HUANG KUAN PO

@ gerber861017@gmail.com   **Personal Website: https://nobel861017.github.io/**   **github: https://github.com/nobel861017**   **Google Scholar**

PHD. RESEARCH GOAL

- Robustness analysis of speech technologies and development of robust approaches against adverse noisy conditions or code-switching scenarios for many different speech processing tasks.
- Using speech representations obtained with self-supervised learning to perform downstream speech processing tasks following the recipe of the Speech Processing Universal Performance Benchmark (SUPERB).

PUBLICATIONS

- **IMPACT: Iterative Mask-based Parallel Decoding for Text-to-Audio Generation with Diffusion Modeling (ICML 2025) (project page)**
  - Sota text-to-audio generation model on AudioCaps.
- **Generative Audio Language Modeling with Continuous-valued Tokens and Masked Next-Token Prediction (ICML 2025)**
  - Decoder-only text-to-audio generation model.
- **Towards Fast Inference Speed with One-step Sampling in Continuous Space for Text-to-audio Generation (submit to ICLR 2026)**
- **Zero Resource Code-switched Speech Benchmark Using Speech Utterance Pairs For Multiple Spoken Languages (ICASSP 2024)**
  - Evaluate the code-switching ability of SSL speech models by language modeling on discrete units. (code)
  - Proposed new code-switch speech datasets (fr-en, zh-en, es-en).
- **Dataset-Distillation Generative Model for Speech Emotion Recognition (InterSpeech 2024 under review)**
  - Speech representation generation using WGAN.
  - Reduce data size, preserve privacy
- **Do Prompts Really Prompt? Exploring the Prompt Understanding Capability of Whisper**
- **Investigating Zero-Shot Generalizability on Mandarin-English Code-Switched ASR and Speech-to-text Translation of Recent Foundation Models with Self-Supervision and Weak Supervision (ICASSP 2024 workshop)**
  - Investigation of prompting Whisper for code-switch speech.
  - Prompting, Prompt-conditioned fine-tuning, speech in-context learning
- **Ensemble Knowledge Distillation of Self-Supervised Speech Models (ICASSP 2023)**
  - Distil multiple SSL teacher models into one student model.
- **Improving Distortion Robustness of Self-supervised Speech Processing Tasks with Domain Adaptation (Interspeech 2022)**
  - Continual train HuBERT with noisy data. (model)
- **Improving generalizability of distilled self-supervised speech processing models under distorted settings (SLT 2022)**
  - Distil HuBERT with noisy speech data. (robust teacher model)
- **Multi-accent Speech Separation with One Shot Learning (ACL 2021 metaNLP workshop)**
- **One shot learning for speech separation (ICASSP 2021)**

CONFERENCE PAPER REVIEWER

Interspeech 2021, 2022, ICASSP 2022, 2023, 2024, ASRU 2021, ACL2021, ACL 2021 metaNLP workshop, ISCSLP 2022

ACHIEVEMENTS

- 2019 Verizon Media (Yahoo!) The Best Internship Award
- 2021 Forward-looking Speech Technology Scholarship

OTHER EXPERIENCES

Applied Scientist Intern
Amazon AGI Speech & Audio

📅 Jul - Oct 2024, Apr - Jun 2025   📍 Bellevue, Washington

Delivered a state-of-the-art diffusion-based model that excels in both performance and inference speed for text-to-audio generation. (Paper accepted by ICML 2025)
Improved inference speed 20x while preserving performance. (Submit to ICLR 2026)

Ph.D. Student
AICS Ph.D. Program

📅 start from Sep 2022   📍 Taipei, Taiwan

Search Team Backend Developer Intern
Verizon Media (Yahoo!)

📅 Jul 2019 – Aug 2019   📍 Taipei, Taiwan

- Designed a sentence ranking system to determine which sentence best fits for being the summary among the whole news document.

TEACHING EXPERIENCES

Teaching Assistant / Head Teaching Assistant
Machine Learning Course of Prof. Hung-Yi Lee

📅 Nov 2019 - Jun 2021   📍 National Taiwan University

- Mandarin and English bilingual course "Machine Learning" offered by Prof. Hung-yi Lee.
- Developed a video translation tool and team working pipeline.

Teaching Assistant
Digital Signal Processing by Prof. Lin-shan Lee

📅 Sep 2020 - Jan 2021   📍 National Taiwan University

EDUCATION

National Taiwan University
Graduate Institute of Computer Science and Information Engineering (Ph.D. direct admission)

📅 Sep 2022 - present   📍 Taipei, Taiwan

- Advisor: Prof. Lin-shan Lee, Prof. Hung-yi Lee

National Taiwan University
Graduate Institute of Computer Science and Information Engineering

📅 Sep 2020 - Jun 2022   📍 Taipei, Taiwan

National Taiwan University
Department of Computer Science and Information Engineering Undergraduate

📅 Sep 2016 - Jun 2020   📍 Taipei, Taiwan