Xuanjun Chen

Ph.D. Student · National Taiwan University

🎓 CHEN, XUANJUN | 🏠 xjchen.tech | 💻 xjchenGit | 💼 jun-ntu | ✉ d12942018@ntu.edu.tw

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:** Professor **Hung-yi Lee** and **Jyh-Shing Roger Jang**

National Taiwan University of Science and Technology (Taiwan Tech)   *Taipei Taiwan*

B.S. in Computer Science and Information Engineering, GPA: 4.11/4.30   *Sep. 2018 - Jun. 2020*

Shenzhen Institute of Information Technology (SZIIT)   *Shenzhen, China*

J.C. in Electronic Information Engineering Technology   *Sep. 2015 - Jul. 2018*

Publications and Preprints († indicates equal contribution)

Publications

[6] Codec-SUPERB: An In-Depth Analysis of Sound Codec Models   *Findings of ACL 2024*

Haibin Wu, Ho-Lam Chung, Yi-Cheng Lin, Yuan-Kuei Wu, **Xuanjun Chen**, Yu-Chi Pai, Hsiu-Hsuan Wang,   *[arXiv] [Code] [Leaderboard]*
Kai-Wei Chang, Alexander H Liu, Hung-yi Lee   *[Huggingface]*

[5] Multimodal Transformer Distillation for Audio-Visual Synchronisation   *IEEE ICASSP 2024*

Xuanjun Chen, Haibin Wu, Chung-Che Wang, Hung-yi Lee, and Jyh-Shing Roger Jang   *[IEEE] [arXiv] [Code] [Poster]*

[4] Singing Voice Graph Modeling for SingFake Detection   *Interspeech 2024*

Xuanjun Chen, Haibin Wu, Jyh-Shing Roger Jang and Hung-yi Lee   *[ISCA] [arXiv] [Code]*

[3] Neural Codec-based Adversarial Sample Detection for Speaker Verification   *Interspeech 2024*

Xuanjun Chen†, Jiawei Du†, Haibin Wu, Jyh-Shing Roger Jang and Hung-yi Lee   *[ISCA] [arXiv] [Code]*

[2] Push-Pull: Characterizing the Adversarial Robustness for Audio-Visual Active Speaker Detection   *IEEE SLT 2022*

Xuanjun Chen†, Haibin Wu†, Helen Meng, Hung-yi Lee, and Jyh-Shing Roger Jang   *[IEEE] [arXiv] [Demo] [Poster]*

[1] Adversarial Speaker Distillation for Countermeasure Model on Automatic Speaker Verification   *ISCA SPSC 2022*

Yen-Lun Liao†, **Xuanjun Chen**†, Chung-Che Wang, and Jyh-Shing Roger Jang   *[ISCA] [arXiv]*

Preprints

[2] Towards audio language modeling-an overview   *Preprints 2024*

Haibin Wu, **Xuanjun Chen**, Yi-Cheng Lin, Kai-wei Chang, Ho-Lam Chung, Alexander H Liu, Hung-yi Lee   *[arXiv] [Awesome List]*

[1] Singer separation for karaoke content generation   *Preprints 2021*

Hsuan-Yu Chen, **Xuanjun Chen**, Jyh-Shing Roger Jang   *[arXiv] [Project]*

Honors and Awards

2024	**Google Conference Scholarship (Asia–Pacific)**, Google LLC	*Travel Grant*
2020-2023	**Distinguished Academic Record Award (4 years)**, Taipei Kwong Tong Community Associations	*Scholarships*
2018-2020	**Certificate of Achievement (3 semesters)**, Department of CSIE, Taiwan Tech	*Top 5%*
2021	**Ranked 3rd of the logical access track of the ASVspoof 2021 challenge**, Interspeech 2021	*3rd/42 Worldwide*
2017	**National Bronze Award and Guangdong Provincial Gold Award**, the 3rd China College Students' "Internet Plus" Innovation and Entrepreneurship Competition	*3rd Nationwide*
2017	**3rd Prize of Academic Award**, Department of EIET, SZIIT	*Top 20%*
2016	**National Encouragement Scholarship**, Chinese Ministry of Education	*Only 3%*

