Baotong Tian

+1 (585)-451-6253 • baotongtian1@gmail.com • user-tian.github.io
user-tian

Education

University of Rochester **Aug, 2024 – Aug, 2029 (Expected)**
Ph.D. in Electrical and Computer Engineering
GPA 4.0/4.0

Tsinghua University **Sep, 2020 – Jun, 2024**
Bachelor of Engineering in Automation
GPA 3.84/4.0

Research Interests

My research interests lie in AI for Music Information Retrieval and speech processing. In music, I focus on developing AI tools for musicians and producers (expressive and controllable music rendering, AI mixing, etc). In speech, I'm interested in voice conversion and controllable text-to-speech synthesis.

Scholarships & Awards

- **2021, 2022 Tsinghua Arts and Culture Merit Scholarship** (Awarded to undergraduate students for great contribution to cultural activities within the department)
- **2022, 2023 Tsinghua Academic Excellence Scholarship** (Awarded to undergraduate students for excelling GPA throughout the school year)
- **2023 Tsinghua Innovation Award of Science and Technology** (Awarded to undergraduate students with excellent research potential)
- **2025 WIMIR Grant from ISMIR**
- Zhang, Yu*, **Baotong Tian***, and Zhiyao Duan. "Conan: A Chunkwise Online Network for Zero-Shot Adaptive Voice Conversion." Proc. ASRU 2025. [Arxiv], [Demo Page], [Code]
- You, Zhang*, **Baotong Tian***, Lin Zhang, Zhiyao Duan. "PartialEdit: Identifying Partial Deepfakes in the Era of Neural Speech Editing." Proc. Interspeech 2025. [Paper], [Project Page], [Dataset]
- Borui, Zhang*, **Baotong Tian***, Wenzhao Zheng, Jie Zhou, and Jiwen Lu. "Exploring Unified Perspective For Fast Shapley Value Estimation." arXiv preprint arXiv:2311.01010 (2023). [Arxiv]

Research Experience

Diffusion-based Violin Transcription and Synthesis **Jun, 2025 - Present**
Research Assistant, Advisor: Zhiyao Duan
Audio Information Research (AIR) Lab, Department of Electrical and Computer Engineering, University of Rochester
- Created a paired dataset by annotating expressive MIDI with note-level violin techniques and rendering corresponding audio with a commercial Kontakt sample library.

PartialEdit **Sep, 2024 - Jun, 2025**
Research Assistant, Advisor: Zhiyao Duan
Audio Information Research (AIR) Lab, Department of Electrical and Computer Engineering, University of Rochester
- Created the partially deepfake dataset **PartialEdit** using advanced neural speech-editing models.
- Evaluated state-of-the-art utterance-level detectors and segment-level localization methods on this dataset, showing that models trained on traditional partially spoofed audio fail to detect speech edited by neural speech-editing models.
- Discussed how artifacts introduced by codec models affect the performance of models.

Streaming Voice Conversion **Sep, 2024 - Present**
Research Assistant, Advisor: Zhiyao Duan
Audio Information Research (AIR) Lab, Department of Electrical and Computer Engineering, University of Rochester
- Co-developed **Conan**, a chunk-wise online model for zero-shot voice conversion.
- Adapted to model timbre and style separately, and proposed a pixel-shuffle HifiGAN for streaming applications.

Controllable Sheet Music Generation **Jun, 2021 - Dec, 2022**

Research Assistant, Advisors: Xin Jin & Duo Xu & Song-chun Zhu
AI Music Group, Beijing Institute for General Artificial Intelligence

- Learned music psychology and emotion perception in music.
- Learned about the generative theory and basic structural analysis of tonal music.
- Participated in the implementation of Music And-or-Graph.
- Developed a genetic algorithm for **music variation** in Python.

Explainable Machine Learning **Feb, 2023 - Jun, 2024**

Research Assistant, Advisors: Jiwen Lu & Jie Zhou
Intelligent Vision Group, Department of Automation, Tsinghua University

- **Unified** current approaches of Shapley value estimation.
- Proposed **SimSHAP** as a simple and fast amortized Shapley value estimator.

Work Experience

Automatic Chord Recognition for Pop songs (Internship) **Jun, 2023 - Aug, 2023**

Algorithm Engineer, Advisor: Jian Wu
R&D Department, Beijing Deepmusic Technology Co.

- Applied Bi-LSTM for multi-task learning of bass and chord quality.
- Applied **Conditional Random Field** (CRF) as a post-filtering technique to enhance chord recognition task accuracy by incorporating context information.

Teaching

Teaching Assistant:

- ECE208/408 The Art of Machine Learning — Spring 2025
- ECE270 Probability for Electrical Engineers — Fall 2025

Professional Services

Reviewer

Journal:

- IEEE Transactions on Audio, Speech, and Language Processing (TASLP) — 2025

Conference:

- IEEE Automatic Speech Recognition and Understanding Workshop — 2025

Volunteer

- ISMIR 2025 — 2025

Music Background and Interests

Singing

- Joined the school choir at primary school.
- Won the 4th place in the Singing Carnival in High School.

Instrument Playing

- Electric Piano (age 5 to 12, got the Amateur Highest Performance Level at the age of 12)
- Guitar (age 13 to present)

Music Composition & Production

- Learned basic music production skills such as music arrangement and mixing in college.
- Selected for the Class of 2021-2022 in **Tsinghua University's Music Dream Program**, which is designed to nurture and develop campus musicians.
- Released 8 songs on QQ music & The NetEase's Cloud music, produced by **Logic Pro X**.

Links

- Solo and Choir performance can be seen [here].
- My musician page on [The NetEase's Cloud Music], [QQ Music], and [SoundCloud].

Programming & Language Skills

Programming Skills:

- Proficient: Python, PyTorch, Markdown, LaTeX, Git
- Familiar: MATLAB, C/C++, Linux, TensorFlow, HTML, etc

Languages:

- TOEFL iBT: 104/120 (Reading 26, Listening 29, Speaking 24, Writing 25)
- GRE: 328/340 + 4.0/6.0 (Verbal 158, Quantitative 170, Analytical Writing 4.0)