YOU (NEIL) ZHANG

https://yzyouzhang.com

📞 585-732-2916 ✉ you.zhang@rochester.edu in LinkedIn ⌗ GitHub 🎓 Google Scholar

EDUCATION

University of Rochester (UR)   **Aug 2019 – Aug 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, Spatial Audio, Audio-Visual Analysis, Virtual and Augmented Reality, Deep Learning

ACADEMIC & INDUSTRIAL RESEARCH EXPERIENCE

University of Rochester – *Audio Information Research Lab*   **Aug 2019 – Present**
Research Assistant, Advisor: Prof. Zhiyao Duan   *Rochester, NY*
- **Personalized Head-Related Transfer Function (HRTF) Modeling for Spatial Audio**
  * Proposed learning **neural field representations** for unifying measured HRTFs across existing databases
  * Mentored a student and proposed a novel **position-dependent normalization strategy** that effectively mitigates the influence of cross-database differences to improve the learned representation further
  * Built a deep learning system to predict the **spherical harmonic coefficients** from anthropometric measurements and scanned head geometry of subjects for HRTF personalization
- **Speaker Verification Anti-Spoofing / Audio Deepfake Detection**
  * Generalization Ability: Proposed **one-class learning** for better detecting unseen spoofing attacks; Mentored a master student working on extending the one-class idea with speaker attractor multi-center one-class learning to maintain speaker diversity in real speech
  * Channel Robustness: Hypothesized and verified that channel effect is a primary reason for cross-dataset performance degradation, and proposed training strategies to improve the channel robustness for anti-spoofing; Mentored an undergraduate student working on the effect of phase perturbation to improve robustness to unseen channels
  * Joint Optimization: Proposed a probabilistic fusion framework for spoofing aware speaker verification
  * Singing Voice Deepfake Detection: Proposed novel task and identified challenges with collected SingFake dataset.
- **Audio-Visual Rendering and Analysis**
  * Emotional Talking Face Generation: Implemented and evaluated a baseline method and took charge of the subjective evaluation section, including the Amazon Mechanical Turk (AMT) setup, survey design, and data analysis
  * Audio-Visual Speaker Diarization: Mentored a student working on audio-visual synchronization and off-screen speakers problem for audio-visual speaker diarization

IngenID – *Voice Biometrics Group*   **Aug 2019 – Present**
Research Associate, CEO: Peter Soufleris   *Rochester, NY*
- **Development and Deployment of Backend Algorithms for Speech Biometrics Analysis**

Microsoft – *Applied Sciences*   **May 2023 – Aug 2023**
Research Intern, Mentor: Dr. Kazuhito Koishida   *Redmond, WA*
- **Audio-Visual Segmentation by Prompting Segment Anything Model**

Tencent America – *Tencent AI Lab*   **May 2022 – Aug 2022**
Research Intern, Mentor: Dr. Shi-Xiong Zhang   *Bellevue, WA*
- **Multi-Channel Audio-Visual Speaker Diarization with Spatial Features**

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

PUBLICATIONS   (* equal contribution, ‡ student mentored)

Book Chapters

[1] **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. [DOI][code]

Journals

[5] **You Zhang**, Yutong Wen‡, Yuxiang Wang, Mark Bocko, and Zhiyao Duan. "Grid-Agnostic Personalized Head-Related Transfer Function Modeling with Implicit Neural Representations", 2023. (work in progress)

[4] 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*, 2023. (under review) [link][code]

[3] **You Zhang***, Ge Zhu*, Julia M. Soto*, Samantha E. Lettenberger*, Maryam Zafar, Peggy Auinger, Abigail Arky, Emma Waddell, Kelsey Spear, Rajbir Toor, Grace Nkrumah, Emily A. Hartman, Jacob Epifano, Michael J. Hasselberg, Anton P. Porsteinsson, Rich Christie, Zhiyao Duan, Aaron J. Masino, and E. Ray Dorsey, "Words Spoken Daily among Individuals with Neurodegenerative Conditions: A Pilot Study", 2023. (* equal contribution) (in submission)

[2] 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. [DOI][code][project webpage]

[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. [DOI][code][video][project webpage]

Conferences and Workshops

[12] Yongyi Zang*‡, **You Zhang***, Mojtaba Heydari, and Zhiyao Duan. "SingFake: Singing Voice Deepfake Detection", 2023. (under review in ICASSP 2024) [link][code][project webpage]

[11] Enting Zhou‡, **You Zhang**, and Zhiyao Duan. "Learning Arousal-Valence Representation from Categorical Emotion Labels of Speech", 2023. (under review in ICASSP 2024) [link][code]

