

YOGESH KULKARNI

yogkul2000.github.io ykulka10@asu.edu [LinkedIn](#) [Google Scholar](#)

Research Interests

I am interested in research on **multimodal reasoning**, particularly **video-language understanding**. My PhD thesis focuses on enhancing **multimodal LLMs** for video reasoning through **reinforcement learning post-training** and **preference optimization**. This includes spatiotemporal alignment via **DPO**, off-policy **GRPO** with temporal credit assignment for **audio-video reasoning**, self-supervised preference learning, and structured planning for **egocentric procedural reasoning**.

Education

Arizona State University, Tempe, AZ <i>Doctor of Philosophy, Computer Science</i>	Aug 2024 – May 2028
University of Southern California, Los Angeles, CA <i>Master of Science, Computer Science</i>	May 2024
Pune Institute of Computer Technology, Pune, India <i>Bachelor of Engineering, Computer Engineering</i>	July 2022

Publications

- AVATAR: Reinforcement Learning to See, Hear, and Reason Over Video**
Yogesh Kulkarni, Pooyan Fazli *CVPR 2026*
- VideoPASTA: 7K Preference Pairs That Matter for Video-LLM Alignment**
Yogesh Kulkarni, Pooyan Fazli *EMNLP 2025*
- VideoSAVi: Self-Aligned Video Language Models without Human Supervision**
Yogesh Kulkarni, Pooyan Fazli *COLM 2025*
- EgoVITA: Learning to Plan and Verify for Egocentric Video Reasoning**
Yogesh Kulkarni, Pooyan Fazli *Under submission*
- ReGATE: Learning Faster and Better with Fewer Tokens in MLLMs**
Chaoyu Li, Yogesh Kulkarni, Pooyan Fazli *Under submission*
- EnsembleNTLDetect: An Intelligent Framework for Electricity Theft Detection in Smart Grid**
Yogesh Kulkarni, Sayf Hussain, Krithi Ramamritham, Nivethitha Somu *ICDM 2021*
- Kryptonite: An Adversarial Attack Using Regional Focus**
Yogesh Kulkarni, Krisha Bhambani *ACNS 2021*
- Intensive Image Malware Analysis and Least Significant Bit Matching Steganalysis**
Yogesh Kulkarni, Anurag Gorkar *BigData 2020*

Experience

Graduate Researcher | PeRL Lab, Arizona State University Aug 2024 – Present
Video-Language Alignment and Reasoning (EMNLP'25, COLM'25), Audio-Video Reasoning (CVPR'26), and Egocentric Video Understanding (Under submission).

Research Scientist Intern | Nokia Bell Labs, New Providence, NJ May 2023 – Aug 2023
Implemented automatic model parallelism for GPT-3 and LLaMA models, increasing training throughput by 15% across heterogeneous clusters. Designed communication/compute-efficient inter-node pipeline parallelism for training LLMs on geo-distributed GPUs.

Graduate Researcher | USC ICT, Los Angeles, CA Jan 2023 – Mar 2024
Built 3D style transfer pipeline (ViT, CLIP, Gaussian Splatting) for text-guided point cloud coloring. Explored diffusion models (DDPM) and GANs (Pix2Pix, CycleGAN) with PointNet/KPConv/SparseUNet for 3D colorization.

Academic & Technical Details

Languages & Frameworks: Python, C/C++, SQL, PyTorch, VeRL, SWIFT, L^AT_EX

Voluntary Service: Reviewer for CVPR, ICCV, ECCV, ARR, AAAI, TPAMI