

ZIZHAO HU

zizhao-hu.github.io | 213-675-4878 | zizhaoh@usc.edu | linkedin.com/in/zizhaohu/

SKILLS

- Research expertise in Multimodal Learning, Multimodal Fusion Architecture, VLM, MLLM
- Programming Language: English (LLM), Python, Java, C++, C, Javascript, SQL
- Developer Tools: PyTorch, TensorFlow, CUDA, Git, SLURM, GNU/Linux, Unity
- Familiar Open-source Model Families: Llama, Qwen, DeepSeek, CLIP, LLaVA, InternVL, Stable Diffusion

EDUCATION

University of Southern California Los Angeles, CA
PhD Candidate in Computer Science August 2022-Present

- Multimodal Learning, Reinforcement Learning, Continual Learning, Federated Learning, Vision Transformers, Chain of Thoughts, Mixture of Expert, Multimodal fusion, LoRA, Supervised Fine-tuning

University of Southern California Los Angeles, CA
Master of Science in Computer Science August 2020-May 2022

- Machine Learning, Artificial Intelligence, Neural Networks, Computer Vision, Natural Language Processing, Reinforcement Learning

Georgia Institute of Technology Atlanta, GA
Bachelor of Science in Physics August 2016-May 2020

- Key courses/concepts: Remote Sensing, Photonics, Acoustics, Optic Signal Processing, Statistics, Quantum Computing

RESEARCH EXPERIENCE

USC Information Science Institute Los Angeles, CA
Research Assistant February 2022-Present

- Optimized Neural Architectures such as Transformers, ViT, VAE, CNNs, MLP-Mixers, and MoE for Vision and Vision-Language tasks. Improved state-of-the-art performance by up to 10% and efficiency by up to 40% for image classification, generation, segmentation
- Improved Crossmodal alignment in Multimodal architectures such as text-to-image Diffusion Models and Vision-Language Models. Achieved >10% preference margin on existing alignment benchmarks with 20% reduced inference cost

USC iLab Los Angeles, CA
Research Assistant February 2021-February 2022

- Developed new Reinforcement Learning algorithms for safe AI agents in real-time environments, presented works at AISTATS
- Designed Intrinsic Reward for AI agents to learn physics of environments, opened a new arena for human-inspired RL

Georgia Tech Nanophotonics Laboratory Atlanta, GA
Research Assistant August 2019-August 2020

- Designed Autoencoders to reverse-design Photonic Metamaterials, first in industry
- Achieved human-expert-level design quality, and reduced design cost from 6+ hours to 3 seconds per setting

Georgia Tech Agile System Laboratory Atlanta, GA
Research Assistant August 2017-August 2018

- Automated experimental data processing pipeline with anomaly detection Classifiers and OOD detection with Clustering. Reduced experiment turnaround time from 1-2 days to <30 minutes
- Designed CNN Segmentation Network for Video Processing pipeline for flight profile of Hawk Moth

PUBLICATIONS

- Multimodal Alignment: "Intermediate Adapter: Efficient Alignment of Text in Diffusion Models" (2024)
- Efficient Vision Architecture: "Static Key Attention in Vision" (2024)
- Efficient Multimodal Architecture: "Lateralization MLP: A Simple Brain-inspired Architecture for Diffusion" (2024)
- Multimodal Generation: "Cognitively Inspired Cross-Modal Data Generation Using Diffusion Models" (2023)
- Generative Replay: "Class-incremental learning using generative experience replay based on time-aware regularization" (2023)
- Latent Space: "Encoding binary concepts in the latent space of generative models for enhancing data representation" (2023)
- RL: "GalilAI: Out-of-Task Distribution Detection using Causal Active Experimentation for Safe Transfer RL" (2022)

LEADERSHIP & INVOLVEMENT

Project Supervisor & Mentors

- Supervised 1000+ graduate students on Deep Learning projects in computer Vision, NLP, etc.
- Mentored 10+ junior researchers on AI publications in Computer Vision, NLP, etc.

Conference Reviewers

- NeurIPS (2024), ICML (2024, 2025), ICLR (2024, 2025)

PROJECTS

Semeval 2022

August 2021-December 2021

- Led a group of experts on a survey benchmarking pretrained language models on a novel cloze-filling task for instructional texts. Served as an initial benchmark for community

App Development

August 2018-May 2022

- Led 4 teams in developing 5 apps (Android, React Native, Flutter, Web) with backend Android Activity Recognition API, Magenta audio generation AI, database, traffic prediction ML model, and LLM-based scoring system for Yelp business comparison.

AWARDS

- 2 golds at 27th & 28th Chinese Mathematics Olympiads
- Joyce M. and Glenn A. Burdick Prize @ Georgia Tech