Keshav Gupta

+91 7087068290 | iamkeshav06@gmail.com | 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 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*
  Research Assistant - 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 involved 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 was in 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*
  Research Assistant - Advised by Dr. Charu Sharma and Dr. Avinash Sharma — Hyderabad, India
  - Using **Gaussian splatting** for 3D scene reconstruction and novel view synthesis where we worked on **interpretable** methods for compression in Gaussian Splatting, by utilizing local **repetitions** and **symmetries** within the scene to enable efficient storage compression. Work submitted to **AAAI (Association for the Advancement of Artificial Intelligence) 2026**.

- **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.

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)**
Deepti Rawat[1], **Keshav Gupta***, Aryamaan Basu Roy, Ravi Kiran Sarvadevabhatla*
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. We introduced a novel cross-segmentation algorithm, a novel multiple object tracking (MOT) method, and a comprehensive video dataset *Ridesafe-400*.

[2] **Diffusion-FS: Multimodal Free-Space Prediction via Diffusion for Autonomous Driving International Conference on Robots and Systems (IROS 2025)**
Keshav Gupta, Tejas Stephen Stanley, Pranjal Paul, Arun K. Singh, K. Madhava Krishna
We propose a novel self-supervised methodology for predicting freespace multimodal segments from a dataset of raw driving logs. We also introduce ContourDiff, a specialized diffusion-based architecture that denoises over contour points rather than relying on binary mask representations, enabling structured and interpretable freespace predictions.

---
[1]* denotes authors contributed equally.

SELECTED PROJECTS

- **Computer Vision Projects and Paper Implementations**
  Concepts: [2D and 3D Object Detection, Neural Networks]
  - 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. [⟳]
  - Implemented the paper **DETR3D** focusing on 3D object detection using transformers. [⟳]
  - Implemented the paper **Center Net** focusing on 2D object detection, by modelling an object through its center and corresponding offsets. [⟳]
  - Implemented the paper **"Learning Transferable Visual Models From Natural Language Supervision" aka CLIP** focusing on self supervised pretraining. The details are described in a blog. [☰]
  - Implemented **Monocular Depth Estimation** on NYUv2 dataset. The details are described in a blog. [☰]
  - Re-implemented **Polygon YOLO** (object detection with a generic bounding box) by extending the YOLOv5 repository. [⟳]

- **Robotics**
  Concepts: [Mobile Robotics, Reinforcement Learning, SLAM]
  - Implemented a non-linear weighted least squares optimization approach for **2D SLAM pose-graph**, improving robot trajectory estimation. [⟳]
  - Developed a feature-based **visual odometry pipeline** by estimating essential matrices with RANSAC. [⟳]
  - Implemented a Rubik's Cube Solver using **DQN**. [⟳]

- **Miscellaneous Deep Learning Projects**
  Concepts: [Deep Learning, Neural Machine Translation, LibTorch]
  - 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 **"Neural Machine Translation" by Bahdanau et al.,** focusing on sequence-to-sequence models. [⟳]
  - Implemented Distributed Training in PyTorch in **C++** using **LibTorch** API and **Open MPI**. [⟳]

- **Window Manager and Compositor for Linux**
  Concepts: [Operating Systems, Computer Graphics]
  - 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 multiple **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, Game Development]
  - Created a **2D** Game - Jeptpack Joyride Clone in **OpenGL C++** [⟳]
  - Created a **3D** Car Racing Game in **OpenGL C++** [⟳]
  - Created a short 3D animated film in **Blender**. [▶]

- **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** in C capable of executing standard system commands as in bash. [⟳]
  - 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.

SKILLS

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

COURSES

- Computer Vision, Mobile Robotics, Digital Image Processing, Linear Algebra, Statistical Methods in AI, Advanced Natural Language Processing, Reinforcement Learning, Computer Graphics, Distributed Systems, Data Structures and Algorithms, Operating Systems and Networks, GPU Programming, Optimization Methods, Machine Learning for Natural Sciences, Quantum Information and Computation

EDUCATION

- **University of California, San Diego**   *September 2025 -*
Masters in Computer Science (CSE)   San Diego, California
  ∘ GPA: NA

- **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
  ∘ JEE Mains All India Rank - 723, JEE Advanced All India Rank - 1196 out of 1,000,000 students

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