ZHIKANG NIU

(+86) 186-2861-8289 · zhikangniu@sjtu.edu.cn · GitHub @ZhikangNiu ·

EDUCATION

Shanghai Jiao Tong University (SJTU), Shanghai, China   2024.9 – Present

Ph.D. Student in Department of Computer Science and Engineering (CS), expected March 2029

Xidian University, Shaanxi, China   2020.9 – 2024.6

B.S. in Artificial Intelligence (AI). National Scholarship (2021-2022, rank 2/138)

INTERNSHIP

Microsoft Research Asia General Artificial Intelligence Group, Beijing, China   2023.8 – 2024.8

Research Intern   Manager: Shujie Liu and Long Zhou

Under the supervision of Dr. Shujie Liu and Dr. Long Zhou, explored audio discretization models (Audio Codec) and para-linguistic controllable speech synthesis models (Text to Speech).

- Independently completed the training, inference, and evaluation code for the audio codec model based on the fairseq framework. Explored robust audio discretization models on this framework. **One paper**, accepted by 2024 IEEE Spoken Language Technology Workshop (SLT 2024).
- Conducted research on existing speech synthesis work and built a Chinese speech synthesis dataset with para-linguistic annotations. Explored expressive and controllable speech synthesis models based on decoder-only models such as VALLE and MELLE.

PROJECT

Thorough PyTorch (Open-Source Chinese PyTorch Tutorial)   2021.7 - Present

- Thorough PyTorch is an open-source Chinese PyTorch tutorial and has received over 2,300 stars on GitHub.
- **Project Link**: https://github.com/datawhalechina/thorough-pytorch

High Fidelity Neural Audio Compression (Paper Reproduction)   2023.2 - 2023.9

- We reproduced the training process of EnCodec on the LibriTTS dataset. The model outperformed the open-source EnCodec on PESQ, STOI, Mel Distance, and STFT Distance speech quality metrics on the test set. The code and weights have been open-sourced, gaining 127 stars
- **Project Link**: https://github.com/ZhikangNiu/encodec-pytorch

PAPER LIST

NDVQ: Robust Neural Audio Codec with Normal Distribution-Based Vector Quantization.

- **Zhikang Niu**, Sanyuan Chen, Long Zhou, Ziyang Ma, Xie Chen, Shujie Liu*
- **IEEE Spoken Language Technology Workshop (SLT)**, 2024.

Fast-HuBERT: An Efficient Training Framework for Self-Supervised Speech Representation Learning.

- Guanrou Yang, Ziyang Ma, Zhisheng Zheng, Yakun Song, **Zhikang Niu**, Xie Chen*
- **IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)**, 2023.

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.
- Submitted to **IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP)**

A Controllable Emotion Voice Conversion Framework with Pre-trained Speech Representations.

- Tianrui Wang, Meng Ge, **Zhikang Niu**, Chunyu Qiang, Cheng Gong, Ziyang Ma, Xiaobao Wang, Xie Chen, Longbiao Wang, Dangjian Wu*
- Submitted to **IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP)**

HONORS AND AWARDS

Stars of Tomorrow, Microsoft Research Asia	2024
National Scholarship, Ministry of Education in China.	2022
Meritorious Winner, Interdisciplinary Contest In Modeling.	2021
The First Prize Scholarship, Xidian University.	2021,2023