Zhong-Qiu Wang

Contact Info
E-mail: wang.zhongqiu41@gmail.com / wangzq3@sustech.edu.cn
Homepage: `https://zqwang7.github.io/`
Lab webpage: `https://sustechailab.github.io/`
Google Scholar: `https://scholar.google.com/citations?user=fGUzTN8AAAAJ`

Research Interests

I am interested in broad speech/audio signal processing and artificial intelligence problems, aiming at building machine listening systems that can robustly perceive and understand speech/audio in reverberant environments with multiple concurrent sound sources. My current research focuses on deep learning based approaches for speech enhancement, speaker separation, speech dereverberation and robust automatic speech recognition based on a single microphone or an array of microphones, targeting at solving the *cocktail party problem*.

Work Experiences

Southern University of Science and Technology (SUSTech, 南方科技大学) Shenzhen, Guangdong, China
• *Tenure-Track Associate Professor*, Department of Computer Science and Engineering 2024.07 - present

Carnegie Mellon University Pittsburgh, Pennsylvania , USA
• *Post-doctoral Research Associate*, Language Technologies Institute 2021.09 - 2024.07

Mitsubishi Electric Research Laboratories Cambridge, Massachusetts, USA
• *Visiting Research Scientist*, Speech & Audio Group 2020.06 - 2021.08

Google AI Perception Cambridge, Massachusetts, USA
• *Research Intern*, Sound Understanding Team 2019.05 - 2019.08

Mitsubishi Electric Research Laboratories Cambridge, Massachusetts, USA
• *Research Intern*, Speech & Audio Group 2017.05 - 2017.08

Microsoft Research Redmond, Washington, USA
• *Research Intern*, Audio and Acoustics Research Group 2016.05 - 2016.08

Education

The Ohio State University Columbus, Ohio, USA
• *Ph.D. & M.Sc.*, Department of Computer Science and Engineering 2013.08 - 2020.05

Harbin Institute of Technology Harbin, Heilongjiang, China
• *B.Eng.*, School of Computer Science and Technology 2009.09 - 2013.07

Journal Publication

J23. S. Cornell, C. Boeddeker, T. Park, H. Huang, D. Raj, M. Wiesner, Y. Masuyama, X. Chang, **Z.-Q. Wang**, S. Squartini, P. Garcia, and S. Watanabe, "Recent trends in distant conversational speech recognition: A review of CHiME-7 and 8 DASR challenges", in *Computer Speech & Language (**CSL**)*, vol. 97, issue 101901, pp. 1-36, 2026.

J22. Y. Masuyama, X. Chang, W. Zhang, S. Cornell, **Z.-Q. Wang**, N. Ono, Y. Qian, and S. Watanabe, "An end-to-end integration of speech separation and recognition with self-supervised learning representation", in ***CSL***, vol. 95, issue 101813, pp. 1-18, 2026.

J21. P. Shen, K. Chen, S. He, P. Chen, S. Yuan, H. Kong, X. Zhang, and **Z.-Q. Wang**, "Listen to extract: Onset-prompted target speaker extraction", in *IEEE Transactions on Audio, Speech and Language Processing (**TASLPRO**)*, vol. 32, pp. 4832-4843, 2025.

J20. **Z.-Q. Wang**, "ctPulSE: Close-talk, and pseudo-label based far-field, speech enhancement", in *Journal of The Acoustical Society of America (**JASA**)*, vol. 158, issue 4, pp. 2849-2862, 2025.

J19. **Z.-Q. Wang**, "SuperM2M: Supervised and mixture-to-mixture co-learning for speech enhancement and noise-robust ASR", in *Neural Networks (**NN**)*, vol. 188, issue 107408, pp. 1-16, 2025.

J18. **Z.-Q. Wang**, "USDnet: Unsupervised speech dereverberation via neural forward filtering", in *IEEE/ACM Transactions on Audio, Speech, and Language Processing (**IEEE/ACM TASLP**)*, vol. 32, pp. 3882-3895, 2024.

J17. **Z.-Q. Wang**, "Mixture to mixture: Leveraging close-talk mixtures as weak-supervision for speech separation", in *IEEE Signal Processing Letters (**IEEE SPL**)*, vol. 31, pp. 1715-1719, 2024.

