

I am a Computer Vision Engineer with experience in training object detection, segmentation, pose estimation models and developing new techniques for multi-view 3D reconstruction.

## Education

- **M.Sc. Computer Science (Research)**, Saarland University, Germany. (converted GPA: 9.12/10.0) **2021 - 2024**
- **B.Eng. Electronics and Communication**, B.M.S College of Engineering, India. (GPA: 9.2/10.0) **2014 – 2018**

## Technologies and Coursework

- Languages: Python, C++, Java, JS
- Libraries: OpenCV, Pytorch, Pytorch3D, Huggingface-transformers, Sklearn, Pandas, Numpy
- Courses: Machine Learning, Neural Networks, High-Level Computer Vision, Computer Graphics, 3D Object Representation and Reconstruction with Machine Learning, Reinforcement Learning etc.

## Work Experience

**Research Engineer** **Perspectiv Labs Pvt Ltd.** **July 2024 - Present**  
Bengaluru, India

- Developed a photogrammetry pipeline to reconstruct road surfaces from stereo videos and calculate pothole volume.
- Developed an object counting system for a variety of industrial screws and nuts.
  - Trained a detection model (YOLOX) on a synthetic dataset generated using Isaac Omniverse and fine-tuned it on real-world data.
  - Quantised and deployed the model to run Raspberry Pi AI HAT.

**Research Assistant** **DFKI (German Research Center for AI)** **2022 – Feb 2024**  
Saarland, Germany

- Developed a pipeline for 6DOF Pose estimation of Industrial objects.
  - Generated synthetic dataset for industrial objects using BlenderProc(3000 images for each object).
  - Trained YOLOv8 2D object detection model and obtained a test mAP(50-95) of 97%.
  - Trained ZerbaPose (CVPR2022) to predict 6DOF pose and got test average recall ADD(-S) of 78%.
  - Deployed as a docker container for real-time prediction for RGB-D input from Intel Realsense camera.
- Developed a pipeline for natural interaction with the humanoid robot Pepper.
  - Developed a custom Speech-Recognition pipeline using Voice Activity detection (VAD) and WhisperJax, using huggingface libraries.
  - Integrated this with a local quantised LLM (NeuralChat-7B) and text-to-speech module, to get conversation-style interaction.
- Developed a real-time compressed point-cloud streaming and rendering for Unity VR with less than 1s latency (at 50fps).

**Student Developer** **JdeRobot, Google Summer of Code** **Jun 2021 - Aug 2021**  
Remote

- Developed a web-based block-programming tool for Computer Vision and Robotic applications – that converts interconnected blocks of Python code (custom or pre-built) into a coherent multi-process application.
- Implemented around 20 building blocks in the tool for Sensors (using ROS topics), Control (like PID) and Image processing.
- Demonstrated the tool's usefulness by building a demo Drone application (Gazebo simulation) using these blocks.

- Developed custom front-end and backend features for multiple airline booking systems like Air Canada, Japan Airlines, Southwest etc.
- Decreased the overall product memory usage by 10% by fixing memory leaks in DOM.
- Involved in the initial design and prototyping of a new tool for converting XML schemas to Typescript classes.

## Projects

---

- **Master Thesis: Separable multi-view 3D Human-Object Reconstruction (3DV 2025) – Max Planck Institute for Informatics, Saarland**
  - Developed a novel markerless approach for high-quality separable 3D reconstruction of a human and an object while interacting, from multi-view RGB inputs; also supporting separable free-viewpoint novel view renderings.
  - Collected a dataset of human-object interaction videos for 3 identities (people) and 6 objects – each consisting of 120 views and 30 seconds.
  - Developed a preprocessing pipeline for scene segmentation using GroundingDINO and Segment Anything models.
  - Project page: <https://vcai.mpi-inf.mpg.de/projects/separable-recon/>
- **Automated Guided Vehicle Survey**
  - Developed a prototype for autonomous navigation of the MIR100 robot and detected the 3D position of pre-defined objects for scene surveying.
  - Trained Deep Object Pose (2021) for estimating 3D positions of different container-type objects and built a ROS pipeline with Unity VR simulation for demonstrating the use case of real-time digital twin.
  - Github: <https://github.com/Suhas-G/agv-survey>
- **RayTracer**
  - Developed a multi-threaded CPU-based raytracer from scratch in C++ as part of the Computer Graphics course.
  - It includes support for triangular meshes, BVH acceleration, different types of cameras (Perspective, Orthographic, Fisheye), lights (Point, Spot, Area), materials (Lambertian, Phong, Dielectric, Mirror) and simple textures and normal maps.
  - Github: <https://github.com/Suhas-G/RayTracer.git>
- **Bachelor Thesis: Signature Fraud Recognition**
  - Developed a multi-task model with a shared backbone to predict signature forgeries and the user it belongs to.
  - Trained the model on SigComp2009 and SigComp2011 datasets and obtained a forgery prediction accuracy of 96% and user prediction accuracy of 95%.
  - Published as ICNEWS conference proceedings: <https://ieeexplore.ieee.org/document/8903995>

## Other Activities

---

- **Google Summer of Code 2022 - Mentor** for VisualCircuit.
- Gave a **ROS World 2021 - Lightning talk** about VisualCircuit <https://vimeo.com/635607300#t=2667s>.
- **Event Co-ordinator in Phase Shift, 2016** - tech fest of B.M.S College of Engineering.
- Participated in **ITeach**: An initiative of the Rotaract club of B.M.S College of Engineering to teach kids in government schools in the campus vicinity.