Xuanjun Chen

✉ d12942018@ntu.edu.tw | ⌂ xjchen.tech | ⊙ xjchenGit | in jun-ntu | 🎓 Google Scholar

Education

National Taiwan University (NTU)   *Taipei Taiwan*
PH.D. STUDENT IN GRADUATE INSTITUTE OF COMMUNICATION ENGINEERING, GPA 4.30/4.30   *Sep. 2023 - Present*
M.S. IN COMPUTER SCIENCE AND INFORMATION ENGINEERING, GPA 4.19/4.30   *Sep. 2020 - Jan. 2023*
- **Advisor:** Prof. Jyh-Shing Roger Jang and Prof. Hung-yi Lee

National Taiwan University of Science and Technology (NTUST)   *Taipei Taiwan*
B.S. IN COMPUTER SCIENCE AND INFORMATION ENGINEERING, GPA 4.11/4.30   *Sep. 2018 - Jun. 2020*
Shenzhen Univ. of Information Technology (f.k.a. Shenzhen Institute of Information Technology, SZIIT)   *Shenzhen China*
JUNIOR COLLEGE IN ELECTRONIC INFORMATION ENGINEERING TECHNOLOGY   *Sep. 2015 - Jul. 2018*

Selected Publications († indicates equal contribution.)

[1] Chien-yu Huang, Wei-Chih Chen, Shu-wen Yang, Andy T. Liu, Chen-An Li, Yu-Xiang Lin, Wei-Cheng Tseng, Anuj Diwan, Yi-Jen Shih, Jiatong Shi, William Chen, **Xuanjun Chen**, et al., "Dynamic-SUPERB Phase-2: A Collaboratively Expanding Benchmark for Measuring the Capabilities of Spoken Language Models with 180 Tasks". *in International Conference on Learning Representations (ICLR), Singapore, April 2025.*

[2] Chih-Kai Yang, Yu-Kuan Fu, Chen-An Li, Yi-Cheng Lin, Yu-Xiang Lin, Wei-Chih Chen, Ho Lam Chung, Chun-Yi Kuan, Wei-Ping Huang, Ke-Han Lu, Tzu-Quan Lin, Hsiu-Hsuan Wang, En-Pei Hu, Chan-Jan Hsu, Liang-Hsuan Tseng, I-Hsiang Chiu, Ulin Sanga, **Xuanjun Chen**, Po-chun Hsu, Shu-wen Yang, Hung-yi Lee, "Building a Taiwanese Mandarin Spoken Language Model: A First Attempt". *Tech Report, Nov. 2024.*

[3] Wenze Ren, Haibin Wu, Yi-Cheng Lin, **Xuanjun Chen**, Rong Chao, Kuo-Hsuan Hung, You-Jin Li, Wen-Yuan Ting, Hsin-Min Wang, Yu Tsao, "Leveraging Joint Spectral and Spatial Learning with MAMBA for Multichannel Speech Enhancement". *Submitted to ICASSP 2025, Sep. 2024.*

[4] **Xuanjun Chen**, Haibin Wu, Chung-Che Wang, Hung-yi Lee, Jyh-Shing Roger Jang "Multimodal Transformer Distillation for Audio-Visual Synchronization". *in the 2024 International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2024), Seoul, Korea, April 2024.*

[5] **Xuanjun Chen**, Haibin Wu, Jyh-Shing Roger Jang, Hung-yi Lee "Singing Voice Graph Modeling for SingFake Detection". *in Interspeech 2024, Kos Island, Greece, Sept 2024.*

[6] **Xuanjun Chen**†, Jiawei Du†, Haibin Wu, Jyh-Shing Roger Jang, Hung-yi Lee "Neural Codec-based Adversarial Sample Detection for Speaker Verification". *in Interspeech 2024, Kos Island, Greece, Sept 2024.*

