

Hankun Wang

王翰坤

PhD Candidate, X-LANCE Lab
School of Computer Science
Shanghai Jiao Tong University
Supervisor: Prof. Kai Yu

wang.hankun@icloud.com
+86 15855127061
<https://whk-sjtu.github.io/>
Google Scholar

FIELDS OF SPECIALIZATION

Speech synthesis, speech tokenization, and speech language models, with an emphasis on robust, efficient, and controllable speech generation.

EXPERIENCE

rednote (Xiaohongshu) Sep 2025 – Present

Research Intern, HiLab (dots studio)

Work on industrial speech synthesis, speech editing, full-duplex models, and the speech representations that underpin them.

dots.tts *Core Contributor*

Demo: <https://studio-dots-ai.github.io/dots.tts-demo/>

Developed an industrial text-to-speech foundation model: a 2B-parameter continuous autoregressive system supporting expressive voice cloning, multilingual synthesis, and low-latency streaming generation.

dots.tts.edit *Project Lead and Primary Contributor*

Demo: <https://huggingface.co/spaces/dots-studio/dots.tts.edit-demo>

Led the development of an end-to-end model for precise speech editing. The system supports local text, emotion, prosody, pause, free-form style, and compositional edits while preserving the surrounding speech and speaker identity; the project also introduces doteBench for evaluating instruction following, local preservation, and audio quality.

ByteDance Jul – Oct 2021

Intern, Ad Infrastructure Group

Worked on architectural optimizations for the ad-request serving path of a PID-based bidding system, to improve throughput and reduce response latency and timeout rates.

EDUCATION

Shanghai Jiao Tong University 2022 – Present

PhD Candidate in Computer Science and Technology
School of Computer Science; Supervisor: Prof. Kai Yu
Research on speech generation and representation.

Harbin Institute of Technology 2018 – 2022

Bachelor's degree in Data Science and Big Data Technology
Faculty of Computing; Supervisor: Prof. Dongjing Miao
GPA: 91.45/100 Rank in Major: 2/35

Hefei No. 1 High School 2015 – 2018

AWARDS AND HONORS

2021	People's Scholarship, Harbin Institute of Technology (Top 3%)
2020	Outstanding Student, Harbin Institute of Technology
2019	Silver Medal, International Collegiate Programming Contest (ICPC), Asia Xuzhou Regional
2019	Gold Medal, Northeast Collegiate Programming Contest (CCPC)
2019	Gold Medal (2nd Place), Heilongjiang Provincial Collegiate Programming Contest (CCPC)
2017	Bronze Medal, Asia-Pacific Informatics Olympiad (APIO)
2016	First Prize, National Olympiad in Informatics in Provinces (NOIP), Advanced Group
2015	First Prize, National Olympiad in Informatics in Provinces (NOIP), Advanced Group

