Xuanjun Chen

 CHEN, XUANJUN | xjchen.tech | xjchenGit | jun-ntu | d12942018@ntu.edu.tw

Research Interest

- My research interests include deep learning and speech processing, especially in audio coding [1, 2], model compression[3, 4], multimodal learning [3, 5], and robustness [4, 5]. I am honored to receive a **Google Conference Scholarship** in 2024.

Education

National Taiwan University (NTU)   *Taipei Taiwan*

PH.D. STUDENT IN GRADUATE INSTITUTE OF COMMUNICATION ENGINEERING   *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

* indicates equal contribution

[1] Towards audio language modeling-an overview   *Submitted to SPL 2024*

HAIBIN WU, **XUANJUN CHEN**, YI-CHENG LIN, KAI-WEI CHANG, HO-LAM CHUNG, ALEXANDER H LIU, HUNG-YI LEE   *[arXiv]*

[2] Codec-SUPERB: An In-Depth Analysis of Sound Codec Models   *Submitted to 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]*

[3] 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]*

[5] 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]*

[4] 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]*

Research Experience

[~] indicates the research topics

Speech Processing and Machine Learning Lab, NTU   *Taipei Taiwan*

RESEARCH ASSISTANT, ADVISOR: PROF. HUNG-YI LEE   *Sep. 2023 - Present*

- **[Audio coding]** Neural codecs overview survey and benchmarking [1, 2].
- **[Robustness]** Design neural codec-based adversarial sample detection method based on Descript-audio-codec, surpassing seven prior SOTA detection methods. (**Submitted to Interspeech 2024**)
- **[Robustness]** Propose 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%. (**Submitted to Interspeech 2024**)

Multimedia Information Retrieval Lab, NTU   *Taipei Taiwan*

RESEARCH ASSISTANT, ADVISOR: PROF. JYH-SHING ROGER JANG   *Oct. 2020 - Sep. 2023*

- **[Compression]** Proposed multimodal Transformer distillation for the audio-visual synchronization model, reducing teacher parameters by 83.52%, achieving competitive performance and providing comprehensive analysis. [3]
- **[Compression]** 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. [4]
- **[Robustness]** Revealed the vulnerability of the audio-visual active speaker detection model in many ways, such as single- and multimodal attacks, three-attack algorithms, white- and black-box attackers, and training-aware and inference-aware scenarios.[5]
- **[Robustness]** 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. [5]

Industry Cooperation

TAIHUCAIS: TAIwan HUmanities Conversational AI Knowledge Discovery System *Taipei Taiwan*
TEAM LEADER, ADVISOR: PROF. HUNG-YI LEE AND PROF. JYH-SHING ROGER JANG (SPONSORED BY NSTC) *Jan. 2024 - Present*
- **Goal:** Build a retrieval augmented generation system based on large language models for Taiwanese humanistic knowledge
- **Responsibilities:** Lead and mentor a team of around 10 members in conducting sub-research projects, which include prompt engineering, dataset construction, retrieval system development, and instruction fine-tuning.

Advanced Technologies for Designing Trustable AI Services *Taipei Taiwan*
TEAM LEADER, ADVISOR: PROF. HUNG-YI LEE AND PROF. JYH-SHING ROGER JANG (SPONSORED BY NSTC) *Sep. 2023 - Present*
- **Goal:** To conduct research on trustworthy artificial intelligence and develop a biometric identification system.
- **Responsibilities:** Lead and mentor a team of 5 master's students in conducting sub-research projects, which include the construction of an audio-visual deepfake dataset and the development of a model to detect deepfake singing voices.

Honors and Awards

SCHOLARSHIP

2024	**Google Conference Scholarship (Asia–Pacific)**, Google Inc.	*Travel Grant*
2020-2023	**Distinguished Academic Record Award (4 years)**, Taipei Kwong Tong Community Associations	*Acad. Scholarship*
2018-2020	**Certificate of Achievement (3 semesters)**, Department of CSIE, Taiwan Tech	*Top 5%*
2017	**3rd Prize of Academic Award**, Department of EIET, SZIIT	*Top 20%*
2016	**National Encouragement Scholarship**, Chinese Ministry of Education	*Only 3%*

COMPETITION

2021	**The logical access track of the ASVspoof 2021 challenge**, Interspeech 2021	*3rd/42 Teams Worldwide*
2017	**National Bronze Award and Guangdong Provincial Gold Award**, the 3rd China College Students' "Internet +" Innovation and Entrepreneurship Competition	*3rd Nationwide*

Professional Services

INTERNATIONAL SERVICES

2024	**Organizer**, Codec-SUPERB Challenge at IEEE 2024 Spoken Language Technology Workshop (SLT 2024)	
2024	**Reviewer**, The 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024)	
2024	**Reviewer**, The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)	
2023-2024	**Reviewer**, The 2023-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'23'24)	

TEACHING ASSISTANT

2024 Spr. **EE5200: Introduction to Generative AI**, National Taiwan University

Cross-functional Leadership and Services

2023-2024	**Administrative Assistant**, NVIDIA-NTU Artificial Intelligence Joint Research Center, National Taiwan University	
2021-2022	**Vice Director**, Mainland Alumni Association, National Taiwan University	
2019	**Head Volunteer**, The Fourth Cross-Strait Youth Maker Competition, Hosted by Tongji University, Shanghai, China	
2019	**Propaganda Minister and Teaching Material Leader**, Asia-Pacific University Volunteers Association, Taiwan Tech	
2017-2018	**Team Leader**, Started a photography studio business with 5 members while studying at SZIIT	
2016-2017	**Minister**, Led 20 members of the propaganda department of SZIIT's largest Speaking Club, akin to Toastmasters Clubs	

Skills

Programming	Python, C, C++, LaTeX
Machine Learning Related	PyTorch, PyTorch Lightning, Numpy, Pandas, Matplotlib
Languages	Mandarin, Cantonese, English