Zhisheng Zheng

Shanghai Jiao Tong University, Shanghai, China
matrixzheng01@gmail.com
👤 Homepage | 𝕾 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 Electrical Engineering — Shanghai, China
School of Electronic Information and Electrical Engineering, SJTU; GPA: 3.8 — *Sept. 2020 - June 2024 (expected)*

REPRESENTATIVE PUBLICATIONS

- **BAT: Learning to Reason about Spatial Sounds with Large Language Models**
  Zhisheng Zheng, Puyuan Peng, Ziyang Ma, Xie Chen, Eunsol Choi, David Harwath   *—preprint*

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

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

- **Emotion2vec: Self-Supervised Pre-Training for Speech Emotion Representation**
  Ziyang Ma, **Zhisheng Zheng**, Jiaxin Ye, Jinchao Li, Zhifu Gao, Shiliang Zhang, Xie Chen.   *—preprint*

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

- **EAT: Self-Supervised Pre-Training with Efficient Audio Transformer**
  Wenxi Chen, Yuzhe Liang, Ziyang Ma, **Zhisheng Zheng**, Xie Chen.   *—preprint*

RESEARCH EXPERIENCE

Speech, Audio, and Language Technologies (SALT) Lab, UT-Austin — Austin, USA
Research Intern, Advised by Prof. David Harwath and Eunsol Choi — *May, 2023 - Jan, 2024*

- **LLM-based Spatial Audio Understanding**
  we introduced BAT, which merges binaural acoustic scene analysis with large language model reasoning, overcoming the challenge of limited in-the-wild spatial sound data by creating a novel dataset SPATIALSOUNDQA. Utilizing SPATIAL-AST and integrating it with LLaMA-2 7B, BAT excels in spatial sound perception and reasoning tasks.

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

- **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.

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).