# YOU (NEIL) ZHANG

https://yzyouzhang.com

585-732-2916 you.zhang@rochester.edu LinkedIn GitHub Google Scholar

## EDUCATION

**University of Rochester (UR)** **Aug 2019 – Dec 2024 (Expected)**
*Ph.D., Electrical and Computer Engineering* *Rochester, NY*

**University of Rochester (UR)** **Aug 2019 – May 2021**
*M.S., Electrical and Computer Engineering* *Rochester, NY*

**University of California, Berkeley (UCB)** **Jan 2018 – Jan 2019**
*Undergraduate Exchange Studies, Electrical Engineering and Computer Science* *Berkeley, CA*

**University of Electronic Science and Technology of China (UESTC)** **Sep 2015 – Jun 2019**
*B.Eng., Automation* *Chengdu, Sichuan, China*

## RESEARCH INTERESTS

Speech & Audio Processing, Deepfake Detection, Spatial Audio, Audio-Visual Rendering and Analysis, Deep Learning

## ACADEMIC & INDUSTRIAL RESEARCH EXPERIENCE

**University of Rochester –** *Audio Information Research Lab* **Aug 2019 – Present**
*Research Assistant,* Advisor: Prof. Zhiyao Duan *Rochester, NY*

- **HRTF Modeling for Spatial Audio in Virtual and Augmented Reality**
  - Proposed a deep learning system to predict the personalized head-related transfer functions (HRTF) employing anthropometric measurements and scanned head geometry of subjects.
  - Proposed **neural field representations** for unifying measured HRTFs across existing databases. We also proposed a **generative model** for such representation and applied it to HRTF interpolation and generative tasks.
- **Speaker Verification Anti-Spoofing / Audio Deepfake Detection**
  - Improved the **generalization ability** to unseen spoofing attacks with proposed **one-class learning**.
  - Hypothesized and verified that channel effect is a primary reason for **cross-dataset** performance degradation. We proposed training strategies to improve the **channel robustness** for anti-spoofing.
  - Jointly optimized speaker verification and anti-spoofing with a proposed **probabilistic framework**.
  - Extended the one-class idea with speaker attractor **multi-center one-class learning** to maintain speaker diversity in real speech.
- **Emotional Talking Face Generation**
  - Implemented and evaluated the baseline method and took charge of the **subjective evaluation** section, including the Amazon Mechanical Turk (AMT) setup, survey design, and **data analysis**, and proved the proposed method exceeds the baseline.

**Microsoft** – *Applied Sciences* **May 2023 – Aug 2023**
*Research Intern,* Mentor: Dr. Kazuhito Koishida *Redmond, WA*

**Tencent America** – *Tencent AI Lab* **May 2022 – Aug 2022**
*Research Intern,* Mentor: Dr. Shi-Xiong Zhang *Bellevue, WA*

- **Multi-Channel Audio-Visual Speaker Diarization**
  - Proposed a probabilistic framework to incorporate the spatial information from multi-channel audio, speaker characteristics, and visual information to perform **speaker diarization**.

**Bytedance / Tiktok** – *Speech, Audio & Music Intelligence* **May 2021 – Aug 2021**
*Research Intern,* Mentor: Dr. Ming Tu *Mountain View, CA*

- **Audio-Visual Active Speaker Detection**
  - Implemented state-of-the-art active speaker detection methods and adapted them to real-world data on short-video platforms with a **semi-supervised learning** method, noisy student training.

**Tencent** – *Tencent Media Lab* **Jun 2019 – Aug 2019**
*Research Intern,* Mentor: Dr. Yannan Wang *Shenzhen, Guangdong, China*

- **Perceptual Loss Design for Mask-based Speech Enhancement**
  - Improved the perceptual quality of the enhanced speech using **multi-task learning** with the implementation of several perception-inspired losses using **uncertainty**.

## PUBLICATIONS

[14] Yutong Wen, **You Zhang**, and Zhiyao Duan. "Mitigating cross-database differences for learning unified HRTF representation", Under review, 2023.

[13] Yongyi Zang, **You Zhang**, and Zhiyao Duan. "Phase Perturbation Improves Channel Robustness for Speech Spoofing Countermeasures", accepted by *Interspeech*, 2023. [link][code]

[12] Siwen Ding, **You Zhang**, and Zhiyao Duan. "SAMO: Speaker Attractor Multi-Center One-Class Learning for Voice Anti-Spoofing", in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023. [link][code]

[11] **You Zhang**, Yuxiang Wang, and Zhiyao Duan. "HRTF Field: Unifying Measured HRTF Magnitude Representation with Neural Fields", in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023. [link][code] (**Recognized as one of the top 3% of all papers accepted at ICASSP 2023)**

