Hsi-Che Lin

hsichelin@gmail.com | Personal Page | GitHub | LinkedIn

Education

National Taiwan University *Taipei, Taiwan*

MS in Graduate Institute of Communication Engineering *Sep 2024. - present*

- Advisor: Yu-Chiang Frank Wang
- Research Topics: Multi-modal large language models, Efficient AI

National Taiwan University *Taipei, Taiwan*

BS in Electrical Engineering (GPA: 4.17/4.30, Rank: 19/202) *Sep. 2021 - Jun. 2024*

Publications

[1] EMLoC: Emulator-based Memory-efficient Fine-tuning with LoRA Correction **[NeurIPS'25]**

Hsi-Che Lin, Yu-Chu Yu, Kai-Po Chang, Yu-Chiang Frank Wang

• Developed a novel fine-tuning method reducing memory usage by a low-rank compressed alternative of the model.

• Enable fine-tuning **38B VLM** and **12B Diffusion Model** with just **24GB VRAM**, outperforming existing baselines.

[2] Improving Speech Emotion Recognition in Under-Resourced Languages via Speech-to-Speech Translation with Bootstrapping Data Selection **[ICASSP'25]**

Hsi-Che Lin*, Yi-Cheng Lin*, Huang-Cheng Chou, Hung-yi Lee

• Propose a pipeline using **synthesized data** and **bootstrapping selection** to improve under-resourced SER.

• Outperform other augmentation methods and demonstrate generalizability across diverse models and languages.

[3] QuAVF: Quality-aware Audio-Visual Fusion for Ego4D Talking to Me Challenge **[CVPR'23 Workshop]**

Hsi-Che Lin, Chien-Yi Wang, Min-Hung Chen, Szu-Wei Fu, Yu-Chiang Frank Wang

• Designed a **multi-modal framework (vision, audio, and facial landmarks)** to achieve robust detection results.

• Secured 1st place on the leaderboard, achieving state-of-the-art performance.

[4] Dynamic-SUPERB Phase-2: A Collaboratively Expanding Benchmark for Measuring the Capabilities of Spoken Language Models with 180 Tasks **[ICLR'25]**

Chien-yu Huang, Wei-Chih Chen, ..., **Hsi-Che Lin**, ..., Hung-yi Lee

• Introduce several evaluation tasks beyond classification, including regression and sequence generation tasks.

• The largest and most comprehensive open-source benchmark for spoken language models.

[5] ReXTime: A Benchmark Suite for Reasoning-Across-Time in Videos **[NeurIPS'24 D&B]**

Jr-Jen Chen, Yu-Chien Liao, **Hsi-Che Lin**, Yu-Chu Yu, Yen-Chun Chen, Yu-Chiang Frank Wang

• Develop an automated pipeline for generating a large-scale temporal reasoning question-answer data.

• The first benchmark for video understanding focusing on **cause and effect reasoning**.

[6] On the social bias of speech self-supervised models **[INTERSPEECH'24]**

Yi-Cheng Lin, Tzu-Quan Lin, **Hsi-Che Lin**, Andy T. Liu, Hung-yi Lee

• Conduct a comprehensive analysis of various factors that influence social bias and methods to debias.

Experience

Deep Learning for Computer Vision (120 students), Teaching assistant *Fall 2023-2025*

- Lead a team of 12 TAs (2025), Design homework about VLM (2024), SSL and segmentation (2023)

Machine Learning (650 tudents), Teaching assistant *Spring 2025*

- Design homework assignment about diffusion model customization

Introduction to Generative AI (1000 students), Teaching assistant *Spring 2024*

- Design homework assignment about LLM fine-tuning

Projects

E.SUN AI Open Competition 2021 Winter–Credit Card Product Category Prediction *Top 18% of 859 Teams*

- LSTM model ensemble for time-series data with feature engineering to reduce the input dimension and complexity.

Kaggle ”G-Research Crypto Forecasting” Challenge *Top 9% of 1946 Teams*

- Developed a redictive model (Light Gradient Boosting Machine) with feature selection.

Skills

Programming languages: Python

- Libraries: PyTorch, transformers, diffusers
- Models & Architectures: CLIP, InternVL, Llama, Diffusion models