Yu Wu

☎ 551-229-5305   ✉ yw828@rutgers.edu   ⊞ https://yuwuofrutgers.github.io

EDUCATION

Rutgers, The State University of New Jersey, New Jersey, USA   *Aug. 2020-present*

PhD candidate in Electrical and Computer Engineering (Advisor: Prof. Anand D. Sarwate)

University of Science and Technology of China (USTC), Anhui, China   *Jun. 2017-Jun. 2020*

Master's in Electronic Engineering and Information Science (Advisor: Prof. Bin Liu)

University of Science and Technology of China (USTC), Anhui, China   *Jul. 2013-Jun. 2017*

Bachelor's in Information Security

WORK EXPERIENCE

• **Online Learning Algorithm for Audio Echo Cancellation** | least mean square, adaptive filtering, acoustic signal processing

- *Researcher (Intern) at **Nokia Bell Lab***   *Jun. 2022-Aug. 2022*

- *Design and analysis on online learning algorithm for audio echo cancellation system.* [paper]

SELECTED RESEARCH PROJECTS

• **Interactive LLM Cascade** | LLM, RAG, In-context Learning, PyTorch, Knowledge Distillation, scalable, modular

- ***Project Leader**, Research Assistant at Rutgers*   *Sep. 2024- present*

- *To achieve higher accuracy and save tokens in **multi-LLM collaboration**, we propose an online **RAG**-augmented system where weaker LLM can be assisted by stronger LLM both in short and long term.*

- *Inter-Cascade improves the accuracy of the weaker model by **33.06%** and saves **49.63%** cost of using stronger LLM.* [paper]

• **Collaborative Machine Learning for Edge-cloud System** | real-time inference, distributed system, computer vision, ViT, PyTorch

- ***Project Leader**, Research Assistant at Rutgers*   *Sep. 2020- Sep. 2024*

- *To enhance legacy ML models, we propose the Learning to Help diagram to jointly utilize external machine or human expert on hybrid system. Our method is **Bayes optimal** and increase the system overall accuracy by **4%~12%**.* [paper1, paper2, code]

• **Enhancing Model-Based Reinforcement Learning With Data Filter** | Out-of-distribution, RL, MuJoCo, MBPO

- *Project collaborator, Research Assistant at Rutgers*   *Jun. 2024 - Sep. 2024*

- *To bridge model-free and model-based RL, we propose **Out-of-distribution** data filter, which adaptively improves the quality of generated data. We provide **tighter bound** for estimated error and save up to **25%** epochs to reach convergence.* [paper]

• **Anti-interference for WiFi-based Human Activity Recognition (HAR)** | CSI, non-intrusive sensing, machine learning, PyToch

- *Research Assistant at EEIS department of USTC*   *Sep. 2017-May 2020*

- *Propose interference mitigation algorithms for WiFi signals. **Improves 16%** on accuracy and **9× speed**.* [paper1, paper2]

SELECTED PUBLICATIONS

• **Y Wu**, S Wu, et al., "Not only a helper, but also a teacher: Interactive **LLM** Cascade." *Under review by ICLR 2026*

• **Y Wu**, Y Li, et al., "Learning to Help in Multi-Class Settings." *ICLR 2025*

• **Y Wu**, and Anand Sarwate, "Learning to Help: Training Models to Assist Legacy Devices." *ISIT 2024 Workshop IT-TML*

• J Huang, B Liu, C Miao, Y Lu, **Y Wu**, et al., "PhaseAnti: An anti-interference WiFi-based activity recognition system using interference-independent phase component." *IEEE Transactions on Mobile Computing 2021*

• J Huang, B Liu, P Liu, C Chen, N Xiao, **Y Wu**, et al., "Towards anti-interference WiFi-based activity recognition system using interference-independent phase component." *INFOCOM 2020*

• J Gong, **Yu Wu**, et al., "Tessutivo: Contextual interactions on interactive fabrics with inductive sensing." *UIST 2019*

- **Preprint**

W Zhang, Y Li, Z Dong, **Y Wu**, et al., "Enhancing **LLM**-Based Code Generation in Large-Scale Projects." *arXiv (2024)*

Y Li, Z Dong, E Luo, **Y Wu**, et al., " Enhancing Dyna-Style Model-Based **Reinforcement Learning** With Data Filter." *arXiv (2024)*

CURRENT RESEARCH FOCUS

• Efficient and reliable AI for hybrid system (Learning to Defer, Reinforcement Learning, distributed optimization)

• LLM Collaboration (LLM Route, human–computer interaction, In-context Learning)

SKILLS

• Python, C, C++, MATLAB, SQL, Java, R, PyTorch, Pandas, scikit-learn, Slurm, Git, LoRA, HuggingFace