

# 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 | 📺 rohitguptahpf

## Education

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

## Selected Publications

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

## Work Experience

<b>Applied Scientist Intern, Amazon Nile, Rufus Multi-Modal Team</b>	May-Nov 2024, Seattle
• Built an unified foundation model consisting of a Video LLM with the capabilities of outputting text as well as embeddings.	
<b>Applied Scientist Intern, Amazon Search Science and AI, M5 Team</b>	May-Nov 2023, Palo Alto
• Worked on multi-label open vocabulary video classification; recognizing objects and actions in videos not present in training data.	
<b>Research Intern, SRI International</b>	May- Aug 2022, Menlo Park (remote)
• Developed multi-label, multi-modal prototype contrastive learning to solve fine-grained video content understanding.	
<b>Research Engineer, Computer Vision, Conduent Labs (erstwhile Xerox Research)</b>	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.	
<b>Data Scientist, Fuzzy Logicx</b>	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	<b>Oral Presentation</b> , Workshop on Large Scale Holistic Video Understanding, CVPR
2021	<b>1st Place and Jury Prize</b> , VI-Priors Action Recognition Challenge, ICCV
2019	<b>Fellowship</b> , ORCGS Doctoral Fellowship, UCF
2018	<b>1st Place</b> , MediaEval 2018: Predicting Media Memorability Task
2010	<b>National Rank 433 (Top 0.1%)</b> , Joint Entrance Exam, Indian Institutes of Technology

## Service

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

## Recent Research Projects

<b>Joint Generative and Embedding Video LLM</b>	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.	
<b>Multi-Label Open Vocabulary Video Classification</b>	Amazon
INTERNSHIP AND GRADUATE RESEARCH PROJECT	2023
• Developed a video classification model capable of recognizing object and action classes not seen during training.	
<b>Multi-Label Contrastive Learning for Fine-Grained Educational Video Classification</b>	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)	
<b>Robustness of Contrastive Self-Supervised Representations</b>	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%	