You (Neil) Zhang

https://yzyouzhang.com

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

Research Interests: Applied Machine Learning in Speech, Acoustics, and Audio Signal Processing

* **Spatial Audio AI**: Sound Source Localization, Detection, Separation; Head-Related Transfer Function (HRTF) Personalization

* **Security and Privacy in Speech & Audio**: Speech Anti-Spoofing, Singing Voice Deepfake Detection; Audio Watermarking

* **Multimodal Learning**: Audio-Visual Segmentation, Diarization and Separation; Emotion Understanding and Rendering

Education

University of Rochester (UR) **Aug 2019 – Apr 2026**
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**
Bachelor's Reciprocity Student, 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*

Honors & Awards

IEEE WASPAA Best Student Paper Award (*3 awarded out of 95 accepted papers*) and Travel Grant (*$462*) [link] Fall 2025

Open Scholarship Award @ OSC Rochester (*11 awardees in 2025 among all UR graduate students*) [link] Spring 2025

IEEE Signal Processing Society (SPS) Scholarship (*45 international recipients in 2024*) [link] Fall 2024

National Institute of Justice (NIJ) Graduate Research Fellowship (*24 national awardees in 2023*) [link] Fall 2023

Top 3% of all papers accepted at IEEE ICASSP 2023 (*75 awarded out of 2722 accepted papers*) [link] Summer 2023

IEEE ICASSP Rising Stars in Signal Processing (*24 international awardees in 2023*) [link] Summer 2023

Signal Processing at the ASA Student Paper Award - Second Place (*$200*) Spring 2023

Travel Grant from AS&E Graduate Student Association @ UR (*$500 each*) Fall 2021 & Summer 2022

Travel Grant from NSF-NRT AR/VR Training Program (*$1000*) Spring 2022

Outstanding Graduate @ UESTC (*top 1% in the same year of graduation*) Spring 2019

Renmin Scholarship (*top 3% in the same grade and major*) Fall 2016 & Fall 2017 & Fall 2018

Academic & Industrial Research Experience

Dolby Laboratories – *Multimodal Experiences Lab* **Jun 2025 – Present**
Sr. Researcher – Spatial Audio and Multimodal AI, Manager: Dr. Lie Lu *Atlanta, GA*

- **Spatial Audio and Multimodal AI**
  - Building spatial generative models to support sound localization, captioning, and separation
  - Research on multimodal spatial audio understanding across audio and vision

University of Rochester – *Audio Information Research Lab* **Aug 2019 – Apr 2026**
PhD Student, Committee: Prof. Zhiyao Duan (Advisor), Prof. Mujdat Cetin, Prof. Chenliang Xu *Rochester, NY*

- **Audio Deepfake Detection / Speaker Verification Anti-Spoofing**
  - Generalization Ability: Developed **one-class learning** for better detecting unseen spoofing attacks; Extended the one-class learning idea with speaker attractor multi-center one-class (SAMO) learning to maintain speaker diversity in real speech
  - Channel Robustness: Established that **channel effect** is a primary reason for cross-dataset performance degradation, and developed training strategies to improve the channel robustness for anti-spoofing
  - Joint Optimization: Developed a **probabilistic fusion framework** and expanded SAMO for spoofing aware speaker verification
  - Singing Voice Deepfake Detection (SVDD): Proposed novel SVDD task and identified challenges with the collected SingFake dataset; Organized SVDD 2024 Challenge at IEEE SLT2024 and MIREX@ISMIR2024

  - Algorithm Deployment: Impact real-world by working with IngenID to deploy the developed anti-spoofing algorithms

- **Personalized Head-Related Transfer Function (HRTF) Modeling for Spatial Audio**
  - Established learning **neural field representations** to unify measured HRTFs databases for upsampling and personalization, and proposed a **generative model** for such representation and applied it to HRTF interpolation and generative tasks.
  - Developed 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

