

Rohit Gupta

Citizen of India, Currently present in US on F1 visa status

☎ (+1) 4079905985 | ✉ rohitgupta.hp@gmail.com | 🏠 rohit-gupta.github.io | 📱 rohit-gupta | 📞 0WukQpMAAAJ | 📧 rohitguptahp

Education

2019 - Ongoing	PhD. in Computer Science , University of Central Florida, Orlando
2015 - 2017	M.Tech. in Computer Science and Engineering , IIT Kanpur, Kanpur, India
2010 - 2014	B.Tech. in Electrical Engineering , IIT Kanpur, Kanpur, India

Selected Publications

Under Review	ViLLaGE: Video LLM for Generative and Embedding Tasks R Gupta , J Unnikrishnan, S Tran, R Hamid, M Shah (Private Pre-Print (Google Drive): https://drive.google.com/file/d/1H6vGhaEPmrRkK-VcAbbRwLam8v8NFE7t)
Under Review	BrailleVision: Text Instruction Tuning of LLMs to Improve Visual Skills R Gupta , MN Rizve, P Tirupattur, M Shah (Private Pre-Print (Google Drive): https://drive.google.com/file/d/1dCeQsyb_qe49nmF0ikx6m8ZIW5Yx-dIR)
ECCV 2024	Open Vocabulary Multi-Label Video Classification R Gupta , MN Rizve, A Tawari, J Unnikrishnan, S Tran, M Shah (Link: https://www.ecva.net/papers/eccv_2024/papers_ECCV/papers/05599.pdf)
CVPR 2023 Citations: 15	Class Prototypes based Contrastive Learning for Classifying Multi-Label and Fine-Grained Educational Videos R Gupta , A Roy, S Kim, C Christensen, T Grindal, S Gerard, M Cincebeaux, A Divakaran, M Shah Link: https://openaccess.thecvf.com/content/CVPR2023/html/Gupta_Class_Prototypes_...
AAAI 2023 Citations: 10	Contrastive Self-Supervised Learning Leads to Higher Adversarial Susceptibility R Gupta , N Akhtar, A Mian, M Shah Link: https://ojs.aaai.org/index.php/AAAI/article/view/26733
ICPR 2020 Citations: 110	RescueNet: Joint building segmentation and damage assessment from satellite imagery R Gupta , M Shah Link: https://ieeexplore.ieee.org/document/9412295
CVIU, Jun '22 Citations: 201	TCLR: Temporal Contrastive Learning for Video Representation I Dave, R Gupta , M N Rizve, M Shah Link: https://www.sciencedirect.com/science/article/pii/S1077314222000376

Work Experience

Applied Scientist Intern, Amazon Nile, Rufus Multi-Modal Team

May-Nov 2024, Seattle

- Built an unified foundation model consisting of a Video LLM with the capabilities of outputting text as well as embeddings.

Applied Scientist Intern, Amazon Search Science and AI, M5 Team

May-Nov 2023, Palo Alto

- Worked on multi-label open vocabulary video classification; recognizing objects and actions in videos not present in training data.

Research Intern, SRI International

May- Aug 2022, Menlo Park (remote)

- Developed multi-label, multi-modal prototype contrastive learning to solve fine-grained video content understanding.

Research Engineer, Computer Vision, Conduent Labs (erstwhile Xerox Research)

Sep. 2017 - Jul. 2019, Bangalore, India

- Contributed to a variety of projects in Computer Vision: Video memorability prediction, Analyzing multi-modal data for smart-city applications, Instance recognition and image classification for augmented reality (AR) and appearance based re-identification of cars for traffic flow analysis.

Data Scientist, Fuzzy Logicx

Jul. 2014 - Jun. 2015, Bangalore, India

- Developer on DB Lytix™ suite of machine learning, statistical and financial algorithms embedded into data warehouses like Teradata™ and Netezza™

Achievements

2023	Oral Presentation , Workshop on Large Scale Holistic Video Understanding, CVPR
2021	1st Place and Jury Prize , VI-Priors Action Recognition Challenge, ICCV
2019	Fellowship , ORCGS Doctoral Fellowship, UCF
2018	1st Place , MediaEval 2018: Predicting Media Memorability Task
2010	National Rank 433 (Top 0.1%) , Joint Entrance Exam, Indian Institutes of Technology

Service

2022-present	Reviewer , CVPR, ECCV, ICCV, AAAI, ICLR, IEEE Journals (TIP, TNNLS, TCSVT)
2020, 2022	Mentor , NSF Research Experiences for Undergraduates, UCF-CRCV REU Site

Recent Research Projects

Joint Generative and Embedding Video LLM

Amazon

INTERNSHIP AND GRADUATE RESEARCH PROJECT

2024

- Developed a multi-modal large language model (LLM) capable of generating both embeddings and text to solve video understanding tasks.

Multi-Label Open Vocabulary Video Classification

Amazon

INTERNSHIP AND GRADUATE RESEARCH PROJECT

2023

- Developed a video classification model capable of recognizing object and action classes not seen during training.

Multi-Label Contrastive Learning for Fine-Grained Educational Video Classification

SRI & UCF

INTERNSHIP AND GRADUATE RESEARCH PROJECT

2022

- Achieved state of the art results on a novel dataset of education videos and two prior benchmark datasets (YouTube-8M and COIN)

Robustness of Contrastive Self-Supervised Representations

UCF

GRADUATE RESEARCH AS LEAD RESEARCHER

2021

- Identified root causes of the adversarial vulnerability of contrastive self-supervised models and boosted the robustness of state of the art by about 5%