

# Tingle Li

(86)166-7516-4723  
github.com/Tinglok  
tingle.li@outlook.com  
https://tinglok.netlify.com

RESEARCH ASSISTANT, TSINGHUA UNIVERSITY

## EDUCATION

**Tsinghua University** Beijing, China  
*Post-bachelor* at Institute for Interdisciplinary Information Sciences Jul. 2020-Jul. 2021  
• Research-oriented program during COVID-19, also affiliated to Shanghai Qi Zhi Institute.  
• **Advisor:** Prof. *Hang Zhao*

**Tiangong University** Tianjin, China  
*B.Eng.* in Computer Science and Technology Sep. 2016 - Jun. 2020  
• Pilot class of Artificial Intelligence (an elite program for top 10% students)  
• **GPA:** 3.74/4.0 **Ranking:** 3/124  
• **Advisor:** Prof. *Ming Li* (Duke University)  
• **Selected Courses:** Machine Learning (97), Mathematical Foundation of Artificial Intelligence (98), Data Mining (91), Computer Vision (91), Numerical Analysis (94), Applied Statistics (91), Linear Algebra (93), Probability and Statistics (93), Advanced Mathematics (94), Data Structure (94), Algorithm Design and Analysis (93), High-level Programming Language (95)

## RESEARCH INTERESTS

**Speech Processing:** Source Separation, Voice Conversion, Music Information Retrieval  
**Multimodal Learning:** Audio-Visual Learning, Self-Supervised Learning, Representation Learning

## PUBLICATIONS

**Tingle Li**, Yichen Liu, Chenxu Hu, Hang Zhao. “CVC: Contrastive Learning for Non-parallel Voice Conversion”. Submitted to INTERSPEECH 2021.

**Tingle Li**, Jiawei Chen, Haowen Hou, Ming Li. “Sams-Net: A Sliced Attention-based Neural Network for Music Source Separation”. *In Proc. ISCSLP*, Hong Kong, China, January 2021. (Oral Presentation)

Qingjian Lin\*, **Tingle Li**\*, Lin Yang, Junjie Wang, Ming Li. “Optimal Mapping Loss: A Faster Loss for End-to-End Speaker Diarization”. *In Proc. Speaker Odyssey*, Tokyo, Japan, November 2020.

**Tingle Li**, Qingjian Lin, Yuanyuan Bao, Ming Li. “Atss-Net: Target Speaker Separation via Attention-based Neural Network”. *In Proc. INTERSPEECH*, Shanghai, China, October 2020.

Qingjian Lin, **Tingle Li**, Ming Li. “The DKU Speech Activity Detection and Speaker Identification Systems for Fearless Steps Challenge Phase-02”. *In Proc. INTERSPEECH*, Shanghai, China, October 2020.

## AWARDS & ACHIEVEMENTS

Best Undergraduate Dissertation (**top 1%**) Jun. 2020  
Merit Scholarship for Outstanding Students (**top 5%**) Dec. 2017, 2018, 2019  
1<sup>st</sup> for SID and 3<sup>rd</sup> for SAD among 50 teams, Fearless Steps Challenge Phase-02 May. 2020  
1<sup>st</sup> Prize, China Students Innovation and Entrepreneurship Competition May. 2019  
3<sup>rd</sup> Prize, Lan Qiao Cup National Selection Competition May. 2017

## TECHNICAL SKILLS

**Programming Languages:** Python, Java, C/C++, Shell, HTML/CSS, and MATLAB  
**Tools and Frameworks:** Git, L<sup>A</sup>T<sub>E</sub>X, PyTorch, TensorFlow, Keras, OpenCV, and Kaldi

RESEARCH  
EXPERIENCE

**Institute for Interdisciplinary Information Sciences (IIIS)**

Advisor : Prof. *Hang Zhao*

Tsinghua University

*Jul. 2020 - Present*

*CVC: Contrastive Learning for Non-parallel Voice Conversion*

- Given a speech signal from source speakers, we aim to convert it to the voice of a target speaker while preserving the speech content.
- Exploring a voice conversion model based on noise contrastive estimation (infoNCE) loss, which enables one-way conversion in the non-parallel voice conversion setting, while comparatively improving speech intelligibility and effectively reducing training time.

*(Ongoing) Neural Dubber: Dubbing for Silent Videos According to Scripts*

- Proposed a new audio-visual task: given a silent video and its corresponding script, our goal is to dub it temporal synchronously.
- Exploring soft & hard attention mechanism to align lip motion with generated speech, which allows us to dub naturally even when modifying some words of the script.

**Language Intelligence Team, Speech Lab**

Advisor : Prof. *Ming Li*

Samsung Research Institute Beijing

*Jan. 2020 – Mar. 2020*

*Research on Joint Speech Enhancement and Separation System*

- Implemented a simultaneous speech enhancement and separation system, which explores the feasibility of using it in [Bixby](#).

**Speech and Multimodal Intelligent Information Processing Lab**

Advisor : Prof. *Ming Li*

Duke Kunshan University

*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 music (i.e. vocals, drums, bass and others).
- 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](#).

*Research on Singing Voice Separation for Singer Verification*

- Given a musical utterance, this task aims to identify who is singing.
  - Designed a paradigm that input song to the separation system and separate the vocals from it, then used the speaker verification system to identify singers.
  - Adopted as the third party duplicate checking technique for Guinness Records of the CCTV National Day Celebration Program, where the press coverage can be found [here](#) (in Chinese).
-