

Wuao Liu

13B Brandywine Amherst, MA 01002

☎ (734) 5480507 • ✉ wuaoliu@umass.edu • 🌐 wuaoliu.github.io • 📄 [Wuao652](#)

Education

University of Massachusetts Amherst

Ph.D. in Computer Science

Amherst, MA

Aug 2024 – Dec 2028 (expected)

University of Michigan, Ann Arbor

M.S. in Robotics

Ann Arbor, MI

Aug 2021 – Apr 2023

Zhejiang University (ZJU)

B.Eng. in Robotics

Hangzhou, China

Sep 2017 – Jun 2021

Publications

Mustafa Chasmai, **Wuao Liu**, Subhransu Maji, and Grant Van Horn. Audio Geolocation: A Natural Sounds Benchmark. *Under Review*. [[paper](#)]

Filippos Bellos, Yayuan Li, **Wuao Liu**, and Jason J. Corso. Can Large Language Models Reason About Goal-Oriented Tasks? In *ACL Workshop on the Scaling Behavior of Large Language Models*, 2024. [[paper](#)]

Work Experience

Honda Research Institute

Research Scientist Intern

San Jose, CA

May 2025 – Aug 2025

- Presented a novel framework integrating spatial audio cues with Large Language Models for AR/VR applications.
- Designed a spatial audio encoder optimized for both sound classification and direction-of-arrival (DoA) estimation.
- Introduced the SpatialAudioQA dataset with **872k** perceptual and reasoning questions and answers.

University of Michigan, Ann Arbor

Research Engineer

Ann Arbor, MI

Jun 2023 – Apr 2024

- Designed an egocentric video collection protocol, producing a curated dataset of **224** instructional cooking videos.
- Implemented a Faster R-CNN object detector for **14** kitchen items, achieving **85.99 mAP** on the curated dataset.
- Developed a novel procedural step classification model, achieving **92.54%** accuracy and enabling automated detection of step transition errors.

Research Experience

University of Massachusetts Amherst

Self-supervised Learning for Natural Sounds

Advisor: Prof. Grant Van Horn

Jan 2025 – Present

- Extended the Audio Masked Autoencoder (AudioMAE) pretraining framework to wildlife sounds around the world.
- Enabled self-supervised representation learning via audio spectrogram reconstruction.
- Improved sound classification accuracy on the iNatSounds dataset by **1.6%**.

University of Massachusetts Amherst

Contrastive Learning for Audio Geolocation

Advisor: Prof. Grant Van Horn

Aug 2024 – Dec 2024

- Proposed the first framework for global-scale geolocation using natural audio.
- Developed a retrieval-based AG-CLIP model, integrating natural audio geolocation with species range prediction.
- Achieved a median localization error of **1127 km** and regional-level accuracy of **17.2%** on the iNatSounds dataset.

Technical Skills

Programming Languages: Python, C/C++, MATLAB, $\LaTeX$

Deep Learning Frameworks: PyTorch, TensorFlow2, JAX

Machine Learning Packages: Hugging Face transformers, Scikit-learn, Librosa, OpenCV, Pandas

LLM Training: Verl, Hugging Face TRL