Keshav Gupta

+91 7087068290 | keshav.gupta@students.iiit.ac.in | keshav0306.github.io
keshavgupta06 | keshav0306

Synopsis

I am a Computer Science Engineer, and will be pursuing my **MS in Computer Science** from **University of California, San Diego**, starting the fall of 2025. I have done my **B.Tech in Computer Science** from the **International Institute of Information Technology, Hyderabad (IIIT-H)**, with an **honors research** degree in **Computer Vision**. My research is deeply aligned with leveraging computer vision to interpret **complex and unstructured environments**, building algorithms for modeling **autonomous driving agents** to facilitate safe navigation, and exploring **3D representations** for the same.

Research Experience

- **Center for Visual Information and Technology (CVIT) Lab** *June 2023 - May 2025*
  Honors - Advised by Dr. Ravi Kiran Hyderabad, India
  - Proposed a novel methodology for detecting triple riding and helmet violations on **unstructured, high-density Indian roads** using only a dashcam, i.e., a monocular RGB camera video. This is the **first** approach to address this problem from a **moving** camera, tackling the unique challenges posed by India's unstructured environments, such as high occlusion and ambiguity in rider-to-motorcycle correspondence.
  - Introduced a novel **cross-segmentation** algorithm and **a novel multiple object tracking (MOT) method**, and proposed a **comprehensive video dataset** that includes track-level information, such as bounding boxes and track IDs for rider-motorcycle pairs. This dataset is **unique** among other MOT datasets, and is made for testing tracking and detection capabilities in **unstructured settings.**
  - The work was accepted in **WACV (Winter Conference on Applications of Computer Vision) 2025.**

- **Robotics Research Center (RRC) Lab** *January 2024 - May 2025*
  Independent Study - Advised by Dr. K Madhava Krishna Hyderabad, India
  - Working on developing an **autonomous driving stack** that predicts control commands from RGB perspective images, focusing on detecting **multimodal freespace** segments for safe driving. The training data involves just the video sequences of an agent navigating in the presence of dynamic obstacles (cars and pedestrians), along with its ego trajectories. The challenge of the project is designing a model that could operate with this minimal annotated data, while still ensuring effective navigation and obstacle avoidance based **purely** on the **visual input**.
  - The work was accepted at **IROS (International Conference on Robots and Systems) 2025.**

- **Machine Learning Lab (MLL) Lab** *July 2024 - May 2025*
  Independent Study - Advised by Dr. Charu Sharma and Dr. Avinash Sharma Hyderabad, India
  - Using **Gaussian splatting** for 3D scene reconstruction and novel view synthesis where we are working on **interpretable** methods for compression in Gaussian Splatting, by utilizing local **repetitions** and **symmetries** within the scene to enable efficient storage compression.

- **Cognitive Science Lab** *December 2022 - May 2023*
  Research Assistant - Advised by Dr. Kavita Vemuri Hyderabad, India
  - Designed and built a **Hand Control** Device to assist **disabled** and **handicapped** patients in exercising hand movements, connected to a game for motivation. The device tracks real time progress and sends reports to doctors.

Education

- **International Institute of Information Technology** *August 2021 - May 2025*
  B.Tech in Computer Science (CSE) Hyderabad, India
  - GPA: 8.83/10.0

- **Guru Nanak Public School** *2021*
  High School Education Ludhiana, Punjab
  - 12th CBSE - 96.6%
  - JEE Mains All India Rank - 723, JEE Advanced All India Rank - 1196

Publications

C=Conference, J=Journal, P=Patent, S=In Submission, T=Thesis

[1] **DashCop: Automated E-Ticket Generation for Two-Wheeler Traffic Violations Using Dashcam Videos - Winter Conference on Applications of Computer Vision (WACV 2025) - Accepted** - We propose a novel methodology and working solution for detecting triple riding and helmet violations on unstructured, high-density Indian roads using a dashcam feed.

[2] **Diffusion-FS: Multimodal Free-Space Prediction via Diffusion for Autonomous Driving - International Conference on Robots and Systems (IROS) 2025 - Accepted** - We propose a novel self-supervised methodology for predicting freespace multimodal segments from a dataset of raw driving logs.

Skills

- **Programming Languages:** C, C++, CUDA, Python, Javascript, Bash, x86
- **Data Science & Machine Learning:** Pytorch, CUDA Programming
- **3D Softwares:** Open3D, Blender
- **Web Technologies:** NodeJS, ReactJS, ExpressJS
- **Database Systems:** MySQL, MongoDB

Extra Co-Curricular Activities

- **Literary CLub:** Coordinator
- **Dance Crew:** Video Editing Team Lead
- **Campus Canine Club:** Member
- **Passionate Runner:** Won 2nd prize in 10K - IIIT Sports Day

Scholarships

- **Scholarship of Rs. 2,50,000 ($3000):** By Aakash Institute in High School

Courses

- **Computer Vision, Mobile Robotics, Digital Image Processing, Linear Algebra, Statistical Methods in AI, Advanced Natural Language Processing, Computer Graphics, Data Structures and Algorithms, Operating Systems and Networks, Quantum Information and Computation**

Projects

- **Deep Learning and Neural Network Paper Implementations**
  Concepts: [Deep Learning, Object Detection, Neural Networks, Neural Machine Translation] [] [] [] []
  - Implemented the paper **"Visualizing the Loss Landscape of Neural Nets" by Li et al.** where we were able to plot and visualize the loss landscape of the network around the minima.
  - Implemented the paper **"Image-Space Modal Bases for Plausible Manipulation of Objects in Video" by Abe Davis et al.** where we were able to simulate the movement of any object response, to any virtual force, without knowledge of its material properties.
  - Extended on **"Dynamic 3D Gaussian Splatting" by Luiten et al.** for modeling small movements through a spring mass model.
  - Re-implemented **Polygon YOLO** (object detection with a generic bounding box) by extending the YOLOv5 repository.
  - Implemented the paper **"Neural Machine Translation" by Bahdanau et al.,** focusing on sequence-to-sequence models
- **Window Manager and Compositor for Linux**
  Concepts: [Operating Systems, Computer Graphics] []
  - I developed my own **Window Manager** and **Compositor** completely written in **C**, implementing the userspace **Linux graphics stack** from scratch.
  - Implemented the coordination and compositing of many **framebuffers** into one and sending them to the **GPU/CPU** for rendering from scratch.
  - Implemented various **rasterization algorithms**, built a basic **Raytracing application** on top of the window manager, implemented **Phong Shading** in it, and also created my own **font** for my own text editing applications.
  - Made a **terminal emulator** application for my own window manager.
- **2D and 3D Games**
  Concepts: [Computer Graphics] [] []
  - Created a **2D** Game - Jeptpack Joyride Clone in **OpenGL C++**
  - Created a **3D** Car Racing Game - in **OpenGL C++**
- **Operating Systems and Systems Programming**
  Concepts: [Operating Systems, Assembly Language, Multithreading] [] [] []
  - Implemented a **bootloader** in Intel assembly that parses the kernel's ELF file and loads it to the appropriate address as mentioned in the ELF file.
  - Extended the **MIT Operating System xv6** by adding new system calls and schedulers to the kernel.
  - Built a **UNIX shell** capable of executing any system commands.
  - Designed a **multithreaded** restaurant simulator.
- **Web Applications Development**
  Concepts: [Full Stack Development, Progressive Web Apps (PWA)]
  - Developed a **social media** web application where users can post, react, follow, comment, and make friends on the platform.
  - Worked on **IIIT's Virtual Labs Online Learning Platform**, converting a subset into a Progressive Web Application (PWA) and deploying it on the Android Play Store.