XU LI

http://lixucuhk.github.io
(+86)15712028892 ⋄ xuli@se.cuhk.edu.hk

BIOGRAPHY

I am currently a Ph.D. student in the Department of Systems Engineering and Engineering Management, the Chinese University of Hong Kong. My supervisor is Prof. Helen Meng.

The Chinese University of Hong Kong, Hong Kong SAR, China *August 2017 - Present*
Ph.D. candidate in Dept. of Systems Engineering and Engineering Management;

University of Science and Technology of China, Hefei, Anhui, China *July 2013 - June 2017*
B.E. in Dept. of Information Science and Technology;

MANUSCRIPTS & PUBLICATIONS

Investigating Robustness of Adversarial Samples Detection for Automatic Speaker Verification,
Xu Li, Na Li, Jinghua Zhong, Xixin Wu, Xunying Liu, Dan Su, Dong Yu and Helen Meng, in the Proceedings of Interspeech, 2020

Bayesian x-vector: Bayesian Neural Network based x-vector System for Speaker Verification,
Xu Li, Jinghua Zhong, Jianwei Yu, Shoukang Hu, Xixin Wu, Xunying Liu and Helen Meng, in Speaker Odyssey, 2020

Adversarial Attacks on GMM i-vector based Speaker Verification Systems,
Xu Li, Jinghua Zhong, Xixin Wu, Jianwei Yu, Xunying Liu and Helen Meng, in IEEE ICASSP, 2020.

Unsupervised Discovery of Non-native Phonetic Patterns in L2 English Speech for Mispronunciation Detection and Diagnosis,
Xu Li, Shaoguang Mao, Xixin Wu, Kun Li, Xunying Liu and Helen Meng, in the Proceedings of Interspeech, 2018, pp. 2254-2258

Deep Segmental Phonetic Posterior-grams based Discovery of non-categories in L2 English Speech,
Xu Li, Xixin Wu, Xunying Liu and Helen Meng, arXiv preprint arXiv:2002.00205 (2020).

Integrating Articulatory Features into Acoustic-Phonemic Model for Mispronunciation Detection and Diagnosis in L2 English Speech,
Shaoguang Mao, Zhiyong Wu, **Xu Li**, Runnan Li, Xixin Wu, and Helen Meng, in IEEE ICME, 2018, pp. 1-6

Unsupervised Discovery of an Extended Phoneme Set in L2 English Speech for Mispronunciation Detection and Diagnosis,
Shaoguang Mao, **Xu Li**, Kun Li, Zhiyong Wu, Xunying Liu and Helen Meng, in IEEE ICASSP, 2018, pp. 6244-6248

Applying Multitask Learning to Acoustic-Phonemic Model for Mispronunciation Detection and Diagnosis in L2 English Speech,
Shaoguang Mao, Zhiyong Wu, Runnan Li, **Xu Li**, Helen Meng and Lianhong Cai, in IEEE ICASSP, 2018, pp. 6254-6258

Automatic lexical stress and pitch accent detection for L2 English speech using multi-distribution deep neural networks,
Kun Li, Shaoguang Mao, **Xu Li**, Zhiyong Wu and Helen Meng, in Speech Communication, vol. 96, pp. 28-36, 2018

Comparative Study of Parametric and Representation Uncertainty Modeling for Recurrent Neural Network Language Models,
Jianwei Yu, Max WY Lam, Shoukang Hu, Xixin Wu, **Xu Li**, Yuewen Cao, Xunying Liu and Helen Meng, in the Proceedings of Interspeech, 2019, pp. 3510-3514

End-to-end Code-switched TTS with Mix of Monolingual Recordings,
Yuewen Cao, Xixin Wu, Songxiang Liu, Jianwei Yu, **Xu Li**, Zhiyong Wu, Xunying Liu and Helen Meng, in IEEE ICASSP, 2019, pp. 6935-6939

Speech Emotion Recognition Using Capsule Networks,
Xixin Wu, Songxiang Liu, Yuewen Cao, **Xu Li**, Jianwei Yu, Dongyang Dai, Xi Ma, Shoukang HU, Zhiyong Wu, Xunying Liu and Helen Meng, in IEEE ICASSP, 2019, pp. 6695-6699

RESEARCH EXPERIENCE & PROJECT

07/2019 - Present: **speaker verification**, **anti-spoofing counter-measures**
In the current period, I focus on the development of speaker verification systems, and counter-measures against the followed spoofing threats. Recently, I verified the vulnerability of i-vector and x-vector based speaker embedding systems to adversarial attacks. This work was recently accepted to ICASSP 2020. I also investigated to incorporate DNN x-vector systems with Bayesian neural networks to enhance the system generalization ability across the mismatched evaluation data, accepted to Speaker Odyssey 2020.

08/2017 - 06/2019: **mispronunciation detection and diagnosis**, **second language learning**
In this period, I focused on modeling the non-categorical pronunciations in second language (L2) English speech. The proposed framework captured the non-categories in L2 English speech in an unsupervised manner. These non-categories cannot be described by any native English phoneme. This modeling of non-categories results in a more effective mispronunciation detection and more precise feedback (diagnosis) to language learners.

02/2017 - 06/2017: Research Assistant at **Tsinghua University (Shenzhen)**, supervised by Prof. Zhiyong Wu
Focus on acoustic-phonemic modeling of L2 English speech units.

HONORS & AWARDS

2015: Kwang-Hua Scholarship, Kwang-Hua Educational Foundation
2014: Third Prize of Excellent Undergraduate Scholarship, USTC
2013: Freshman Scholarship, USTC
2015: Third Prize in the Electronics Development Competition, USTC

PROGRAMMING SKILLS

Language: Python > C/C++ > MATLAB > Java
Frameworks: Pytorch, Kaldi, TensorFlow