Daniel Vadranapu
17808 Sherman Way, Reseda, California, 91335
(716) 348-4581 | vadranapudaniel66@gmail.com | linkedin.com/in/daniel-vadranapu | danielvadranapuub.github.io/portfolio/

Summary

AI Engineer with 6+ years of experience developing multimodal systems, fine-tuning LLMs, and deploying deep learning models in production. Specializing in parameter-efficient training, generative AI, and real-time vision pipelines.

Work Experience

iDrive, Los Angeles, California, USA **March 2025 - Present**
AI Engineer - **Python, PyTorch, OpenCV, LLM, FastAPI**

- Collaborated on mimicking Comma.ai's perception and planning stack using open datasets and vision-based **lane detection models** integrated with LLM-based driver assistance.
- Developed real-time pipelines for camera-based **drivable area segmentation** and object detection, optimizing performance for in-vehicle deployment.
- Contributed to interactive driving assistance features by integrating LLMs to reason about road geometry, object context, and route planning.

Nexquared, Los Angeles, USA **Sep 2024 - Present**
Founder and AI Engineer - **Python, PostgreSQL, Ollama, LLMs**

- Launched a **job-matching platform** that uses LLMs to extract skills, visa, and clearance data from job descriptions and match candidates in near real-time.
- Built and deployed a multi-threaded metadata extractor and skill matcher using **Ollama** on consumer GPUs, processing 10,000+ job postings per day.
- Implemented a PostgreSQL backend with full audit tracking and candidate scoring logic, enabling scalable client-side usage with React/Next.js frontend.

University at Buffalo, New York, USA **August 2023 - December 2024**
Research Assistant - **Python, C++, Pytorch, HuggingFace**

- Designed a new approach using deep **cross-attention** frameworks that strengthened connections between images and text formats, resulting in 10% improvement in model.
- Developed a multi-modal architecture for **Visual Question Answering** (VQA) by integrating self-attention and guided attention mechanisms, leading to a 7% improvement in answer accuracy.
- Integrated **LiDAR and GPS** data into the Autoware pipeline, optimizing vehicle localization by increasing precision by 12 % and reducing drift error for NDT scan matcher.

Tata Insights and Quants, Bangalore, India **June 2021 - August 2023**
Deep Learning, Assistant Manager - **Python, Pytorch, Opencv, HuggingFace, AWS**

- Innovated real-time **license plate detection** technology that processed up to 20 vehicles per hour while maintaining a consistent accuracy level above 90%, reducing manual oversight requirements on parking enforcement operations.
- Launched a YOLOv8 model to **enhance safety protocols** on construction sites, achieving 98% accuracy in identifying individuals wearing protective equipment; this implementation led to a significant reduction in safety violations and incidents.
- Optimized workflow efficiency by deploying a two-stage **detection and classification pipeline** for the recognition of handwritten text on metal slabs, processing 1K slabs per hour
- Engineered a real-time surveillance solution to **identify graffiti on walls** while capturing the identity of individuals; enhanced community safety protocols increased detection accuracy by 10%, ensuring safer urban environments.
- **Generated synthetic datasets** for enhancing automatic number plate recognition and detecting instances of littering on roadways, using advanced image processing techniques, resulting in a 15% improvement in detection accuracy.
- Cultivated talent among a cohort exceeding **15 interns** over two years; guided innovative design processes leading to substantial advancements in computer vision solutions with notable impacts on team performance.

Flux Auto, Bangalore, India **June 2019 - June 2021**
Deep Learning Engineer - **Python, C++, Pytorch, Tensorflow, Cuda, ROS**

- Designed, trained, optimized, and tested object detector ConvNet for autonomous driving perception stack inspired by YoloV3 on the Berkley Deep-Drive data set and improved original YoloV3 MAP@75 at 38% .
- Built a **novel architecture** that can simultaneously detect objects and segments using a single backbone, reducing the computational complexity and memory requirements by 50 %.
- Created tools for corrective data sourcing and training pipeline for handling edge cases for **object detection** which improved the model performance and reduced the false positives by 25 %.

- Maximized prediction efficiency by integrating camera and radar data streams, resulting in accurate trajectory forecasts for neighboring vehicles near the ego vehicle with 75% reliability.
- Enhanced performance of the Yolov3 model through optimization using the **TensorRT** library, resulting in a significant increase in fps and deployed on the **Jetson AGX** platform.
- Conducted a thorough performance assessment of various Deep Learning frameworks by employing Convolutional Neural Networks (CNN) on diverse Nvidia GPU architectures.
- Analyzed and benchmarked different **Deep Learning frameworks** through rigorous tests with Convolutional Neural Networks across diverse Nvidia GPU setups; results were instrumental in enhancing processing speed by 17%.

Publications

- Daniel Vadranapu, *"An Old Dog with New Tricks: Lessons from Building an AV Testbed with Autoware"*, IEEE Intelligent Vehicles Symposium (MOST), 2025.
- Daniel Vadranapu, *"MEDVLAT: A Novel Vision Language Attention Transformer for Medical Visual Question Answering in Clinical Decision Support Systems"*, under review at Image and Vision Computing (IVC), Elsevier.

Projects

- **GANS:** Enhanced Latent Diffusion model performance by training the text encoder on enriched contextual data, significantly improving Image generation.
- **Path Planning**: Developed and implemented **RRT and A* algorithms**, reducing path planning computation time by 25% and improving route efficiency in autonomous simulations.
- **Gap Follow**: Enhanced vehicle path-following and obstacle avoidance strategies by applying **Pure Pursuit and Gap Follow methods**, achieving a 20% increase in navigation accuracy and enabling safe autonomous operation in dynamic environments.
- **Localization and Planning**: Leveraged AMCL localization algorithm to improve pose estimation accuracy by 30% in simulation, crucial for enhancing vehicle stability and control in real-time navigation.

Swaayatt Robots, India **April 2018 - August 2018**
Research Intern

- Conducted comprehensive research on Markov Decision Processes and Q-learning, followed by practical implementation of these models within a grid world environment, improving the agent's decision-making performance.
- Constructed a miniature robotic car equipped with a camera, IR sensors, wheel encoders, motor drivers, and a Raspberry Pi, allowing for keyboard-based control and operation.
- Curated a dataset for autonomous vehicles and streamlined annotation procedure by leveraging pre-existing models for automation.

Education

State University of New York at Buffalo, USA **August 2023 - December 2024**
Masters in Robotics

Indian Institute of Information Technology, Guwahati **August 2015 - May 2019**
Bachelor of Technology in Electronics & Communication Engineering

Skills

Python, C++, SQL, PyTorch, Tensorflow, Pandas, Keras, OpenCV, TensorRT, HuggingFace, Nltk, NLP, AI, Transformers, ML, MLOps, CI/CD, PySpark, NLP, LLM, Git, Docker, AWS, GCP, Kafka, FastApi