ANNA MIN

anna.min1754@gmail.com ⋄ Personal Website: annamin0.github.io

EDUCATION

Bachelor of Engineering in Software Engineering, Tsinghua University — 2024 (Expected)

Major GPA: 3.73 /4.00, sophomore: 3.78, junior: 3.86

SKILLS

Programming: Python, C, C++, MATLAB, Rust, Java, Javascript, LaTeX, Verilog, SQL, Docker

Framework & Tools: Pytorch, Fairseq, Soundspace, Detectron, MMDetection, COLMAP, Qt

Languages: Chinese (Native), English (TOEFL 106 (R30+L27+W26+S23))

PUBLICATIONS

[1] **Anna Min**, Ziyang Chen, Hang Zhao, Andrew Owens. Crafting Precision in the Wild: Learning Spatial Sound Localization from Unlabelled Egocentric Videos (*in submission*) [Paper]

[2] **Anna Min**, Chenxu Hu, Yi Ren, Xiang Yin, Hang Zhao. When End-to-End is Overkill: Rethinking Cascaded Speech-to-Text Translation (*in submission to Interspeech2024, under review*) [Paper]

[3] **Anna Min**, Chenxu Hu, Yi Ren, Hang Zhao. A Unit-based System Speech-to-Speech Translation in Expressive Video Dubbing (*selected to AAAI24 workshop on AI for Digital Human Oral Presentation, in submission to Interspeech2024, under review*) [Paper]

RESEARCH EXPERIENCE

Synchronized Video-to-Audio Synthesis for Multi-Style Videos [Media Coverage]

Researcher, Pika   Dec 2023 - Mar 2024

Supervisor: Chenlin Meng, Prof Christopher Manning, Stanford University

- Led the project in designing a latent-diffusion based audio-visual synchronization model for generated videos

Crafting Precision in the Wild: Learning Spatial Sound Localization from Unlabelled Egocentric Videos

Visiting Research Intern   July 2023 - 2024

Advisor: Prof. Andrew Owens, University of Michigan

- Explore weak labels based on the direction of source motion for accurate binaural sound localization prediction
- Introduce the first stereo dataset and benchmark for sound localization in the wild which is gathered from an extensive corpus of 8,000 hours of YouTube binaural sound videos
- Design a pipeline for Sim2Real from synthetic data and models to in-the-wild data

When End-to-End is Overkill: Rethinking Cascaded Speech-to-Text Translation

Research Intern   Mar 2023 - 2023

Advisor: Prof. Hang Zhao, Tsinghua University

- Explore the advantages of integrating multiple ASR candidates and self-supervised speech features into MT
- Propose a method that leverages two different pre-trained auto-regressive models to enhance a cascaded system, addressing the challenges associated with error divergence

Fine-grained Emotion Transfer for Speech-to-Speech Translation in Expressive Video Dubbing

Research Intern   Nov 2022 - 2023

Advisor: Prof. Hang Zhao, Tsinghua University

- Construct the first training set with aligned bilingual audio tracks with the same emotion from movies
- Use wavform to tokenized unit translation and HiFi-GAN-based networks for transferring pitches and rhythms
- Outperform the baseline by a significant 20% improvement in emotional expression

SELECTED PROJECTS

Mini Database [Code]   Feb - May 2023
- Implemented a database that supports basic SQL queries with optimization, transactions, locks, and recovery

Android Chat App [Code]   Feb - May 2023
- Utilized Kotlin/Jetpack Compose along with Material3 to craft the Android frontend
- Built the backend using Django with Channels, enabling the implementation of both HTTP APIs and real-time WebSocket communication which support user searching, following, and chatting

FTP Server & Client [Code]   Sep 2022 - Jan 2023
- Wrote an FTP server in C and FTP client UI in Python supporting most basic commands

Compiler from C++ to LLVM [Code]   Sep 2022 - Jan 2023
- Translated C++ code to LLVM intermediate representation utilizing Python-Lex-Yacc and LLVMLite compilers
- Implemented robust error handling, preprocessing capabilities, multidimensional array operations, and scope mechanisms to ensure comprehensive support s

Mars Online Judge Platform [Code] [Website] [Media Coverage]   Sep 2022 - Jan 2023
A web-based project for MarsOJ organization's students to provide online supplementary training
- Implemented real-time answer battles utilizing Socket.IO, comprehensive review functionalities, and a sophisticated tracking system empowering administrators to closely monitor student progress
- Led backend and contributed to frontend development using Vue3 for the frontend and Flask for the backend HTTP API, along with python-lsp-server for backend language services
- Achieved a user base of over 3000 users and 1000 subscribers, launching both web and mini-program versions

Machine Learning Modeling Services [Code]   Jul 2022 - Aug 2022
One-stop solution based on providing model deployment online
- Adapt to onnx, pmml, keras multiple machine learning models, out of the box, load balancing, complete testing
- Mainly responsible for the back-end, participate in the front-end, familiar with django, docker, Kubernetes, celery, vue, javascript and other technology stacks

HONORS AND COMPETITIONS

Tsinghua Academic Excellence Award (2/102)	2023
Tsinghua Science and Technology Innovation Outstanding Scholarship (2/102)	2023
Tsinghua Science and Technology Innovation Outstanding Scholarship (2/102)	2022
Tsinghua Spark Scientific Innovation Fellowship (50/3900)	2022
The first prize of 2021 National Student Mathematical Modeling Competition (ranked 89 out of 3000)	2021