Yuankun Xie

xieyuankun@cuc.edu.cn | https://xieyuankun.github.io | github.com/xieyuankun

RESEARCH INTERESTS

Audio Deepfake Detection; Domain Generalization; Out-of-Distribution (OOD) Detection; Audio Large Language Models (ALLMs); Neural Codec Attribution

EDUCATION

- **Institute of Automation, Chinese Academy of Sciences (CASIA)** 2023.9 – Present
  Joint Ph.D. Program in Information and Communication Engineering, Advisor: Prof. Jianhua Tao Beijing, China
- **Communication University of China** 2022.9 – Present
  Ph.D. Candidate in Information and Communication Engineering, Advisor: Prof. Long Ye Beijing, China
- **Communication University of China** 2018.9 – 2021.7
  M.Eng. in Information and Communication Engineering, Advisor: Dr. Yutian Wang Beijing, China
- **North China University of Technology** 2013.9 – 2017.7
  B.Eng. in Communication Engineering Beijing, China

RESEARCH EXPERIENCE

- **Open-Domain Audio Deepfake Detection** 2022.9 – Present
  Ph.D. Research
  - **Research Scope:** Global detection, partial manipulation localization, and deepfake algorithm attribution in open-domain scenarios.
  - **Data-centric Contributions:**
    * **Codecfake**: First dataset constructed from neural audio codec generation mechanisms for ALM-based deepfake detection, exceeding 10,000 downloads.
    * **FSD**: First work to formalize *fake singing voice* detection, releasing a Chinese fake song dataset and benchmark.
    * **FSW**: First large-scale real-world Chinese social-media speech deepfake dataset with platform-level evaluation protocols.
    * **ST-Codecfake**: First open-set neural codec source tracing dataset with ID/OOD evaluation protocols.
  - **Algorithmic Contributions:**
    * **ASDG**: Domain-generalized representation learning via real-domain aggregation and fake-domain dispersion.
    * **WPT**: Wavelet Prompt Tuning for all-type audio (speech, sound, singing, music), improving generalization with parameter-efficient training.
    * **TDL**: Efficient frame-level localization for partially spoofed audio, achieving SOTA on ASVspoof19-PS.
    * **REFD**: Two-stage source tracing with score-based OOD detection, achieving SOTA on ADD2023 Track 3.

INDUSTRIAL RESEARCH INTERNSHIPS

- **Ant Group** 2025.10 – 2026.2
  Research Intern Beijing, China
  - Interpretable audio deepfake detection with Audio LLMs.
  - Studied SFT and RFT paradigms for ALLMs and proposed Frequency–Time GRPO (FT-GRPO).
  - Achieved interpretable all-type audio deepfake detection via reinforcement learning on frequency-time annotations.
- **ByteDance** 2025.6 – 2025.7
  Research Intern Beijing, China
  - Spatial audio synthesis for First-Order Ambisonics (FOA).
  - Reproduced SOTA models for FOA generation from 360-degree videos.
  - Introduced text-based spatial descriptions to enhance spatial realism and audio-video alignment; deployed in production.
- **Tencent YouTu Lab** 2024.11 – 2025.4
  Research Intern Beijing, China
  - Built FSW dataset from real social-media platforms for wild deepfake detection.
  - Proposed codec-robust detection methods adaptable to diverse platform compression conditions.
  - Developed open-set neural codec attribution for ALM-generated audio; deployed on Tencent Cloud.
  - Conducted all-type (speech, sound, singing, music) cross-type evaluation and prompt-based SSL training.
- **Tsinghua Qiyuan Lab** 2023.9 – 2024.11
  Research Intern Beijing, China
  - Designed domain-invariant representations via real aggregation and fake dispersion for robust detection.
  - Proposed efficient partial spoof localization with frozen SSL encoders.
  - Conducted deepfake source tracing for unknown generation algorithms.

PUBLICATIONS

First or co-first author of 14 papers in top-tier venues (TIFS, TASLP, ICASSP, INTERSPEECH, AAAI), with 10+ additional collaborative publications.

[J1, TIFS 2024] **Yuankun Xie**, Haonan Cheng, Yutian Wang, Long Ye. *Domain Generalization via Aggregation and Separation for Audio Deepfake Detection*.

[J2, TASLP 2025] **Yuankun Xie**, Yi Lu, Ruibo Fu, Zhengqi Wen, Zhiyong Wang, Jianhua Tao, Xin Qi, Xiaopeng Wang, Yukun Liu, Haonan Cheng, Long Ye, Yi Sun. *The Codecfake Dataset and Countermeasures for Universal Detection of Deepfake Audio*.

[C1, AAAI 2026] **Yuankun Xie**, Ruibo Fu, Zhiyong Wang, Xiaopeng Wang, Songjun Cao, Long Ma, Haonan Cheng, Long Ye. *Detect All-Type Deepfake Audio: Wavelet Prompt Tuning for Enhanced Auditory Perception*.

