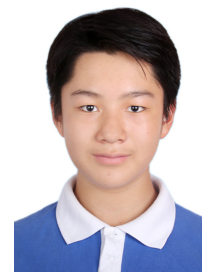


Zeyu Yin (Joey)

E-mail: yinzeyu@yahoo.com * *Telephone number:* +86 13902448055

Place of birth: Guangdong, China



Education

BSc Information and Computing Science

Undergraduate degree (Year 4)

Expected graduation: July 2026

Current GPA: 3.63/4.00

Capstone Project: Listening or Reading? An Empirical Study of Modality Importance Analysis Across AQA Question Types

Xi'an Jiaotong-Liverpool University

September 2022 – Expected July 2026

High School Diploma

High school diploma program

Award: National Second Prize, "Future Engineer" Competition (Smart Invention)

Project: Developed an Arduino-based intelligent system integrating AI technology

Hongling Middle School

September 2019 – July 2022

Project experience

DCASE 2025 (Challenge & Workshop)

Participant / Researcher

2025

Barcelona, Spain

- Participated in **DCASE 2025 Challenge Task 5: Audio Question Answering (AQA)** and the **DCASE 2025 Workshop**.
- Developed an end-to-end AQA system and conducted experiments/ablation studies to analyze modality importance across question types.

Summer Undergraduate Research Fellowship (SURF), XJTU

Undergraduate Researcher

Summer 2024

Suzhou, China

- **Expressive Timing Modelling in Performed Classical Piano Music**

Honors & Awards

Twice National Second Prize, "Future Engineer" Competition (Smart Invention)(2021,2022)

University Academic Excellence Award (Full Scholarship)(2023)

Publications

Listening or Reading? An Empirical Study of Modality Importance Analysis Across AQA Question Types

First Author

Designed the study, implemented modality-importance experiments, analyzed results across question types, and wrote the paper.

DCASE 2025 Workshop

2025

ECHOTWIN-QA: A Dual-Tower BeatsBERT System for DCASE 2025 Task 5 Audio Question Answering

First Author

Built the end-to-end AQA system from scratch, ran training/evaluation pipelines, conducted ablations, and wrote the technical report.

DCASE 2025 Challenge (Task 5)

2025

ADAPTF-SEPNET: AudioSet-Driven Adaptive Pre-training of TF-SEPNet for Multi-device Acoustic Scene Classification

DCASE 2025 Challenge

Co-author

2025

Contribution: Contributed to model development and experimental evaluation; supported results analysis and manuscript preparation.

EmoSound: A Multimodal AI Agent for Emotion-Aware Audio Accompaniment of Emotions

BICS 2025

Co-author

2025

Contribution: Implemented components for the multimodal agent and evaluation pipeline; contributed to experiments and writing.

Technical skills

Programming Languages	Python, Java, JavaScript/TypeScript, SQL, Bash
ML / Data	PyTorch, HuggingFace, scikit-learn, pandas, NumPy
Web / Backend	NestJS, Spring Boot, Node.js, REST APIs
Databases	MySQL, Postgres
DevOps / Systems	Linux, Docker, Git, Cloudflare Tunnel, SLURM (HPC)
Writing / Tools	L ^A T _E X, Markdown

Language proficiencies

Chinese	Native
English	Fluent (academic and professional proficiency)
Spanish	Basic conversational

Research Interests

I am particularly interested in building foundation models that model structured relationships between sound objects and contextual meaning, forming a unified framework for robust audio intelligence. My target focuses on developing Large Audio Models (LAMs) that hierarchically integrate acoustic and semantic representations to enable audio understanding across recognition, generation, and restoration tasks. Rather than relying on brute-force scaling, I aim to achieve strong capability through principled model design, data-efficient learning, and lightweight adaptation strategies.

With the AQA modality importance analysis, along with hands-on experience in cross-modal audio generation and semantic-driven evaluation, I have developed a strong foundation in multimodal representation learning, experimental design, and interpretable audio modeling. I aim to leverage this background to advance scalable, trustworthy, and perceptually aligned audio foundation models.