

XIQUAN LI

[Homepage](#) [Github](#) [Google Scholar](#) [Linkedin](#) [✉ andreasli0912@gmail.com](mailto:andreasli0912@gmail.com)

EDUCATION

- **Shanghai Jiao Tong University** Sep. 2024 - Mar. 2027
ME., Electronic Engineering. Supervisor: Prof. Xie Chen
Shanghai, China
 - GPA: 3.98/4.0
- **Telecom Paris - Polytechnic Institute of Paris** Sep. 2023 - June 2026
ME., Electronic Engineering (Double Degree Master Program). Supervisor: Prof. Slim Essid
Paris, France
 - GPA: 3.88/4.0
- **Shanghai Jiao Tong University** Sep. 2020 - June 2024
BS., Information Engineering; BA., French. Supervisor: Prof. Xie Chen
Shanghai, China
 - GPA: 91.30, ranked 3/85

SELECTED PUBLICATIONS

- [1] [Resonate: Reinforcing Text-to-Audio Generation via Online Feedback from Large Audio Language Models](#)
Xiquan Li, Junxi Liu, Wenxi Chen, Haina Zhu, Ziyang Ma, Xie Chen
Interspeech 2026
- [2] [FineLAP: Taming Heterogeneous Supervision for Fine-grained Language-Audio Pretraining](#)
Xiquan Li, Xuenan Xu, Ziyang Ma, Wenxi Chen, Haolin He, Qiuqiang Kong, Xie Chen
ACL 2026 Main
- [3] [MeanAudio: Fast and Faithful Text-to-Audio Generation with Mean Flows](#)
Xiquan Li, Junxi Liu*, Yuzhe Liang, Zhikang Niu, Wenxi Chen, Xie Chen*
ACL 2026 Main
- [4] [TimuMU: A Compact Audio-Language Model for Music Understanding](#)
Xiquan Li, Aurian Quellenec, Slim Essid
ICASSP 2026
- [5] [DRCap: Decoding CLAP Latents with Retrieval-Augmented Generation for Zero-shot Audio Captioning](#)
Xiquan Li, Wenxi Chen, Ziyang Ma, Xuenan Xu, Yuzhe Liang, Zhisheng Zheng, Qiuqiang Kong and Xie Chen
ICASSP 2025 (Oral)
- [6] [URO-Bench: A Comprehensive Benchmark for End-to-End Spoken Dialogue Models](#)
Ruiqi Yan, Xiquan Li, Wenxi Chen, Zhikang Niu, Chen Yang, Ziyang Ma, Kai Yu, Xie Chen
EMNLP 2025 Findings
- [7] [Towards Reliable Large Audio Language Model](#)
Ziyang Ma, Xiquan Li, Yakun Song, Wenxi Chen, Chenpeng Du, Jian Wu, Yuanzhe Chen, Zhuo Chen, Yuping Wang, Yuxuan Wang, Xie Chen
ACL 2025 Findings
- [8] [SLAM-AAC: Enhancing Audio Captioning with Paraphrasing Augmentation and CLAP-Refine through LLMs](#)
Wenxi Chen, Ziyang Ma*, Xiquan Li, Xuenan Xu, Yuzhe Liang, Zhisheng Zheng, Kai Yu, Xie Chen*
ICASSP 2025
- [9] [SAC: Neural speech codec with semantic-acoustic dual-stream quantization](#)
Wenxi Chen, Xinsheng Wang, Ruiqi Yan, Yushen Chen, Zhikang Niu, Ziyang Ma, Xiquan Li, Yuzhe Liang, Hanlin Wen, Shunshun Yin, Ming Tao, Xie Chen
ACL 2026 Main
- [10] [Measuring Audio’s Impact on Correctness: Audio-Contribution-Aware Post-Training of Large Audio Language Models](#)
Haolin He, Xingjian Du*, Renhe Sun*, Zheqi Dai, Yujia Xiao, Mingru Yang, Jiayi Zhou, Xiquan Li, et al.*
ICLR 2026
- [11] [MMAR: A Challenging Benchmark for Deep Reasoning in Speech, Audio, Music, and Their Mix](#)
Ziyang Ma, Yinghao Ma*, Yanqiao Zhu*, Chen Yang*, Yi-Wen Chao*, Ruiyang Xu*, Wenxi Chen, Yuanzhe Chen, Zhuo Chen, Jian Cong, Kai Li, Keliang Li, Siyou Li, Xinfeng Li, Xiquan Li, et al.*
NeurIPS 2025

