LEYING ZHANG

✉zhangleying@sjtu.edu.cn ✆(+86) 18621098717
3-520 SEIEE Building, Shanghai Jiao Tong University
800 Dongchuan Road, Minhang District, Shanghai, China 200240

RESEARCH INTERESTS

Text-to-speech, Audio generation, Speaker Verification, Multi-modality

EDUCATION

Shanghai Jiao Tong University Sep. 2023 - Present
PhD, Computer Science and Engineering Supervisor: Prof. Yanmin Qian
Shanghai Jiao Tong University Sep. 2021 - 2023
Master, Electronic Information Supervisor: Prof. Yanmin Qian
Télécom Paris (Institut polytechnique de Paris) Sep. 2021 - Feb. 2022
Exchange Student, Data science and Image processing
Shanghai Jiao Tong University Sep. 2017 - Jun. 2021
Bachelor of Information Engineering and French (double degree)

PUBLICATIONS

[C1] **Leying Zhang**, Yao Qian, Long Zhou, Shujie Liu, Dongmei Wang, Xiaofei Wang, Midia Yousefi, Yanmin Qian, Jinyu Li, Lei He, Sheng Zhao, Michael Zeng. "CoVoMix: Advancing Zero-Shot Speech Generation for Human-like Multi-talker Conversations". *38th Annual Conference on Neural Information Processing Systems* (**NeurIPS**),*Dec. 2024*

[C2] **Leying Zhang**, Yao Qian, Linfeng Yu, Heming Wang, Hemin Yang, Shujie Liu, Long Zhou, Yanmin Qian. " DDTSE: Discriminative Diffusion Model for Target Speech Extraction". *IEEE Spoken Language Technology Workshop* **(SLT)**, *Dec. 2024*

[C3] **Leying Zhang**, Zhengyang Chen and Yanmin Qian. "Adaptive Large Margin Fine-tuning for Speaker Verification". *2023 IEEE International Conference on Acoustics, Speech and Signal Processing* (**ICASSP**), *June. 2023*

[C4] **Leying Zhang***, Zhengyang Chen* and Yanmin Qian. "Enroll-Aware Attentive Statistics Pooling for Target Speaker Verification ". *23rd Annual Conference of the International Speech Communication Association* (**InterSpeech**), *Sep. 2022*

[C5] **Leying Zhang**, Zhengyang Chen and Yanmin Qian. "Knowledge Distillation from Multi-Modality to Single-Modality for Person Verification". *22nd Annual Conference of the International Speech Communication Association* (**InterSpeech**), *Sep. 2021*

[C6] Linfeng Yu, Wangyou Zhang, Chenpeng Du, **Leying Zhang**, Zheng Liang, Yanmin Qian. "Generation-Based Target Speech Extraction with Speech Discretization and Vocoder". *2024 IEEE International Conference on Acoustics, Speech and Signal Processing* (**ICASSP**), *April. 2024*

[C7] Yichong Leng, Zhifang Guo, Kai Shen, Xu Tan, Zeqian Ju, Yanqing Liu, Yufei Liu, Dongchao Yang, **Leying Zhang**, Kaitao Song, Lei He, Xiang-Yang Li, Sheng Zhao, Tao Qin, Jiang Bian. "Prompttts 2: Describing and generating voices with text prompt". *The Twelfth International Conference on Learning Representations* (**ICLR**), *May. 2024*

INDUSTRY EXPERIENCE

Research Intern (Remote) — Micrsoft Research Redmond
Supervised by Yao Qian — *Oct. 2024 - Present*

Dialogue Generation: Ongoing Project

Research Intern (Remote) — Micrsoft Research Redmond
Supervised by Yao Qian — *Apr. 2023 - Mar. 2024*

Target speech extraction: Investigated diffusion-based model for target speech extraction. Proposed an efficient approach by combining diffusion and discriminative methods for handling multi- and single-speaker scenarios in both noisy and clean conditions.
Text-to-Dialogue Generation: Investigated Conversational Voice Mixture Generation, a novel model for zero-shot, human-like, multi-speaker, multi-round dialogue speech generation

Research Intern — Microsoft Research Asia
Supervised by Xu Tan — *Nov. 2022 - Mar. 2023*

Audio generation: Implemented vector-quantized diffusion model with classifier-free guidance. Achieved 10% improvement over baseline. Investigated latent diffusion model's effects by fine-tuning Stable diffusion.
Text-to-speech: Utilized vector-quantized diffusion model for text-to-speech on large-scale dataset with different neural audio codecs. Generated high-quality speech and get improvements on zero-shot text-to-speech.

TEACHING EXPERIENCE

Teaching Assistant - Machine Learning — Fall, 2022
Teaching Assistant - Mobile Communication Systems — Fall, 2022

HONORS AND AWARDS

National Scholarship — 2022
NeurIPS 2024 Scholar Award — 2024
First place in CN-Celeb Speaker Recognition Challenge 2022 — 2022
ISCA and Interspeech Travel Grant — 2021
Outstanding Graduates of Shanghai — 2021
Outstanding student leader of SJTU — 2021
Guanghua Scholarship — 2020
SJTU Class B Scholarship — 2019

SKILLS

Programming skills: Python, Pytorch, C/C++, Matlab
Languages: Chinese(native), English(IELTS 7.5), French(DELF B2), Spanish(Beginner)
Extracurriculars: Piano, Yoga, Badminton