



## Basic Information

---

Shuai Wang      Mobile: +86 176 2564 1972      E-mail: [wsstriving@gmail.com](mailto:wsstriving@gmail.com)      Zip Code: 200240  
 Personal page: <https://speechlab.sjtu.edu.cn/members/shuai-wang>  
 Affiliation: School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China

## Education

---

09/2015-Present    **Shanghai Jiao Tong University, School of Electronic Information and Electrical Engineering**  
 Ph.D. Candidate in Computer Science and Engineering  
 09/2014-07/2015    **Shanghai Jiao Tong University, School of Electronic Information and Electrical Engineering**  
 Ph.D. Candidate in Software Engineering  
 09/2010-07/2014    **Northwestern Polytechnical University, School of Software and Microelectronics**  
 B.E in Software Engineering

## Internship Experiences

---

02/2019-10/2019    **Brno University of Technology, Faculty of Information Technology**  
 Research assistant at the speech group  
 Investigate the usage of phonetic information for speaker embedding learning.  
 Work on Challenges such as DIHARD 2019, Voxceleb 2019 and SRE 2019  
 06/2018-12/2018    **AISPEECH**  
 Introduce deep speaker embeddings to the real-word system, which requires less memory and achieves better performance than i-vector.  
 Research outcomes are used in OPPO Reno (Smart phone released in 2019)  
 05/2017-08/2017    **AISPEECH**  
 Co-channel Multi-talker Speaker Identification System, aiming to recognize the identifies of three young Chinese Drama singers (about 9 years old) through their chorus.  
 Recorded one TV show called "Human-AI Competition" on the China Central Television (CCTV)

## Research Topics

---

Topic                **Speaker Diarization**  
 Description        The "Who spoke when" problem  
 Achievements     • A related paper was accepted by Interspeech 2019  
                           • Rank **1st** in all 4 tracks in DIHARD Challenge 2019

Topic                **Robust Deep Speaker Embeddings**  
 Description        Extract fixed-dimension embeddings using deep neural networks, research topics include analysis of different speaker embeddings, metrics for more discriminative speaker embeddings and compensation methods in the embedding space.  
 Achievements     • Related papers were accepted by Interspeech 2017, Interspeech 2018, ICASSP2018, ISCLSP2018, Interspeech 2019 and TASLP  
                           • VoxSRC Challenge 2019: Rank **1st** in both 2 Tracks  
                           • NIST SRE2019: Rank **7th** in CMN Track and **5th** in VAST Track

Topic                **Deep Generative Models for Speaker Recognition**

Description	Generative Adversarial Nets (GAN) and Variational Auto-encoder (VAE) based data augmentation for robust speaker recognition
Achievements	Related papers were accepted by ISCSLP 2018 and Interspeech 2019
Topic	<b>Co-channel Multi-talker Speaker Identification</b>
Description	A sub-problem of the “Cocktail Party Problem”, aiming at recognizing the identities of multiple talkers via their overlapped speech
Achievements	<ul style="list-style-type: none"> <li>Recorded one TV show called “Human-AI Competition” on the China Central Television (CCTV), won the competition and promoted to the final round.</li> <li>A related paper was accepted by ICASSP 2018</li> </ul>
Topic	<b>Speaker Anti-spoofing</b>
Description	Detect the potential attacks (Replay, TTS) on speaker verification systems
Achievements	Related papers were accepted by Interspeech 2019 ASVSpooF 2019: Rank <b>5th</b> in LA Track, <b>8th</b> in PA Track

---

## Awards

2014	Outstanding graduates of Northwestern Polytechnical University (5% )
2011,2012,2013	First Class Scholarship of Northwestern Polytechnical University (15% )
2012,2013	National Scholarship (1%)

---

## Languages

**Native Language** Chinese

**Other Language** English

---

## Paper List

### Journal papers

- Yanmin Qian, Chao Weng, Xuankai Chang, **Shuai Wang** and Dong Yu. *Past Review, Current Progress and Challenges Ahead on Cocktail Party Problem*. *Frontiers of Information Technology & Electronic Engineering* 2018
- **Shuai Wang**, Zili Huang, Yanmin Qian and Kai Yu. *Discriminative Neural Embedding Learning for Short-Duration Text-Independent Speaker Verification*. *IEEE/ACM Transactions on Audio Speech and Language Processing* 2019
- **Shuai Wang**, Yexin Yang, Zhanghao Wu, Yanmin Qian and Kai Yu. *Data Augmentation using Deep Generative Models for Embedding based Speaker Verification*. *Submitted to IEEE/ACM Transactions on Audio Speech and Language Processing*.

### Conference papers

- **Shuai Wang**, Yanmin Qian and Kai Yu. *What Does the Speaker Embedding Encode?* *Interspeech* 2017.
- Xiaowei Jiang, **Shuai Wang**, Xu Xiang, Yanmin Qian. *Integrating Online i-vector into GMM-UBM for Text-dependent Speaker Verification*. *APSIPA* 2017.
- **Shuai Wang**, Yanmin Qian and Kai Yu. *Focal KL-Divergence based Dilated Convolutional Neural Networks for Co-channel Speaker Identification*. *ICASSP 2018 (IEEE Ganesh N. Ramaswamy Memorial Award)*
- Zili Huang, **Shuai Wang** and Yanmin Qian. *Joint i-vector with End-to-End system for Short Duration Text-independent speaker verification*. *ICASSP* 2018

