

Hsiu-Hsuan Wang

 [anthony-wss](#) |  b09902033@ntu.edu.com | Taipei

EDUCATION

B.S. (Computer Science) at National Taiwan University 2024
M.S. (Electrical Engineering, Computer Science group) at National Taiwan University 2026(exp.)

RESEARCH EXPERIENCE

NTU Computer Learning Lab supervised by Prof. Hsuan-Tien Lin Jun 2022 - present

- Currently doing research related to time series forecasting.
- Corwdsourcing Summer Intern: Contribute two datasets – CLCIFAR10 and CLCIFAR20, with complementary labels annotated by 1K workers from Amazon Mechanical Turk, a crowdsourcing platform.

NTU Speech Lab supervised by Prof. Hong-Yi Lee Sep 2022 - present

- Currently doing research on improving the semantics information of speech dialogue generated by Generative Spoken Dialogue Language Modeling.
- Use fairseq library to train a multilingual BERT aligning HuBERT embedding tokens and phonemes to do unsupervised automatic speech recognition tasks.

PREPRINTS

- [1] **Hsiu-Hsuan W.**, Wei-I L., Hsuan-Tien L. “CLCIFAR: CIFAR-Derived Benchmark Datasets with Human Annotated Complementary Labels”. Submitted to NeurIPS Dataset and Benchmark track. 2023. arXiv: [2305.08295 \[cs.LG\]](#).
- [2] Iu-thing K., **Hsiu-Hsuan W.**, Chun-Yi K., Ho-Lam C., Tsu-Yuan H., Kai-Wei C., Peng-Jen C., **Changhan W.**, Ann L., Hongyu G., Hung-yi L. “Voxcentum: Spoken Language Identification for 100+ Languages Expanded to 100+ Hours”. Under review. 2023.
- [3] Jiatong S., William C., Dan B., **Hsiu-Hsuan W.**, Wei Ping H., En Pei H., C., Xuankai C., Yuxun T., Shang-Wen L., Abdelrahman M., Hung-yi L., Shinji W. “Findings of the 2023 ML-SUPERB Challenge: Pre-Training and Evaluation over More Languages and Beyond”. Under review. 2023.
- [4] Xuankai C., Brian Y., Kwanghee C., Jeeweon J., Yichen L., Soumi Maiti S., Jiatong S., Jinchuan T., Shinji W., Yuya F., Takashi M., Pengcheng G., Yao-Fei C., Pavel D., Kohei S., **Hsiu-Hsuan W.** “Exploring Speech Recognition, Translation, and Understanding with Discrete Speech Units: A Comparative Study”. Under review. 2023.

PROJECTS

Speech Crawling Project with Meta(Facebook) Oct 2022 - Aug 2023

- Developed code to validate a spoken language identification model trained on VoxCentum, a 137-language dataset proposed by our team. The paper has been submitted to 2023 ASRU conference.
- Collaborate remotely with the Meta research team from the United States and senior students from the department of electrical engineering, NTU.

Automatic Essay Scoring System

Mar 2023 - present

- Engaged in cross-domain collaboration with Prof. Zhao Ming Gao from the Department of Foreign Languages and Literatures.
- Implemented off-topic essay identification techniques to enhance score prediction accuracy.

AWARDS

2023 Honorable Mention Undergraduate Project Award

2022 Excellent Teaching Assistant Award for course Datastructure and Algorithm, NTU