[10] **You Zhang**, Fei Jiang, Ge Zhu, Xinhui Chen, and Zhiyao Duan. "Generalizing Voice Presentation Attack Detection to Unseen Synthetic Attacks and Channel Variation", *Handbook of Biometric Anti-spoofing (3rd ed.)*, Springer, 2023. [link][code]

[ 9] Abudukelimu Wuerkaixi, Kunda Yan, **You Zhang**, Zhiyao Duan, and Changshui Zhang. "DyViSE: Dynamic Vision-Guided Speaker Embedding for Audio-Visual Speaker Diarization", in *Proc. IEEE 24th International Workshop on Multimedia Signal Processing (MMSP)*, pp. 1-6, 2022. [link][code]

[ 8] Yuxiang Wang, **You Zhang**, Zhiyao Duan, and Mark Bocko. "Predicting Global Head-Related Transfer Functions From Scanned Head Geometry Using Deep Learning and Compact Representations", *arXiv preprint arXiv:2207.14352*, 2022. (submitted) [link][code]

[ 7] Abudukelimu Wuerkaixi, **You Zhang**, Zhiyao Duan, and Changshui Zhang. "Rethinking Audio-visual Synchronization for Active Speaker Detection", in *Proc. IEEE 32nd International Workshop on Machine Learning for Signal Processing (MLSP)*, 2022. [link][code]

[ 6] **You Zhang**, Ge Zhu, and Zhiyao Duan. "A Probabilistic Fusion Framework for Spoofing Aware Speaker Verification", in *Proc. The Speaker and Language Recognition Workshop (Odyssey)*, pp. 77-84, 2022. [link][code]

[ 5] Sefik Emre Eskimez, **You Zhang**, and Zhiyao Duan. "Speech Driven Talking Face Generation from a Single Image and an Emotion Condition", *IEEE Transactions on Multimedia*, vol. 24, pp. 3480-3490, 2021. [link][code][project webpage]

[ 4] Xinhui Chen*, **You Zhang***, Ge Zhu*, and Zhiyao Duan. "UR Channel-Robust Synthetic Speech Detection System for ASVspoof 2021", in *Proc. ASVspoof 2021 Workshop*, pp. 75-82, 2021. (* equal contribution) [link][code][video]

[ 3] **You Zhang**, Ge Zhu, Fei Jiang, and Zhiyao Duan. "An Empirical Study on Channel Effects for Synthetic Voice Spoofing Countermeasure Systems", in *Proc. Interspeech*, pp. 4309-4313, 2021. [link][code][video]

[ 2] Yuxiang Wang, **You Zhang**, Zhiyao Duan, and Mark Bocko. "Global HRTF Personalization Using Anthropometric Measures", in *Audio Engineering Society (AES) 150th Convention*, 2021. [link][code][video]

[ 1] **You Zhang**, Fei Jiang, and Zhiyao Duan. "One-class Learning Towards Synthetic Voice Spoofing Detection", *IEEE Signal Processing Letters*, vol. 28, pp. 937-941, 2021. [link][code][video][project webpage]

## AWARDS & PROFESSIONAL SERVICES

**Awarded**

- ICASSP Rising Stars in Signal Processing — Summer 2023
- Travel Grant from AS&E Graduate Student Association @ UR — Fall 2021 & Summer 2022
- Travel Grant from NSF-NRT AR/VR Training Program — Spring 2022
- Outstanding Graduate @ UESTC — Spring 2019
- Renmin Scholarship — Fall 2016 & Fall 2017 & Fall 2018

**Leadership**

- Western New York Virtual and Augmented Reality Mini-Conference 2022 & 2023 Co-chair [link] — Spring 2022 & Spring 2023
- IEEEXtreme 16.0 Ambassador [link] — Fall 2022

**Reviewer**

- **Journals**: IEEE Open Journal of Signal Processing (OJSP), IEEE Transactions on Computational Imaging (TCI), IEEE Access
- **Conferences**: Interspeech 2023, Audio Engineering Society (AES) 152nd, 153rd, 154th Convention

## TEACHING

**Teaching Assistant**

- ECE 208 / 408 The Art of Machine Learning — Spring 2022 & Spring 2023
- ECE 440 Introduction to Random Processes — Fall 2022
- ECE 272 / 472 Audio Signal Processing — Spring 2020 & Spring 2021
- ECE 477 Computer Audition — Fall 2020

**Students Mentored / Mentoring**

- Yutong Wen — CS undergraduate @ UR — Spring 2023 - Present
- Yongyi Zang — AME undergraduate @ UR — Summer 2022 - Summer 2023
- Siwen Ding — DS master @ Columbia University — Summer 2022 - Fall 2022
- Abudukelimu Wuerkaixi — PhD student @ Tsinghua University — Fall 2021 - Summer 2022
- Xinhui Chen — CS master @ UR — Spring 2021 - Summer 2021