J16. Y.-J. Lu, X. Chang, C. Li, W. Zhang, S. Cornell, Z. Ni, Y. Masuyama, B. Yan, R. Scheibler, **Z.-Q. Wang**, Y. Tsao, Y. Qian, and S. Watanabe, "Software design and user interface of ESPnet-SE++: Speech enhancement for robust speech processing", in *Journal of Open Source Software (**JOSS**)*, vol. 8, iss. 91, 5403, 2023.

J15. **Z.-Q. Wang**, S. Cornell, S. Choi, Y. Lee, B.-Y. Kim, and S. Watanabe, "TF-GridNet: Integrating full- and sub-band modeling for speech separation", in ***IEEE/ACM TASLP***, vol. 31, pp. 3221-3236, 2023.

J14. D. Petermann, G. Wichern, A. Subramanian, **Z.-Q. Wang**, and J. Le Roux, "Tackling the cocktail fork problem for separation and transcription of real-world soundtracks", in ***IEEE/ACM TASLP***, vol. 31, pp. 2592-2605, 2023.

J13. **Z.-Q. Wang**, G. Wichern, S. Watanabe, and J. Le Roux, "STFT-domain neural speech enhancement with very low algorithmic latency", in ***IEEE/ACM TASLP***, vol. 31, pp. 397-410, 2022.

J12. **Z.-Q. Wang** and S. Watanabe, "Improving frame-online neural speech enhancement with overlapped-frame prediction", in ***IEEE SPL***, vol. 29, pp. 1422-1426, 2022.

J11. K. Tan, **Z.-Q. Wang**, and D.L. Wang, "Neural spectrospatial filtering", in ***IEEE/ACM TASLP***, vol. 30, pp. 605-621, 2022.

J10. **Z.-Q. Wang**, G. Wichern, and J. Le Roux, "Convolutive prediction for monaural speech dereverberation and

noisy-reverberant speaker separation", in ***IEEE/ACM TASLP***, vol. 29, pp. 3476-3490, 2021.
J9. **Z.-Q. Wang**, G. Wichern, and J. Le Roux, "On the compensation of magnitude and phase in speech separation", in ***IEEE SPL***, vol. 28, pp. 2018-2022, 2021.
J8. **Z.-Q. Wang**, P. Wang, and D.L. Wang, "Multi-microphone complex spectral mapping for utterance-wise and continuous speech separation", in ***IEEE/ACM TASLP***, vol. 29, pp. 2001-2014, 2021.
J7. **Z.-Q. Wang**[*], P. Wang[*], and D.L. Wang, "Complex spectral mapping for single- and multi-channel speech enhancement and robust ASR", in ***IEEE/ACM TASLP***, vol. 28, pp. 1778-1787, 2020. ([*]denotes equal contribution)
J6. H. Taherian, **Z.-Q. Wang**, J. Chang, and D.L. Wang, "Robust speaker recognition based on single-channel and multi-channel speech enhancement", in ***IEEE/ACM TASLP***, vol. 28, pp. 1293-1302, 2020.
J5. **Z.-Q. Wang** and D.L. Wang, "Deep learning based target cancellation for speech dereverberation", in ***IEEE/ACM TASLP***, vol. 28, pp. 941-950, 2020.
J4. **Z.-Q. Wang** and D.L. Wang, "Combining spectral and spatial features for deep learning based blind speaker separation", in ***IEEE/ACM TASLP***, vol. 27, pp. 457-468, 2019.
J3. **Z.-Q. Wang**, X. Zhang, and D.L. Wang, "Robust speaker localization guided by deep learning based time-frequency masking", in ***IEEE/ACM TASLP***, vol. 27, pp. 178-188, 2019.
J2. Y. Zhao, **Z.-Q. Wang**, and D.L. Wang, "Two-stage deep learning for noisy-reverberant speech enhancement", in ***IEEE/ACM TASLP***, vol. 27, pp. 53-62, 2019.
J1. **Z.-Q. Wang** and D.L. Wang, "A joint training framework for robust automatic speech recognition", in ***IEEE/ACM TASLP***, vol. 24, pp. 796-806, 2016.