Research Experience ([∼] indicates the research proficiency)

Speech Processing and Machine Learning Lab, NTU
Research Assistant, Advisor: Prof. Hung-yi Lee

Taipei Taiwan

Sep. 2023 - Present

- **[Lead research]** Researched on **Singing Voice DeepFake Detection**
  - Proposed a SingGraph model for singing voice deepfake detection, which improves EER relatively for seen singers by 13.2%, for unseen singers by 24.3%, and unseen singers using different codecs by 37.1%. [Interspeech 2024]
- **[Lead research]** Researched on **Adversarial Robustness of Automatic Speaker Verification**
  - Designed neural codec-based adversarial sample detection method based on Descript-audio-codec, surpassing seven prior SOTA detection methods. [Interspeech 2024]
- **[Collaboration]** Neural codecs overview survey [arXiv] and benchmarking [Findings of ACL 2024].
- **[Mentorship]** Led 10 students on **TAIHUCAIS: TAIwan HUmanities Conversational AI Knowledge Discovery System** project.

Multimedia Information Retrieval Lab, NTU
Research Assistant, Advisor: Prof. Jyh-Shing Roger Jang

Taipei Taiwan

Oct. 2020 - Present

- **[Lead research]** Researched on leveraging **Cross-attention Distillation for Audio-Visual Learning**.
  - Proposed multimodal Transformer distillation for the audio-visual synchronization model, reducing teacher parameters by 83.52%, achieving competitive performance and providing comprehensive analysis. [ICASSP 2024]
  - Received the Google Conference Scholarship (Asia-Pacific) of Google LLC
- **[Lead research]** Researched on the **Adversarial Robustness for Audio-Visual Learning** [IEEE SLT 2022]
  - Revealed the vulnerability of the audio-visual active speaker detection model in many ways, such as single- and multi-modal attacks, three-attack algorithms, white- and black-box attackers, and training-aware and inference-aware scenarios.
  - Proposed audio-visual interaction loss enables inter-class dispersion and intra-class compactness, which outperforms the adversarial training by 33.14 mAP (%) under multi-modal attacks.
- **[Collaboration]** Researched on **Automatic Speaker Verification Spoofing Countermeasures (ASVspoof)**
  - Proposed ASD-ResNetSE model combines generalized end-to-end pre-training and adversarial fine-tuning, achieving competitive performance with only 22.5% teacher parameters and 19.4% teacher MACs [ISCA SPSC 2022].
  - Ranked **3rd out of 42 submissions** in the logical access track of the ASVspoof 2021 challenge of **Interspeech 2021**.
- **[Mentorship]** Led 5 students on **Advanced Technologies for Designing Trustable AI Services** industry cooperation project.

Communication and Leadership

PROFESSIONAL

2023 - Now	**Reviewer**, ACL (2024), ICASSP (2023, 2024), LREC-COLING (2024), MLSP (2024), IALP (2024), ECCV AVGenL (2024)
2024	**Technical Committee**, Codec-SUPERB Challenge at IEEE 2024 Spoken Language Technology Workshop (SLT 2024)
2024 Spr.	**Teaching Assistant**, EE5200: Introduction to Generative AI, NTU

CROSS-FUNCTIONAL

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, Taiwan Tech
2017-2018	**Team Leader**, Started a photography studio with 5 members while studying at SZIIT
2016-2017	**Propaganda Minister**, Led 20 members in Speaking Club's Propaganda Department of SZIIT

Skills

- **Programming languages & Frameworks:** Python, C/C++, LaTeX, PyTorch, PyTorch Lightning, Numpy, Pandas, Matplotlib
- **Languages:** Mandarin, Cantonese, English