Shaojin Ding

shjd@tamu.edu, +1-(979)-412-5190
LinkedIn: https://www.linkedin.com/in/shaojin-ding-5b9092162/

Research Interests

Speech synthesis, Speech recognition, Person Re-ID, Machine learning

Key Skills

Programming Languages: **Python**, **C/C++**, MATLAB, HTML, Javascript
Toolkits: TensorFlow, Pytorch, Kaldi, Caffe, Django

Education

Texas A&M University, College Station, TX, USA
Ph.D. program in the Department of Computer Science and Engineering 2016 – 2021 (expected)
Xi'an Jiaotong University, Xi'an, Shaanxi, China
B.S. in Electronic and Information Engineering 2011 – 2015
Special Class for the Gifted Youth 2009 – 2011

Selected Projects

Personal Voice Activity Detection (Google intern project) May 2019 – Aug. 2019
Host: Dr. Quan Wang

- For on-device speech processing modules, the computational resources such as CPU, memory, and battery are typically limited. In such cases, we wish to run the computationally intensive components such as speech recognition only when the target user (e.g., the user of the mobile phone) is talking to the device.
- To fill the gap, we propose a personal Voice Activity Detector (VAD) to detect the voice activity only for a target speaker . We propose four different architectures to implement personal VAD.
- We conducted experiments on Librispeech dataset. Results show that personal VAD can effectively detect the voice activity for the target speaker with a lightweight neural network. In addition, there is no degradation if we replace standard VAD with personal VAD for standard speech/non-speech VAD task. This work has been submitted to ICASSP 2019.

Voice Conversion and Foreign Accent Conversion (Ph.D. dissertation project) Fall 2016 – present
Advisor: Dr. Ricardo Gutierrez-Osuna

- Developed various speech synthesis systems that convert the speech from a source speaker to sound as if a target speaker had produced them.
- Focused on sparse representation based methods and deep AutoEncoder based methods (see publications for detailed methods).
- Achieved over 3.6 Mean Opinion Score acoustic quality on ARCTIC dataset and VCTK dataset.

Attentive But Diverse Person Re-ID Fall 2018 – Spring 2019
Advisor: Dr. Zhangyang Wang

- Although local attention mechanisms achieved reasonable results on Person Re-ID tasks, they are

likely to overfit local noise. As a result, we will need to obtain more diverse features to span a compact and representative feature space.

- To address the problem, we designed the network according to an attentive but diverse paradigm. We propose an Attentive but Diverse Network (ABD-Net), that strives to integrate attention modules and diversity regularization and enforces them throughout the entire network.
- ABD-Net achieved new state-of-the-art performance on Market-1501 (95.60% top-1/88.28% mAP), DukeMTMC-Re-ID (89.00% top-1/78.59% mAP), and MSMT17 (82.30% top-1/60.80% mAP) benchmarks. This work has been accepted by ICCV 2019.

Experiences

Research Intern May 2019 – Aug. 2019

Google, New York City, New York

Host: Dr. Quan Wang

- Conducted research on Personal Voice Activity Detection.

Research Assistant Fall 2016 – present

PSI Lab, Department of Computer Science and Engineering, TAMU.

Advisor: Dr. Ricardo Gutierrez-Osuna

- Conducted research on speech synthesis, voice conversion, and accent conversion.

Research Assistant Fall 2014 – Spring 2016

Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University.

Advisor: Dr. Jinjun Wang

- Conducted research on face detection and verification.

Intern Jul. 2014 – Sep. 2014

Wireless and network group, Microsoft Research Asia, Beijing, China

Advisor: Lead Researcher Jacky Shen

- Conducted research on Typing correction model for Touchscreen Keyboards.

Publications

- **Shaojin Ding**, Quan Wang, Shuo-yiin Chang, Li Wan, Ignacio Lopez Moreno "Personal VAD: Speaker-Conditioned Voice Activity Detection", submitted to ICASSP, 2019.
- Tianlong Chen, **Shaojin Ding**, Jingyi Xie, Ye Yuan, Wuyang Chen, Yang Yang, Zhou Ren, Zhangyang Wang "ABD-Net: Attentive but Diverse Person Re-Identification", accepted by ICCV, 2019.
- **Shaojin Ding**, Ricardo Gutierrez-Osuna "Group Latent Embedding for Vector Quantized Variational Autoencoder in Non-Parallel Voice Conversion", accepted *by INTERSPEECH, 2019*.
- Guanlong Zhao, **Shaojin Ding**, Ricardo Gutierrez-Osuna "Foreign Accent Conversion by Synthesizing Speech from Phonetic Posteriorgrams", accepted *by INTERSPEECH, 2019*.
- **Shaojin Ding**, Christopher Liberatore and Ricardo Gutierrez-Osuna "Learning Structured Dictionaries for Exemplar-based Voice Conversion," in *Proceedings of INTERSPEECH, 2018.*
- **Shaojin Ding**, Guanlong Zhao, Christopher Liberatore and Ricardo Gutierrez-Osuna "Improving Sparse Representations in Exemplar-Based Voice Conversion with a Phoneme-Selective Objective Function," in *Proceedings of INTERSPEECH, 2018.*
- **Shaojin Ding**, Christopher Liberatore, Guanlong Zhao, Sinem Sonsaat, Evgeny Chukharev-Hudilainen, John Levis and Ricardo Gutierrez-Osuna "Golden Speaker Builder: an interactive online tool for L2 learners to build pronunciation models," *in Proceedings of Pronunciation in Second Language Learning and Teaching, 2017.*