[7] Haibin Wu, **Xuanjun Chen**, Yi-Cheng Lin, Kai-wei Chang, Ho-Lam Chung, Alexander H. Liu, Hung-yi Lee "Towards audio language modeling-an overview". *Overview Report, Feb. 2024.*

[8] Haibin Wu, Ho-Lam Chung, Yi-Cheng Lin, Yuan-Kuei Wu, **Xuanjun Chen**, Yu-Chi Pai, Hsiu-Hsuan Wang, Kai-Wei Chang, Alexander H. Liu, Hung-yi Lee "Codec-SUPERB: An In-Depth Analysis of Sound Codec Models". *in Findings of the Association for Computational Linguistics ACL 2024, Bangkok, Thailand, Aug 2024.*

[9] Haibin Wu, **Xuanjun Chen**, Yi-Cheng Lin, Jiawei Du, Kai-Wei Chang, Ke-Han Lu, Alexander Liu, Ho-Lam Chung, Yuan-Kuei Wu, Dongchao Yang, Songxiang Liu, Yi-Chiao Wu, Xu Tan, James Glass, Shinji Watanabe, and Hung-yi Lee "Codec-SUPERB @ SLT 2024: A lightweight benchmark for neural codec models". *in the 2024 IEEE Spoken Language Technology Workshop (IEEE SLT 2024), Macao, China, 2024.*

[10] Jiawei Du, I-Ming Lin, I-Hsiang Chiu, **Xuanjun Chen**, Haibin Wu, Wenze Ren, Yu Tsao, Hung-yi Lee, Jyh-Shing Roger Jang "DFADD: The Diffusion and Flow-Matching based Audio Deepfake Dataset". *in the 2024 IEEE Spoken Language Technology Workshop (IEEE SLT 2024), Macao, China, 2024.*

[11] Hsuan-Yu Lin, **Xuanjun Chen**, Jyh-Shing Roger Jang "Singer separation for karaoke content generation". *in The 27th Oriental COCOSDA Conference (O-COCOSDA 2024), Hsinchu, Taiwan, Oct 2024.*

[12] **Xuanjun Chen**†, Haibin Wu†, Helen Meng, Hung-yi Lee, Jyh-Shing Roger Jang "Push-Pull: Characterizing the Adversarial Robustness for Audio-Visual Active Speaker Detection". *in the 2022 IEEE Spoken Language Technology Workshop (IEEE SLT 2022), Doha, Qatar, Jan 2023.*

[13] **Xuanjun Chen**†, Yen-Lun Liao†, Chung-Che Wang, Jyh-Shing Roger Jang "Adversarial Speaker Distillation for Countermeasure Model on Automatic Speaker Verification". *in Proc. 2nd Symposium on Security and Privacy in Speech Communication (SPSC 2022), Incheon, Korea, Sept 2022.*

Awards

2024	**Google Student Travel Grant,** Google LLC	*Travel Grant*
2024	**National Science and Technology Council (NSTC) Conference Subsidies,** NSTC Taiwan	*Travel Grant*
2024	**CTCI Foundation Bursary Award for Overseas Students** Taiwan CTCI Foundation	*Bursary Award*
2020-2025	**Distinguished Academic Record Award (5 years),** Taipei Kwong Tong Community Associations	*Scholarship*
2021	**Ranked 3rd of the logical access track of the ASVspoof 2021 challenge,** Interspeech 2021	*3rd/42 Worldwide*
2018-2020	**Certificate of Achievement (Top 5%, 3 semesters),** NTUST	*Scholarship*
2017	**National Bronze & Guangdong Provincial Gold Awards,** 3rd China College Internet Entrep. Comp.	*3rd Nationwide*
2016-2017	**National Encouragement Scholarship** (Only 3%) **& 3rd Prize of Academic Award** (Top 20%)	*SZIIT Scholarship*

Research Experiences

Speech Processing and Machine Learning Lab, NTU, *Advisor: Prof. Hung-yi Lee*   *Sept. 2023 - Present*

