Jiarui Hai

 haidog-yaqub   ✉ jhai2@jhu.edu
 Google Scholar   in Jiarui's linkedin
 Personal Page   📞 (+1) 224-866-2166

Education

Johns Hopkins University	Doctor of Philosophy
Electrical and Computer Engineering	Aug 2022 – Present
Advisor: Prof. Mounya Elhilali	
Tsinghua University	Master
M.E. in Civil Engineering	Aug 2020 – Jun 2022
Member of Big Data Program	
Tsinghua University	Undergraduate
B.S. in Civil Engineering	Aug 2016 – Jun 2020
B.E. in Business Analytics	

Selected Projects

EzAudio: Enhancing Text-to-Audio Generation with Efficient Diffusion Transformer [1] — May 2024 – Aug 2024
Intern at Tencent AI Lab — 📻 EzAudio Page
Mentors: Dr. Yong Xu, Dr. Hao Zhang, Dr. Dong Yu

- Proposed a text-to-audio generation model featuring an efficient diffusion transformer and a multi-stage training strategy, surpassing open-source text-to-audio generators in both subjective and objective evaluations.
- Developed editing tools and integrated ControlNet for real-world sound and music production applications.
- Selected as Hugging Face's Daily Paper.

DreamVoice: Text-Guided Voice Conversion [4] — Feb 2024 – Apr 2024
Johns Hopkins University | Supported by Myshell — 👪 DreamVoice Page
Advisor: Prof. Mounya Elhilali

- Created a dataset for speaker timbre characterization.
- Proposed two models:
  - An end-to-end diffusion model for text-guided voice conversion.
  - A plugin model that adapts pre-trained voice conversion systems to support text-guided voice conversion.

Diff-Pitcher: Diffusion-based Singing Voice Pitch Correction [9] — Mar 2023 – May 2023
Johns Hopkins University — 🎤 Diff-Pitcher Page
Advisor: Prof. Mounya Elhilali

- Proposed a diffusion-based automatic pitch correction system that generates natural-sounding singing voices.
- Received high praise from Hollywood music producer David Kershenbaum.

Music Production

- Co-founder, InnerSup Records, a Chinese record label and record producer collective — *2021*
- Honorary President, Rap/HipHop Association at Tsinghua University — *2020–2021*
- Founder and Leader, HipHop Production Group at Tsinghua University — *2018–2022*
- Featured producer in a rap song on *HIPHOP BANK*, a Chinese TV show — *August 2021*
- Delivered a lecture on music production at Modern Sky, China's leading music company — *June 2021*
- Won Top 1% New Producer at the BeatsHome Hip-hop Production Contest — *April 2021*
- My music video ranked in the Top 10 on Bilibili's trending board — *October 2020*

Pre-prints

- 1. **Jiarui Hai**, Yong Xu, Hao Zhang, Chenxing Li, Helin Wang, Mounya Elhilali, Dong Yu. EzAudio: Enhancing Text-to-Audio Generation with Efficient Diffusion Transformer. Submitted to ICASSP 2025.
- 2. Helin Wang*, **Jiarui Hai***, Yen-Ju Lu, Karan Thakkar, Mounya Elhilali, Najim Dehak. SoloAudio: Target Sound Extraction with Language-oriented Audio Diffusion Transformer. Submitted to ICASSP 2025.
- 3. Helin Wang, Meng Yu, **Jiarui Hai**, Chen Chen, Yuchen Hu, Rilin Chen, Najim Dehak, Dong Yu. SSR Speech: Towards Stable, Safe, and Robust Zero-shot Text-based Speech Editing and Synthesis. Submitted to ICASSP 2025.

Selected Publications

- 4. **Jiarui Hai***, Karan Thakkar*, Helin Wang, Zengyi Qin, Mounya Elhilali. DreamVoice: Text-Guided Voice Conversion. Interspeech 2024.
- 5. Helin Wang, Jesus Villalba, Laureano Moro-Velazquez, **Jiarui Hai**, Thomas Thebaud, Najim Dehak. Noise-robust Speech Separation with Fast Generative Correction. Interspeech 2024.
- 6. **Jiarui Hai***, Helin Wang*, Dongchao Yang, Karan Thakkar, Najim Dehak, Mounya Elhilali. DPM-TSE: A Diffusion Probabilistic Model for Target Sound Extraction. ICASSP 2024.
- 7. Karan Thakkar, **Jiarui Hai**, Mounya Elhilali. Investigating Self-Supervised Deep Representations for EEG-based Auditory Attention Decoding. ICASSP 2024.
- 8. **Jiarui Hai***, Yu-jeh Liu*, Mounya Elhilali. Boosting Modality Representation with Pre-trained Models and Multi-task Training for Multimodal Sentiment Analysis. ASRU 2023.
- 9. **Jiarui Hai**, Mounya Elhilali. Diff-Pitcher: Diffusion-based Singing Voice Pitch Correction. WASPAA 2023.
- 10. Rui Lu, Baigong Zheng, **Jiarui Hai**, Fei Tao, Zhiyao Duan, Ji Liu. Progressive Teacher-Student Training Framework for Music Tagging. ICASSP 2022.