# Srijay KOLVEKAR

## Machine Learning Engineer

linkedin.com/in/srijayjk github.com/srijayjk
+49 1764 3255 739 srijayjk@gmail.com
Munich, Germany


An enthusiastic Machine Learning Engineer passionate about solving engineering problems. Expertise includes object detection, tracking, classification, and Embedded AI using 2D and 3D data. possesses a solid understanding of the fundamentals of machine learning, **embedded AI**, signal processing, and adheres to clean programming principles to ensure the development of robust, efficient and production-grade solutions with utmost priority to **AI safety** and **explainable** AI models.

## Skills

| | |
|---|---|
| Programming | C, RUST, C++, Python, Matlab |
| ML Frameworks | Pytorch, Sklearn, Tensorflow, Keras |
| ML Operations | ML Flow, Hydra, Docker, Docker Compose |
| ML Deployment | FASTAPI, FLASK, AWS, Lambda, EC2, SageMaker |
| Embedded AI | ONNX,TensorRT, Jetson, RTOS, QNX, snapdragon q-drive |

## Experience

**present – Feb 2023** — Machine Learning Engineer, MAGNA ELECTRONICS, Munich

- Engineered advanced 3D perception models for auto labeling using VoxelNext to reduce labeling time by **40% to 40 min per file**.
- Spearheaded the design and implementation of an **end-to-end** ML framework for radar object and environment classification, managing the entire lifecycle from data ingestion, extensive feature engineering to SoC deployment for developing **trustworthy** and **Explainable AI** model.
- Deployed optimized light-weight deep learning models on embedded System-on-Chips (SoCs) TI C66 and Arm R5. Achieved a **30% reduction** in model size and a **10% improvement** in runtime performance **(2ms for 30 objects)** with Knowledge distillation and low-rank adaption.
- Deployed complex DL models on Jetson and **Snapdragon Ride platforms** within a QNX real-time operating system.
- Developed and applied **Bayesian modeling** techniques to **quantify uncertainty** in object classification, providing more robust and interpretable predictions.

Radar | Lidar | Embedded AI | Deployment | Quantization | Data Analysis | DL Accelerator | CI/CD | ROS2

**Nov 2022 – Sept 2021** — Robotics Machine Learning (Internship and Work Student) , PHENO-INSPECT GMBH, Stuttgart

- Developed **advanced computer vision models, for semantic segmentation, object, and keypoint detection**, for phenotyping solutions such as weed and plant detection.
- Deployed light-weight ML models on Jetson nano to detect weeds. Did post training quantization and pruning to reduce the memory footprint by 20% while achieving 24 FPS.

CNN | OpenCV | Semantic Segmentation | Key-Point Detection | ML-Ops | Object Detection | Computer Vision

**April 2021 – Mar 2020** — Embedded Programmer - Work Student , MEANWAVE GMBH, Stuttgart

- Create POC on inter-processor communication to send data packet for Xilinx Zynq-700 using Rust on OpenAMP protocol.

embedded system | ZYNQ | RUST

## Education

**Oct 2022 – Oct 2019** — MSc Electrical Engineering, UNIVERSITY OF STUTTGART, Stuttgart

- Mathematics, Detection and Pattern Recognition, DL, ML, Computer Vision, Problem Solving

ML | Software Engineering | Software Architecture Design | Presentation | Communication Skills

## Language

English (c1) ●●●●●
German (B1) ●●●○○

**Project**: LLM Agent Development, Fine Tuning and Quantization (LoRA) 2025

GitHub Link

Developed multiple AI agents, including a data analysis agent and a weather application agent. Focused on prompt engineering and creating custom tools for Large Language Models (LLMs) using frameworks like LangChain.

AI Agents | LangChain | Prompt Engineering | TinyLlama

**Internal Patent Draft**: Machine Learning Based Ego Motion Estimation 2024

-

Received a internal innovation award for the work on ego motion estimation using ML. Developed novel architecture for using point cloud data to estimation ego motion.

**Project**: AI-Assisted Semi-Automatic Labeling of Eye Contact Data Using Contrastive Learning 2022

Thesis

Created a semi-automatic labeling method for eye contact data using contrastive learning and active learning. Enhanced feature space embeddings by distinguishing similar and dissimilar eye contact clusters.

Contrastive Learning | Pose Estimation | Key-Point Detection | Active Learning