RESEARCH ASSISTANT

- [**Research**] Researched on **Neural Audio Codec Model** and **Multimodal Large Language Model**
  - Co-built a large-scale spoken language model benchmark, ensuring standardized evaluation tasks. [ICLR '25]
  - Co-built a neural audio codec benchmark [ACL Findings '24] and co-organize a neural audio codec challenge [SLT '24].
  - Contributed to an audio language modeling report comparing methodologies systematically. [arXiv '24]
  - Contributed to a technical report on spoken language models for neural audio codec evaluation. [arXiv '24]
- [**Project**] Leading a student AI team in **TAIHUCAIS: TAIwan HUmanities Conversational AI Knowledge Discovery System**
  - Developing a multi-agent system for humanities research scenarios based on Large Language Models [Demo].

Multimedia Information Retrieval Lab, NTU, *Advisor: Prof. Jyh-Shing Roger Jang*   *Oct. 2020 - Present*

RESEARCH ASSISTANT

- [**Research**] Researched on **Multimodal Biometric Authentication**. Co-Advisor: Prof. Hung-yi Lee; Mentor: Dr. Haibin Wu
  - Proposed a effective & efficient audio-visual active speaker detection model via Transformer distillation [ICASSP 2024]
  - Proposed a SingGraph model using music and speech foundation models for SingFake detection [Interspeech 2024]
  - Proposed neural codec-based adversarial sample detection method for speaker verification [Interspeeceh 2024]
  - Built a diffusion and flow-matching based deepfake speech dataset [SLT 2024]
  - Analyzed model vulnerabilities and proposed a Push-Pull loss function against adversarial attack [SLT 2022]
  - Proposed an effective and efficient anti-spoofing model via two-stage distillation schema [ISCA SPSC 2022].
  - Ranked **3rd out of 42 submissions** in the logical access track of the ASVspoof 2021 challenge of **Interspeech 2021**.
- [**Research**] Researched on **Speech Enhancement** and **Source Separation**
  - Proposed a MAMBA-based multichannel speech enhancement model using spectral-spatial information [arXiv 2024].
  - Built a singer separation dataset for karaoke content generation [O-COCOSDA 2024].

Leadership and Services

PROFESSIONAL SERVICES

2023-Now	**Reviewer / Program Committee,** EMNLP (2025), ACL (2024), ICASSP (2023-2025), LREC-COLING/COLING (2024-2025), MLSP (2024), IALP (2024), ECCV AVGenL (2024)
2024	**Invited Talker,** *Topic:* "Singing Voice Graph Modeling for SingFake Detection", Special Session: SVDD @ IEEE SLT 2024
2024	**Technical Committee,** Codec-SUPERB Challenge at IEEE 2024 Spoken Language Technology Workshop (SLT 2024)
2024-2025	**Teaching Assistant,** EE5200: Introduction to Generative AI (2024 Spr.) and EE5184: Machine Learning (2025 Spr.), NTU

CROSS-FUNCTIONAL SERVICES

2023-Now	**Administrative Assistant,** NVIDIA-NTU Artificial Intelligence Joint Research Center, NTU
2021-2022	**Vice Director,** Mainland Alumni Association, NTU
2019	**Head Volunteer,** The Fourth Cross-Strait Youth Maker Competition, Tongji University, Shanghai & Suzhou, China
2019	**Propaganda Minister and Teaching Material Leader,** Asia-Pacific University Volunteers Association, NTUST
2016-2018	**Team Leader,** founded a photography studio and led the propaganda department of the speaking club at SZIIT

Open Source and Skills

Open Source	Official repository of MTDVocaLiST model for Audio-Visual Synchronization task (Starred 23)
Open Source	Official repository of SingGraph model for SingFake Detection task (Starred 22, Fork 4)
PLs and Toolkits	Python, LaTeX, Shell; PyTorch, Huggingface, PyTorch Lightning, Numpy, Pandas, Matplotlib.