Conference Publication in ML/AI

A3. Z. Xu, X. Fan, **Z.-Q. Wang**, X. Jiang, and R. Roy Choudhury, "Unsupervised blind speech separation with a diffusion prior", in *International Conference on Machine Learning (**ICML**)*, 2025.
A2. **Z.-Q. Wang**, A. Kumar, and S. Watanabe, "Cross-talk reduction", in *International Joint Conference on Artificial Intelligence (**IJCAI**)*, pp. 1715-1719, 2024.
A1. **Z.-Q. Wang** and S. Watanabe, "UNSSOR: Unsupervised neural speech separation by leveraging over-determined training mixtures", in *Advances in Neural Information Processing Systems (**NeurIPS**)*, pp. 34021-34042, 2023.

Conference Publication in Speech/Audio

C57. **Z.-Q. Wang** and R. Pang, "Mixture to beamformed mixture: Leveraging beamformed mixture as weak-supervision for speech enhancement and noise-robust ASR", in *IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP**)*, 2026.
C56. S. He and **Z.-Q. Wang**, "VM-UNSSOR: Unsupervised neural speech separation enhanced by higher-SNR virtual microphone arrays", in ***ICASSP***, 2026.
C55. J. Sun, S. He, R. Pang, and **Z.-Q. Wang**, "Neural forward filtering for speaker-image separation", in ***ICASSP***, 2026.
C54. P. Shen, S. He, X. Zhang, and **Z.-Q. Wang**, "LExTra: Folded prompt and split-role attention for target speaker extraction", in ***ICASSP***, 2026.
C53. T. Ling, S. He, P. Shen, and **Z.-Q. Wang**, "MC-LExt: Multi-channel target speaker extraction with onset-prompted speaker conditioning mechanism", in ***ICASSP***, 2026.
C52. P. Lu, P. Zhou, X. Chen, J. Wang, and **Z.-Q. Wang**, "UJCodec: An end-to-end UNet-style Codec for joint speech compression and enhancement", in ***ICASSP***, 2026.
C51. Y. Zhu, J. Jin, X. Luo, W. Yang, **Z.-Q. Wang**, G. Huang, J. Chen, and J. Benesty, "Forward convolutive prediction for frame online monaural speech dereverberation based on Kronecker product decomposition", in ***ICASSP***, 2026.
C50. R. Sachdev, **Z.-Q. Wang**, and C.-H. H. Yang, "Evolutionary prompt design for LLM-based post-ASR error correction", in *IEEE Workshop on Signal Processing Systems (**SiPS**)*, pp. 131-135, 2025.
C49. Y. Wu, Z. Xu, J. Chen, **Z.-Q. Wang**, and R. Roy Choudhury, "Unsupervised multi-channel speech dereverberation via diffusion", in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (**WASPAA**)*, 2025.
C48. P. Shen, X. Zhang, and **Z.-Q. Wang**, "ARiSE: Auto-regressive multi-channel speech enhancement", in *Annual Conference of the International Speech Communication Association (**Interspeech**)*, pp. 1183-1187, 2025.
C47. F. Zhao, X. Zhang, and **Z.-Q. Wang**, "Multi-channel acoustic echo cancellation based on direction-of-arrival estimation", in ***Interspeech***, pp. 629-633, 2025.
C46. L. Fu, Y. Liu, Z. Liu, Z. Yang, **Z.-Q. Wang**, Y. Li, and H. Kong, "AuralNet: Hierarchical attention-based 3D binaural localization of overlapping speakers", in ***Interspeech***, pp. 938-942, 2025.
C45. A. Shoko, I. Nobutaka, R. Haeb-Umbach, G. Wichern, **Z.-Q. Wang**, and Y. Mitsufuji, "30+ years of source separation research: Achievements and future challenges", in ***ICASSP***, 2025.
C44. H. Chen, S. Wu, C. Wang, J. Du, C.-H. Lee, S. Siniscalchi, S. Watanabe, J. Chen, O. Scharenborg, **Z.-Q. Wang**, B.-C. Yin, and J. Pan, "Summary on the Multimodal Information-based Speech Processing (MISP) 2023 challenge", in ***ICASSP Workshops***, pp. 123-124, 2024.
C43. S. Wu, C. Wang, H. Chen, Y. Dai, C. Zhang, R. Wang, H. Lan, J. Du, C.-H. Lee, J. Chen, S. Watanabe, S. Siniscalchi, O. Scharenborg, **Z.-Q. Wang**, J. Pan, and J. Gao, "The Multimodal Information based Speech Processing (MISP) 2023 challenge: Audio-visual target speaker extraction", in ***ICASSP***, pp. 8351-8355, 2024.
C42. Y. Lee, S. Choi, B.-Y. Kim, **Z.-Q. Wang**, and S. Watanabe, "Boosting unknown-number speaker separation with transformer decoder-based attractor", in ***ICASSP***, pp. 446-450, 2024.
C41. K. Saijo, W. Zhang, **Z.-Q. Wang**, S. Watanabe, T. Kobayashi, and T. Ogawa, "A single speech enhancement

