Yen-Ju Lu

+1-(626)215-5052
ylu125@jhu.edu
https://github.com/neillu23
https://scholar.google.com/citations?user=emtNw84AAAAJ

Education

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

Research Experiences

- **Meta Superintelligence Labs (MSL)** *Menlo Park, CA*
  Research Scientist Intern — Mentors: Duc Le, Srinivasan Iyer, Mike Lewis May 2025 – Oct 2025
  – Drove large-scale pretraining of an 8B-parameter multimodal foundation model across 32 nodes (256×H100), achieving stable scaling and compute-efficient multimodal learning at distributed scale.
  – Invented the Latent Speech-Text Transformer (LST), a patch-based framework that improved scaling laws and reduced multimodal compute imbalance at billion-parameter scale (***ICLR 2026 Oral***).
  – Delivered 20% compute savings and narrowed the speech–text gap by 30%; early product-facing tests showed consistent gains and strong integration potential.
- **Apple Machine Learning Research** *Pittsburgh, PA*
  Research Intern — Mentors: Ting-Yao Hu, Hema Swetha Koppula, Raviteja Vemulapalli May 2024 – Aug 2024
  – Developed Mutual Reinforcing Data Synthesis (MRDS) with DPO for joint alignment of LLM dialogue generation and summarization, enabling self-improving preference supervision without human annotation (*NAACL 2025*).
  – Outperforms 300-shot performance using only 100 shots with improved human evaluation; integrated into internal evaluation pipelines for production-scale testing.
- **AI2AI, Johns Hopkins University & Amazon** *Baltimore, MD*
  Research Collaboration & Fellowship — Mentors: Jing Liu, Ariya Rastrow Aug 2022 – Present
  – Proposed Paired by the Teacher (PbT) framework, converting unpaired data into high-fidelity pairs for summarization and question generation, achieving near human-level performance. (*EMNLP 2025*).
  – Developed CA-SSLR, a condition-aware multilingual SSLR, reducing LID errors by 10%, ASR CER by 37%, and SV EER by 27% with <5% trainable parameters (*NeurIPS 2024*).
- **Carnegie Mellon University & Academia Sinica** *Pittsburgh, PA & Taipei, Taiwan*
  Visiting Scholar / Research Assistant — Advisors: Shinji Watanabe, Yu Tsao Mar 2020 – Jun 2022
  – Invented a theoretically rigorous conditional diffusion model (CDiffuSE), achieving SOTA speech enhancement and widely adopted in follow-up research. (*ICASSP 2022*)
  – Built and open-sourced speech enhancement frameworks, integrating SE with ASR/ST/SLU and impaired-speech modeling (*Interspeech 2022; TASLP 2023*).
  – Led a 7-member team to 1st place (98.4%) in the ICASSP 2022 L3DAS22 Grand Challenge, demonstrating technical leadership and publishing the winning open-source pipeline.
- **MediaTek Computational AI Group** *Hsinchu, Taiwan*
  Machine Learning Engineer Feb 2018 – Feb 2020
  – Optimized CNN/UNet models for edge AI DSPs, enabling low-latency multimodal inference across devices.
  – Built automated deployment toolkit, adopted in 2020 departmental AOP, and awarded 2018 Performance Award.

Honors & Awards

- **ICLR Oral Presentation (1.18% acceptance rate)** *2026*
- **Amazon AI2AI Fellow, Johns Hopkins University & Amazon** *2025 – 2026*
  Selected as 1 of 3 Ph.D. candidates university-wide
  – Award package includes $50,000 stipend, 20% tuition remission, and $3,000 travel grant.
  – Granted AWS credits of $30,000/quarter ($120,000 annually) for large-scale AI research.
- **Taiwan Government Scholarship for Overseas Study** *2023 – 2024*
  Selected as 1 of 16 nationwide recipients in Electrical Engineering and Computer Science
- **ICASSP Grand Challenge Champion** *2022*
  Team Leader – 1st place (98.4%) out of 17 teams in the L3DAS22 Challenge

Selected Publications

- **Yen-Ju Lu**, Yashesh Gaur, Wei Zhou, Benjamin Muller, Jesus Villalba, Najim Dehak, Luke Zettlemoyer, Gargi Ghosh, Mike Lewis, Srini Iyer, Duc Le, "Latent Speech-Text Transformer", In **ICLR 2026 (Oral)**.
- **Yen-Ju Lu**, Thomas Thebaud, Laureano Moro-Velazquez, Najim Dehak, Jesus Villalba, "Paired by the Teacher: Turning Unpaired Data into High-Fidelity Pairs for Low-Resource Text Generation", In **EMNLP 2025**.
- **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 **NAACL 2025**.
- **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 **NeurIPS 2024**.
- **Yen-Ju Lu**, Chia-Yu Chang, Cheng Yu, Ching-Feng Liu, Jeih-weih Hung, Shinji Watanabe, Yu Tsao, "Improving Speech Enhancement Performance by Leveraging Contextual Broad Phonetic Class Information," In **TASLP 2023**.
- **Yen-Ju Lu**, Zhong-Qiu Wang, Alexander Richard, Yu Tsao, Shinji Watanabe "CDiffuSE: Conditional Diffusion Probabilistic Model for Speech Enhancement," In **ICASSP 2022**.
- **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 **Interspeech 2022**.
- **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 **ICASSP 2022**.

Leadership Experience

- **Master's Research Mentor,** Johns Hopkins University *2024-2025*
  - Supervised creation of EmoDialog-Sum, the first spoken dialogue summarization dataset with emotion & paralinguistic labels (+**28% ROUGE-L** over cascaded baselines).
- **Cross-Institution Research Lead,** Johns Hopkins University & Northwestern University *2023-2024*
  - Initiated Johns Hopkins–Northwestern collaboration, improving multimodal speech-language models adaptation efficiency by **41%;** resulted in ICML workshop paper.
- **ICASSP Grand Challenge Champion,** Carnegie Mellon University *2022*
  - Led **7-member team to 1st place (98.4%) in L3DAS22**; published and released open-source SE pipeline.
- **Intern Research Mentor,** Academia Sinica *2020-2022*
  - Supervised an intern researcher on speech enhancement, leading to **co-first authorship** on a TASLP paper.

Academic Service

- **Session Chair:** CISS 2025
- **Reviewer:** NeurIPS, ICLR, ICASSP, Interspeech, ASRU, SLT, TASLP

Programming Skills

- **Frameworks:** PyTorch, TensorFlow, HuggingFace Transformers, fairseq, ESPnet
- **Languages:** Python, C++/C, MATLAB, OpenCL