PUBLICATIONS

1. dots.tts.edit: Precisely Controlled Speech Editing with a Continuous Autoregressive Model
Hankun Wang, Bohan Li, Shi Lian, Xiaoyu Gu, Jing Peng, Da Zheng, Yiwei Guo, Colin Zhang, Shuai Wang, Kai Yu.
2026.
2. dots.tts Technical Report
Shi Lian, Changtao Li, Bohan Li, **Hankun Wang**, Da Zheng, Junfeng Tian, Yufeng Ma, Colin Zhang, Kai Yu.
2026.
3. CodecSlime: Temporal Redundancy Compression of Neural Speech Codec via Dynamic Frame Rate
Hankun Wang, Yiwei Guo, Chongtian Shao, Bohan Li, Kai Yu.
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2026.
SPS Travel Grant recipient.
4. Why Do Speech Language Models Fail to Generate Semantically Coherent Outputs? A Modality Evolving Perspective
Hankun Wang, Haoran Wang, Yiwei Guo, Zhihan Li, Chenpeng Du, Kai Yu.
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2026.
Oral presentation.
5. Attention-Constrained Inference for Robust Decoder-Only Text-to-Speech
Hankun Wang, Chenpeng Du, Yiwei Guo, Shuai Wang, Xie Chen, Kai Yu.
IEEE Spoken Language Technology Workshop (SLT), 2024.
6. HoliTok: A Continuous Holistic Tokenization with Robust Dual Capabilities of Speech Generation and Understanding
Bohan Li, Shi Lian, **Hankun Wang**, Yiwei Guo, Yu Xi, Zhihan Li, Da Zheng, Colin Zhang, Kai Yu.
Conference on Empirical Methods in Natural Language Processing (EMNLP), 2026.
Oral presentation, main conference.
7. Read What You Hear: Reference-Free Hypotheses Evaluation with Acoustic Discrepancy
Zhihan Li, **Hankun Wang**, Yiwei Guo, Bohan Li, Xie Chen, Kai Yu.
Interspeech 2026.
8. Robust and Efficient Autoregressive Speech Synthesis with Dynamic Chunk-wise Prediction Policy
Bohan Li, Zhihan Li, Haoran Wang, Hanglei Zhang, Yiwei Guo, **Hankun Wang**, Kai Yu.
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2026.
9. Recent Advances in Discrete Speech Tokens: A Review
Yiwei Guo, Zhihan Li, **Hankun Wang**, Bohan Li, Chongtian Shao, Hanglei Zhang, Chenpeng Du, Xie Chen, Shujie Liu, Kai Yu.
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2026.

10. AHAMask: Reliable Task Specification for Large Audio Language Models Without Instructions
Yiwei Guo, Bohan Li, **Hankun Wang**, Zhihan Li, Shuai Wang, Xie Chen, Kai Yu.
AAAI Conference on Artificial Intelligence, 2026.
11. A Survey on Speech Large Language Models for Understanding
Jing Peng, Yucheng Wang, Bohan Li, Yiwei Guo, **Hankun Wang**, Yangui Fang, Yu Xi, Haoyu Li, Xu Li, Ke Zhang, Shuai Wang, Kai Yu.
IEEE Journal of Selected Topics in Signal Processing (JSTSP), 2026.
12. Towards General Discrete Speech Codec for Complex Acoustic Environments: A Study of Reconstruction and Downstream Task Consistency
Haoran Wang, Guanyu Chen, Bohan Li, **Hankun Wang**, Yiwei Guo, Zhihan Li, Xie Chen, Kai Yu.
IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), 2025.
13. LSCodec: Low-Bitrate and Speaker-Decoupled Discrete Speech Codec
Yiwei Guo, Zhihan Li, Chenpeng Du, **Hankun Wang**, Xie Chen, Kai Yu.
Interspeech 2025.
14. Fast and High-Quality Auto-Regressive Speech Synthesis via Speculative Decoding
Bohan Li, **Hankun Wang**, Situo Zhang, Yiwei Guo, Kai Yu.
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2025.
15. VALL-T: Decoder-Only Generative Transducer for Robust and Decoding-Controllable Text-to-Speech
Chenpeng Du, Yiwei Guo, **Hankun Wang**, Yifan Yang, Zhikang Niu, Shuai Wang, Hui Zhang, Xie Chen, Kai Yu.
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2025.
16. vec2wav 2.0: Advancing Voice Conversion via Discrete Token Vocoder
Yiwei Guo, Zhihan Li, Junjie Li, Chenpeng Du, **Hankun Wang**, Shuai Wang, Xie Chen, Kai Yu.
2024.

TALKS

Towards More Stable Fully Continuous Autoregressive Speech Synthesis

Nov 2026

NCMMSC 2026, Youth Forum

An overview of a series of studies—dots.tts, HoliTok, and dots.tts.edit—covering an industrial speech synthesis foundation model, holistic speech representations, and precise speech editing.

ACADEMIC SERVICE

Challenge Co-organizer

ICASSP 2027 Audio Editing Challenge

Conference Reviewer

ICLR, ICASSP, Interspeech, and SLT.