- **Shuai Wang\***, Zili Huang\* and Kai Yu. *Angular Softmax for Short-Duration Text-independent Speaker Verification. (\* Joint First Author) Interspeech 2018 (ISCA Travel Grant)*
- **Shuai Wang**, Heinrich Dinkel, Yanmin Qian and Kai Yu. *Covariance Based Deep Feature for Text-dependent Speaker Verification. IScIDE 2018*
- Yexin Yang, **Shuai Wang**, Man Sun, Yanmin Qian and Kai Yu. *Generative Adversarial Networks based X-vector Augmentation for Robust Probabilistic Linear Discriminant Analysis in Speaker Verification. ISCSLP 2018.*
- **Shuai Wang**, Zili Huang and Kai Yu. *Deep Discriminant Analysis for i-vector Based Robust Speaker Recognition. ISCSLP 2018.*
- **Shuai Wang**, Yexin Yang, Tianzhe Wang, Yanmin Qian and Kai Yu. *Knowledge Distillation for Small Foot-print Deep Speaker Embedding. ICASSP 2019.*
- **Shuai Wang**, Johan Rohdin, Lukáš Burget, Oldřich Plchot, Yanmin Qian, Kai Yu and Jan Černocký. *On the Usage of Phonetic Information for Text-independent Speaker Embedding Extraction. Interspeech 2019.*
- Zhanghao Wu, **Shuai Wang**, Yanmin Qian and Kai Yu. *Data Augmentation using Variational Auto-encoder for Embedding based Speaker Verification. Interspeech 2019.*
- Hongji Wang, Heinrich Dinkel, **Shuai Wang**, Yanmin Qian and Kai Yu. *Cross-domain replay spoofing attack detection using domain adversarial training. Interspeech 2019. (ISCA Travel Grant)*
- Mireia Diez, Lukáš Burget, **Shuai Wang**, Johan Rohdin, Jan Černocký. *Bayesian HMM based x-vector clustering for Speaker Diarization. Interspeech 2019.*
- Yexin Yang, Hongji Wang, Heinrich Dinkel, Zhengyang Chen, **Shuai Wang**, Yanmin Qian and Kai Yu. *The SJTU Robust Anti-spoofing System for the ASVspoof 2019 Challenge. Interspeech 2019.*
- Xu Xiang, **Shuai Wang**, Houjun Huang, Yanmin Qian and Kai Yu. *Margin Matters: Towards More Discriminative Deep Neural Network Embeddings for Speaker Recognition. [arXiv:1906.07317v1](https://arxiv.org/abs/1906.07317v1) APSIPA ASC 2019*
- Hossein Zeinali, **Shuai Wang**, Anna Silnova, Pavel Matějka, Oldřich Plchot. *BUT System Description to VoxCeleb Speaker Recognition Challenge 2019 . [arXiv:1910.12592](https://arxiv.org/abs/1910.12592)*
- Yefei Chen, **Shuai Wang**, Yanmin Qian and Kai Yu. *End-to-End Speaker-Dependent Voice Activity Detection. Presented at NCMMSC2019*
- **Shuai Wang**, Johan Rohdin, Oldřich Plchot, Lukáš Burget, Kai Yu and Jan Černocký. *Investigation of SpecAugment for deep speaker embedding learning. ICASSP 2020 (SPS Travel Grant)*
- **Shuai Wang\***, Yexin Yang\*, Xun Gong, Yanmin Qian and Kai Yu. *Text adaptation for speaker verification with speaker-text factorized embeddings. (\* Joint First Author) ICASSP 2020*
- Zhengyang Chen, **Shuai Wang**, Yanmin Qian and Kai Yu. *Channel Invariant Speaker Embedding Learning With Joint Multi-task and Adversarial Training. ICASSP 2020*
- Federico Landini, **Shuai Wang**, Mireia Diez, Lukáš Burget, Pavel Matějka, Kateřina Žmolíková, Ladislav Mošner, Anna Silnova, Oldřich Plchot, Ondřej Novotný, Hossein Zeinali and Johan Rohdin. *BUT System for DIHARD Speech Diarization Challenge 2019. ICASSP 2020*
- Mireia Diez, Lukáš Burget, Federico Landini, **Shuai Wang**, Jan Černocký. *Optimizing Bayesian HMM based x-vector Clustering for the Second DIHARD Speech Diarization Challenge. ICASSP 2020.*

- XX, **Shuai Wang**, XX. *(A very long author list ordered alphabetically). Analysis of ABC Submission to NIST SRE 2019 CMN and VAST Challenge. Accepted to Odyssey 2020.*
- **Shuai Wang**, Yexin Yang, Yanmin Qian and Kai Yu. *Revisiting the Statistics Pooling Layer in Deep Speaker Embedding Learning. Submitted to Interspeech 2020.*
- Zhengyang Chen, **Shuai Wang** and Yanmin Qian. *Adversarial Domain Adaptation for Speaker Verification using Partially Shared Network. Submitted to Interspeech 2020.*
- Zhengyang Chen, **Shuai Wang** and Yanmin Qian. *Multi-modality Matters: A Performance Leap on VoxCeleb. Submitted to Interspeech 2020.*
- Xun Gong, Zhengyang Chen, Yexin Yang, **Shuai Wang** and Yanmin Qian. *Speaker Embedding Augmentation with Noise Distribution Matching. Submitted to Interspeech 2020.*
- Hongji Wang , Heinrich Dinkel, **Shuai Wang**, Yanmin Qian and Kai Yu. *Dual-adversarial domain adaptation for generalized replay attack detection. Submitted to Interspeech 2020.*
- Xuenan Xu, Heinrich Dinkel, **Shuai Wang**, Mengyue Wu and Kai Yu. *Noise Robust Speaker Verification by Denoising x-vector with Autoencoder. Submitted to Interspeech 2020.*