- **Audio-Visual Generation 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: Alleviated audio-visual **synchronization** and **off-screen speakers** problem
  - Audio-Visual Deepfake Detection: Developed a multi-stream fusion framework complemented with **one-class learning** to improve the generalization ability for audio-visual deepfake detection
  - Audio-Visual Segmentation: Proposed using language space as a bridge to unify audio–visual semantics and improved segmentation performance through cross-modal foundation models

Meta – *Reality Labs Research Audio* **May 2024 – Dec 2024**
Research Intern (part-time since Sep 2024), Mentor: Dr. Ishwarya Ananthabhotla *Redmond, WA*

- **Perceptual Head-Related Transfer Function (HRTF) Representation Learning**
  - Proposed a **perception-informed HRTF representation learning** framework that integrates perceptual loss functions and analyzed the learned latent space through alignment with computational auditory models.

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

- **Audio-Visual Segmentation by Prompting Segment Anything Model**
  - Proposed a **prompting framework** to augment a vision foundation model (Segment Anything Model, SAM) with auditory understanding capabilities, enabling it to **simultaneously localize and segment** sounding sources in video frames.

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

- **Multi-Channel Audio-Visual Speaker Diarization with Spatial Features**
  - 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 with Noisy Student Training**
  - 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**.

External Grants Awarded

I contributed substantially to the **proposal development and writing** of the following externally awarded grants, including independently developing my own graduate fellowship proposal. I was typically the primary PhD student researcher on these projects.

Attributable Watermarking and Deepfake Detection for Responsible Audiobox (PI: Zhiyao Duan) **Dec 2024 – Nov 2025**
Meta Audiobox Responsible Generation Grant *$50,000*

Audio Deepfake Detection for Forensics and Security (Awarded Fellow: You Zhang) **Jan 2024 – Jul 2025**
National Institute of Justice (NIJ) Graduate Research Fellowship *$64,003*

Safeguarding Generative AI by Audio-Visual Deepfake Detection (PI: Zhiyao Duan) **Jul 2024 – Jun 2025**
National AI Research Resource (NAIRR) Pilot *10,000 GPU hours*

Toward Noise Resilient Voice Spoofing Countermeasures (PI: Zhiyao Duan) **Jan 2024 – Dec 2024**
New York State Center of Excellence in Data Science *$60,000*

Training Audio-Visual Foundation Models to Capture Fine-Grained Dependencies (PI: Zhiyao Duan) **Nov 2023 – Jun 2024**
Microsoft Accelerate Foundation Models Research (AFMR) initiative *20,000 Azure credits*

Developing and Deploying Spoofing Aware Speaker Verification Systems (PI: Zhiyao Duan) **Jan 2023 – Dec 2023**
New York State Center of Excellence in Data Science *$59,989*

Personalized Immersive Spatial Audio with Neural Field (PIs: Zhiyao Duan and Mark Bocko) **Nov 2022 – Oct 2023**
University of Rochester Goergen Institute for Data Science seed funding program *$20,000*

Media Coverage

Why AI-generated audio is so hard to detect [link] NBC News

News10NBC Investigates: Here's what happened when we did a deep fake on Berkeley Brean's voice [link] WHEC-TV Rochester

Audio deepfake detective developing new sleuthing techniques [link] UR News Center

Invited Talks

[6] From Neural Fields to Perception-Informed Learning: Scalable and Perceptually Grounded HRTF Personalization [video]
C4DM Seminar Series, Queen Mary University of London, UK – Online Feb 2026

[5] Generalizing Audio Deepfake Detection
Carnegie Mellon University (CMU) Speech Lunch, USA – Online Nov 2024

[4] Audio Deepfake Detection [video]
Generative AI Spring School & Global AI Bootcamp, Ukraine – Online Mar 2024

[3] Improving Generalization Ability for Audio Deepfake Detection
Learning And Mining from DatA (LAMDA) Lab, Nanjing University, China Dec 2023

[2] Generalizing Voice Presentation Attack Detection to Unseen Synthetic Attacks
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) Yamagishi Lab, Japan – Online Jan 2021

