

# Yen-Ju Lu

+1-(626)215-5052

ylu125@jhu.edu

<https://github.com/neillu23>

<https://scholar.google.com/citations?user=emtNw84AAAAJ>

## RESEARCH INTEREST

My research focuses on advancing multimodal language models integrating speech and text for real-world applications in content understanding and generation. Specifically, I aim to improve spoken dialogue understanding by capturing paralinguistic cues with advanced encoders. Recently, I explored techniques to optimize large language models (LLMs) for more accurate and informative dialogue summarization.

## EDUCATION

- Johns Hopkins University (JHU)** *Baltimore, MD*  
*Ph.D. in Electrical and Computer Engineering* 2022 - 2026 (expected)  
– Advisors: Prof. Najim Dehak and Prof. Jesús Villalba
- National Taiwan University (NTU)** *Taipei, Taiwan*  
*M.S. in Electrical Engineering and Computer Science* 2014 - 2017  
– Advisors: Prof. Lin-Shan Lee and Prof. Hung-Yi Lee  
*B.S. in Electrical Engineering* 2010 - 2014

## RESEARCH EXPERIENCES

- Generative AI, Meta** *Menlo Park, CA*  
*Research Scientist Intern (expected)* May. 2025 – Aug. 2025  
– Mentors: Ozlem Kalinli and Jinxi Guo  
– To develop multimodal large language models for improving language and speech integration.
- Machine Learning Research, Apple** *Pittsburgh, PA*  
*Research Scientist Intern* May. 2024 – Aug. 2024  
– Mentors: Ting-Yao Hu, Hema Swetha Koppula, and Oncel Tuzel  
– Developed mutually reinforcement via data synthesis (MRDS) to improve LLM dialogue summarization, achieving 1.5% ROUGE and 0.3% BERT scores gains, outperforming LLaMA3. [10]
- Multimodal Foundation Models Collaboration, Northwestern University** *Remote*  
*Project Lead* Sep. 2023 – Current  
– Led outlier-free SpeechLM project, improving adaptation by 41%, quantization by 45%, and training speed by 1.33x on OPT-1.3b. [8]
- Artificial General Intelligence AI2AI, Amazon** *Remote*  
*Student Researcher* Sep. 2022 – May. 2024  
– Mentors: Jing Liu and Ariya Rastrow  
– Designed condition-aware SSLR (CA-SSLR) model for generalist speech processing, reducing LID errors by 10%, ASR CER by 37%, and SV EER by 27% with minimal parameters. [9]
- Language Technology Institute, Carnegie Mellon University** *Pittsburgh, PA*  
*Visiting Scholar* Aug. 2021 – Dec. 2021  
– Advisor: Prof. Shinji Watanabe  
– Pioneered DiffuSE, a diffusion model for noise reduction, and extended it with CDiffuSE for non-Gaussian noise, achieving state-of-the-art results. [2, 3]  
– Contributed as a core developer to ESPnet, integrating state-of-the-art SSE models and enabling flexible combinations of SE front-ends with ASR, ST, and SLU tasks. [5, 6]
- Biomedical Acoustic Signal Processing Lab, Academia Sinica** *Taipei, Taiwan*  
*Research Assistant* Mar. 2020 – June 2022  
– Advisor: Prof. Yu Tsao  
– Developed SE framework using broad phonetic class predictions, enhancing denoising, dereverberation, and impaired speech processing. [1, 7]
- Computational AI, MediaTek Inc.** *Hsinchu, Taiwan*  
*Machine Learning Engineer* Feb. 2018 – Feb. 2020  
– Developed and optimized deep learning models (CNN, UNet) for AI DSPs, contributing to research and deployment of low-latency solutions.  
– Led the development of a custom model-generator tool to automate neural network deployment, earning a 2018 performance award and adoption into the 2020 departmental annual operating plan (AOP).

## LEADERSHIP EXPERIENCE

---

- **Master Students Research Mentor**, Johns Hopkins University 2024-2025  
—Supervised two master’s students, designing research topics and providing technical guidance in speech generation research projects.
- **Research Collaboration Lead**, Johns Hopkins University & Northwestern University 2023-2024  
—Initiated and led a cross-institutional collaboration on multimodal speech and language models, achieving a 41% improvement in adaptation efficiency and a publication in ICML workshop. [8]
- **ICASSP Grand Challenge Lead**, Carnegie Mellon University 2022  
—Led a team of seven researchers to first place in the ICASSP L3DAS22 challenge with an overall score of 98.4%, coordinating technical strategies and publishing a main conference paper. [4]
- **Intern Research Mentor**, Academia Sinica 2020-2022  
—Supervised an intern researcher on speech enhancement, culminating in co-first authorship on a TASLP paper. [7]

