Ketul Shah

kshah33.github.io

Email : kshah33@jhu.edu
Mobile : +1-240-938-5815

Research Interests

Multimodal LLMs, LLM Agents, Action Recognition, Multi-View Vision, Synthetic Data, Domain Adaptation, Domain Generalization

Education

Johns Hopkins University   **2020–Present**
- *Ph.D. in Electrical and Computer Engineering*   *Baltimore, MD, USA*
  Advisor: Prof. Rama Chellappa

University of Maryland, College Park   **2018–2020**
- *M.S. in Electrical and Computer Engineering; GPA: 3.67/4.0*   *College Park, MD, USA*

Indian Institute of Technology, Madras   **2013–2018**
- *B.Tech & M.Tech in Electrical Engineering; GPA: 8.30/10.0*   *Chennai, India*
  Minor in Operations Research
  Advisor: Prof. Kaushik Mitra

Professional Experience

Amazon Web Services   **Summer 2025**
- *Applied Scientist Intern. Mentor: Dr. Mayank Bansal*   *Santa Clara, CA*
  - Developed VideoMimic, an approach for video-based task replay for UI Agents.

Adobe Research   **Summer 2024**
- *Research Scientist Intern. Mentor: Dr. Fabian Caba Heilbron*   *San Jose, CA*
  - Designed a self-refining agent for multimodal video retrieval; created benchmarks for this novel task.

Amazon Just Walk Out   **Summer 2023**
- *Applied Scientist Intern. Mentor: Dr. Robert Crandall*   *Seattle, WA*
  - Worked on large-scale pre-training using multi-view videos.

Amazon Web Services   **Summer 2021**
- *Applied Scientist Intern. Mentors: Dr. Kaustav Kundu, Dr. Xinyu Li*   *Seattle, WA (Remote)*
  - Worked on using self-supervised representation learning for improving batch active learning.

University of Maryland   **Summer 2019**
- *Research Assistant under Prof. Rama Chellappa*   *College Park, MD*
  - Worked on spatio-temporal action detection in untrimmed videos, specifically for infrared videos. Improved performance of the proposal-based method by 13.5% on infrared modality. Exploring GAN formulations for leveraging multiple modalities available at training time.

NVIDIA Graphics Pvt. Ltd.   **Summer 2016**
- *Software Engineering Intern*   *Bengaluru, India*
  - Part of a multi-team venture to extract usage data of CPU farm from logged data; streamlined load allocation to the CPU farm, also incorporated peak load times.

Works Under Submission

- **MV2MAE: Multi-View Video Masked Autoencoders**
  Ketul Shah, Robert Crandall, Jie Xu, Peng Zhou, Marian George, Vipin Pillai, Mayank Bansal, Rama Chellappa

- **MultLFG: Training-free Multi-LoRA composition using Frequency-domain Guidance**
  Aniket Roy, Maitreya Suin, Ketul Shah, Rama Chellappa

Peer-Reviewed Publications

- **VRAgent: Self-Refining Agent for Multimodal and Interactive Video Retrieval (WACV 2026)**
  Ketul Shah, *Pankaj Nathani, Rama Chellappa, Fabian Caba Heilbron*

- **Diffuse2Adapt: Controlled Diffusion for Synthetic-to-Real Domain Adaptation (ICIP 2025)**
  Ketul Shah, *Arushi Sinha, Arun Reddy, Aniket Roy, Rama Chellappa*

- **AeroGen: Ground-to-Air Generalization for Gesture Recognition (FG 2025)**
  Ketul Shah, *Anshul Shah, Arun Reddy, Aniket Roy, Arushi Sinha, Celso M. de Melo, Rama Chellappa*

- **Multi-View Action Recognition using Contrastive Learning (WACV 2023)**
  Ketul Shah, *Anshul Shah, Chun Pong Lau, Celso M. de Melo, Rama Chellappa*

- **Synthetic-to-Real Domain Adaptation for Action Recognition: A Dataset and Baseline Performances (ICRA 2023)**
  Ketul Shah*, *Arun Reddy*, William Paul, Rohita Mocharla, Judy Hoffman, Kapil D. Katyal, Dinesh Manocha, Celso M. de Melo, Rama Chellappa*

- **Improved Modeling of 3D Shapes with Multi-view Depth Maps (3DV 2020, Oral Presentation)**
  Ketul Shah*, *Kamal Gupta*, Susmija Jabbireddy*, Abhinav Shrivastava, Matthias Zwicker*

- **DiffNat: Improving Diffusion Image Quality Using Natural Image Statistics (TMLR 2025)**
  Aniket Roy, Maitreya Suin, Anshul Shah, Ketul Shah, *Jiang Liu, Rama Chellappa*

- **Cap2Aug: Caption guided Image to Image data Augmentation (WACV 2025)**
  Aniket Roy, Ketul Shah*, *Anshul Shah*, Anirban Roy, Rama Chellappa*

- **Unsupervised Video Domain Adaptation with Masked Pre-Training and Collaborative Self-Training (CVPR 2024)**
  Arun Reddy, William Paul, Corban Rivera, Ketul Shah, *Celso M de Melo, Rama Chellappa*

- **HaLP: Hallucinating Latent Positives for Skeleton-based Self-Supervised Learning of Actions (CVPR 2023)**
  Anshul Shah, Ketul Shah*, *Aniket Roy*, Shlok Mishra, David Jacobs, Anoop Cherian, Rama Chellappa*

- **FeLMi : Few shot Learning with hard Mixup (NeurIPS 2022)**
  Aniket Roy, Anshul Shah, Ketul Shah, *Prithviraj Dhar, Anoop Cherian, Rama Chellappa*

- **Photorealistic Image Reconstruction from Hybrid Intensity and Event based Sensor**
  Prasan A Shedligeri, Ketul Shah, *Dhruv Kumar, Kaushik Mitra*

Voluntary Service

- TPAMI, WACV 2026, ICCV 2025, AAAI 2025, CVPR 2024, WACV 2024, ICLR 2024, NeurIPS 2024, NeurIPS 2023, ICLR 2022, NeurIPS 2022, ECCV 2020

Teaching Experience

- **Machine Perception**   **Fall 2024, Fall 2022, Fall 2020**
  Course Assistant   *Johns Hopkins University*

- **Advanced Data Structures**   **Spring 2019**
  Teaching Assistant   *University of Maryland, College Park*

Programming Skills

- **Libraries and Tools:** PyTorch, Amazon SageMaker, PyTorch Lightning, Weights & Biases, Tensorboard, Caffe, TensorFlow, OpenCV, Blender, AWS

- **Languages:** Python, C, C++, MATLAB, Java