

YANG (YOLANDA) GAO

244 Roberts Engineering Hall ◊ Carnegie Mellon University

(412) · 577 · 8984 ◊ yanggao@andrew.cmu.edu

EDUCATION

Carnegie Mellon University (CMU)

PhD candidate, in Dept. of Electrical and Computer Engineering Expected grad. date: 05/2021

- Advisors: Rita Singh, and Bhiksha Raj

Master of Science, in Dept. of Biomedical Engineering 05/2015

- Advisor: Yongjie (Jessica) Zhang

Huazhong University of Science and Technology (HUST) 06/2013

Bachelor of Science, in Dept. of Biological Science and Technology

- Graduated with Honor

RESEARCH INTERESTS

Machine Learning, Deep Learning, Speech Signal Analysis, Signal Processing

PROGRAMMING SKILLS

Language: Python, Matlab, C/C++, Lua Toolbox: Pytorch, Tensorflow, Caffe

RESEARCH EXPERIENCES

Understanding Voice Disguise in DeepFake Era 09/2019 - now

PI: Rita Singh, Bhiksha Raj

- Collecting dataset for voice impersonation and deepfakes speech.
- Developing and evaluating features and models for deepfakes speech detection.

Multi-modal Learning of Faces and Voices 09/2018 - 05/2019

PI: Rita Singh, Bhiksha Raj

- Learned the facial structure's correlation with voice.
- Built a multi-domain GANs model with WaveNet vocoder.
- ***The Rembrandt Tutorials Project Demo***

Speech Style Transformation and Feature Learning 06/2017 - 05/2018

PI: Rita Singh, Bhiksha Raj

- Built an end-to-end deep learning based speech synthesis model for voice transformation.
- Advanced the CycleGAN model to meet speech transformation purpose.
- ***Results Demo***

Speaker Identification using Short Phonemes and Breath Sounds 01/2017 - 05/2017

PI: Rita Singh, Bhiksha Raj

- Built i-vector baselines for speaker identification using human breath sounds and short phonemes, such as segmented phoneme AH. The results was compared with the proposed model using Constant-Q features and CNNs.

Large-Scale Road Scene Understanding using Neural Networks 08/2015 - 01/2017

PI: Gary Overett, Bhiksha Raj

- Proposed four methods to optimize the object proposal network of Faster-RCNN structure on Kitti Dataset for advanced small vehicle detection.
- Collected road scene data using lab built comprehensive data collection platform, consisting of Light Detection and Ranging (LIDAR), Cameras, and Inertial Measurement Unit (IMU).

WORK EXPERIENCES

Interactive Text-to-Speech (TTS) with Style Transfer Learning 05/2019 - 08/2019
Research Intern, Facebook AI Speech Team Facebook, Menlo Park, CA

- Extracted target style feature from ASR datasets and transferred the style learning to TTS system to produce target style related function.
- **Results Demo**

Expressive TTS with Unsupervised Prosody Learning 05/2018 - 08/2018
Research Intern, Mentor: Jitong Chen Baidu Research, Sunnyvale, CA

- Built an end-to-end TTS model for expressive speech synthesis based on Baidu's Deep Voice 3 system, which can transfer reference utterance's prosody to synthesized utterance.
- Factorized speech style tokens to disentangle the possible prosodies.

PUBLICATIONS

Gao, Y. et al. (2020). *Interactive Text-to-Speech Synthesis via Semi-supervised Style Learning*. arXiv preprint arXiv:2002.06758 (2020).

Zhao, W., **Gao, Y.**, Memon, S. A., Raj, B., Singh, R. (2019). *Hierarchical Routing Mixture of Experts*, arXiv preprint arXiv:1903.07756.

Gao, Y. Singh, R., & Raj, B. (2018). *Voice Impersonation Using Generative Adversarial Networks*. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2018.

Zhao, W., **Gao, Y.**, & Singh, R. (2017). *Speaker identification from the sound of the human breath*. arXiv preprint arXiv:1712.00171.

Gao, Y., Guo, S., Huang, K., Chen, J., Gong, Q., Zou, Y., & Overett, G (2017). *Scale optimization for full-image-CNN vehicle detection*. 2017 Intelligent Vehicles Symposium (IV), 2017 IEEE (pp. 785-791). IEEE, 2017.

Gao, Y., Zhang, Y., Menon, P. G. (2015). *3D Shape Comparison of Cardiac Geometries Using a Laplace Spectral Shape Matching Approach*. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 2015.

Gao, Y., Liao, T., Zhang, Y., Menon, P. G. (2014). Abstract of *Development of Active Shape Models for Left Ventricular Function in Ischemic Cardiomyopathy*, CompIMAGE 2014

COURSE PROJECTS

- Does CurveBall Use Acceleration? - A Contrastive Study of Minimal Newton Solvers (10725 Convex Optimization) **Best Spotlight Award (1/48)**
- Accelerate the Implementation of Gaussian Mixture Model in GrabCut Segmentation Algorithm (18645 How to Write Fast Code) **Top Score (3%)**
- Video Segmentation from Cluttered Background using GrabCut with Intensity/Difference Cues. (16720 Computer Vision) **Top Score (5%)**

TEACHING EXPERIENCES

TA for 10471/671 Intermediate Deep Learning, Instructor: Ruslan Salakhutdinov 09/2019 - 01/2020

TA for 10701 Introduction to Machine Learning, Instructors: Pradeep Ravikumar, Ziv Bar-Joseph 09/2018 - 01/2019

TA for 16720J Computer Vision, Instructor: Gary Overett 09/2015 - 01/2016

SERVICES & ACTIVITIES

Reviewer for IEEE Intelligent Vehicle, Neurocomputing 2017, 2018, 2019
Session Chair for IEEE Intelligent Vehicle Conference (2017) 06/2017

HONORS & AWARDS

Best project award (1/50), CMU Optimization Symposium (COPTS)	2018
ECE Department Fellowship, CMU	2015-2016
Gradation with Honor, HUST	2013
Scholarship for Excellent Individual, HUST	2011, 2012

COURSES TAKEN

10725 Convex Optimization, 10701 Introduction to Machine Learning, 10807 Deep Learning, 10705 Intermediate Statistics, 18751 Applied Stochastic Processing, 18793 Image and Video Processing, 16720 Computer Vision, 18660 Numerical Methods, 18640 How to Write Fast Code, 15513 Introduction to Computer System, 18491 Digital Signal Processing