## SELECTED PUBLICATIONS

---

- [10] **Yen-Ju Lu**, Ting-Yao Hu, Hema Swetha Koppula, Hadi Pouransari, Jen-Hao Rick Chang, Yin Xia, Xiang Kong, Qi Zhu, Xiaoming Simon Wang, Oncel Tuzel, Raviteja Vemulapalli, "Mutual Reinforcement of LLM Dialogue Synthesis and Summarization Capabilities for Few-Shot Dialogue Summarization," In *Findings of the Association for Computational Linguistics (NAACL)*, 2025.
- [9] **Yen-Ju Lu**, Jing Liu, Thomas Thebaud, Laureano Moro-Velazquez, Ariya Rastrow, Najim Dehak, Jesus Villalba "CA-SSLR: Condition-Aware Self-Supervised Learning Representation for Generalized Speech Processing" In *Conference on Neural Information Processing Systems (NeurIPS)*, 2024 (Poster).
- [8] Shang Wu\*, **Yen-Ju Lu\***, Haozheng Luo\*, Jerry Yao-Chieh Hu, Jiayi Wang, Najim Dehak, Jesus Villalba, and Han Liu. "Fast Adaptation and Robust Quantization of Multi-Modal Foundation Models from Associative Memory: A Case Study in SpeechLM," In **ICML Workshops on Efficient Systems for Foundation Models II**, 2024.
- [7] **Yen-Ju Lu\***, Chia-Yu Chang\*, Cheng Yu, Ching-Feng Liu, Jieh-weih Hung, Shinji Watanabe, Yu Tsao, "Improving Speech Enhancement Performance by Leveraging Contextual Broad Phonetic Class Information," In *IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)*, 2023.
- [6] **Yen-Ju Lu**, Xuankai Chang, Chenda Li, Wangyou Zhang, Samuele Cornell, Zhaoheng Ni, Yoshiki Masuyama, Brian Yan, Robin Scheibler, Zhong-Qiu Wang, Yu Tsao, Yanmin Qian, Shinji Watanabe "Software Design and User Interface of ESPnet-SE++: Speech Enhancement for Robust Speech Processing" In *Journal of Open Source Software (JOSS)*, 2023.
- [5] **Yen-Ju Lu**, Xuankai Chang, Chenda Li, Wangyou Zhang, Samuele Cornell, Zhaoheng Ni, Yoshiki Masuyama, Brian Yan, Robin Scheibler, Zhong-Qiu Wang, Yu Tsao, Yanmin Qian, Shinji Watanabe "ESPnet-SE++: Speech Enhancement for Robust Speech Recognition, Translation, and Understanding," In *Conference of the International Speech Communication Association (Interspeech)*, 2022.
- [4] **Yen-Ju Lu**, Samuele Cornell, Xuankai Chang, Wangyou Zhang, Chenda Li, Zhaoheng Ni, Zhong-Qiu Wang, Shinji Watanabe "Towards Low-distortion Multi-channel Speech Enhancement: The ESPNET-SE submission to the L3DAS22 challenge" In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2022.
- [3] **Yen-Ju Lu**, Zhong-Qiu Wang, Alexander Richard, Yu Tsao, Shinji Watanabe "CDiffuSE: Conditional Diffusion Probabilistic Model for Speech Enhancement," In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2022.
- [2] **Yen-Ju Lu**, Yu Tsao, Shinji Watanabe "DiffuSE: A Study on Speech Enhancement Based on Diffusion Probabilistic Model," In *Asia Pacific Signal and Information Processing Association (APSIPA)*, 2021.
- [1] **Yen-Ju Lu**, Chien-Feng Liao, Xugang Lu, Jieh-Wei Hung, Yu Tsao "Incorporating Broad Phonetic Information for Speech Enhancement," In *Conference of the International Speech Communication Association (Interspeech)*, 2020.

## AWARDS

---

- **Government Scholarship to Study Abroad**, Ministry of Education, Taiwan *2023-2024*
- **IEEE ICASSP Grand Challenge Champion, Team Leader**, L3DAS22 Challenge *2022*
- **Ph.D. Scholarship**, Johns Hopkins University *2022*
- **Young Scientist Granted**, InterSpeech *2020*

## ACADEMIC SERVICE

---

- **Session Chair:** Conference on Information Sciences and Systems (CISS) *2025*
- **Journal Reviewer:** TASLP, SpeechCom, IEEE SPL, IEEE J-STSP, IJCNN, IJIG
- **Conference Reviewer:** ICASSP (2023-2025), ASRU (2023), InterSpeech (2023-2025), SLT (2022, 2024)
- **Challenge Host:** Multimodal Information Based Speech Processing (MISP) *2021*

## PROGRAMMING SKILLS

---

- **Programming Languages:** Python, C/C++, MATLAB, OpenCL
- **Machine Learning Frameworks:** PyTorch, TensorFlow

## TEACHING & ACTIVITIES

---

- **Tutor of Physics**, Coached a Gold Medalist, IPHO (International Physics Olympiad) *2019*
- **Lecturer of High School Physics**, Classes of 50-300 students, Chen-Li Educational Group *2013-2018*
- **Homewood Representative**, JHU Taiwanese Student Association *2023*
- **Host of Spring Feast**, MediaTek Inc. *2018-2019*
- **Event Planning Department**, NTU Baking Club *2016-2017*
- **Director of Academic Team**, NTU Guitar Club *2014-2015*