Akashkumar Jain

📍 Nuremberg, DE ✉ bhansaliakash84@gmail.com 📞 +49 17273 12394 in Akash Bhansali ○ ASBJ021

Profile

I am a motivated AI and Robotics Engineer pursuing an MSc in AI and Robotics at the University of Technology Nuremberg. Experienced in machine learning, computer vision, autonomous systems, and robotic simulation. Proficient in Python, C++, PyTorch, TensorFlow, and Huggingface Transformers with hands-on experience in LLM/VLM fine-tuning (LoRA, PEFT), and building model optimization pipelines and AI Agents.

Interested Areas: AI , Robotics, Large Language Model (LLMs), Computer Vision, Foundation Models, Generative AI, Convolutional Neural Networks/Vision Transformers/Diffusion Models, Vision Language Models, VLAs.

Skills

- **Programming Languages:** Python , C++ (Basic)

- **ML Frameworks and Tools:** PyTorch, TensorFlow, Huggingface, Transformers, Large Language Model (LLMs), Vision Language Model (VLMs), LoRA, PEFT, Quantization, Prompt Engineering, Pandas, NumPy, SciPy, scikit-learn, SQL, experiment design, statistical evaluation

- **Robotics & Simulations:** ROS, MuJoCo, GYM, URDF, RCSS

- **Computer Vision:** OpenCV, Faster-RCNN, Mask-RCNN, DETR, SAM2, Vision Transformer

- **Perception & Localization:** SLAM, Kalman Filters, TagSLAM, Sensor Fusion

- **Tools:** Git, Linux, Jupyter Notebook, VS Code, LaTex, Langchain, BabyAI, Ollama, SmolAgents, Docker, RunPod

- **Languages:** English (Fluent), German (A2-Limited working proficiency), Hindi (Native)

Education

MSc	**University of Technology Nuremberg, DE** **AI and Robotics**	Oct 2023 – Present
	• Advanced coursework and projects in AI, robotics, machine learning, Computer Vision, LLMs/VLMs, & Multimodal Foundation Models.	
BE	**KLE Technological University, IN** **Automation and Robotics**	Aug 2016 – June 2020
	• CGPA: 8.02/10 • Specialization in Automation and Robotics.	

Experience

Machine Learning Freelancer
<div align="right">Mumbai, India
Mar 2021 – Oct 2023</div>

- Developed and deployed ML-based applications tailored to client needs.
- Key Projects:
  - Pickleball Tracking using Computer Vision and object detection
  - Eye-Gaze Based Cursor Control for Assistive HCI.
  - Created Stop Sign Detection with CO2 Monitoring, integrating IoT and ROS.
  - Robot Simulation for pick-and-place using PyBullet and OpenAI Gym
  - **Tools:** Python, Pytorch, Tensorflow, RaspberryPi, OAK-D camera, ROS, HCI, Arduino, CO2 sensor, MuJoCO, GYM, PyBullet

Machine Learning Engineer, Nautilus Hearing Solutions — Hubli, India
Sep 2021 – Sep 2022
- Developed CNN models for diagnosing ear diseases from auditory data.
- Integrated ML pipelines into Android platforms, enhancing diagnostic accessibility.

Research Intern, DRDO (CAIR) — Bangalore, India
Feb 2020 – May 2020
- Contributed to Fiducial Markers based SLAM (TagSLAM) project in collaboration with senior researchers.
- Worked with ROS and localization frameworks for map generation and pose estimation.

Projects (Master's)

Machine Learning
- **Object Detection Models:** Implemented Faster-RCNN, Mask RCNN, and DETR for visual object recognition..
  Tools: Python, PyTorch, OpenCV, Jupyter Notebook
- **Autonomous Driving:** Trained ML models on collected sensor and image data to simulate end-to-end self-driving capabilities in a virtual environment (Udacity simulator).
  Tools: Python, PyTorch, OpenCV, Jupyter Notebook, Udacity Simulator
- **Generative Hyperparameter Optimization:** Improving training efficiency and performance of models by exploring generative models for hyperparameter optimization.
  Tools: Python, Diffusion models, pytorch

Computer Vision
- **Captcha Solver:** Developed an end-to-end solution for recognizing alphanumeric CAPTCHAs with variable-length sequences, multiple degradation types, and challenging distortions.
- **Tools:** Python, Pillow, Data Augmentation, Data Generation, ResNet, CNN, Transformer Encoder and Parallel slot query decoder, PyTorch, hydra

Robotics
- **Playing Jenga with Franka:** Created a JengaEnv wrapper for gym, Programmed the Franka Emika 3 (FR3) robot to manipulate Jenga tower using precision actions in both simulation and real-world settings. Built a pipeline to collect data using simulation in MuJoCo.
- **Tools:** Python, MuJoCo, GYM, RCSS, CadQuery, XML, URDF

Large Language Models
- **UTN Chatbot:** Fine-tuned an LLM to answer student FAQs using LoRA and PEFT techniques and evaluated response accuracy and compiled technical report.
- **Tools:** Python, Huggingface, Transformers, PyTorch, LoRA, PEFT, Prompt Engineering

Multimodal Foundation Models
- **Optimized Tarsier:** Applied post-training quantization (INT8 and INT4) and reduced inference latency and memory footprint.
  Tools: Python, Tarsier2, Huggingface, Transformers, Bitsandbytes
- **Dash Grid:** Created and evaluated an LLM based agent to solve grid-world environments using the BabyAI platform, leveraging prompt engineering strategies.
  Tools: Python, BabyAI, Qwen2.5-7B-Instruct, gpt-4o-mini, Prompt Engineering

Personal
- **Video Background Remover:** Built a tool using SAM2 (Segment Anything Model) to remove background in videos.
  Tools: Python, SAM2, OpenCV, Fine-tunning, API Hosting, RunPod, Celery