model unifying dereverberation, denoising, speaker counting, separation, and extraction", in *IEEE Automatic Speech Recognition and Understanding Workshop (**ASRU**)*, 2023.
C40. W. Zhang, K. Saijo, **Z.-Q. Wang**, S. Watanabe, and Y. Qian, "Toward universal speech enhancement for diverse input conditions", in ***ASRU***, 2023.
C39. S. Cornell, M. Wiesner, S. Watanabe, D. Raj, X. Chang, P. Garcia, Y. Masuyama, **Z.-Q. Wang**, S. Squartini, and S. Khudanpur, "The CHiME-7 DASR Challenge: Distant meeting transcription with multiple devices in diverse scenarios", in *Proceedings of CHiME Workshop*, pp. 1-6, 2023.
C38. Y. Masuyama, X. Chang, W. Zhang, S. Cornell, **Z.-Q. Wang**, N. Ono, Y. Qian, and S. Watanabe, "Exploring the integration of speech separation and recognition with self-supervised learning representation", in ***WASPAA***, 2023.
C37. **Z.-Q. Wang**, S. Cornell, S. Choi, Y. Lee, B.-Y. Kim, and S. Watanabe, "TF-GridNet: Making time-frequency domain models great again for monaural speaker separation", in ***ICASSP***, 2023.
C36. **Z.-Q. Wang**, S. Cornell, S. Choi, Y. Lee, B.-Y. Kim, and S. Watanabe, "Neural speech enhancement with very low algorithmic latency and complexity via integrated full- and sub-band modeling", in ***ICASSP***, 2023.
C35. S. Cornell, **Z.-Q. Wang**, Y. Masuyama, S. Watanabe, M. Pariente, N. Ono, and S. Squartini, "Multi-channel speaker extraction with adversarial training: The WAVlab submission to the Clarity ICASSP 2023 grand challenge", in ***ICASSP***, 2023.
C34. S. Cornell, **Z.-Q. Wang**, Y. Masuyama, S. Watanabe, M. Pariente, and N. Ono, "Multi-channel target speaker extraction with refinement: The WAVLab submission to the second Clarity Enhancement Challenge", in *Proceedings of Clarity Challenge*, 2022. **[Winner of The 2nd Clarity Enhancement Challenge]**
C33. S. Choi, Y. Lee, J. Park, H. Kim, B.-Y. Kim, **Z.-Q. Wang**, and S. Watanabe, "An empirical study of training mixture generation strategies on speech separation: Dynamic mixing and augmentation", in *Asia-Pacific Signal and Information Processing Association - Annual Summit and Conference (**APSIPA-ASC**)*, pp. 1071-1076, 2022.
C32. Y.-J. Lu, X. Chang, C. Li, W. Zhang, S. Cornell, Z. Ni, Y. Masuyama, B. Yan, R. Scheibler, **Z.-Q. Wang**, Y. Tsao, Y. Qian, and S. Watanabe, "ESPnet-SE++: Speech enhancement for robust speech recognition, translation, and understanding", in ***Interspeech***, pp. 5458-5462, 2022.
C31. **Z.-Q. Wang** and D.L. Wang, "Localization based sequential grouping for continuous speech separation", in ***ICASSP***, pp. 281-285, 2022.
C30. Y.-J. Lu, **Z.-Q. Wang**, S. Watanabe, A. Richard, C. Yu, and Y. Tsao, "Conditional diffusion probabilistic model for speech enhancement", in ***ICASSP***, pp. 7402-7402, 2022.
C29. Y.-J. Lu, S. Cornell, X. Chang, W. Zhang, C. Li, Z. Ni, **Z.-Q. Wang**, and S. Watanabe, "Towards low-distortion multi-channel speech enhancement: The ESPNet-SE submission to the L3DAS22 challenge", in ***ICASSP***, pp. 9201-9205, 2022. **[Winner of The L3DAS22 Speech Enhancement Challenge]**
C28. D. Petermann, G. Wichern, **Z.-Q. Wang**, and J. Le Roux, "The cocktail fork problem: Three-stem audio separation for real-world soundtracks", in ***ICASSP***, pp. 526-530, 2022.
C27. O. Slizovskaia, G. Wichern, **Z.-Q. Wang**, and J. Le Roux, "Locate this, not that: Class-conditioned sound event DOA estimation", in ***ICASSP***, pp. 711-715, 2022.
C26. **Z.-Q. Wang**, G. Wichern, and J. Le Roux, "Convolutive prediction for reverberant speech separation", in ***WASPAA***, pp. 56-60, 2021.
C25. G. Wichern, A. Chakrabarty, **Z.-Q. Wang**, and J. Le Roux, "Anomalous sound detection using attentive neural processes", in ***WASPAA***, pp. 186-190, 2021.
C24. **Z.-Q. Wang** and D.L. Wang, "Count and separate: Incorporating speaker counting for continuous speech separation", in ***ICASSP***, pp.11-16, 2021.
C23. **Z.-Q. Wang**, H. Erdogan, S. Wisdom, K. Wilson, D. Raj, S. Watanabe, Z. Chen, and J. R. Hershey, "Sequential multi-frame neural beamforming for speech separation and enhancement", in *IEEE Spoken Language Technology Workshop (**SLT**)*, pp. 905-911, 2021.
C22. **Z.-Q. Wang** and D.L. Wang, "Multi-microphone complex spectral mapping for speech dereverberation", in ***ICASSP***, pp. 486-490, 2020.
C21. H. Taherian, **Z.-Q. Wang**, and D.L. Wang, "Deep learning based multi-channel speaker recognition in noisy and reverberant environments", in ***Interspeech***, pp. 4070-4074, 2019.
C20. **Z.-Q. Wang**, K. Tan, and D.L. Wang, "Deep learning based phase reconstruction for speaker separation: A trigonometric perspective", in ***ICASSP***, pp. 71-75, 2019.
C19. **Z.-Q. Wang** and D.L. Wang, "Integrating spectral and spatial features for multi-channel speaker separation", in ***Interspeech***, pp. 2718-2722, 2018.
C18. **Z.-Q. Wang**, X. Zhang, and D.L. Wang, "Robust TDOA estimation based on time-frequency masking and deep neural networks", in ***Interspeech***, pp. 322-326, 2018.
C17. **Z.-Q. Wang** and D.L. Wang, "All-neural multi-channel speech enhancement", in ***Interspeech***, pp. 3234-3238, 2018.
C16. **Z.-Q. Wang**, J. Le Roux, D.L. Wang, and J. R. Hershey, "End-to-end speech separation with unfolded iterative phase reconstruction", in ***Interspeech***, pp. 2708-2712, 2018.
C15. **Z.-Q. Wang**, J. Le Roux, and J. R. Hershey, "Multi-channel deep clustering: Discriminative spectral and spatial embeddings for speaker-independent speech separation", in ***ICASSP***, pp. 1-5, 2018. **[Best Student Paper Award]**
C14. **Z.-Q. Wang**, J. Le Roux, and J. R. Hershey, "Alternative objective functions for deep clustering", in ***ICASSP***, pp. 686-690, 2018.