[C2, ICASSP 2026] **Yuankun Xie**, Ruibo Fu, Xiaopeng Wang, Zhiyong Wang, Ya Li, Zhengqi Wen, Haonan Cheng, Long Ye. *Fake Speech Wild: Detecting Deepfake Speech on Social Media Platforms*.

[C3, ICASSP 2024] **Yuankun Xie**, Haonan Cheng, Yutian Wang, Long Ye. *An Efficient Temporal Deepfake Localization Approach Based on Embeddings for Partially Spoofed Audio Detection*.

[C4, ICASSP 2024] **Yuankun Xie**, Jingjing Zhou, Xiaolin Lu, Zhenghao Jiang, Yuxin Yang, Haonan Cheng, Long Ye. *FSD: An Initial Chinese Dataset for Fake Song Detection*.

[C5, INTERSPEECH 2024] **Yuankun Xie**, Ruibo Fu, Zhengqi Wen, Zhiyong Wang, Xiaopeng Wang, Haonan Cheng, Long Ye, Jianhua Tao. *Generalized Source Tracing: Detecting Novel Audio Deepfake Algorithms with Real Emphasis and Fake Dispersion Strategy*.

[C6, INTERSPEECH 2024] Yi Lu[†], **Yuankun Xie**[†], Ruibo Fu, Zhengqi Wen, Jianhua Tao, Zhiyong Wang, Xin Qi, Xuefei Liu, Yongwei Li, Yukun Liu, Xiaopeng Wang, Shuchen Shi. *Codecfake: An Initial Dataset for Detecting LLM-based Deepfake Audio*.

[C7, INTERSPEECH 2024 (ASVspoof 5)] **Yuankun Xie**, Xiaopeng Wang, Zhiyong Wang, Ruibo Fu, Zhengqi Wen, Haonan Cheng, Long Ye. *Temporal Variability and Multi-Viewed Self-Supervised Representations to Tackle the ASVspoof 5 Deepfake Challenge*.

[C8, INTERSPEECH 2023] **Yuankun Xie**, Haonan Cheng, Yutian Wang, Long Ye. *Learning a Self-Supervised Domain-Invariant Feature Representation for Generalized Audio Deepfake Detection*.

[C9, IJCAI 2023 Workshop] **Yuankun Xie**, Haonan Cheng, Yutian Wang, Long Ye. *Single Domain Generalization for Audio Deepfake Detection*.

[C10, ICME 2023] Yutian Wang[†], **Yuankun Xie**[†], Kun Zhao, Hui Wang, Qin Zhang. *Unsupervised Quantized Prosody Representation for Controllable Speech Synthesis*.

[C11, ISCSLP 2024] **Yuankun Xie**, Chenxu Xiong, Xiaopeng Wang, Zhiyong Wang, Yi Lu, Xin Qi, Ruibo Fu, Yukun Liu, Zhengqi Wen, Jianhua Tao, et al. *Does Current Deepfake Audio Detection Models Effectively Detect ALM-based Deepfake Audio?*.

[S1, ACL 2026] **Yuankun Xie**, Xiaoxuan Guo, Jiayi Zhou, Tao Wang, Jian Liu, Ruibo Fu, Xiaopeng Wang, Haonan Cheng, Long Ye, et al. *Interpretable All-Type Audio Deepfake Detection with Audio LLMs via Frequency–Time Reinforcement Learning*.

[S2, IJCAI 2026] Xiaoxuan Guo[†], **Yuankun Xie**[†], Haonan Cheng, Jiayi Zhou, Jian Liu, Hengyan Huang, Long Ye, Qin Zhang, et al. *Towards Explicit Acoustic Evidence Perception in Audio LLMs for Speech Deepfake Detection*.

[S3, Neurocomputing] **Yuankun Xie**, Xiaopeng Wang, Zhiyong Wang, Ruibo Fu, Zhengqi Wen, Songjun Cao, Long Ma, Chenxing Li, Haonan Cheng, Long Ye. *Neural Codec Source Tracing: Toward Comprehensive Attribution in Open-Set Conditions*.

COMPETITIONS

[1] ICASSP 2026 ESDD: Track 1 (1/24), Track 2 (1/15)

[2] Alibaba Tianchi 2025 Global AI Security Challenge (Audio Deepfake): Preliminary 1/360, Final 3/360

[3] INTERSPEECH 2024 ASVspoof 5 (Progress): 2/48

[4] IJCAI 2024 FinVolution Cup: Preliminary 2/202, Final 8/30

[5] IJCAI 2023 ADD Challenge Track 1.2: 5/52

[6] IJCAI 2023 ADD Challenge Track 2: 6/17