

ROHAN WADHAWAN

+1(424)977-9469 • Los Angeles, CA • rwadhawan7@cs.ucla.edu • Portfolio • LinkedIn • Google Scholar

EDUCATION

University of California, Los Angeles (UCLA)	Los Angeles, CA
MS in Computer Science (GPA 3.96/4.0)	Sep 2024
Netaji Subhas Institute of Technology, University of Delhi	New Delhi, India
B.E. in Computer Engineering (GPA 8.94/10.00, Graduated with Distinction)	Jul 2020

SKILLS

Core Competencies: Natural Language Processing, Computer Vision, Multimodal & Generative AI, Prompt & Data Engineering
Languages & Tools: C++ • Python • SQL • Pytorch • HuggingFace • WandB • Lightning AI • Flask • MongoDB • Git • GCP • Linux

RESEARCH EXPERIENCE & PROJECTS

University of California Los Angeles Los Angeles, CA
Graduate Student Researcher, Advised by Prof. Violet Peng Jun 2023 - Present

- **ConTextual: Evaluating Context-Sensitive Text-Rich Visual Reasoning in Large Multimodal Models**
Established a novel benchmark comprising instructions designed explicitly to evaluate the context-sensitive reasoning of LMMs on the text and visual elements in text-rich images.
- Performed exhaustive quantitative (human & automatic evaluation like GPT4 eval) and qualitative analysis of **13 models**: closed source LMMs (**GPT-4V(ision), Gemini-Pro-Vision**), open-source LMMs (**LLaVA-1.5-13B, Instruct-Blip-2**, etc.) and augmented LLMs (**GPT4 + layout OCR + image caption**)
- **GPT-4V(ision)**, the best performing LMM, has a performance gap of **30.8%** to human performance.
- **TRiViS: Visual Instruction Tuning for Text-in-Image Comprehension**: Developed a dataset with complex instructions for text-rich images (~70K samples) across documents, websites, mobile apps, and simple OCR scenes.
- Utilized PEFT (LoRA) to fine-tune an LLM (mPlug-Owl), which surpassed the base model's performance in OCR and text-based VQA tasks by 16% and 7%, respectively, achieving a 58% preference rate in human evaluations. **Presented at the SoCAL NLP Symposium 2023.**

Indian Institute of Technology, Delhi, Neurocomputing Lab New Delhi, India
Research Assistant, Advised by Prof. Tapan K. Gandhi Jul 2020 - Jan 2021

- Architected a Human-inspired, Landmark-aware Ensemble Facial Expression Recognition CNN & increased SOTA accuracies on the CK+ & JAFFE datasets by **0.51%** & **5.34%** with only **3.28 MFLOPs** required for inference.
- Research work published in **IEEE Transactions on Artificial Intelligence, 2022, [IF: 4.9]**.
- Invented a Spatio-temporal (CNN-LSTM) deep learning pipeline for water stress phenotyping of Chickpea plant that achieved a ceiling level classification performance of **98.52%** on JG-62 & **97.78%** on Pusa-372 chickpea plant datasets.
- Research work published in **IEEE Transactions on Instrumentation and Measurement, 2021, [IF: 5.3]**.

WORK EXPERIENCE

University of California Los Angeles Los Angeles, CA
Teaching Associate, CS162-Natural Language Processing Jan - Mar 2024

- Spearheaded the creation and ongoing management of the **Fact or Flawed?** course project, aimed at evaluating the Factuality and Fairness of Large Language Models like **LLaMA-2, Phi-2, Gemma**, for **130 undergraduate students**.
- Facilitating regular hands-on sessions and contributing to the development of assignments and examinations.

Teaching Assistant, CS32, PIC10A & Break Through Tech AI Program Sep 2022 - Jun 2023

- Conducted weekly hands-on sessions on Data Structures and Object Orientation in **C++** for **150 undergraduate students**.
- Created an automation script that significantly reduced grading errors and workload of teaching assistants.
- Delivered workshops on machine learning and mentored **26 students belonging to underrepresented groups in tech**

Samsung R&D Institute, Visual Intelligence Team Bengaluru, India
Senior Software Engineer Jan 2021 - Jun 2022

- Enhanced tone & saturation of images rendered by the Expert Raw Application via integration of a new deep learning module to the camera pipeline in flagship smartphone models (S22).
- Commercialized **5** camera solutions like Video Stabilization, Hyperlapse, & Single Take on **> 10** smartphones and tablets like S22, A22e, M53 & Tab S8. **Skills used: C++, Python, Android, Version Control, Object Orientation**