGMOD Programming Contest 
- **National Gold Medal, Ranked 11/442** Dec 2023  
48th International Collegiate Programming Contest (ICPC) Regional Hangzhou
- **National 2nd runner-up** Oct 2023  
9th China Collegiate Programming Contest (CCPC) Guilin Site 
- **National Gold Medal, Ranked 16/406** Oct 2023  
48th International Collegiate Programming Contest (ICPC) Regional Xi'an
- **Provincial Champion** May 2023  
20th Guangdong Collegiate Programming Contest (GDCPC)
- **National Silver Medal** July 2021  
38th National Olympiad in Informatics (NOI2021)

## SKILLS

---

- **Programming Languages:** C/C++, CUDA, Python, Rust, GoLang, Java, SQL
- **Tools & Frameworks:** PyTorch, Git, Linux, Docker
- **Languages:** Chinese (Native), English (Professional working)

## COURSE PROJECTS

---

- **Virtio-GPU Implementation for Asterinas OS Kernel** 2024  
SUSTech CS334 Operating Systems (H)
- **GAN-based Data Augmentation for Self-Driving Object Detection** 2024  
SUSTech CS329 Machine Learning (H)
- **Reliable Data Transfer Implementation over UDP Sockets** 2024  
SUSTech CS305 Computer Networks
- **Pipelined CPU Implementation in Verilog** 2024  
SUSTech CS214 Computer Organization (H)
- **Bullet Comments System Backend in Java with PostgreSQL** 2023  
SUSTech CS213 Principles of Database Systems (H)
- **FPGA Controller for Game Overcooked!** 2023  
SUSTech CS211 Digital Logic (H)

## SERVICES

---

- **Problem Setter and Organizer** Dec 2023  
SUSTech Programming Competition (On-Site)
- **30+ Hours Volunteering** 2022-2025  
Shenzhen, China