C13. **Z.-Q. Wang** and D.L. Wang, "On spatial features for supervised speech separation and its application to beamforming and robust ASR", in ***ICASSP***, pp. 5709-5713, 2018.
C12. **Z.-Q. Wang** and D.L. Wang, "Mask weighted STFT ratios for relative transfer function estimation and its application to robust ASR", in ***ICASSP***, pp. 5619-5623, 2018.
C11. I. J. Tashev, **Z.-Q. Wang**, and K. Godin, "Speech emotion recognition based on Gaussian mixture models and deep neural networks", in *Information Theory and Applications Workshop (**ITA**)*, pp. 1-4, 2017.
C10. Y. Zhao, **Z.-Q. Wang**, and D.L. Wang, "A two-stage algorithm for noisy and reverberant speech enhancement", in ***ICASSP***, pp. 5580-5584, 2017.
C9. X. Zhang, **Z.-Q. Wang**, and D.L. Wang, "A speech enhancement algorithm by iterating single- and multi-microphone processing and its application to robust ASR", in ***ICASSP***, pp. 276-280, 2017.
C8. **Z.-Q. Wang** and D.L. Wang, "Recurrent deep stacking networks for supervised speech separation", in ***ICASSP***, pp. 71-75, 2017.
C7. **Z.-Q. Wang** and I. J. Tashev, "Learning utterance-level representations for speech emotion and age/gender recognition using deep neural networks", in ***ICASSP***, pp. 5150-5154, 2017.
C6. **Z.-Q. Wang** and D.L. Wang, "Unsupervised speaker adaptation of batch normalized acoustic models for robust ASR", in ***ICASSP***, pp. 4890-4894, 2017.
C5. **Z.-Q. Wang**, Y. Zhao, and D.L. Wang, "Phoneme-specific speech separation", in ***ICASSP***, pp. 146-150, 2016.
C4. **Z.-Q. Wang** and D.L. Wang, "Robust speech recognition from ratio masks", in ***ICASSP***, pp. 5720-5724, 2016.
C3. D. Bagchi, M. I. Mandel, **Z. Wang**, Y. He, A. Plummer, and E. Fosler-Lussier, "Combining spectral feature mapping and multi-channel model-based source separation for noise-robust automatic speech recognition", in ***ASRU***, pp. 496-503, 2015.
C2. **Z.-Q. Wang** and D.L. Wang, "Joint training of speech separation, filterbank and acoustic model for robust automatic speech recognition", in ***Interspeech***, pp. 2839-2843, 2015.
C1. Y. Liu, **Z. Wang**, M. Guo, and P. Li, "Hidden conditional random field for lung nodule detection", in *IEEE International Conference on Image Processing (**ICIP**)*, pp. 3518-3521, 2014.

