

# ZIANG ZHOU

☎ (424) 439-0856 | ✉ ziangz@cs.cmu.edu | 🌐 GitHub | 🏠 Homepage | 🔗 LinkedIn | 📍 Pittsburgh, PA

## EDUCATION

### Carnegie Mellon University - School of Computer Science

*M.Sc. in Intelligent Information Systems*

Pittsburgh, PA

Aug 2022 – Dec 2023

- **Coursework:** Search Engines, Speech Recognition & Understanding, Advanced NLP, Deep Learning System, etc

### Duke University & Duke Kunshan University

*Dual B.Sc. in Data Science*

Kunshan, China & Durham, NC

Aug 2018 – May 2022

- **GPA:** 3.834/4.00, **Major GPA:** 3.94/4.00
- **Dean's List w/ Distinction:** 2018 Fall, 2020 Fall, 2020 Spring

## WORK EXPERIENCE

### Pinterest

*Machine Learning Engineer Intern*

(Remote) Pittsburgh, PA

May 2023 – Aug 2023

- Developed TabNet-based **User Match Prediction** model for ads conversions to address future 3rd-party cookie ban
- Trained in **dual-phase:** self-supervised pretraining (Phase-I) & supervised finetune with encoder weights (Phase-II)
- Created **Airflow** (Spinner) workflow to sample conversions w/wo Source of Truth (SoT) as training data for Phase-I
- Designed Unit Tests in **pytest** for multiple Phase-I modules (e.g. FeatureObfuscator), auto-triggered in CI pipeline
- Achieved notable **8%** Match Rate boost in post-ban and **4%** in pre-ban scenarios compared with online prod models

### ByteDance (TikTok) AI Lab

*Machine Learning Engineer Intern*

Shanghai, China

May 2021 – Aug 2021

- Developed Mandarin-Cantonese translation module using crafted tri-branch transformer model, applied Beam Search on autoregressive decoder, and achieved **67 BLEU** score on internal test set, outperforming SOTA by **40%+**
- Optimized torch model into ONNX graph, compressing model by **10x** factor to deploy-friendly size, less than **10MB**
- Deployed on text-to-speech (TTS) frontend in new **Capcut** app release, with **10M+** cumulative downloads in China

## RESEARCH EXPERIENCE

### Watanabe's Audio and Voice (WAV) Lab

*Research Assistant*

Pittsburgh, PA

Aug 2022 – May 2023

#### Compositional End-to-End Models for Cyclic ASR-TTS Systems

- Enabled dual-learning paradigm in Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) systems
- Proposed end-to-end differentiable system that composes ASR & TTS modules trained in semi-supervised approach

### Speech and Multimodal Intelligent Information Processing (SMIIP) Lab

*Undergraduate Research Assistant*

Kunshan, China

2019 – 2022

#### Research 1: Large-Scale Bird Sound Detection and Classification in China Region [\[pdf\]](#)

- Designed three tasks on 120, 332, and 821 species, achieved 89.8%, 83.69% and 78.70% utterance-level accuracies
- Deployed detection & classification modules on linux-driven ARM64 Raspberry pi. Registered patent for our system

#### Research 2: IEEE ICASSP Signal Processing Cup 2022 - Synthetic Speech Attribution [\[pdf\]](#)

- Proposed an end-to-end multi-system fusion solution with ResNet backbone and ranked **#6** among 25+ teams
- Performed gaussian disturbance with MATLAB and train/dev mismatching strategy to augment data for robustness

#### Research 3: NCMSC 2022 - Escalation Assessment from Dialogues [\[pdf\]](#)

March 2021

- Devised low-resource transfer learning from pre-trained emotion recognition model in speech-textual pipeline
- Bridged conversation escalation assessment with emotion recognition task by proving underlying reciprocity

## SELECTED PROJECTS

### ESPnet - ASR x SLU Recipe for Multimodal EmotionLines Dataset (MELD)

[Huggingface](#) | [GitHub](#)

- Implemented unimodal ESPnet recipe on MELD dataset with ASR x SLU multitasking, achieved **0.39 F1** on test set
- Contributed recipe for ESPnet, released `espnet/realzza-meld-asr-hubert-transformer` model on huggingface

### XenoPy: Python Library for Xeno-canto API 2.0

[GitHub](#) | [Zenodo](#)

- Boosted efficiency in metadata and recording retrieval utilizing multiprocessing, hugely improved data traffic
- Increased retrieval stability by capturing HTTPErrors with Attempt Buffer, compared to previous API wrappers

### EmoTag: Emotion Recognition on Chinese Audios

[Paper](#) | [GitHub](#)

- Proposed a deep learning recipe for emotion recognition on Emotional Speech Dataset (ESD) with **80%+** accuracy
- Tailored inference script to tag emotions in audio streams, applied as Emotion Scoring Module in **ICPR 2022** [paper](#)

## SKILLS

**Programming & Frameworks:** Python · C · Java · Spark SQL · Javascript · PyTorch · Keras · ONNX · Node.js · PyTest

**Software:** Git · Vim · Linux · Slurm · Docker · Airflow · Jenkins · Jira · Grafana · Kafka · S3 · k8s · MLflow · Presto · Hadoop