- [12] **SLAM-Omni: Timbre-Controllable Voice Interaction System with Single-Stage Training**
 Wenxi Chen, Ziyang Ma, Ruiqi Yan, Yuzhe Liang, **Xiquan Li**, Ruiyang Xu, Zhikang Niu, Yanqiao Zhu, Yifan Yang, Zhanxun Liu, Kai Yu, Yuxuan Hu, Jinyu Li, Yan Lu, Shujie Liu, Xie Chen
ACL 2025 Findings
- [13] **EmoBox: Multilingual Multi-corpus Speech Emotion Recognition Toolkit and Benchmark**
 Ziyang Ma*, Mingjie Chen*, Hezhao Zhang, Zhisheng Zheng, Wenxi Chen, **Xiquan Li**, Jiaxin Ye, Xie Chen, Thomas Hain
Interspeech 2024

EXPERIENCE

- **Seed Speech Team - Bytedance** Shanghai, China
 Research scientist intern. Supervisor: *Dr. Zhuo Chen and Dr. Jian Wu* Nov. 2025 - Now
 – Worked on post-training of large audio-language models, detailed audio captioning, and unified representation for understanding & generation.
- **Hunyuan Team - Tencent** Shanghai, China
 Research scientist intern. Supervisor: *Jim Zhou, led by Zhao Zhong* July 2025 - Nov. 2025
 – Worked on detailed audio-visual captioning and universal audio generation.
- **ADASP Group - Telecom Paris** Paris, France
 Research intern. Supervisor: *Prof. Slim Essid* Sep. 2024 - June. 2025
 – Proposed an efficient audio-language model for music understanding, leading to a first-author publication at ICASSP 2026 ([4]).
- **DSP Lab - The Chinese University of Hong Kong** Hong Kong SAR, China
 Research assistant. Supervisor: *Prof. Qiuqiang Kong* June 2024 - Sep. 2024
 – Worked on contrastive language-audio pre-training and zero-shot audio captioning. Published paper [5] at ICASSP 2025.
- **X-LANCE Lab - Shanghai Jiao Tong University** Shanghai, China
 Research intern. Supervisor *Prof. Xie Chen* Feb. 2023 - Now
 – Worked on audio understanding and generation. Published papers at ICASSP, ACL, Interspeech, NeurIPS, and ICLR.

OPEN-SOURCE PROJECTS

- **SLAM-LLM (1k+ stars)**
 Core Contributor
 – SLAM-LLM is an open-source toolkit for training multimodal large language models (MLLMs) across speech, audio, and music domains. It provides a scalable pipeline for training a large audio language model.
- **MeanAudio (140 stars)**
 Project Owner
 – MeanAudio is an open-source repository for training text-to-audio generation models based on MeanFlow and Flow Matching. It provides a reproducible framework for fast and faithful audio synthesis.
- **Resonate (46 stars)**
 Project Owner
 – Resonate is an open-source repository that supports pre-training, SFT, DPO, and GRPO for text-to-audio generation. It provides a unified pipeline for building a strong audio generation system from scratch.

SELECTED AWARDS

- **DCASE 2024 Task 6: Automated Audio Captioning** Ranked 3rd DCASE Org. (2024)
- **ICASSP 2024 ICMC-ASR Grand Challenge** Ranked 7/36 IEEE Org. (2024)
- **SPEIT Excellent Scholarship (First Class)** Top 3% SPEIT, SJTU (2022, 2023)
- **Mathematical Contest in Modeling** Finalist Winner - Top 2% World Wide COMAP Org. (2022)
- **ARDIAN Scholarship** Top 3% ARDIAN Cop. (2021)

TECHNICAL SKILLS

Languages: English (CET6: 624), French (DELFB2), Chinese (Native)
Programming: Python, MATLAB, L^AT_EX, SQL, PyTorch, C/C++