Tutorials

[4] From Neural Fields to Personalized Spatial Audio: A Hands-on Tutorial on HRTF Modeling (Co-present with Yoshiki Masuyama)
International Conference on Digital Audio Effects (DAFx), Cambridge, MA Sep 2026

[3] Machine Learning for Acoustics (Co-presented with Ryan McCarthy, Samuel A. Verburg, Peter Gerstoft)
Acoustical Society of America (ASA) 187th Meeting, Online Nov 2024

[2] Multimedia Deepfake Detection (Co-presented with Menglu Li, Luchuan Song,
Co-organized with Xiao-Ping Zhang, Chenliang Xu, Zhiyao Duan)
IEEE International Conference on Advanced Visual and Signal-Based Systems (AVSS), Taiwan – Online Aug 2025
IEEE International Conference on Multimedia and Expo (ICME), Niagara Falls, Canada Jul 2024

[1] Machine Learning for Personalized Head-Related Transfer Functions (HRTFs) Modeling in Gaming [slides]
AES 6th International Conference on Audio for Games, Tokyo, Japan Apr 2024

Special Sessions and Challenges Co-organized

[3] Partially Edited Audio: Perspectives from Synthesis and Defense *Special Session*
IEEE Spoken Language Technology Workshop (SLT), Palermo, Sicily, Italy Dec 2026

[3] Frontiers in Deepfake Voice Detection and Beyond *Special Session*
IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), Honolulu, HI Dec 2025

[2] Singing Voice Deepfake Detection (SVDD) *Challenge and Special Session*
IEEE Spoken Language Technology Workshop (SLT), Macao, China Dec 2024

[1] Singing Voice Deepfake Detection (SVDD) Task in MIREX *Challenge*
International Society for Music Information Retrieval (ISMIR) Conference, San Francisco, CA Nov 2024

Teaching

I performed regular TA duties, and additionally **delivered guest lectures** and **designed new homework assignments** for multiple courses.

- ECE 411 **Selected Topics in Augmented and Virtual Reality** TA, Spring 2024
  Guest lecture: Audio Deepfake Detection and Watermarking Spring 2024
  Homework: "Reverberation and Spatial Audio"
  "Source Separation and Voice Cloning"
- ECE 277 / 477 **Computer Audition** TA, Fall 2023 & Fall 2020
  Guest lecture: Room Acoustics and Spatial Audio Fall 2024
  Guest lecture: Python Programming for Audio Fall 2023
  Speech Technology
  Speech Anti-Spoofing
  Homework: "Speaker Diarization"
  Guest lecture: Generative Adversarial Networks (GAN) Spring 2022
  Guest lecture: Introduction to Speech Technology Fall 2020
- ECE 208 / 408 **The Art of Machine Learning** TA, Spring 2023 & Spring 2022
  Guest lecture: Support Vector Machines (SVM) Spring 2023
  Neural Network Training
  Homework: 4 assignments reinforcing machine learning concepts
  Guest lecture: Generative Adversarial Networks (GAN) Spring 2022
- ECE 440 **Introduction to Random Processes** TA, Fall 2022
- ECE 272 / 472 **Audio Signal Processing** TA, Spring 2021 & Spring 2020
  Homework: "Audio Machine Learning" Spring 2020
- ECE 216 **Microprocessor & Data Conversion** TA, Fall 2019

Mentoring

Undergraduate Students Mentored

- Kyungbok Lee — CS undergraduate @ UR — Audio-Visual Segmentation and Deepfake Detection
- Yanyu Zhou — EE undergraduate @ UESTC — Audio-Visual Speaker Diarization
- 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

Graduate Students Mentored

- Ye In (Brynn) Lee — DS master @ UR — Audio Deepfake Detection
- Siwen (Sivan) 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

Research Interns Co-Mentored @ Dolby

- S Sakshi — PhD student @ UMD — Spatial Audio Understanding with LLM
- Mengyu Yang — PhD student @ Georgia Tech — Spatial Audio-Visual Understanding

PUBLICATIONS (* Equal contribution (EC), ‡ Student mentored) [Google Scholar Profile]

I publish primarily within IEEE Signal Processing Society (SPS), International Speech Communication Association (ISCA), and relevant research communities, and I am strongly committed to **open science**—over 90% of my work is released with open-source code and data.

