

Raja Kumar

PhD Student in Computer Science
University of Southern California

Los Angeles, CA

✉ kumarraj@usc.edu

in [raja-kumar-58971010a](#)

🔗 [raja-kumar](#)

🎓 [Google Scholar](#)

Summary

Research Interest Vision Language Model (VLM), Post-training RL, Fine-grained Image and Video Understanding, on-device AI, 3D vision

Summary Computer Science PhD student at USC researching vision-language foundation models and post-training for fine-grained visual understanding. Applied-science internship experience at Amazon AGI and Flawless AI spans reinforcement learning, benchmark evaluation, and multimodal synthetic data. Previously developed data-free model quantization for efficient on-device inference at Samsung Research.

Education

Aug. 2024–
April 2028 (expected) **University of Southern California,**
PhD Computer Science,
GPA: 4.0/4.0
Advisor: [Prof. Ram Nevatia](#)

Sep. 2022–March 2024 **University of California Santa Cruz,**
MS in Computer Science,
GPA: 4.0/4.0
Advisor: [Prof. James Davis](#)

July 2015–May 2019 **Indian Institute of Technology (IIT) Kharagpur,**
B.Tech in Electrical Engineering
Advisor: [Prof. Pranab K Dutta](#)

Work Experience

- May 2026–Aug. 2026 **Applied Scientist Intern, [Amazon AGI Foundation for AWS, Santa Clara](#)**
- **SPLIT-RL: Staged perception-language reasoning with claim-level advantages:** Developed a staged reinforcement-learning framework that decouples visual reasoning from language reasoning in vision-language models. Introduced component-level rewards and visual-type-grouped advantages, improving average accuracy by 2.38 points over GRPO and reducing response length by up to 20.9% across six benchmarks.
- April 2024–July 2024 **Applied Science Intern, [Flawless AI, Santa Monica](#)**
- **3D Multimodal Facial Synthetic Data Generation:** Optimized 3D facial synthetic data generation pipeline and Advanced multi-modal understanding by augmenting synthetic data generation to integrate depth maps and positional encoding, creating a richer dataset for enhanced geometric cues

Jun. **Machine Learning Engineer**, *Samsung Research, Bangalore*
2019–Feb. 2022 ○ **Efficient Deep Learning via Model Quantization:** Developed data-free quantization techniques for deploying deep learning models on edge devices, improving inference efficiency by 10–20% while preserving model accuracy and reducing dependence on original training data.

Publications

Under review, [\[Paper\]](#) **SPLIT-RL: Staged perception-language reasoning with claim-level advantages.**
Raja Kumar, Rajat Koner, Ritwick Chaudhry, ... , Yifan Xing
ICLR 2026, [\[Paper\]](#) [\[Code\]](#) **DiVE-k: Differential visual reasoning for fine-grained image recognition.**
Raja Kumar, Arka Sadhu, Ram Nevatia
ECCVW 2024 [\[Paper\]](#) [\[Code\]](#) **Few-shot Novel View Synthesis using Depth Aware 3D Gaussian Splatting.**
Raja Kumar, Vanshika Vats
ICCVW 2023, [\[Paper\]](#) **Disjoint Pose and Shape for 3D Face Reconstruction.**
Raja Kumar, Jiahao Luo, Alex Pang, James Davis
ArXiv 2023, [\[Paper\]](#) **3D Face Reconstruction: Is Model-Based Classic Passive Stereo Competitive.**
Jiahao Luo, **Raja Kumar**, Alex Pang, James Davis
ArXiv 2023, [\[Paper\]](#) [\[Code\]](#) **Task Oriented Conversational Modelling With Subjective Knowledge.**
Raja Kumar
IJCNN 2021, [\[Paper\]](#) [\[Code\]](#) **Hybrid and Non-Uniform Quantization Methods Using Retro Synthesis Data for Efficient Inference.**
Raja Kumar, Tejpratap GVSL, Pradeep NS