Technical Report

TR1. F. Wu and **Z.-Q. Wang**, "TS-TFGridNet: Extending TF-GridNet for label-queried target sound extraction via embedding concatenation", in ***DCASE Challenge***, 2025.

Manuscripts

M3. F. Zhao and **Z.-Q. Wang**, "Why not put a microphone near the loudspeaker? A new paradigm for acoustic echo cancellation", in *arxiv preprint arXiv:2511.03244*, 2025.
M2. K. Li, G. Chen, W. Sang, Y. Luo, Z. Chen, S. Wang, S. He, **Z.-Q. Wang**, A. Li, Z. Wu, and X. Hu, "Advances in speech separation: Techniques, challenges, and future trends", in *arxiv preprint arXiv:2508.10830*, 2025.
M1. **Z.-Q. Wang**, G. Wichern, and J. Le Roux, "Leveraging low-distortion target estimates for improved speech enhancement", in *arXiv preprint arXiv:2110.00570*, 2021.

Patents

P4. **Z.-Q. Wang**, G. Wichern, and J. Le Roux, "Method and system for audio signal enhancement with reduced latency", *US Patent Application 18/045,380*, 2023.
P3. G. Wichern, A. Chakrabarty, **Z.-Q. Wang**, and J. Le Roux, "Method and system for detecting anomalous sound", *US Patent 11,978,476 B2*, 2024.
P2. **Z.-Q. Wang**, G. Wichern, and J. Le Roux, "Method and system for dereverberation of speech signals", *US Patent 11,790,930 B2*, 2023.
P1. J. Le Roux, J. R. Hershey, **Z. Wang**, and G. P. Wichern, "Methods and systems for end-to-end speech separation with unfolded iterative phase reconstruction", *US Patent 10,529,349 B2*, 2020.