Under Review / Preprint

[U4] Anonymous Authors. "Learning Spatially Aware Audio-Visual Features from Videos with Mono Audio", 2026.

[U3] S Sakshi, Vaibhavi Lokegaonkar, **You Zhang**, Ramani Duraiswami, Sreyan Ghosh, Dinesh Manocha, and Lie Lu. "SPUR: A Plug-and-Play Framework for Integrating Spatial Audio Understanding and Reasoning into Large Audio-Language Models", *arXiv preprint arXiv:2511.06606*, 2026. [arXiv]

[U2] Lin Zhang, Johan Rohdin, Xin Wang, Junyi Peng, Tianchi Liu, **You Zhang**, Hieu-Thi Luong, Shuai Wang, Chengdong Liang, Anna Silnova, and Nicholas Evans. "WeDefense: A Toolkit to Defend Against Fake Audio", *arXiv preprint 2601.15240*, 2026. [arXiv] [code]

[U1] 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 2207.14352*, 2022. [arXiv] [code]

Book Chapters

[B1] **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

[J4] Ryan A McCarthy, **You Zhang**, Samuel A Verburg, William F Jenkins, and Peter Gerstoft. "Machine Learning in Acoustics: A Review and Open-Source Repository", *npj Acoustics*, vol. 1, pp. 18, 2025. [DOI] [arXiv] [code]

[J3] Xin Wang, Héctor Delgado, Hemlata Tak, Jee-weon Jung, Hye-jin Shim, Massimiliano Todisco, Ivan Kukanov, Xuechen Liu, Md Sahidullah, Tomi Kinnunen, Nicholas Evans, Kong Aik Lee, Junichi Yamagishi, Myeonghun Jeong, Ge Zhu, Yongyi Zang, **You Zhang**, Soumi Maiti, Florian Lux, Nicolas Müller, Wangyou Zhang, Chengzhe Sun, Shuwei Hou, Siwei Lyu, Sébastien Le Maguer, Cheng Gong, Hanjie Guo, Liping Chen, and Vishwanath Singh. "ASVspoof 5: Design, collection and validation of resources for spoofing, deepfake, and adversarial attack detection using crowdsourced speech", *Computer Speech & Language*, vol. 95, pp. 101825, 2025. [DOI] [dataset]

[J2] 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]

[J1] **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]

Peer-Reviewed Conferences and Workshops

[P21] Xuanjun Chen, Chia-Yu Hu, I-Ming Lin, Yi-Cheng Lin, I-Hsiang Chiu, **You Zhang**, Sung-Feng Huang, Yi-Hsuan Yang, Haibin Wu, Hung-yi Lee, and Jyh-Shing Roger Jang. "How Does Instrumental Music Help SingFake Detection?", *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2026. [DOI] [arXiv]

[P20] Kun Zhou, **You Zhang**, Dianwen Ng, Shengkui Zhao, Hao Wang, and Bin Ma. "Emotional Dimension Control in Language Model-Based Text-to-Speech: Spanning a Broad Spectrum of Human Emotions", *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2026. [DOI] [arXiv] [demo]

[P19] **You Zhang**, Andrew Francl, Ruohan Gao, Paul Calamia, Zhiyao Duan, and Ishwarya Ananthabhotla. "Towards Perception-Informed Latent HRTF Representations", *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2025. [DOI] [arXiv] [video] [slides] [poster] (**I won the IEEE WASPAA Best Student Paper Award and a travel grant.**)

[P18] **You Zhang***, Baotong Tian* (EC), Lin Zhang, and Zhiyao Duan. "PartialEdit: Identifying Partial Deepfakes in the Era of Neural Speech Editing", *Proc. ISCA Interspeech*, 2025. [DOI] [arXiv] [dataset] [demo page]

