Tingle Li

Research Assistant

(86) 16675164723 | tingle.li@outlook.com | tinglok.netlify.com | tinglok

Education

Tiangong University *Tianjin, China*

B.Eng. in Computer Science and Technology, GPA: 3.74/4.0 (top 5%) *Sep. 2016-Jun. 2020*

- Honors Program of Artificial Intelligence (an elite program for top 10% students)
- Advisor: Prof. Ming Li (Duke University), Prof. Rize Jin (Tiangong University)

Research Interests

Source Separation, Music Information Retrieval, Audio-Visual Interaction, Multimodal Learning

Publications

Optimal Mapping Loss: A Faster Loss for End-to-End Speaker Diarization

- Qingjian Lin*, **Tingle Li***, Lin Yang, Junjie Wang, and Ming Li.
- *In Proc. Speaker Odyssey,* Tokyo, Japan, November 2020.

Atss-Net: Target Speaker Separation via Attention-based Neural Network

- **Tingle Li**, Qingjian Lin, Yuanyuan Bao, and Ming Li.
- *In Proc. INTERSPEECH,* Shanghai, China, October 2020.

The DKU Speech Activity Detection and Speaker Identification Systems for Fearless Steps Challenge Phase-02

- Qingjian Lin, **Tingle Li**, and Ming Li.
- *In Proc. INTERSPEECH,* Shanghai, China, October 2020.

Sams-Net: A Sliced Attention-based Neural Network for Music Source Separation

- **Tingle Li**, Jiawei Chen, Haowen Hou, and Ming Li.
- Submitted to *ISCSLP 2021*.

Selected Honors & Awards

Merit Scholarship for Outstanding Students (top 7%), *May. 2017, 2018, 2019*
3rd* Prize, the "Lan-Qiao" Cup National Selection Competition, *May. 2017
1st* Prize, the China Students Innovation and Entrepreneurship Competition (top 7.9%), *May. 2019
Rank *1st* for SID and *3rd* for SAD among the world, Fearless Steps Challenge Phase-02, *May. 2020*
Best Undergraduate Dissertation (top 1%), *Jun. 2020*

Research Experiences

Institute for Interdisciplinary Information Sciences (IIIS) *Tsinghua University, China*

Advisor: Prof. Hang Zhao *Jul. 2020 - Present*

- *(On-going) A Contrastive Learning Based Adversarial Approach for Non-Parallel Voice Conversion*
  - Given two corpora from two speakers, this task aims to convert one's voice to the others.
  - Proposed a voice conversion model based on noise contrastive estimation (NCE) loss, which enables one-sided conversion in the non-parallel voice conversion setting, while improving speech quality and reducing training time.
- *(On-going) Research on Lip Reading with Talking Head Generation*
  - Explored methods of improving the performance of audio-visual alignment with attention mechanism.

Speech and Multimodal Intelligent Information Processing (SMIIP) Lab *Duke Kunshan University, China*

ADVISOR: PROF. MING LI *Jul. 2019 - Jul. 2020*

- *Atss-Net: Target Speaker Separation via Attention-based Neural Network*
  - Given a referenced utterance of the target speaker, and a mixed utterance containing the target speaker, this task aims at filtering the target speaker's voice from the mixed utterance.
  - Proposed a target speaker separation model based on attention neural network, which leveraged the attention mechanism to fuse the mixed spectrogram and the target speaker embedding.
  - Collaborated with Xiaomi Corporation to land in application, where demos are available here.
- *Sams-Net: A Sliced Attention-based Neural Network for Music Source Separation*
  - Given a musical utterance, our goal is to recover the individual stems from the mixed signal (i.e., vocals, drums, bass and other).
  - Proposed a new attention mechanism called Sliced Attention, where the scope of attention is narrowed down to the intra-chunk features that are most likely to affect each other.
  - Our model has achieved the state-of-the-art performance, although it contained fewer parameters compared with baselines, where demos are available here.
- *Singing Voice Separation for Singer Verification*
  - Given a musical utterance, this task aims to identify who is singing.
  - Designed a paradigm that input the musical utterance into the separation system to separate the vocals of it, then used the speaker verification system to identify the singer.
  - Adopted as the third party duplicate checking technique for Guinness Records of the CCTV National Day Celebration Program, where the news can be found here (in Chinese).

Tianjin Key Laboratory of Autonomous Intelligence Technology and Systems *Tiangong University, China*

ADVISOR: PROF. REZE JIN *Sep. 2018 - Jun. 2019*

- *Research on Deep Learning-based Music Information Retrieval*
  - Mastered basic knowledge about audio processing, such as short-time fourier transform (STFT), GMM-HMM, filter-bank and MFCC, etc.
  - Explored methods of improving the performance of singing voice separation model via Transformer.