Wuao Liu

CONTACT

140 Governors Dr
Amherst, MA 01002

Email: wuaoliu52@gmail.com
Web: https://wuao652.github.io/

RESEARCH INTERESTS

Computer vision and machine learning, with a focus on learning visual representations from multimodal supervision (such as audio, video, text, etc.). Problems of interest include fine-grained classification and detection, video understanding tasks such as captioning, reasoning, and question answering.

EDUCATION

University of Massachusetts Amherst — Aug. 2024 – Apr. 2029 (expected)
Ph.D. in Computer Science — Amherst, MA
Advisor: Dr. Grant Van Horn

University of Michigan — Aug. 2021 – Apr. 2023
M.S. in Robotics — Ann Arbor, MI
Advisor: Dr. Jason Corso

Zhejiang University (ZJU) — Sep. 2017 – Jun. 2021
B.Eng. in Automation — Hangzhou, China

PUBLICATIONS

1. Mustafa Chasmai, **Wuao Liu**, Subhransu Maji, and Grant Van Horn. **Audio Geolocation: An Investigation with Natural Sounds**. *Under Submission*.
2. Filippos Bellos, Yayuan Li, **Wuao Liu**, and Jason Corso. **Can Large Language Models Reason About Goal-Oriented Tasks?** In *Proceedings of ACL Workshop on the Scaling Behavior of Large Language Models*, pages 24-34, 2024. [paper]

WORK EXPERIENCE

University of Michigan — Jun. 2023 – Apr. 2024
Research Engineer — Ann Arbor, MI
Mentor: Dr. Jason Corso
Adopted foundation models (OpenAI's Whisper and GPT Models) for a VR/AR system to assist humans in cooking tasks. Responsibilities included training action recognition models and developing an automatic speech recognition module. [project]

Tencent AI Lab — Jun. 2021 – Aug. 2021
Machine Learning Engineer Intern — Shenzhen, China
Mentor: Dr. Peilin Zhao
Developed a multi-agent reinforcement learning algorithm for optimizing electric vehicle charging and navigation systems using real-world data.

RESEARCH EXPERIENCE

UMass Vision Lab — Aug. 2024 – Dec. 2024
Advisor: Dr. Grant Van Horn — Amherst, MA
Project: **Contrastive Learning of Audio and Geographical Data**

- Proposed the first work to geolocate natural audio on a global scale.
- Investigated the potential of species vocalizations as geolocation cues with strong oracles and proof of concepts.
- Implemented an audio geolocation approach integrating species range prediction with retrieval-based methods.
- Enabled multimodal geolocation (audio, image, text) with cross-modal contrastive learning.

TECHNICAL SKILLS

- Programming Languages: Python, C/C++, MATLAB, LaTeX.
- Deep Learning Frameworks: PyTorch, TensorFlow, Hugging Face, Scikit-learn.
- Robotics Frameworks: ROS, GTSAM.