[10] Yutong Wen‡, **You Zhang**, and Zhiyao Duan. "Mitigating Cross-Database Differences for Learning Unified HRTF Representation", in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2023. [DOI][code][video]

[9] Yongyi Zang‡, **You Zhang**, and Zhiyao Duan. "Phase Perturbation Improves Channel Robustness for Speech Spoofing Countermeasures", in *Proc. Interspeech*, pp. 3162-3166, 2023. [DOI][code]

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

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

[6] 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. [DOI][code]

[5] 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. [DOI][code]

[4] **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. [DOI][code]

[3] 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. [DOI][code][video]

[2] **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. [DOI][code][video]

[1] 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]

Conference Abstracts

[3] **You Zhang**, Yuxiang Wang, Mark Bocko, and Zhiyao Duan. "Grid-Agnostic Personalized Head-Related Transfer Function Modeling with Neural Fields", in *Acoustical Society of America (ASA) 184th Meeting*, 2023. [DOI] (**Recognized by Signal Processing at the ASA Student Paper Award - Second Place**)

[2] Samantha E. Lettenberger, Maryam Zafar, Julia M. Soto, **You Zhang**, Ge Zhu, Aaron J. Masino, Grace Nkrumah, Emma Waddell, Kelsey Spear, Abigail Arky, Rajbir Toor, Emily Hartman, Jacob Epifano, Rich Christie, Zhiyao Duan, and Ray Dorsey. "Words Spoken Daily: A Novel Measure of Cognition", in *International Congress of Parkinson's Disease and Movement Disorders (MDS)*, 2023. [DOI]

[1] Yuxiang Wang, **You Zhang**, Zhiyao Duan, and Mark Bocko. "Employing Deep Learning Method to Predict Global Head-Related Transfer Functions from Scanned Head Geometry", in *Acoustical Society of America (ASA) 181st Meeting*, 2021. [DOI]

INVITED TALKS

[2] Generalizing Voice Presentation Attack Detection to Unseen Synthetic Attacks [link]
   ISCA Special Interest Group (SIG) - Security and Privacy in Speech Communication (SPSC) webinar – Online   Feb 2023

[1] One-class Learning Towards Synthetic Voice Spoofing Detection
   National Institute of Informatics (NII), Japan – Online   Jan 2021

TEACHING

Teaching Assistant

• ECE 277 / 477	Computer Audition	Fall 2020 & Fall 2023
• 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 216	Microprocessor & Data Conversion	Fall 2019

Students Mentored

• Yutong Wen	AME undergraduate @ UR	HRTF Personalization with Neural Fields
• Enting Zhou	CS undergraduate @ UR	Speech Emotion Representation Learning
• Yongyi Zang	AME undergraduate @ UR	Audio Deepfake Detection
• Siwen Ding	DS master @ Columbia University	Audio Deepfake Detection
• Abudukelimu Wuerkaixi	PhD student @ Tsinghua University	Audio-Visual Speaker Diarization
• Xinhui Chen	CS master @ UR	Audio Deepfake Detection

HONORS & AWARDS

• Top 3% of all papers accepted at ICASSP 2023 [link]	Summer 2023
• ICASSP Rising Stars in Signal Processing [link]	Summer 2023
• Signal Processing at the ASA Student Paper Award - Second Place	Spring 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 (Top 3%)	Fall 2016 & Fall 2017 & Fall 2018

PROFESSIONAL SERVICES

Leadership

• Executive Committee Member for AR/VR PhD training program	Fall 2023 – Spring 2024
• Western New York AR/VR Mini-Conference Co-chair [link]	Spring 2022 & Spring 2023
• Diversity, Equity, and Inclusion (DEI) Committee Member for ECE Department [link]	Fall 2022 – Spring 2023
• IEEEXtreme 16.0 Ambassador [link]	Fall 2022

Reviewer

- **Journals**:
  - ∗ IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)
  - ∗ IEEE Open Journal of Signal Processing (OJSP)
  - ∗ EURASIP Journal on Advances in Signal Processing
  - ∗ Transactions of the International Society for Music Information Retrieval (TISMIR)
  - ∗ IEEE Transactions on Computational Imaging (TCI)
  - ∗ IEEE Access
- **Conferences and Workshops**:
  - ∗ ICASSP 2024
  - ∗ Interspeech 2023
  - ∗ Audio Engineering Society (AES) 152nd, 153rd, 154th, 155th Convention
  - ∗ IJCAI 2023 Workshop on Deepfake Audio Detection and Analysis (DADA 2023)
  - ∗ CVPR 2023 6th Multimodal Learning and Applications Workshop (MULA 2023)