Rigved Sanku

Advanced Robotics & AI Expert: Pioneering Solutions in Machine Learning, Computer Vision, and Robot Perception

+1 774-502-4227 | San Jose, CA | sankurigved@gmail.com | LinkedIn | GitHub

CAREER SUMMARY

Over the past 5 years I've gained extensive experience in deep learning models for **real-time computer vision tasks** like detection, tracking, 3D reconstruction, pose estimation, and sensor fusion. From designing robust CV pipelines to optimizing inference on edge/cloud platforms, I love turning research into real-world solutions. Proficient in **Python, C++, PyTorch, TensorFlow, and OpenCV**, I'm now actively seeking full-time opportunities as a computer vision engineer, where I can continue solving complex real-world problems.

AREAS OF EXPERTISE AND SKILLS

- **Core Competencies:** Image Processing | Object Detection and Tracking | Scene understanding | Activity Recognition | GAN | Instance and Semantic Segmentation | 2D/3D Pose Estimation | Vision-Language Model (VLMs) | Transformers | Diffusion Models | GANs | Multimodal Learning | Image Captioning | Reinforcement Learning | ADAS | Sensor Fusion | LiDARs | Edge AI | ADAS
- **Technical Skills Proficiency:** Expertise in Python, C, C++, MATLAB and ROS | Advanced utilization of PyTorch, TensorFlow, OpenCV, OpenAI Gym, and Hugging Face | CUDA, Git, Docker and AWS | On-device deep learning – TensorRT and ONNX

PROFESSIONAL EXPERIENCE

ADAS Computer Vision Intern | *Honda Research Institute, San Jose, USA* **Jan 2025 – Present**

- Developed a real-time **human intent prediction** system for autonomous vehicles using 2D and 3D body pose and head pose models achieving 95% accuracy, reduced inference latency to <30ms using **multi-threading**. Filed a **provisional US patent** (63/782,414**).**
- Optimized and deployed computer vision algorithms as **TensorRT** engines with FP16 precision and **ONNX** graph cleanup, achieving a 3x speed up in inference time and reducing GPU memory usage by 30%, enabling seamless **edge deployment.**

Computer Vision Intern | *MUSCO Vision, Iowa, USA* **May 2024 – Aug 2024**

- Engineered comprehensive crowd monitoring system using **object detection** and **tracking** models like YOLO, SORT and ResNet50 for facility entry/exit crowd counting. Implemented advanced tracking algorithms like ByteTrack, improving tracking accuracy by 15%
- Developed and integrated custom deep learning model for **Person Re-Identification** task, achieving 98.1% Rank-1 accuracy in real-time tracking. Created custom dataset generation pipeline for re-training, enhancing mean Average Precision (mAP) by 10%

Computer Vision Intern | *Indian Space Research Organisation (ISRO), India* **Dec 2022 – May 2023**

- Spearheaded the improvement of real-time **human action prediction** accuracy from 74% to 96%**.** Leveraged cutting-edge 2**D pose estimation** models like OpenPose for improved performance in dynamic environments, work published in **Springer 2024** [Link]
- Implemented Bi-Directional **LSTM** with **Attention Mechanism** and Ensemble Learning to capture key spatiotemporal features, resulting in a state-of-the-art model for action-predicting in live video data feeds, improving performance in dynamic environments.

Deep Learning Intern | *Dept. of Persons with Disabilities, India* **Dec 2021 – May 2022**

- Engineered an advanced audio-enabled navigation assistant for the visually impaired, integrating **Vision Transformers** and LiDAR for obstacle detection. Additionally, developed an efficient IMU-based fall detection system with **Sensor Fusion**. [Video] [Website]
- Attained 85% accuracy for **Multimodal Vision-Language tasks** like **Image Captioning** and **Visual Q&A** by leveraging BERT-based encoders and **LLM** decoders, optimized for real-time performance on edge devices like **Nvidia Jetson**.

RESEARCH EXPERIENCE

Computer Vision Research Assistant | *VIS Lab, Robotics Dept, WPI, USA* **Aug 2024 – Jan 2025**

- Designed a **Multimodal** framework for **3D LiDAR point cloud** shape completion, fusing 3D point cloud geometric and 2D image features to improve object **reconstruction** accuracy by 10% across diverse datasets, worked submitted to **IROS 2025** [Link] [Poster]
- Engineered a novel multi-level **cross-attention** mechanism integrating PointNet++ and ResNet-18 features for multimodal fusion.

Computer Vision Research Assistant | *Computer Science Dept, WPI, USA* **Sept 2023 – Jan 2024**

- Designed novel **transformer**-based architecture to enhance **object detection** in adverse weather for autonomous vehicles, achieving 4.423 mAP and 9.289 AP50 improvement on BDD100K dataset, work submitted to **ICCV 2025** [Link]
- Engineered information mixing between adverse and rich feature embeddings, by integrating SAM features with Faster R-CNN Conducted thorough ablation studies to optimize performance for real-world **ADAS** applications.

PATENTS & PUBLICATIONS

- Sanku, R. et al, **2025** "Driver Intent Prediction for Automatic Car Entry and Exit System," **Provisional Patent** Application #63/782,414, with Honda Research Institute using multimodal vision-based system to predict driver intent. (filed)
- Sanku, R. et al, **ICCV 2025** "Improving Object Detection in Adverse Weather Conditions for autonomous driving" [Link] (submitted)
- Sanku, R. et al, **IROS 2025** "FusionNet: Enhancing Point Cloud Reconstruction with Hierarchical Attentions." [Link] (submitted)
- Sanku, R. et al, **Springer 2024** "Real-Time Human Action Prediction using Pose Estimation with Attention-based LSTM Network." Signal, Image and Video Processing, **Springer Nature (Peer-reviewed Journal)** [Link] (published)

RELEVANT PROJECTS

Tesla Vision with Blender Visualization [Report] **March 2024 – May 2024**

- Engineered comprehensive ADAS computer vision stack integrating DETIC for **instance segmentation** for traffic and vehicle lights, YOLO3D for **6D-pose estimation**, UniDepth for **depth estimation**, Mask R-CNN for **lane detection**, RAFT **optical flow** for movement analysis, OSX for 3D human mesh, rendering all elements dynamically in Blender for visualization like **Tesla dashboard**

3D Scene Reconstruction, Structure from Motion & Neural Rendering [Code] [Report] **Jan 2024 – March 2024**

- Implemented 3D reconstruction pipeline, with Structure from Motion **(SfM)** and cutting-edge Neural Radiance Fields **(NeRF)**. Integrated **feature matching**, **camera pose estimation**, and **novel view synthesis** for high-fidelity scene modeling.

Quadrotor Waypoint Tracking with Deep Reinforcement Learning [Code] [Report] **Sept 2023 – Dec 2023**

- Implemented **Proximal Policy Optimization** (**PPO)** for quadrotor waypoint tracking, designing custom reward functions and simulation environment. Achieved efficient navigation using a simplified neural network, outperforming traditional PID controllers.

EDUCATION HISTORY

M.S. in Robotics and Computer Science, Worcester Polytechnic Institute, MA **Aug 2023 – May 2025**

CGPA: 3.83/4.0 | **Focus Areas:** Advanced Computer Vision, Generative AI, Reinforcement Learning

B. Tech in Mechatronics and Computer Science, National Institute of Technology, Trichy, India **July 2019 – May 2023**

CGPA: 8.47/10.0 | **Focus Areas:** Computer Vision, NLP, Deep Learning, Machine Learning, Linear Algebra and Calculus, Probability