Sheng Zhang

1+2028646946 | [Github](#) | [Website](#) | sheng.zhang@mbzuai.ac.ae | eyisheng@outlook.com

EDUCATION

University of Maryland, College Park   Aug. 2024 – Present
Ph.D. in Computer Science, US Maryland
Advisor: [Prof. Heng Huang](#)

MBZUAI   Aug. 2021 – Aug. 2024
Master of Science/Graduate Research Assistant, UAE Abu Dhabi
Advisor: [Prof. Salman Khan](#); Co-advisor: [Dr. Zhiqiang Shen](#)

Tongji University   Sep. 2016 – June 2021
Bachelor of Software Engineering, China Shanghai

RESEARCH INTERESTS

○ Experiential Agents, (Multimodal) Foundation Models, Open-World Learning.
○ Concept & Knowledge & Ontology & Representation.

PUBLICATIONS

1. GenFlowRL: Shaping Rewards with Generative Object-Centric Flow in Visual Reinforcement Learning.
   Kelin Yu, **Sheng Zhang***, Harshit Soora, Furong Huang, Heng Huang, Pratap Tokekar, Ruohan Gao.*
   ICCV 2025, ICRA FMNS Workshop 2025 (**Spotlight**). [Paper](#).

2. Towards Realistic Zero-Shot Classification via Self Structural Semantic Alignment.
   Sheng Zhang, *Muzammal Naseer, Guangyi Chen, Zhiqiang Shen, Salman Khan, Kun Zhang, Fahad Khan.*
   AAAI 2024 (**Oral**). [Paper](#) [Code](#).

3. PromptCAL: Contrastive Affinity Learning via Auxiliary Prompts for Generalized Novel Category Discovery.
   Sheng Zhang, *Salman Khan, Zhiqiang Shen, Muzammal Naseer, Guangyi Chen, Fahad Khan.*
   CVPR 2023. [Paper](#) [Code](#).

4. A Channel Attention Based Deep Neural Network for Automatic Metallic Corrosion Detection.
   Sheng Zhang, *Xinling Deng, Yumin Lu, Shaozheng Hong, Zhengyi Kong, Yongli Peng, Ye Luo.*
   Journal of Building Engineering, 2021. [Paper](#).

5. From Comparison to Composition: Understanding Cognition of Unseen Categories.
   Minghao Fu, **Sheng Zhang***, Guangyi Chen, Zijian Li, Yifan Shen, Fan Feng, Shaoan Xie, Kun Zhang.*
   NeurIPS 2025 CogInterp Workshop, Other Conference Under Review. [Paper](#).

RESEARCH EXPERIENCE

Benchmarking and Probing Conceptual Structure in Multimodal LLM   Mar. 2025 – Present
My individual research advised by Prof. Salman Khan, Dr. Zhiqiang Shen, Prof. Heng Huang.   *UMD, US*

- ➢ `Intro`: Inspired by cognitive and epistemological principles, propose a diagnosis-oriented benchmark that assesses whether multimodal LLMs genuinely align visual features with category–attribute and inter-category relations beyond memorization in various downstream tasks.
- Developed a multi-agent framework which generates and curates ∼3K tasks with reasoning, nomenclature, and web-grounded sources integration.
- Developed an evaluation backend with 60+ frontier SOTA MLLMs.
- Proposed a novel evaluation metric and conducted statistical analysis on inference results.

GenFlowRL: Flow Guided Robot Learning   Oct. 2024 – Mar. 2025
My collaborative course project advised by Prof. Pratap Tokekar and Prof. Ruohan Gao   *UMD, US*

- ➢ `Intro`: A framework that shapes RL rewards using generative object-centric flow learned from diverse cross-embodiment demonstrations, enabling robust visuomotor policy learning.
- Implemented and adapted the foundation generative video model to generate the visual conditioned delta-flow representation for manipulation policy learning.
- Prepared experiments, baseline ablations, and writing to communicate results and methodology.

Realistic Zero-Shot Classification with Vision-Language Learning   Jan. 2023 – June. 2023
My individual research advised by Prof. Salman Khan, Dr. Zhiqiang Shen, Prof. Fahad Khan.   *MBZUAI, UAE*

- ➢ `Intro`: Formulated a novel problem, *Realistic Zero-Shot Classification*, which relaxes the *ideal vocabulary* assumption of the ground-truth target label set. Proposed a semantic structural alignment method in our framework with calibrated image-category alignments by leveraging Large Language Models.
- Adapted and reproduced multiple previous strong baselines to our setting for performance comparisons.
- Developed and benchmarked our method as SOTA performance on six generic and fine-grained datasets evaluated with transductive/inductive, unsupervised/semi-supervised, and out-of-vocabulary setups. In the standard setup, our method surpasses CLIP by over 20% absolute accuracies.

Generalized Novel Category Discovery   June 2022 – Dec. 2022
My individual research and master thesis advised by Prof. Salman Khan   *MBZUAI, UAE*

- ➢ `Intro`: Proposed a visual prompt-based contrastive affinity learning framework to address the generalized category discovery problem, which aims to categorize both known and novel classes with known class annotations. Our method can learn semantic discriminative clusters via contrastive learning on diffused affinity graphs.
- Designed and implemented a novel visual prompt regularization technique to enhance backbone semantic discriminativeness. Demonstrated its superiority over naive visual prompt tuning.
- Achieved SOTA performance on seven challenging benchmarks including fine-grained StanfordCars and CUB, significantly surpassing previous methods, e.g., with nearly 11% cluster accuracy on CUB and 9% on ImageNet-100. Conducted further evaluations in transductive/inductive and few-annotation scenarios.

Weakly-Supervised Semantic Segmentation on Medical Images   Jan. 2021 – June 2021
My bachelor's thesis advised by Dr. Ye Luo.   *Tongji University, China*

- ➢ `Intro`: Developed an adversarial self-supervised approach to address semi-supervised domain adaptation problem.
- Designed a multi-scale patch rotation prediction pretext to enhance the representation transferability on downstream low-annotation segmentation tasks, surpassing RotNet and JigSaw pretexts by over 3% average IoU.
- Evaluated methods by comparing with RotNet and Jigsaw representation learning baselines in the semi-supervised setting of various labeling ratios and in the domain-shit scenario.

ACADEMIC SERVICES

- ○ Received the NIPS 2025 Registration Award as a Top-8% Reviewer
- ○ Delivered a presentation on our work *Towards Realistic Zero-Shot Classification* at AAAI'24
- ○ Serving as a reviewer at NIPS, ICLR, CVPR, ICCV, ECCV, IJCV, TIP, AISTATS, ACM Computing Surveys
- ○ Serving as a conference volunteer at AAAI'24, EMNLP'23

SKILLS

Language: Chinese (native), English (fluent): TOEFL iBT 106
Programming: Python, Java
DL Tools: Pytorch, Numpy, transformers, timm
Others: Triton, Hadoop

GRADUATE COURSEWORK

CMSC818B Decision-Making for Robotics (A)
CMSC848K Multimodal Foundation Models (A+)
MATH858M Stochastic Methods with Applications (A)