

Tinglong Zhu

Phone: (412)-713-8565 Email: tinglonz@andrew.cmu.edu
LinkedIn: [linkedin.com/in/tinglonz](https://www.linkedin.com/in/tinglonz) Personal Page: <https://kurisujhin.github.io>

EDUCATION

Carnegie Mellon University

Pittsburgh, PA

- Master of Science in Intelligent Information Systems (MIIS-21), expected May 2024 GPA: 3.87/4.0

Duke Kunshan University / Duke University (Dual degree)

Kunshan, China

- Bachelor of Science, graduated May 2022; Major: Data Science GPA: 3.89/4.0

PUBLICATIONS

[1] **Tinglong Zhu**, Xingming Wang, Xiaoyi Qin, Ming Li (2022) Source Tracing: Detecting Voice Spoofing. *Asia-Pacific Signal and Information Processing Association (APSIPA) SS03: Security Techniques of Speaker Recognition 2022*

[2] **Tinglong Zhu**, Xiaoyi Qin, Ming Li (2021) Binary Neural Network for Speaker Verification. *Conference of the International Speech Communication (Interspeech) 2021*

[3] Xingming Wang, Xiaoyi Qin, Ming Li, **Tinglong Zhu**, Chao Wang, Shilei Zhang (2021) The DKU-CMRI System for the ASVspoof 2021 Challenge: Vocoder Based Replay Channel Response Estimation. *ASVspoof 2021 Workshop*

INTERNSHIP EXPERIENCE

Alibaba

Hangzhou, China

Research Intern, DAMO Academy Speech Team

May 2023 – Now

- Take advantage of Residual Vector Quantization (RVQ), Neural Codec, and Large Language Model (LLM) techniques to design and implement better Text-to-Speech (TTS) system
- Reimplement SoundStream codec (SEANet based neural codec) and SoundStorm model (350M Conformer)
- Build pipeline for training and evaluating new TTS systems on PAI clusters

Microsoft

Beijing, China

Research Intern, Cloud + AI Speech Team

May 2021 – Aug 2021

- Conducted joint training of ASR TTS ASV models to enlarge data sources of small languages, and enhanced human-computer interaction for more languages and dialects worldwide
- Built pipeline (Listen, Attend and Spell) for Automatic Speech Recognition based on ESPnet, pipeline for Automatic Speaker Verification (ASV) based on SEResNet34 and pipeline for Text-to-Speech (Tacotron2)
- Trained models on Azure clusters
- Achieved following performance improvement: TTS system: Word Error Rate (WER): Reduced from 16.5% to 15.5%. Mean Cosine Similarity: Improves from 0.55 to 0.73

RESEARCH EXPERIENCE

Project: Source Tracing: Detecting Voice Spoofing, Duke Kunshan University

Feb 2022 – Aug 2022

Proposed a Training Strategy for Classifying Attributes of Spoof Systems (RawNet2, ResNet34 based systems)

- Achieved 88.41%, 77.54%, 51.46% accuracy for Conversion, Waveform Generator, and Speaker Representation attributes, respectively on RawNet2 based system
- Improved the anti-spoof system performance by 20% and enhance the robustness of anti-spoof systems at the same time
- Submitted the research outcome to *APSIPA ASC 2022 Special Session 03* [1]

Project: Binary Neural Networks for Speaker Verification, Duke Kunshan University

Jan 2021 – Apr 2021

Applied Binary Weight Neural Network to Automatic Speaker Verification (ASV) systems (based on ResNet34).

- Verified theoretically computational cost down by 50% for inferencing and training network and theoretically network downsize 32X less than original network while keeping an acceptable performance: 5.355% EER for text-independent ASV, while 3.098% EER for text-dependent ASV
- Published the research outcome in *Interspeech 2021* [2]

SKILLS

Programming Languages: Python, C/C++, Java, Bash, MATLAB, R

Software & Open-Source Toolkits: PyTorch, ESPnet, Kaldi toolkit, Linux, TensorFlow, Azure, AWS, SciPy toolkit, Keras, Matplotlib