Zhisheng Zheng

Shanghai Jiao Tong University, Shanghai, China
matrixzheng01@gmail.com

 Homepage | Scholar | Github | twitter

EDUCATION

Visiting Scholar of Computer Science — Austin, USA
College of Natural Science, The University of Texas at Austin — *May 2023 - Jan 2024*

Member of Zhiyuan Honors Program of Engineering — Shanghai, China
Zhiyuan College, Shanghai Jiao Tong University; Top 5% — *Sept. 2020 - June 2024 (expected)*

Bachelor of Information Engineering — Shanghai, China
School of Electronic Information and Electrical Engineering, SJTU; GPA: 3.79 — *Sept. 2020 - June 2024 (expected)*

PUBLICATIONS

- **MT4SSL: Boosting Self-Supervised Speech Representation Learning by Integrating Multiple Targets** —*INTERSPEECH 2023* **Best student paper shortlist**
Ziyang Ma, **Zhisheng Zheng**, Changli Tang, Yujin Wang, Xie Chen.

- **Unsupervised Active Learning: Optimizing Labeling Cost-Effectiveness for Automatic Speech Recognition** —*INTERSPEECH 2023*
Zhisheng Zheng, Ziyang Ma, Yu Wang, Xie Chen.

- **Pushing the Limits of Unsupervised Unit Discovery for SSL Speech Representation** —*INTERSPEECH 2023*
Ziyang Ma, **Zhisheng Zheng**, Guanrou Yang, Yu Wang, Chao Zhang, Xie Chen.

- **Fast-HuBERT: An Efficient Training Framework for Self-Supervised Speech Representation Learning** —*ASRU 2023*
Guanrou Yang, Ziyang Ma, **Zhisheng Zheng**, Yakun Song, Zhikang Niu, Xie Chen.

- **Exploring Effective Distillation of Self-Supervised Speech Models for Automatic Speech Recognition** —*ASRU 2023*
Yujin Wang, Changli Tang, Ziyang Ma, **Zhisheng Zheng**, Xie Chen, Wei-Qiang Zhang.

- **Front-End Adapter: Adapting Front-End Input of Speech based Self-Supervised Learning for Speech Recognition** —*ICASSP 2023*
Xie Chen, Ziyang Ma, Changli Tang, Yujin Wang, **Zhisheng Zheng**.

RESEARCH EXPERIENCE

MoE Key Lab of Artificial Intelligence, AI Institute, X-LANCE Lab, SJTU — Shanghai, China
Research Intern, Advised by Prof. Xie Chen — *Dec. 2021 - Present*

- **Improve ASR Performance Through Self-Supervised and Unsupervised Learning**
Utilizing the *fairseq* framework, replicated mainstream Self-Supervised Learning (SSL) models such as wav2vec 2.0, HuBERT, data2vec, and Wav2vec-U 2.0. By synergistically integrating the unique features of these models, further boosted their performance in Automatic Speech Recognition (ASR).

- **Unsupervised Active Learning for Automatic Speech Recognition**
This work enhances SSL's capability to further reduce labeling costs using active learning. Through unsupervised derivation of speech units and a contrastive data selection method, achieve an over 11% improvement in word error rate (WER) with equivalent labeled data or halve the labeling cost while maintaining the same WER, compared to random selection.

Speech, Audio, and Language Technologies (SALT) Lab, UT-Austin — Austin, USA
Research Intern, Advised by Prof. David Harwath and Eunsol Choi — *May, 2023 - Present*

- **Audio and Language Understanding (LLM)**
Leveraging advanced language models (LLaMA) for spatial audio perception, which demonstrating superior performance in complex audio processing tasks, setting a new standard in spatial audio research.

Selected Awards

- SenseTime Scholarship for Undergraduate AI Researchers (30 winners nationwide each year)   2023
- Rongchang Science and Technology Innovation Scholarship (<0.1%)   2023
- Tencent Scholarship (Top 2%)   2021
- Zhiyuan College Honors Scholarship (Top 5%)   2021, 2022, 2023

Skills

- **Coding**: Python (Pytorch), C/C++, Bash.

- **Languages**: Chinese (Native), English (TOEFL 104).