Patents

US Patent, 2024 **Method and Electronic Device for Efficiently Reducing Dimensions of Image Frame.**
Tejpratap GVSL, **Raja Kumar**, Dewashish Avinash Dharkar.
Patent No: [US20230252602A1](#)
US Patent, 2022 **Method and Electronic Device for Quantizing DNN Model.**
Tejpratap GVSL, Arun Abraham, **Raja Kumar**, ..., Prasen Kumar Sharma.
Patent No: [US20230068381A1](#)

Academic Experience

Reviewer Expert Systems With Applications, ICLR 2026, NeurIPS 2026, ICLR 2027
Teaching **USC:** Advanced Computer Vision, Introduction to AI, Database Systems, Introduction to Programming Systems Design.
UCSC: Analysis of algorithms, Computer Architecture, Computer Systems and Assembly Language.

Skills

Languages Python, C++, Assembly, Javascript, SQL
Libraries PyTorch, Pytorch3D, Tensorflow, Caffe
Tools git, AWS, OpenCV, OpenGL

Relevant Courses

Advanced Computer Vision Deep Learning and its application Computer Vision
ML for NLP Adversarial and Trustworthy Foundation Models Neural Computation

Projects

- Jan. **Test-Time Defense Against Jailbreaks in Vision-Language Models, USC**
2026–May 2026 Developed a label-free GRPO-based framework for multimodal jailbreak defense, showing that self-supervised rollout-consistency rewards collapse to universal refusal and that decoupling the reward with an external safety classifier improves defense success while reducing false refusals on benign inputs.
- Jan. **Parameter Merging for Continual Adaptation in Vision-Language-Action Models, USC**
2026–May 2026 Studied continual adaptation of vision-language-action models by comparing sequential fine-tuning with post-hoc parameter merging on SmolVLA; developed a LIBERO-90 pretraining and LIBERO-10 evaluation pipeline and analyzed catastrophic forgetting and ID/OOD generalization.
- Aug. **In-Context Learning as a Query-Adaptive Parameter Update, USC**
2025–Dec 2025 Proved that in-context learning in Transformers with standard non-linear attention implements a query-adaptive “fast weight” update rather than a static implicit gradient step, extending the theoretical framework of Dai et al. beyond the linear attention assumption.
- Sep. **Dense Face Alignment, [GitHub](#)**
2022–Jan. 2023 Improved dense 2D facial landmark detection by integrating MediaPipe face detection with DAD-3DHeads, enabling robust 445-point landmark prediction on in-the-wild images where faces are only a part of the scene.
- Jul. **Automatic Classification and Localization of Defects in Hot-Rolled Steel Surfaces, IIT Kharagpur, Advisor: Prof. P.K. Dutta**
2018–Dec. 2018 Developed a DNN pipeline using transfer learning (InceptionV3) and Inception-SSD for classifying and localizing defects on steel surfaces, achieving 93.88% classification accuracy and 82.4 mAP@0.5 for localization.
- Jul. **User Classification and Authentication using Mouse and Keystroke Dynamics, IIT Kharagpur, Advisor: Prof. S. Mukhopadhyay**
2018–Dec. 2018 Collected mouse and keystroke dynamics data from daily user interactions and trained ANN models for user classification (84% accuracy) and authentication (89% accuracy).
- May **Identifying Patterns of Safety-Related Incidents in a Steel Plant, IIT Kharagpur, Advisor: Prof. J. Maity**
2017–Jul. 2017 Applied Association Rule Mining (Apriori algorithm) on incident investigation reports to discover patterns among accident factors and incident types in a steel plant.

Awards & Achievements

- Apr. 2021 Received MD's choice award at Samsung Research for notable contribution to GS22 project.
- Nov. 2021 Received SPOT award at Samsung Research for contribution to quantization research and its commercialization.
- 2015–2019 Awarded Merit-cum-Means (MCM) Scholarship at IIT Kharagpur.