Butchi Venkatesh Adari

butchivenkatesh.a@gmail.com +1(774)-670-7192 github.com/VenkateshRoshan linkedin.com/in/abven venkateshroshan.github.io

Technical Skills

Languages: Python, C++, Java, C, SQL, CUDA
Frameworks: PyTorch, TensorFlow, Scikit-learn, Hugging Face, LangChain, LangGraph, FastAPI, ROS, PostgreSQL, PySpark
Tools & Infra: Docker, Kubernetes, MLflow, DVC, Git, Prometheus, Grafana, AWS, GCP, Azure
ML Ops & Deployment: Model Deployment, CI/CD, Inference Optimization, Vector Databases (FAISS)
Course Knowledge: Motion Planning, Computer Vision, DL, LLM, VLLM, MLOps, Foundations of Robotics, Robot Dynamics

Experience

Graduate Researcher *Worcester, Massachusetts*
ELPIS LAB | *Worcester Polytechnic Institute* *Aug 2023 - May 2025*

- Conducted a comparative evaluation of Apple's Depth-Pro monocular depth estimation and applied distributed supervised fine-tuning (SFT) with custom loss tuning to reduce RMSE by 70%, enhancing depth accuracy for robotic grasping.
- Designed a novel Grasp Transformer architecture for joint depth, pose, and heatmap prediction of graspable locations, using LangSAM for object-centric segmentation.
- Achieved ±1–2 cm depth estimates and successfully grasped objects under 30 cm where RealSense depth returned no data by leveraging monocular predictions with 1 cm error.Maintained FP16 performance and a minimal compute footprint.
- Implemented and deployed an end-to-end PyTorch and ROS2 pipeline for real-time grasp inference at 20 FPS, exposing modular APIs for downstream control and testing.
- Developed hand-eye calibration routines, tf broadcasters, and ROS drivers to streamline perception-to-motion integration.

Machine Learning Engineer *Hyderabad, India*
Tata Consultancy Services *July 2021 - June 2023*

- Developed a real-time document intelligence pipeline using TrOCR and LayoutLM with Conditional Random Fields, processing 1000+ noisy scanned forms/hr for fraud-resilient parsing, verification, and structured information extraction.
- Attained 94% structured data accuracy and reduced manual data entry time by 40% through the automated extraction.
- Engineered a deep learning pipeline using YOLOv5 and DeepSORT to detect and track people in real-time from CCTV footage, achieving 25FPS throughput on edge GPUs with CUDA and ONNX + TensorRT optimization.
- Generated zone-wise foot-traffic heatmaps from the tracking outputs, driving a 30% improvement in store layout planning.
- Implemented a dwell-time estimation system using ROI-based tracking to pinpoint customer interest zones, boosting promotional-shelf conversion rates by 20%.

Projects

Agent based Web Data Extractor for RAG Systems | AI, LLMs, Web Scraping *Feb 2025 - Mar 2025*

- Engineered an autonomous web intelligence agent for RAG systems, improving data extraction accuracy by 35% through advanced filtering and structuring techniques, leveraging agentic multimodal web navigation and task simulation.
- Architected a scalable, multi-threaded crawler with content-aware extraction, boosting retrieval accuracy by 25% through real-time synthesis logic deployment.

Research Paper Analysis System with RAG Architecture and MLOps *Oct 2024 - Dec 2024*

- Orchestrated a research paper analysis system using RAG and rolled out on GCP Vertex AI with CI/CD pipelines.
- Integrated ChromaDB, T5 & GPT-2 into a scalable ML pipeline,achieving 1.3s latency on CPU for real-time document QA.

Real-Time Customer Support Chatbot | LLM, NLP, CI/CD, AWS SageMaker *Aug 2024 - Sep 2024*

- Launched a scalable LLM-powered chatbot using SageMaker for real-time inference and continuous fine-tuning via CI/CD.
- Established pipeline for automated deployment and continuous model improvement using AWS SageMaker and Lambda.

Image Captioning with Vision Transformer and GPT-2 | VLLM, NLP *May 2024 - Jun 2024*

- Trained and fine-tuned a ViT,GPT-2 pipeline for image captioning using PyTorch, achieving 90% semantic relevance.
- Hosted the model on Hugging Face Spaces and streamlined deployment pipelines with GitHub Actions and AWS.

High-Fidelity 3D Scene Reconstruction Using NeRF | Computer Vision *Mar 2024 - Apr 2024*

- Reconstructed 3D scenes from 2D images using NeRF, improving scene accuracy by 25% with positional encoding, tested against models without it.

Tesla Vision | Deep Learning, Computer Vision *Jan 2024 - Feb 2024*

- Simulated a 3D autonomous driving dashboard using YOLO3D for vehicle detection, a custom lane recognition model, and ZoeDepth for monocular depth perception in an interactive driving scene in Blender with motion-prediction and visualization.

Indoor Robot Navigation | Motion Planning, Perception *Sep 2023 - Dec 2023*

- Evaluated traditional planners (RRT, RRT*), RL-inspired methods for autonomous indoor navigation in Habitat Simulator.

Self-Driving Car | Deep Learning, Computer Vision *Jan 2021 - May 2021*

- Operationalized a lightweight object-detection model on Raspberry Pi 4 to detect traffic signs, pedestrians with 94% accuracy.
- Enabled real-time autonomous navigation in a miniature vehicle using vision-based path planning and collision avoidance.

Education

Worcester Polytechnic Institute *Aug 2023 - May 2025*
Masters in Robotics Engineering - GPA: 3.8/4.0

Anil Neerukonda Institute of Technology and Sciences *July 2017 - May 2021*
Bachelors in Computer Science and Engineering - GPA: 7.77/10