

SHU ZHAO

🌐 <https://shuzhao.me/> ◇ 🌳 Tree-Shu-Zhao ◇ ✉ smz5505@psu.edu

RESEARCH INTERESTS

My research interests lie at the intersection of Multimodal Large Language Models, Information Retrieval, and Long-Context Understanding, with a particular focus on developing advanced Multimodal Retrieval-Augmented Generation (MM-RAG) systems. I am passionate about designing intelligent systems that can effectively retrieve and reason across different modalities to create robust, adaptable, and efficient AI systems.

EDUCATION

The Pennsylvania State University–University Park Ph.D. in Computer Science	<i>Aug. 2022 - May. 2026 (Expected)</i> <i>Pennsylvania, USA</i>
---	---

University of Chinese Academy of Sciences M.E. in Computer Technology	<i>Sep. 2018 - Jun. 2021</i> <i>Beijing, China</i>
---	---

Anhui University B.E. in Information Security	<i>Sep. 2013 - Jun. 2017</i> <i>Anhui, China</i>
---	---

PUBLICATIONS

- HRScene: How Far Are VLMs from Effective High-Resolution Image Understanding?**
International Conference on Computer Vision (ICCV), 2025.
Yusen Zhang, Wenliang Zheng, Aashrith Madasu, Peng Shi, Ryo Kamoi, Hao Zhou, Zhuoyang Zou, **Shu Zhao**, Sarkar Snigdha Sarathi Das, Vipul Gupta, Xiaoxin Lu, Nan Zhang, Ranran Haoran Zhang, Avitej Iyer, Renze Lou, Wenpeng Yin, Rui Zhang.
- Insect Agent: Improving Insect Recognition via Dynamic Knowledge Augmentation Using Multimodal Large Language Models.**
IEEE Computer Society Annual Symposium on VLSI (ISVLSI), 2025.
Shu Zhao, Ajay Narayanan Sridhar, Harland Patch, Vijaykrishnan Narayanan.
- KALAHASH: Knowledge-Anchored Low-Resource Adaptation for Deep Hashing.**
AAAI Conference on Artificial Intelligence (AAAI), 2025.
Shu Zhao, Tan Yu, Xiaoshuai Hao, Wenchao Ma, Vijaykrishnan Narayanan.
- MapFusion: A Novel BEV Feature Fusion Network for Multi-modal Map Construction.**
Information Fusion. 2025.
Xiaoshuai Hao, Yunfeng Diao, Mengchuan Wei, Yifan Yang, Peng Hao, Rong Yin, Hui Zhang, Weiming Li, **Shu Zhao**, Yu Liu.
- NEUCORE: Neural Concept Reasoning for Composed Image Retrieval.**
UniReps Workshop @ Neural Information Processing Systems (NeurIPS Workshop), 2023.
Shu Zhao, Huijuan Xu.
- Rescuing Deep Hashing from Dead Bits Problem.**
International Joint Conference on Artificial Intelligence (IJCAI), 2021.
Shu Zhao, Dayan Wu, Yucan Zhou, Bo Li, Weiping Wang.
- Technical Report for EPIC-KITCHENS-100 2021 Multi-Instance Retrieval Challenge.**
EPIC-Kitchens Challenges @ Computer Vision and Pattern Recognition (CVPR Workshop, Ranked 1st), 2021.
Xiaoshuai Hao, Wanqian Zhang, Dejie Yang, **Shu Zhao**, Dayan Wu, Bo Li, Weiping Wang.
- Asymmetric Deep Hashing for Efficient Hash Code Compression.**
ACM International Conference on Multimedia (ACM MM, Oral), 2020.
Shu Zhao, Dayan Wu, Wanqian Zhang, Yu Zhou, Bo Li, Weiping Wang.

PREPRINTS

1. Less is More: Toward Zero-Shot Local Scene Graph Generation via Foundation Models.

arXiv:2310.01358. 2023.

Shu Zhao, Huijuan Xu.

RESEARCH EXPERIENCE

Machine Learning Engineer Intern. NVIDIA, Santa Clara, CA

May. 2025 - July. 2025

Advised by Dr. Tan Yu

Topic: Deep Search Agent with Reinforcement Learning with Verifiable Rewards

Research Intern. Baidu Research, Beijing

Apr. 2022 - Jul. 2022

Advised by Dr. Tan Yu

Topic: Large Multimodal Models for Image Retrieval

Research Assistant. IIE, Chinese Academy of Sciences, Beijing

Mar. 2019 - Jun. 2021

Advised by Prof. Dayan Wu and Prof. Bo Li

Topic: Deep Image Hashing

INTERDISCIPLINARITY

InsectEye

2024

Collaboration with Entomology Department, Penn State

Description: we are developing an AI system for automated insect recognition and classification as part of the InsectEye project, a non-lethal biodiversity monitoring initiative at Penn State's Arboretum. We are working on engineering robust computer vision algorithms to overcome challenging environmental conditions, including variable lighting and background noise while achieving accurate identification of small insects with subtle morphological differences.

Textural Threshold: Dreadlock

2023

Collaboration with Architecture Department, Penn State

Description: a user's hair texture is acquired through a camera, and similar and dissimilar images are retrieved through deep retrieval techniques and stitched together to form an art painting. Our work was presented at the Central Pavilion Exhibition, the 18th Architectural Venice Biennale.

REVIEW SERVICES

Computer Vision and Pattern Recognition (CVPR)

2024-2025

International Conference on Computer Vision (ICCV)

2025

ACM International Conference on Multimedia (ACM MM)

2022-2025

Annual Conference on Neural Information Processing Systems (NeurIPS)

2023-2025

TEACHING ASSISTANT

Deep Learning for Computer Vision, Undergraduate

2024

Vision and Language, Graduate

2023

SKILLS

Python, PyTorch, \LaTeX , Linux, MATLAB, C++