Dissertation

• **Z.-Q. Wang**, "Deep learning based array processing for speech separation, localization, and recognition", The Ohio State University, Apr. 2020.

Challenge Rankings

CR3. 3rd place [TR1], DCASE2025 Challenge Task 4 - Spatial Semantic Segmentation of Sound Scenes 2025.07
CR2. 1st place [C34], The 2nd Clarity Enhancement Challenge 2022.12
CR1. 1st place [C29], The L3DAS22 3D Speech Enhancement Challenge 2022.01

Awards

A3. World's Top 2% Scientists - Single-Year Impact 2025.09
A2. Graduate Research Award, Department of CSE at The Ohio State University 2020.04
A1. Best Student Paper Award [C15], IEEE ICASSP 2018 2018.04

Invited Talks

T13, 串讲抑制, 媒体松鼠会，华为中央媒体技术院 2025.10
T12. *Speech separation and its generalization*, School of Computer Science and Technology, Anhui University 2025.05
T11. *Speech separation and its generalization*, NERC-SLIP, University of Science and Technology of China 2025.05
T10. *Speech separation and its generalization*, Electronic Information School, Wuhan University 2025.04
T9. *Speech separation and its generalization*, Huawei Technologies Co., Ltd. 2024.11
T8. *Deep learning based speech separation*, ECE/CS, UIUC 2023.04
T7. *Can you hear me? A deep learning approach to speech separation*, LTI, CMU 2023.02
T6. *Can you hear me? A deep learning approach to speech separation*, CS, NUS 2023.02
T5. *Neural spectrospatial filtering*, IEEE CONECCT 2022.07
T4. *Deep learning based speech separation*, CSE, SUSTech 2022.03
T3. *Convolutive prediction for noisy-reverberant speech separation*, MERL 2021.08

T2. *Integrating spectral and spatial processing for deep learning based speech separation and dereverb*, MERL 2020.01
T1. *Emotion, gender, and age recognition from speech utterances using neural networks*, Microsoft Research 2016.08

Professional Services

Professional Membership:

- Committee Member, Speech Dialogue and Auditory Processing Technical Committee (CCF TCSDAP, 语音对话与听觉专业委员会), China Computer Federation, 2024.08 - now
- Committee Member, Audio and Acoustic Signal Processing Technical Committee (AASP-TC), IEEE Signal Processing Society, 2023.01 - 2025.12.

Journal Editorships:

- Action Editor, Neural Networks, 2026.01 - 2028.12

Conference Chair:

- Area Chair, "Speech Coding and Enhancement", Interspeech, 2024 and 2025
- Area Chair, "Audio and Speech Source Separation", ICASSP, 2024, 2025 and 2026
- Area Chair, WASPAA, 2025
- Meta Reviewer, WASPAA, 2023
- Challenge Organizer, "CHiME-7 Task 1: Distant automatic speech recognition with multiple devices in diverse scenarios", CHiME workshop, 2023.
- Special Session Chair, "Resource-efficient real-time neural speech separation", ICASSP, 2023.

Journal Reviewer:

- IEEE/ACM TASLP
- Neural Networks
- Speech Communication
- Journal of The Acoustical Society of America
- IEEE SPL
- IEEE Open Journal of Signal Processing
- Journal of Signal Processing Systems
- EURASIP Journal on Audio, Speech, and Music Processing
- Pattern Recognition Letters
- Digital Signal Processing
- IET Signal Processing
- Electronics Letters

Conference Reviewer:

- ICASSP
- Interspeech
- SLT
- ASRU
- WASPAA
- CHiME Workshop
- Detection and Classification of Acoustic Scenes and Events (DCASE) workshop
- International Joint Conference on Neural Networks (IJCNN)
- International Symposium on Chinese Spoken Language Processing (ISCSLP)
- International Conference on Asian Language Processing (IALP)
- NeurIPS

Teaching

- SUSTech CS307 - Principles of Database Systems, Fall 2024 and 2025