Pranay Gupta

Email: pranaygu@andrew.cmu.edu | +1 412-626-9880 | Google Scholar | LinkedIn | Github | Homepage

Education

Carnegie Mellon University — Pittsburgh, PA
Master of Science in Robotics, GPA ; 4.08/4 — *Aug. 2022 – Aug 2024*

International Institute of Information Technology — Hyderabad, India
Bachelor and Master of Science in Computer Science, GPA ; 8.03/10 — *Aug. 2016 – Jul 2021*

Experience

Research Assistant — Oct 2022 – Present
The Robotics Institute, Carnegie Mellon University — *Pittsburgh, PA*
- Devised a counterfactual reasoning based approach to identify important objects in driving scenarios.
- Validated our approach with HOIST, a novel dataset with multi-modal sensor data for driving scenarios with human-annotated importance labels for vehicles and pedestrians. Published at IEEE RA-L.

Predoctoral Apprentice — May 2021 – July 2022
TCS research — *Delhi, India*
- Implemented 3-D CNN to approximate implicit functions for 3-D single view reconstruction (SVR).
- Employed an energy based out-of-distribution (OOD) detection classifier to increase robustness for SVR.

Undergraduate Research Assistant — June 2018 – June 2021
Center for Visual Information Technology(CVIT), IIIT-H — *Hyderabad, Telengana*
- Investigated the problem of skeleton based action recognition. Explored new frontiers by studying into-the-wild and out-of-context action recognition. Accepted at IJCV.
- Devised a VAE backed approach which learned syntactically aware embeddings for zero shot skeleton action recognition. Achieved SOTA results on the NTU-60 and 120 datasets. Accepted at 2021 IEEE ICIP.

Applied Scientist Intern — Jun. 2020 – Aug.2020
Amazon India — *Bengaluru, Karnataka*
- Leveraged a siamese network with an LLM to estimate semantic similarity between query and product description.
- Applied transfer learning for multilingual data. Fine-tuned models trained on English data with German data.

Google Summer of Code Intern — Jun 2018 – Aug 2018
Purr-Data — *Remote*
- Successfully updated purr-data's core and the external libraries from single precision float to double precision.

Projects

Leveraging VLMs for Zero-Shot, Personalization of Multi-Object Rearrangement
- Performed in-context learning with GPT4V to enable preference aligned household task planning.
- Demonstrated feasibility through successful implementation in a one-step table setting task.
- Presented at the Human-LLM interaction workshop at HRI-24

News-KVQA
- Curated a new large scale video question answering dataset (12k videos, 1 million QA pairs). Automated question generation using videos, subtitles and knowledge base facts.
- Proposed a multi-modal LLM based approach that processed visual, textual and factual data for question answering. Published at PAKDD-22.

3D Scene Reconstruction using Monocular Image
- Developed a pipeline that systematically performed object segmentation and pose estimation using YOLOv3 and cube RCNN and 3D reconstruction and localization of each detected object using pixelNERF and iNERF.

Improving PlanT
- Improved the results from the CORL 2022 paper "PlanT: Explainable Planning Transformers via Object-Level Representations" by ensuring consistency of the frame of reference across the inputs and using a history of states.

Distributed Attendance System
- Automatic Attendance System based on Face Detection and Recognition via Facenet. Distributed System, worked simultaneously with multiple cameras.

Bash Shell
- Implemented as a part of a course project in Operating Systems course. Developed a unix shell in C. Implemented basic functionalities like killing a process, input/output redirection, piping and signal handling

Technical Skills

Research Areas: Computer Vision, Multimodal LLMs, 3D Computer Vision, Autonomous Driving, NLP
Languages: Python, Matlab, C/C++, HTML/CSS
Developer Tools: Git, Vim, VS Code, AWS
Libraries: CARLA, Pytorch, Pytorch-3D, Open3D, Opencv, Scikit-Learn, Pandas, NumPy, Matplotlib