[P17] Kyungbok Lee‡, **You Zhang**, and Zhiyao Duan. "Audio Visual Segmentation Through Text Embeddings", *Proc. IEEE International Conference on Image Processing (ICIP)*, 2025. [DOI] [arXiv] [code]

[P16] Jiatong Shi, Hyejin Shim, Jinchuan Tian, Siddhant Arora, Haibin Wu, Darius Petermann, Jia Qi Yip, **You Zhang**, Yuxun Tang, Wangyou Zhang, Dareen Alharthi, Yichen Huang, Koichi Saito, Jionghao Han, Yiwen Zhao, Chris Donahue, and Shinji Watanabe. "VERSA: A Versatile Evaluation Toolkit for Speech, Audio, and Music", *Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) – System Demonstration Track*, 2025. [DOI] [arXiv] [code]

[P15] **You Zhang**, Yongyi Zang, Jiatong Shi, Ryuichi Yamamoto, Tomoki Toda, and Zhiyao Duan. "SVDD 2024: The Inaugural Singing Voice Deepfake Detection Challenge", *Proc. IEEE Spoken Language Technology Workshop (SLT)*, 2024. [DOI] [arXiv] [code]

[P14] Kyungbok Lee‡, **You Zhang**, and Zhiyao Duan, "A Multi-Stream Fusion Approach with One-Class Learning for Audio-Visual Deepfake Detection", *Proc. IEEE 26th International Workshop on Multimedia Signal Processing (MMSP)*, 2024. [DOI] [code]

[P13] Yongyi Zang‡, Jiatong Shi, **You Zhang**, Ryuichi Yamamoto, Jionghao Han, Yuxun Tang, Shengyuan Xu, Wenxiao Zhao, Jing Guo, Tomoki Toda, and Zhiyao Duan. "CtrSVDD: A Benchmark Dataset and Baseline Analysis for Controlled Singing Voice Deepfake Detection", *Proc. ISCA Interspeech*, 2024. [DOI] [code] [dataset]

[P12] Yongyi Zang*‡, **You Zhang* (EC)**, Mojtaba Heydari, and Zhiyao Duan. "SingFake: Singing Voice Deepfake Detection", in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024. [DOI] [code] [project webpage]

[P11] Enting Zhou‡, **You Zhang**, and Zhiyao Duan. "Learning Arousal-Valence Representation from Categorical Emotion Labels of Speech", in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024. [DOI] [code]

[P10] 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]

[P9] Yongyi Zang‡, **You Zhang**, and Zhiyao Duan. "Phase Perturbation Improves Channel Robustness for Speech Spoofing Countermeasures", in *Proc. ISCA Interspeech*, pp. 3162-3166, 2023. [DOI] [code]

[P8] 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]

[P7] **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 IEEE ICASSP 2023**)

[P6] 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]

[P5] 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]

[P4] **You Zhang**, Ge Zhu, and Zhiyao Duan. "A Probabilistic Fusion Framework for Spoofing Aware Speaker Verification", in *Proc. ISCA The Speaker and Language Recognition Workshop (Odyssey)*, pp. 77-84, 2022. [DOI] [code]

[P3] Xinhui Chen*‡, **You Zhang* (EC)**, Ge Zhu*, and Zhiyao Duan. "UR Channel-Robust Synthetic Speech Detection System for ASVspoof 2021", in *Proc. ISCA ASVspoof 2021 Workshop*, pp. 75-82, 2021. [DOI] [code] [video]

[P2] **You Zhang**, Ge Zhu, Fei Jiang, and Zhiyao Duan. "An Empirical Study on Channel Effects for Synthetic Voice Spoofing Countermeasure Systems", in *Proc. ISCA Interspeech*, pp. 4309-4313, 2021. [DOI] [code] [dataset] [video]

[P1] Yuxiang Wang, **You Zhang**, Zhiyao Duan, and Mark Bocko. "Global HRTF Personalization Using Anthropometric Measures", in *Proc. Audio Engineering Society (AES) 150th Convention*, 2021. [DOI] [code] [video]

Technical Reports

[T3] **You Zhang**, Yongyi Zang, Jiatong Shi, Ryuichi Yamamoto, Jionghao Han, Yuxun Tang, Tomoki Toda, and Zhiyao Duan. "SVDD Challenge 2024: A Singing Voice Deepfake Detection Challenge Evaluation Plan", 2024. [link] [challenge webpage]

[T2] **You Zhang***, Ge Zhu*, Julia M. Soto*, Samantha E. Lettenberger*(EC), 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. [link]

[T1] **You Zhang**, Ge Zhu, and Zhiyao Duan. "UR Spoofing Aware Speaker Verification System for the SASV Challenge", 2022. [link]

Conference Abstracts

[C4] **You Zhang**, Andrew Francl, Ruohan Gao, Paul Calamia, Zhiyao Duan, and Ishwarya Ananthabhotla. "Perception-Informed Alignment of Learning Personalized Head-Related Transfer Function Latent Representations", in *Acoustical Society of America (ASA) 189th Meeting—joint with the Acoustical Society of Japan*, 2025. [DOI] (**Invited Paper**)

[C3] **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**)

[C2] 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]

[C1] 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]

Professional Services

Leadership

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

Reviewer

- **Journals**:
  - IEEE Transactions on Audio, Speech, and Language Processing (TASLP) 2023-2026
  - IEEE Transactions on Multimedia (TMM) 2025-2026
  - IEEE Transactions on Information Forensics and Security (TIFS) 2026
  - IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 2024
  - IEEE Journal of Selected Topics in Signal Processing (JSTSP) 2024
  - IEEE Signal Processing Letters 2024-2026
  - IEEE Open Journal of Signal Processing (OJSP) 2022-2026
  - Transactions of the International Society for Music Information Retrieval (TISMIR) 2022-2026
  - ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 2024-2026
  - EURASIP Journal on Information Security 2025-2026
  - EURASIP Journal on Audio, Speech, and Music Processing 2024
  - EURASIP Journal on Advances in Signal Processing 2023-2025
  - Computer Speech & Language 2024
  - ACM Computing Surveys 2024
  - Neural Networks 2023-2025
  - IEEE Access 2023-2026
  - IEEE Transactions on Computational Imaging (TCI) 2021
- **Peer-Reviewed Conferences and Workshops**:
  - IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2024-2026
  - ISCA Interspeech 2023-2026
  - IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) 2025
  - IEEE Spoken Language Technology Workshop (Meta-reviewer for SVDD special session) 2024
  - IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU) 2025
  - ACM Multimedia 2025
  - Audio Engineering Society (AES) 152nd - 158th Convention 2022-2025
  - International Joint Conference on Neural Networks (IJCNN) 2025
  - ISCA Automatic Speaker Verification and Spoofing Countermeasures (ASVspoof) Workshop 2024
  - ISCA Symposium on Security and Privacy in Speech Communication 2024-2025
  - CVPR Multimodal Learning and Applications Workshop (MULA) 2023-2025
  - IJCAI Workshop on Deepfake Audio Detection and Analysis (DADA) 2023

Membership

- IEEE Signal Processing Society (SPS) Student Member 2021-2026
- IEEE Graduate Student Member 2021-2026
- Acoustical Society of America (ASA) Student Member 2023-2025
- Association for Computing Machinery (ACM) Student Member 2022-2025
- Audio Engineering Society (AES) Student Member 2019-2025

Skills

Programming: Python (PyTorch, Numpy, Pandas), MATLAB, Java, R, VHDL, C, LaTeX, Markdown

Platforms: Linux, Git, Jupyter Notebook, Slurm, Visual Studio Code, PyCharm, IntelliJ, Xilinx Vivado, Multisim

Languages: English (Fluent), Mandarin Chinese (Native)